



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Instituto de Ciencias Matemáticas

“ANÁLISIS ESTADÍSTICO DEL NIVEL DE CONOCIMIENTOS DE
MATEMÁTICAS Y LENGUAJE, DE LOS ALUMNOS DEL DÉCIMO
AÑO DE EDUCACIÓN BÁSICA, DE LOS COLEGIOS FISCALES
RURALES DE GUAYAQUIL”

TESIS DE GRADO

Previa a la obtención del título de:

INGENIERO EN ESTADÍSTICA INFORMATICA

Presentado por:

LOURDES DEL ROCÍO BENÍTEZ ESPINOZA

GUAYAQUIL-ECUADOR

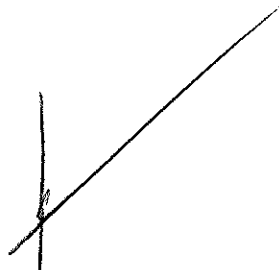
AÑO

2001

AGRADECIMIENTO

Al ingeniero Gaudencio Zurita Herrera, por su invaluable ayuda; a la familia Ronquillo-Franco; y a todas las personas que de uno u otro modo ayudaron a la realización de esta tesis.

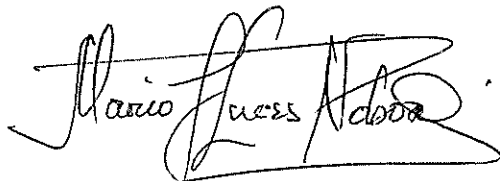
TRIBUNAL DE GRADUACION



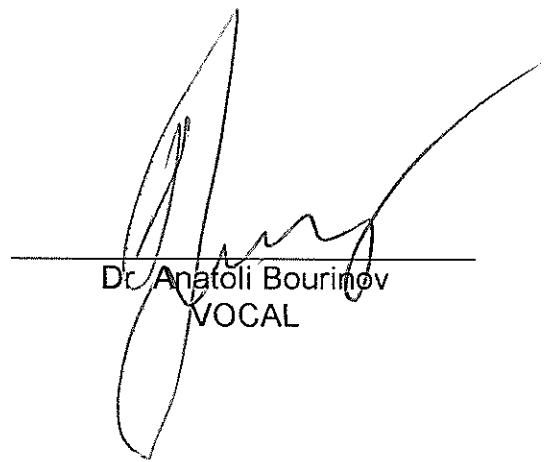
Ing. Felix Ramirez Cruz
DIRECTOR DEL ICM



Ing. Gaudencio Zurita Herrera
DIRECTOR DE TESIS



Ing. Mario Luces Noboa
VOCAL



Dr. Anatoli Bourinov
VOCAL

DECLARACION EXPRESA

"La responsabilidad del contenido de esta Tesis de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOS POLITECNICA DEL LITORAL"

(Reglamento de graduación de la ESPOL)



Lourdes Benítez Espinoza

RESUMEN

La investigación realizada en esta tesis es un análisis estadístico del nivel de conocimiento de matemáticas y lenguaje de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil en el año 2000. Este estudio forma parte de un conjunto de tesis que se desarrollaron, para analizar el nivel de conocimiento en las materias antes mencionadas, de estudiantes de séptimo, décimo y último año de bachillerato, en las zonas urbana y rural, en establecimientos educativos fiscales y particulares.

En el primer capítulo de esta tesis se realiza una revisión de los hechos más destacados en la historia de la educación y desarrollo del Ecuador, en el cual se resaltan los acontecimientos que se presentaron fundamentalmente desde la época de la colonización. Además se presentan datos estadísticos de la situación de la educación en el país en la década pasada.

En el segundo capítulo se plantean los temas y subtemas que forman parte de las pruebas de matemáticas y lenguaje, con las cuales se midió el nivel de conocimiento de los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil. Además se describe y codifica cada variable aleatoria que se utilizará para la evaluación.

En el tercer capítulo se realiza el análisis estadístico de cada una de las variables aleatorias, con el objetivo de conocer el comportamiento y las características más significativas de estas variables estudiadas. En éste análisis se determina los parámetros poblacionales, así como los histogramas y diagramas de cajas de cada variable aleatoria.

En el capítulo cuatro se realiza el análisis multivariado de las características investigadas. El cual contiene el análisis de correlación lineal para determinar la dependencia lineal entre las variables, el análisis de contingencia por medio del cual se determina la dependencia lineal o no, que existe entre las características de investigación, además de otras técnicas multivaridas que se desarrollaron como el análisis de componentes principales y el análisis de correlación canónica.

INDICE GENERAL

RESUMEN.....	II
INDICE GENERAL	III
INDICE DE GRAFICOS.....	IV
INDICE DE TABLAS	V
INTRODUCCIÓN.....	6
Capítulo I Educación y desarrollo en el Ecuador	8
1.1 Antecedentes	8
1.2 Educación y desarrollo: definiciones y generalidades.....	10
1.3 Historia y cultura de los primeros habitantes del Ecuador	11
1.4 La educación y cultura en tiempos de la Colonia.....	13
1.5 Desarrollo y educación del Ecuador durante la Gran Colombia	18
1.6 La educación durante la vida republicana del Ecuador.....	21
1.6.1 Primera Etapa	22
1.6.2 Segunda Etapa	26
1.6.3 Tercera Etapa	28
1.6.4 Cuarta Etapa	28
1.6.5 Quinta Etapa	31
1.7 El sistema educativo Ecuatoriano	35
1.8 Educación Rural en el Ecuador	38
1.9 Datos estadísticos sobre la educación en el Ecuador.....	39
Capítulo II Diseño del cuestionario y codificación de variables.....	63
2.1 Introducción	63
2.2 Definiciones	64
2.3 Marco censal	65
2.4 Diseño del cuestionario	69
2.4.1 Temas y subtemas para la evaluación.....	69
2.5 Definición de variables	72
2.5.1 Variables generales	73
2.5.2 Variables de la prueba de lenguaje	74
2.5.3 Variables de la prueba de matemáticas	74
Capítulo III Análisis Estadístico Univariado.....	99
3.1 Introducción	99
3.2 Análisis univariado de las variables generales	99
3.2.1 Variable aleatoria: Edad	100
3.2.2 Variable aleatoria sexo	104
3.3 Análisis univariado de las variables aleatorias correspondientes a la prueba de matemáticas	106
3.3.1 Variable aleatoria suma de números enteros	106
3.3.2 Variable aleatoria suma de números racionales	109
3.3.3 Variable aleatoria resta de números enteros	111

3.3.4	Variable aleatoria resta de fracciones	114
3.3.5	Variable aleatoria multiplicación de enteros	116
3.3.6	Variable aleatoria multiplicación de fracciones	119
3.3.7	Variable aleatoria división de enteros	122
3.3.8	Variable aleatoria división de fracciones	124
3.3.9	Variable aleatoria definiciones de números	126
3.3.10	Variable aleatoria desigualdad	128
3.3.11	Variable aleatoria relación de orden.....	131
3.3.12	Variable aleatoria potenciación.....	134
3.3.13	Variable aleatoria divisibilidad	137
3.3.14	Variable aleatoria ejercicio de proporcionalidad.....	140
3.3.15	Variable aleatoria ejercicio de interés simple.....	142
3.3.16	Variable aleatoria ejercicio de regla de tres	145
3.3.17	Variable aleatoria sistema métrico.....	148
3.3.18	Variable aleatoria operaciones de conjuntos	150
3.3.19	Variable aleatoria proposiciones lógicas	153
3.3.20	Variable aleatoria evaluación de funciones	156
3.3.21	Variable aleatoria perímetro del cuadrado	159
3.3.22	Variable aleatoria área del triángulo.....	162
3.3.23	Variable aleatoria área del círculo	164
3.3.24	Variable aleatoria teorema de pitágoras	167
3.3.25	Variable aleatoria factorización de dos polinomios	169
3.3.26	Variable aleatoria factorización de un polinomio.....	172
3.3.27	Variable aleatoria ecuación con una incógnita	175
3.3.28	Variable aleatoria sistema de dos ecuaciones con dos incógnitas	177
3.3.29	Variable aleatoria probabilidad	179
3.3.30	Variable aleatoria estadística.....	182
3.3.31	Variable aleatoria calificación de matemáticas	185
3.4	Análisis univariado de las variables de lenguaje.....	188
3.4.1	Variable aleatoria identificación de diptongos.....	188
3.4.2	Variable aleatoria identificación de triptongos	191
3.4.3	Variable aleatoria identificación de hiatos	194
3.4.4	Variable aleatoria identificación de palabras agudas	197
3.4.5	Variable aleatoria identificación de palabras graves	200
3.4.6	Variable aleatoria identificación de palabras esdrújulas	203
3.4.7	Variable aleatoria identificación de palabras sobreesdrújulas	206
3.4.8	Variable aleatoria sentido de la oración.....	209
3.4.9	Variable aleatoria identificación de sinónimas	212
3.4.10	Variable aleatoria identificación de antónimas	215
3.4.11	Variable aleatoria identificación de palabras aumentativas	218
3.4.12	Variable aleatoria identificación de palabras diminutivas	220
3.4.13	Variable aleatoria identificación de palabras despectivas	223
3.4.14	Variable aleatoria identificación de simples nombres	225

3.4.15 Variable aleatoria vocabulario	227
3.4.16 Variable aleatoria conjugación de verbos	230
3.4.17 Variable aleatoria identificación del sujeto	233
3.4.18 Variable aleatoria identificación del predicado	236
3.4.19 Variable aleatoria identificación del sustantivo	239
3.4.20 Variable aleatoria identificación del verbo	241
3.4.21 Variable aleatoria identificación de frases	244
3.4.22 Variable aleatoria identificación de oraciones	247
3.4.23 Variable aleatoria clasificación de oraciones	250
3.4.24 Variable aleatoria corrección de sintaxis	253
3.4.25 Variable aleatoria palabras tildadas correctamente	256
3.4.26 Variable aleatoria lectura comprensiva.....	259
3.4.27 Variable aleatoria calificación de lenguaje	262
3.4.28 Variable aleatoria calificación de total	265
3.5 Determinación las preguntas más fáciles y más difíciles de las pruebas de matemáticas y lenguaje	268
Capítulo IV Análisis Multivariado	275
4.1 Introducción	275
4.2 Definiciones	276
4.3 Análisis de la matriz de correlación	280
4.3.1 Análisis de los mayores coeficientes de correlación.....	280
4.3.2 Análisis de los menores coeficientes de correlación	285
4.4 Análisis de contingencia	291
4.4.1 Análisis de contingencia entre las variables edad y calificación de lenguaje.....	295
4.4.2 Pruebas entre las variables sexo y calificación de lenguaje	297
4.4.3 Pruebas entre las variables ubicación del establecimiento educativo y calificación de lenguaje.....	298
4.4.4 Pruebas entre las variables jornada de trabajo del establecimiento educativo y calificación de lenguaje	300
4.4.5 Pruebas entre las variables edad y calificación de matemáticas ..	302
4.4.6 Pruebas entre las variables sexo y calificación de matemáticas ...	303
4.4.7 Pruebas entre las variables ubicación del establecimiento educativo y calificación de matemáticas	305
4.4.8 Pruebas entre las variables jornadas del establecimiento educativo y calificación de matemáticas	307
4.4.9 Pruebas entre las variables calificación de lenguaje y calificación de matemáticas	308
4.4.10 Resultados del análisis de contingencia entre otras variables aleatorias	310
4.5 Análisis de varianza	312

4.5.1	Análisis de varianza de la variable aleatoria calificación de lenguaje con respecto a la ubicación de los planteles educativos fiscales rurales del cantón Guayaquil.	320
4.5.2	Análisis de varianza de la variable aleatoria calificación de lenguaje con respecto a la jornada de los planteles educativos fiscales rurales del cantón Guayaquil.	322
4.5.3	Análisis de varianza de la variable aleatoria calificación de matemáticas con respecto al factor ubicación.	325
4.5.4	Análisis de varianza de la variable aleatoria calificación de matemáticas con respecto al factor jornada de trabajo de los colegios fiscales rurales del cantón Guayaquil.	328
4.5.5	Análisis de varianza de la variable aleatoria calificación total con respecto al factor ubicación de los colegios fiscales rurales del cantón guayaquil.	331
4.5.6	Análisis de varianza de la variable aleatoria calificación total con respecto al factor jornada de los colegios fiscales rurales del cantón Guayaquil.	333
4.6	Análisis de componentes principales	336
4.6.1	Determinación de las componentes principales	337
4.6.2	Construcción de componentes principales aplicadas a las variables originales	340
4.6.3	Construcción de componentes principales aplicadas a las variables estandarizadas	343
4.6.4	Rotación de las componentes principales obtenidas en base de las variables aleatorias estandarizadas	348
4.6.5	Rotulación de las componentes principales aplicadas a las variables aleatorias estandarizadas y rotadas	351
4.6.6	Aplicación de la prueba de Bartlett	364
4.7	Análisis de correlación canónica.....	366
4.7.1	Determinación de las variables canónicas	369
4.7.2	Cálculo de las correlaciones canónicas para los conjuntos de variables de la prueba de lenguaje y de la prueba de matemáticas	373
	CONCLUSIONES.....	378
	RECOMENDACIONES	392
	ANEXOS	
	BIBLIOGRAFIA	

Indice de Tablas

	Pag.
Tabla I Cronología de la fundación de las universidades y escuelas politécnicas estatales en el Ecuador hasta 1999.....	19
Tabla II Cronología de la fundación de las universidades y escuelas politécnicas particulares que reciben subvención del estado hasta 1997	29
Tabla III Porcentaje de analfabetismo en la población de 15 años y más.....	33
Tabla IV Cronología de la fundación de las universidades y escuelas politécnicas particulares que no reciben subvención del estado hasta 2000.....	34
Tabla V Porcentaje del presupuesto del estado destinado a la educación	40
Tabla VI Total de planteles por nivel en el Ecuador	43
Tabla VII Total de profesores por nivel en el Ecuador	44
Tabla VIII Total de alumnos por nivel en el Ecuador.....	45
Tabla IX Planteles por sostenimiento y por zona en el Ecuador	47
Tabla X Profesores por sostenimiento y por zona en el Ecuador	49
Tabla XI Alumnos por sostenimiento y por zona en el Ecuador.....	50
Tabla XII Planteles por sostenimiento en el Ecuador	52
Tabla XIII Causas de deserción en el Ecuador	54
Tabla XIV Alumnos por sostenimiento en el Ecuador	55
Tabla XV Causas de deserción en la zona urbana de todo el país	56
Tabla XV Causas de deserción en la zona rural de todo el país.....	58
Tabla XVII Causas de no promoción en el nivel medio de todo el país	59
Tabla XVIII Causas de no promoción en el nivel medio en la zona urbana de todo el país	61
Tabla XIX Causas de no promoción en el nivel medio en la zona rural de todo el país	62
Tabla XX Ubicaión y número de alumnos de los colegios fiscales del cantón Guayaquil.....	65
Tabla XXI Clasificación por género del alumnado de los colegios fiscales rurales del cantón Guayaquil.....	67
Tabla XXII Clasificación por actividad del personal que labora en los colegios fiscales rurales del cantón Guayaquil	68
Tabla XXIII Temas y subtemas para la evaluación de matemáticas	70
Tabla XXIV Temas y subtemas para la evaluación de lenguaje.....	71
Tabla XXV Parámetros poblacionales de la variable aleatoria edad de los estudiantes	100
Tabla XXVI Frecuencias de la variable aleatoria edad de los estudiantes	102
Tabla XXVII Parámetros poblacionales de la variable aleatoria sexo.....	104
Tabla XXVIII Parámetros poblacionales de la variable aleatoria suma de números enteros	107
Tabla XXIX Parámetros poblacionales de la variable aleatoria suma de números racionales.....	109
Tabla XXX Parámetros poblacionales de la variable aleatoria resta de números enteros ..	111

Tabla XXXI Frecuencias de la variable aleatoria resta de números enteros ..	113
Tabla XXXII Parámetros poblacionales de la variable aleatoria resta de números racionales	115
Tabla XXXIII Parámetros poblacionales de la variable aleatoria multiplicación de números enteros	117
Tabla XXXIV Frecuencias de la variable aleatoria multiplicación de números racionales	119
Tabla XXXV Parámetros poblacionales de la variable aleatoria multiplicación de números racionales	120
Tabla XXXVI Parámetros poblacionales de la variable aleatoria división de números enteros	122
Tabla XXXVII Parámetros poblacionales de la variable aleatoria división de números racionales	125
Tabla XXXVIII Parámetros poblacionales de la variable aleatoria definiciones de números	126
Tabla XXXIX Frecuencias de la variable aleatoria definiciones de números ...	128
Tabla XL Parámetros poblacionales de la variable aleatoria desigualdad	129
Tabla XLI Frecuencias de la variable aleatoria desigualdad	131
Tabla XLII Parámetros poblacionales de la variable relaciones de orden	132
Tabla XLIII Frecuencias de la variable aleatoria relaciones de orden	134
Tabla XLIV Parámetros poblacionales de la variable aleatoria potenciación y radicación	135
Tabla XLV Frecuencias de la variable aleatoria potenciación y radicación	136
Tabla XLVI Parámetros poblacionales de la variable aleatoria divisibilidad	137
Tabla XLVII Frecuencias de la variable aleatoria divisibilidad	139
Tabla XLVIII Parámetros poblacionales de la variable aleatoria ejercicio de proporcionalidad	140
Tabla IL Frecuencias de la variable aleatoria ejercicio de proporcionalidad	141
Tabla L Parámetros poblacionales de la variable aleatoria ejercicio de de interés simple	143
Tabla LI Frecuencias de la variable aleatoria ejercicio de interés simple	144
Tabla LII Parámetros poblacionales de la variable aleatoria ejercicio de regla de tres	146
Tabla LIII Frecuencias de la variable aleatoria ejercicio de regla de tres	147
Tabla LIV Parámetros poblacionales de la variable aleatoria ejercicio sistema métrico	148
Tabla LV Frecuencias de la variable aleatoria sistema métrico	150
Tabla LVI Parámetros poblacionales de la variable aleatoria operaciones de conjuntos	151
Tabla LVII Frecuencias de la variable aleatoria operaciones de conjuntos	153
Tabla LVIII Parámetros poblacionales de la variable aleatoria proposiciones lógicas	154
Tabla LIX Frecuencias de la variable aleatoria proposiciones lógicas	155

Tabla LX Parámetros poblacionales de la variable aleatoria evaluación de funciones	157
Tabla LXI Frecuencias de la variable aleatoria evaluación de funciones	158
Tabla LXII Parámetros poblacionales de la variable aleatoria perímetro del cuadrado	160
Tabla LXIII Frecuencias de la variable aleatoria perímetro del cuadrado.....	161
Tabla LXIV Parámetros poblacionales de la variable aleatoria área del triángulo	163
Tabla LXV Parámetros poblacionales de la variable aleatoria área del círculo	164
Tabla LXVI Frecuencias de la variable aleatoria área del círculo	166
Tabla LXVII Parámetros poblacionales de la variable aleatoria teorema de Pitágoras	167
Tabla LXVIII Frecuencias de la variable aleatoria teorema de Pitágoras	169
Tabla LXIX Parámetros poblacionales de la variable aleatoria factorización de dos polinomios	170
Tabla LXX Frecuencias de la variable aleatoria factorización de dos polinomios	171
Tabla LXXI Parámetros poblacionales de la variable aleatoria factorización de un polinomio.....	172
Tabla LXXII Parámetros poblacionales de la variable aleatoria ecuación de una incógnita.....	175
Tabla LXXIII Parámetros poblacionales de la variable aleatoria sistema de dos ecuaciones con dos incógnitas	177
Tabla LXXIV Frecuencias de la variable aleatoria sistema de dos ecuaciones con dos incógnitas	179
Tabla LXXV Parámetros poblacionales de la variable aleatoria probabilidad ..	180
Tabla LXXVI Frecuencias de la variable aleatoria probabilidad	180
Tabla LXXVII Parámetros poblacionales de la variable aleatoria estadística ..	183
Tabla LXXVIII Frecuencias de la variable aleatoria estadística.....	184
Tabla LXXIX Parámetros poblacionales de la variable calificación de matemáticas	185
Tabla LXXX Parámetros poblacionales de la variable aleatoria identificación de diptongos	188
Tabla LXXXI Frecuencias de la variable aleatoria identificación de diptongos	191
Tabla LXXXII Parámetros poblacionales de la variable aleatoria identificación de triptongos.....	192
Tabla LXXXIII Frecuencias de la variable aleatoria identificación de triptongos	194
Tabla LXXXIV Parámetros poblacionales de la variable aleatoria identificación de hiatos.....	194
Tabla LXXXV Frecuencias de la variable aleatoria identificación de hiatos	196
Tabla LXXXVI Parámetros poblacionales de la variable aleatoria identificación de palabras agudas.....	198

Tabla LXXXVII Frecuencias de la variable aleatoria identificación de palabras agudas	200
Tabla LXXXVIII Parámetros poblacionales de la variable aleatoria identificación de palabras graves	201
Tabla LXXXIX Frecuencias de la variable aleatoria identificación de palabras graves	203
Tabla XC Parámetros poblacionales de la variable aleatoria identificación de palabras esdrújulas	204
Tabla XCI Frecuencias de la variable aleatoria identificación de palabras esdrújulas	205
Tabla XCII Parámetros poblacionales de la variable aleatoria identificación de palabras sobresdrújulas	207
Tabla XCIII Parámetros poblacionales de la variable aleatoria corrección de semántica.....	209
Tabla XCIV Frecuencias de la variable aleatoria corrección de semántica.....	211
Tabla XCV Parámetros poblacionales de la variable aleatoria identificación de sinónimas	212
Tabla XCVI Frecuencias de la variable aleatoria identificación de sinónimas ...	214
Tabla XCVII Parámetros poblacionales de la variable aleatoria identificación de antónimas	215
Tabla XCVIII Frecuencias de la variable aleatoria identificación de antónimas.	217
Tabla IC Parámetros poblacionales de la variable aleatoria identificación de palabras aumentativas.....	218
Tabla C Frecuencias de la variable aleatoria identificación de palabras aumentativas	218
Tabla CI Parámetros poblacionales de la variable aleatoria identificación de palabras diminutivas	221
Tabla CII Frecuencias de la variable aleatoria identificación de palabras diminutivas	222
Tabla CIII Parámetros poblacionales de la variable aleatoria identificación de palabras despectivas	223
Tabla CIV Parámetros poblacionales de la variable aleatoria identificación de simples nombres	225
Tabla CV Frecuencias de la variable aleatoria identificación de simples nombres	226
Tabla CVI Parámetros poblacionales de la variable aleatoria vocabulario.....	228
Tabla CVII Frecuencias de la variable aleatoria identificación vocabulario	230
Tabla CVIII Parámetros poblacionales de la variable aleatoria conjuntos de verbos	231
Tabla CIX Frecuencias de la variable aleatoria identificación conjuntos de verbos	232
Tabla CX Parámetros poblacionales de la variable aleatoria identificación del sujeto.....	234
Tabla CXI Frecuencias de la variable aleatoria identificación del sujeto	235

Tabla CXII Parámetros poblacionales de la variable aleatoria identificación del predicado	237
Tabla CXIII Frecuencias de la variable aleatoria identificación del predicado ..	238
Tabla CXIV Parámetros poblacionales de la variable aleatoria identificación del sustantivo	239
Tabla CXV Frecuencias de la variable aleatoria identificación del sustantivo ..	241
Tabla CXVI Parámetros poblacionales de la variable aleatoria identificación del verbo	242
Tabla CXVII Frecuencias de la variable aleatoria identificación del verbo	244
Tabla CXVIII Parámetros poblacionales de la variable aleatoria identificación de frases	245
Tabla CXIX Frecuencias de la variable aleatoria identificación de frases	247
Tabla CXX Parámetros poblacionales de la variable aleatoria identificación de oraciones	248
Tabla CXXI Frecuencias de la variable aleatoria identificación de oraciones ...	250
Tabla CXXII Parámetros poblacionales de la variable aleatoria clasificación de oraciones	251
Tabla CXXIII Frecuencias de la variable aleatoria clasificación de oraciones ..	252
Tabla CXXIV Parámetros poblacionales de la variable aleatoria corrección de sintaxis	254
Tabla CXXV Frecuencias de la variable aleatoria corrección de sintaxis	255
Tabla CXXVI Parámetros poblacionales de la variable aleatoria palabras tildadas correctamente	257
Tabla CXXVII Frecuencias de la variable aleatoria palabras tildadas correctamente	259
Tabla CXXVIII Parámetros poblacionales de la variable aleatoria lectura comprensiva	260
Tabla CXXIX Frecuencias de la variable aleatoria lectura comprensiva	262
Tabla CXXX Parámetros poblacionales de la variable aleatoria calificación de lenguaje	263
Tabla CXXXI Parámetros poblacionales de la variable aleatoria calificación de total	267
Tabla CXXXII Criterios de clasificación de las variables edad y calificación de lenguaje	295
Tabla CXXXIII Criterios de clasificación de las variables sexo y calificación de lenguaje	297
Tabla CXXXIV Criterios de clasificación de las variables ubicación y calificación de lenguaje	299
Tabla CXXXV Criterios de clasificación de las variables jornada y calificación de lenguaje	301
Tabla CXXXVI Criterios de clasificación de las variables edad y calificación de matemáticas	302
Tabla CXXXVII Criterios de clasificación de las variables sexo y calificación de matemáticas	304

Tabla CXXXVIII Criterios de clasificación de las variables ubicación y calificación de matemáticas	306
Tabla CXXXIX Criterios de clasificación de las variables jornada y calificación de matemáticas.....	307
Tabla CXL Criterios de clasificación de las variables calificación de lenguaje y calificación de matemáticas	309
Tabla CXLI Resultados del análisis de contingencia realizado a otras variables aleatorias	311
Tabla CXLII Análisi de varianza de clasificación en un sentido.....	318
Tabla CXLIII Varianzas y porcentajes de explicación de las varianza de las cinco priemeras componentes principales de los datos originales	342
Tabla CXLIV Varianzas y porcentajes de explicación de las varianza de las veinte priemeras componentes principales de las variables estandarizadas...	345
Tabla CXLV Varianzas y porcentajes de explicación de las varianza de las veinte priemeras componentes principales de las variables estandarizadas y rotadas	349
Tabla CXLVI Resultados de la prueba de Bartlett aplicada a las matrices de correlación de las variables aleatorias de la prueba de matemáticas y de lenguaje.....	366
Tabla CXLVII Correlaciones canónicas	374

Indice de Gráficos

	Pag.
Gráfico 1.1 Total de planteles por nivel en el Ecuador	42
Gráfico 1.2 Total de profesores por nivel en el Ecuador.....	43
Gráfico 1.3 Total de alumnos por nivel en el Ecuador	45
Gráfico 1.4 Planteles por sostenimiento y por zona en el Ecuador	46
Gráfico 1.5 Profesores por sostenimiento y por zona en el Ecuador.....	48
Gráfico 1.6 Alumnos por sostenimiento y por zona en el Ecuador	50
Gráfico 1.7 Planteles por sostenimiento en el Ecuador	51
Gráfico 1.8 Planteles por género en el Ecuador	53
Gráfico 1.9 Causas de deserción en el Ecuador.....	54
Gráfico 1.10 Causas de deserción en la zona urbana de todo el país	56
Gráfico 1.11 Causas de deserción en la zona rural de todo el país	57
Gráfico 1.12 Causas de no promoción en el nivel medio de todo el país	59
Gráfico 1.13 Causas de no promoción en el nivel medio en la zona urbana de todo el país	60
Gráfico 1.14 Causas de no promoción en el nivel medio en la zona rural de todo el país	62
Gráfico 2.1 Ubicación geográfica de las parroquias de los colegios fiscales rurales de cantón Guayaquil	66
Gráfico 3.1 Función de distribución de la variable aleatoria edad.....	101
Gráfico 3.2 Ojiva de la variable aleatoria edad de los estudiantes	103
Gráfico 3.3 Diagrama de cajas de la variable aleatoria edad	103
Gráfico 3.4 Histograma de la variable aleatoria sexo.....	105
Gráfico 3.5 Histograma de la variable aleatoria suma de números enteros.....	108
Gráfico 3.6 Diagrama de cajas de la variable aleatoria suma de números enteros.....	108
Gráfico 3.7 Histograma de la variable aleatoria suma de números racionales	110
Gráfico 3.8 Histograma de la variable aleatoria resta de números enteros	112
Gráfico 3.9 Diagrama de cajas de la variable aleatoria resta de números enteros	113
Gráfico 3.10 Histograma de la variable aleatoria resta de números racionales	116
Gráfico 3.11 Histograma de la variable aleatoria multiplicación de números enteros	118
Gráfico 3.12 Diagrama de cajas aleatoria multiplicación de números enteros	119
Gráfico 3.13 Histograma de la variable aleatoria multiplicación de números racionales	121
Gráfico 3.14 Histograma de la variable aleatoria división de números enteros	123
Gráfico 3.15 Histograma de la variable aleatoria división de números racionales	125
Gráfico 3.16 Histograma de la variable aleatoria definiciones de números	127

Gráfico 3.17 Diagrama de cajas de la variable aleatoria definiciones de números	128
Gráfico 3.18 Histograma de la variable aleatoria desigualdad	130
Gráfico 3.19 Diagrama de cajas de la variable aleatoria desigualdad.....	130
Gráfico 3.20 Diagrama de cajas de la variable relaciones de orden	133
Gráfico 3.21 Histograma de la variable aleatoria relaciones de orden	133
Gráfico 3.22 Histograma de la variable aleatoria potenciación y radicación....	135
Gráfico 3.23 Diagrama de cajas de la variable aleatoria divisibilidad.....	139
Gráfico 3.24 Histograma de la variable aleatoria divisibilidad.....	139
Gráfico 3.25 Histograma de la variable aleatoria ejercicio de proporcionalidad	141
Gráfico 3.26 Diagrama de cajas de la variable aleatoria ejercicio de proporcionalidad	142
Gráfico 3.27 Histograma de la variable aleatoria ejercicio de de interés simple	143
Gráfico 3.28 Diagrama de cajas de la variable aleatoria ejercicio de interés simple	145
Gráfico 3.29 Histograma de la variable aleatoria ejercicio de regla de tres	146
Gráfico 3.30 Diagrama de cajas de la variable aleatoria ejercicio de regla de tres	147
Gráfico 3.31 Histograma de la variable aleatoria ejercicio sistema métrico	149
Gráfico 3.32 Diagrama de cajas de la variable aleatoria sistema métrico.....	150
Gráfico 3.33 Histograma de la variable aleatoria operaciones de conjuntos ...	152
Gráfico 3.34 Diagrama de cajas de la variable aleatoria operaciones de conjuntos	152
Gráfico 3.35 Histograma de la variable aleatoria proposiciones lógicas.....	154
Gráfico 3.36 Diagrama de cajas de la variable aleatoria proposiciones lógicas	156
Gráfico 3.37 Histograma de la variable aleatoria evaluación de funciones	158
Gráfico 3.38 Diagrama de cajas de la variable aleatoria evaluación de funciones	159
Gráfico 3.39 Histograma de la variable aleatoria perímetro del cuadrado.....	161
Gráfico 3.40 Diagrama de cajas de la variable aleatoria perímetro del cuadrado.	162
Gráfico 3.41 Histograma de la variable aleatoria área del triángulo	163
Gráfico 3.42 Diagrama de cajas de la variable aleatoria área del triángulo	164
Gráfico 3.43 Histograma de la variable aleatoria área del círculo	165
Gráfico 3.44 Diagrama de cajas de la variable aleatoria área del círculo	166
Gráfico 3.45 Histograma de la variable aleatoria teorema de Pitágoras.....	167
Gráfico 3.46 Diagrama de cajas de la variable aleatoria teorema de Pitágoras	169
Gráfico 3.47 Histograma de la variable aleatoria factorización de dos polinomios	171

Gráfico 3.48 Diagrama de cajas de la variable aleatoria factorización de dos polinomios	172
Gráfico 3.49 Histograma de la variable aleatoria factorización de un polinomio	173
Gráfico 3.50 Diagrama de cajas de la variable aleatoria factorización de un polinomio.....	174
Gráfico 3.51 Histograma de la variable aleatoria ecuación de una incógnita...	176
Gráfico 3.52 Diagrama de cajas de la variable aleatoria ecuación de una incógnita.....	177
Gráfico 3.53 Histograma de la variable aleatoria sistema de dos ecuaciones con dos incógnitas	178
Gráfico 3.54 Histograma de la variable aleatoria probabilidad.....	181
Gráfico 3.55 Diagrama de cajas de la variable aleatoria probabilidad	182
Gráfico 3.56 Diagrama de cajas de la variable aleatoria estadística.....	183
Gráfico 3.57 Histograma de la variable aleatoria estadística.....	184
Gráfico 3.58 Función de distribución de la variable aleatoria calificación de matemáticas	186
Gráfico 3.59 Diagrama de cajas de la variable aleatoria calificación de matemáticas	187
Gráfico 3.60 Diagrama de cajas de la variable aleatoria identificación de diptongos	189
Gráfico 3.61 Histograma de la variable aleatoria identificación de diptongos ..	190
Gráfico 3.62 Histograma de la variable aleatoria identificación de triptongos ..	193
Gráfico 3.63 Diagrama de cajas de la variable aleatoria identificación de triptongos	193
Gráfico 3.64 Histograma de la variable aleatoria identificación de hiatos.....	196
Gráfico 3.65 Diagrama de cajas de la variable aleatoria identificación de hiatos	197
Gráfico 3.66 Diagrama de cajas de la variable aleatoria identificación de palabras agudas	198
Gráfico 3.67 Histograma de la variable aleatoria identificación de palabras agudas	199
Gráfico 3.68 Histograma de la variable aleatoria identificación de palabras graves.....	202
Gráfico 3.69 Diagrama de cajas de la variable aleatoria identificación de palabras graves	202
Gráfico 3.70 Histograma de la variable aleatoria identificación de palabras esdrújulas	205
Gráfico 3.71 Diagrama de cajas de la variable aleatoria identificación de palabras esdrújulas	206
Gráfico 3.72 Histograma de la variable aleatoria identificación de palabras sobresdrújulas	208
Gráfico 3.73 Histograma de la variable aleatoria corrección de semántica.....	210

Gráfico 3.74 Diagrama de cajas de la variable aleatoria corrección de semántica.....	211
Gráfico 3.75 Histograma de la variable aleatoria identificación de sinónimas ...	213
Gráfico 3.76 Diagrama de cajas de la variable aleatoria identificación de sinónimas	214
Gráfico 3.77 Histograma de la variable aleatoria identificación de antónimas...	216
Gráfico 3.78 Diagrama de cajas de la variable aleatoria identificación de antónimas	217
Gráfico 3.79 Histograma de la variable aleatoria identificación de palabras aumentativas	218
Gráfico 3.80 Diagrama de cajas de la variable aleatoria identificación de palabras aumentativas.....	220
Gráfico 3.81 Histograma de la variable aleatoria identificación de palabras diminutivas	222
Gráfico 3.82 Histograma de la variable aleatoria identificación de palabras despectivas	224
Gráfico 3.83 Histograma de la variable aleatoria identificación de simples nombres	226
Gráfico 3.84 Diagrama de cajas de la variable aleatoria identificación de simples nombres	226
Gráfico 3.85 Histograma de la variable aleatoria vocabulario.....	228
Gráfico 3.86 Diagrama de cajas de la variable aleatoria vocabulario	229
Gráfico 3.87 Histograma de la variable aleatoria conjugación de verbos	231
Gráfico 3.88 Diagrama de cajas de la variable aleatoria conjugación de verbos	232
Gráfico 3.89 Histograma de la variable aleatoria identificación del sujeto.....	234
Gráfico 3.90 Diagrama de cajas de la variable aleatoria identificación del sujeto	235
Gráfico 3.91 Histograma de la variable aleatoria identificación del predicado.	237
Gráfico 3.92 Diagrama de cajas de la variable aleatoria identificación del predicado	238
Gráfico 3.93 Histograma de la variable aleatoria identificación del sustantivo	240
Gráfico 3.94 Diagrama de cajas de la variable aleatoria identificación del sustantivo	240
Gráfico 3.95 Histograma de la variable aleatoria identificación del verbo	243
Gráfico 3.96 Diagrama de cajas de la variable aleatoria identificación del verbo	244
Gráfico 3.97 Histograma de la variable aleatoria identificación de frases	246
Gráfico 3.98 Diagrama de cajas de la variable aleatoria identificación de frases	247
Gráfico 3.99 Histograma de la variable aleatoria identificación de oraciones ..	249
Gráfico 3.100 Diagrama de cajas de la variable aleatoria identificación de oraciones	250
Gráfico 3.101 Histograma de la variable aleatoria clasificación de oraciones ..	251

Gráfico 3.102 Diagrama de cajas de la variable aleatoria clasificación de oraciones	253
Gráfico 3.103 Histograma de la variable aleatoria corrección de sintaxis	255
Gráfico 3.104 Diagrama de cajas de la variable aleatoria corrección de sintaxis .	
.....	256
Gráfico 3.105 Histograma de la variable aleatoria palabras tildadas correctamente	258
Gráfico 3.106 Diagrama de cajas de la variable aleatoria palabras tildadas correctamente	258
Gráfico 3.107 Histograma de la variable aleatoria lectura comprensiva.....	261
Gráfico 3.108 Diagrama de cajas de la variable aleatoria lectura comprensiva	
.....	262
Gráfico 3.109 Histograma de la variable aleatoria calificación de lenguaje	264
Gráfico 3.110 Diagrama de cajas de la variable aleatoria calificación de lenguaje	264
Gráfico 3.111 Histograma de la variable aleatoria calificación total	266
Gráfico 3.112 Ojivas de las variables aleatorias con mayor coeficiente de asimetría	268
Gráfico 3.113 Ojivas de las variables aleatorias con menor coeficiente de asimetría	269
Gráfico 3.114 Ojivas de las variables aleatorias con mayor coeficiente de asimetría	272
Gráfico 3.115 Ojivas de las variables aleatorias con menor coeficiente de asimetría	273
Gráfico 4.1 Diagrama de dispersión de las variables aleatorias identificación de diptongos e identificación de triptongos	281
Gráfico 4.2 Diagrama de dispersión de las variables aleatorias identificación del sujeto de la oración e identificación de predicado de la oración	283
Gráfico 4.3 Diagrama de dispersión de las variables aleatorias multiplicación de números racionales y división de números racionales	284
Gráfico 4.4 Diagrama de dispersión de las variables aleatorias identificación de palabras agudas e identificación de palabras diminutivas.....	287
Gráfico 4.5 Diagrama de dispersión de las variables aleatorias número de relaciones de orden correctas y número de palabras definidas correctamente	288
Gráfico 4.6 Diagrama de dispersión de las variables definición de conjuntos de números y ecuación lineal con una incógnita.....	290
Gráfico 4.7 Medias de los tratamientos del factor ubicación con respecto a al calificación de lenguaje	322
Gráfico 4.8 Medias de los tratamientos del factor jornada con respecto a al calificación de lenguaje	324

Gráfico 4.9 Medias de los tratamientos del factor ubicación con respecto a al calificación de matemáticas	327
Gráfico 4.10 Medias de los tratamientos del factor jornada con respecto a al calificación de matemáticas	330
Gráfico 4.11 Medias de los tratamientos del factor ubicación con respecto a al calificación total	333
Gráfico 4.12 Medias de los tratamientos del factor jornada con respecto a al calificación total	335
Gráfico 4.13 Valores propios de las componentes principales, de los datos originales	341
Gráfico 4.14 Valores propios de las componentes principales de las variables aleatorias estandarizadas	346
Gráfico 4.15 Valores propios de las componentes principales estandarizadas y rotadas.....	350

RESUMEN

La investigación realizada en esta tesis es un análisis estadístico del nivel de conocimiento de matemáticas y lenguaje de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil en el año 2000. Este estudio forma parte de un conjunto de tesis que se desarrollaron, para analizar el nivel de conocimiento en las materias antes mencionadas, de estudiantes de séptimo, décimo y último año de bachillerato, en las zonas urbana y rural, en establecimientos educativos fiscales y particulares.

En el primer capítulo de esta tesis se realiza una revisión de los hechos más destacados en la historia de la educación y desarrollo del Ecuador, en el cual se resaltan los acontecimientos que se presentaron fundamentalmente desde la época de la colonización. Además se presentan datos estadísticos de la situación de la educación en el país en la década pasada.

En el segundo capítulo se plantean los temas y subtemas que forman parte de las pruebas de matemáticas y lenguaje, con las cuales se midió el nivel de conocimiento de los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil. Además se describe y codifica cada variable aleatoria que se utilizará para la evaluación.

En el tercer capítulo se realiza el análisis estadístico de cada una de las variables aleatorias, con el objetivo de conocer el comportamiento y las características más significativas de estas variables estudiadas. En éste análisis se determina los parámetros poblacionales, así como los histogramas y diagramas de cajas de cada variable aleatoria.

En el capítulo cuatro se realiza el análisis multivariado de las características investigadas. El cual contiene el análisis de correlación lineal para determinar la dependencia lineal entre las variables, el análisis de contingencia por medio del cual se determina la dependencia lineal o no, que existe entre las características de investigación, además de otras técnicas multivaridas que se desarrollaron como el análisis de componentes principales y el análisis de correlación canónica.

INTRODUCCIÓN.....	6
Capítulo 1.....	8
1 Educación y desarrollo en el Ecuador	8
1.1 Antecedentes	8
1.2 Educación y desarrollo: definiciones y generalidades.....	10
1.3 Historia y cultura de los primeros habitantes del Ecuador.....	11
1.4 La educación y cultura en tiempos de la Colonia.....	13
1.5 Desarrollo y educación del Ecuador durante la Gran Colombia	18
1.6 La educación durante la vida republicana del Ecuador.....	21
1.6.1 Primera Etapa	22
1.6.2 Segunda Etapa	26
1.6.3 Tercera Etapa.....	28
1.6.4 Cuarta Etapa	28
1.6.5 Quinta Etapa	31
1.7 El sistema educativo Ecuatoriano	35
1.8 Educación Rural en el Ecuador	38
1.9 Datos estadísticos sobre la educación en el Ecuador.....	39

INTRODUCCIÓN

La educación ha demostrado ser el principal factor explicativo en el desarrollo de la humanidad, dadas las diferencias observadas en los distintos países y regiones, en los que se invierte considerablemente en la planificación y mejoramiento de la educación, para conseguir el crecimiento a largo plazo, la reducción de la pobreza y la equidad social; en este sentido podemos decir que la educación es una condición básica para cualquier proceso de desarrollo.

"La calidad del sistema educativo ecuatoriano, atraviesa una grave crisis que se manifiesta en la falta de infraestructura, un limitado presupuesto que incide en paralizaciones permanentes de los educadores"¹ y programas de estudio que generalmente transmiten conocimientos con atrasos considerables y no estimulan la formación integral, la capacidad intelectual o las habilidades técnicas de los estudiantes.

La crisis económica, política y social que vive nuestro país, es tan solo un reflejo de la calidad de la educación, por esta razón, es conveniente realizar una evaluación del nivel de calidad, mediante la medición de los conocimientos adquiridos por los estudiantes, durante su paso por las

escuelas y colegios. En particular en la presente tesis se realiza, el análisis estadístico de la medición del conocimiento, en las áreas de matemáticas y lenguaje, de los alumnos de décimo curso de los colegios fiscales rurales de Guayaquil, la evaluación contemplará aspectos básicos de conocimiento, que se considera deben tener los estudiantes en estas materias, de acuerdo a los programas de estudios vigentes establecidos por el Ministerio de Educación y Cultura.

¹ Saltos N, vasquez L. (1999) Ecuador su realidad, Fundación Jose Peralta, Séptima Edición, Quito Ecuador

Capítulo I

1 Educación y desarrollo en el Ecuador

1.1 Antecedentes

Los primeros establecimientos educativos en el Ecuador fueron fundados en la ciudad de Quito. Luego de la fundación de esta ciudad en el año de 1534, hasta alrededor de 1554, se produjo la primera generación de criollos, mestizos e indios, que necesitaban ser atendidos, en lo educativo, por la nueva sociedad organizada.

La obra educativa, fue entonces llevada a cabo, por los sacerdotes de las diferentes órdenes religiosas que se encontraban en la ciudad. Es destacable, la labor que realizó al principio, la orden de los sacerdotes franciscanos, con la fundación del colegio San Juan Evangelista en el año de 1552 en Quito, en este colegio se enseñaba doctrina cristiana, a leer y escribir y la educación era gratuita. Luego de algún tiempo este colegio se transformo en el colegio San Andrés, el cual para el año de

1568, ya había capacitado a la primera generación de indios para el magisterio, y de mestizos que aspiraban al sacerdocio.

La labor emprendida por las órdenes religiosas en la ciudad de Quito, fue acogida por el afán de cultura, en otras ciudades de la Real Audiencia. A partir de 1570 en las ciudades de Guayaquil, Loja, Cuenca, Riobamba, Ambato, Ibarra y Latacunga los párrocos de las iglesias fundaron escuelas junto a estas, en las cuales se enseñaba a leer, escribir, artes y catequización. Algunos de los indios preparados para el magisterio en el colegio San Andrés de Quito trabajaron como profesores en estos nuevos establecimientos educativos.

La labor realizada por los sacerdotes jesuitas, marca el comienzo de la educación en el Ecuador, con establecimientos muy organizados, en lo que respecta a la admisión de los alumnos, la distribución del tiempo de estudio, metodologías de dirección y enseñanza, obligaciones de los seminaristas y administración de los bienes. Los jesuitas fundaron el colegio Seminario de San Luis, que podía competir en cuanto a la organización con algunos seminarios de España.

En lo referente a la educación superior, se fundaron tres universidades: la de los agustinos llamada de San Fulgencio, la los jesuitas de San

Gregorio Magno y la de los Dominicos llamada de Santo Tomás de Aquino.

1.2 Educación y desarrollo: definiciones y generalidades

La educación es el conjunto de medios y procedimientos, que desarrollan en los individuos, facultades intelectuales basados en el conocimiento, la instrucción y las opiniones.

Es una poderosa maquinaria de conocimiento sistematizado, que transforma a los receptores (materia prima), en seres con capacidad de desarrollo individual y colectivo. El conocimiento sistemático, debe contemplar, la preparación del individuo por etapas, tomando en cuenta su capacidad de aprendizaje y el orden natural del desarrollo del conocimiento.

Los objetivos de la educación se resumen en el desarrollo científico y tecnológico, esencial para el progreso económico-social de una nación, una planificada administración e inversión de los recursos del estado, y la preparación y capacitación de los individuos para disminuir el desempleo y la pobreza.

El desarrollo es el conjunto de planes y acciones, que se realizan para conseguir mejoras, en las condiciones de vida de los habitantes de un país; la planificación debe contemplar los aspectos económicos, sociales, políticos, tecnológicos y culturales.

1.3 Historia y cultura de los primeros habitantes del Ecuador

La primera etapa del estudio de la historia de lo que actualmente es el Ecuador es la etapa pre-incaica, la cual se divide en tres períodos. Las culturas que se conocen, se desarrollaron durante el primer período de la época pre-incaica (desde los 3300 años a.C) fueron Valdivia, Machalilla y Chorrera.

Durante el segundo período de la época pre-incaica (desde 500 años A.C hasta 500 años D.C) se produjo la formación de las primeras ciudades. El avance tecnológico le permitió al hombre andino realizar culturas artísticas de realce. Las culturas de este período fueron Guangala, Tolita y Bahía. En este período se logró un avanzado desarrollo social, el asentamiento humano fue mayor y mejor planificado en la sierra ecuatoriana.

El tercer período pre-incaico se destacó por los pobladores urbanos agrupados en cacicazgos o señoríos étnicos, que eran organizaciones

que mantenían estabilidad en lo político y económico (1000 y 1500 años D.C).

La segunda etapa de estudio la constituye, el período incaico que surge por la expansión que se produce luego de que "las familias Collas conquistaran la región del Títicaca en el siglo XII"², este hecho dio lugar a la formación del imperio incaico denominado Tahuantinsuyo, el cual fue fundado por Pachacutec y cuya capital era la ciudad de Cuzco; este imperio estaba dividido en cuatro regiones; y era regido por una monarquía totalitaria cuyo líder máximo era el Inca.

En el aspecto cultural los Incas no desarrollaron un sistema de escritura, y utilizaron una técnica llamada Quipus para recordar datos y hechos importantes. Elaboraron un calendario basado en las fases de la luna que comenzaba en diciembre con grandes fiestas. Construyeron una amplia red de caminos que cruzaban el imperio con puentes colgantes, la arquitectura llevó a niveles muy elevados con la construcción de palacios, templos y fortalezas.

² Dávila J, Orbe J, Gómez M, 1990, Ciencias Sociales: Historia, Geografía y Cívica.

1.4 La educación y cultura en tiempos de la Colonia

Las primeras instituciones educativas que se implantaron en América luego de la llegada de los españoles en 1492, fueron las escuelas; además de las formas educativas los españoles trajeron al nuevo continente su cultura, su idioma, la religión católica, sus tradiciones y costumbres.

Los establecimientos educativos originalmente estaban destinados a la catequización. El clero era el encargado de ejecutar el proceso educativo, el mismo que debía ser selectivo, guardando las respectivas distancias entre los hijos de los españoles, los criollos y los indios; basándose en esto las escuelas se clasificaron en doctrinales, monacales, abadengas y catedrales.

Por otro lado la educación de la mujer fue restringida únicamente a monasterios o a conventos y solo para aquellas que poseían requisitos de nobleza. Posteriormente se incorpora a la enseñanza religiosa el conocimiento del cálculo, la lectura, artes y oficios.

Después de ser fundada la ciudad de San Francisco de Quito, llegó a esta ciudad en el año de 1535 una expedición pedagógica encabezada por el sacerdote Jacobo Ricke y tres religiosos de la orden seráfica, los

cuales fundaron la primera escuela primaria que tenía el carácter de no discriminatoria, en esta escuela se enseñaba toda clase de ciencias sociales y actividades de oficios. El programa de estudios consistía en la enseñanza de aritmética, castellano además de algunos oficios como: carpintería, sastrería, mecánica y artes como la pintura, dibujo y música. Esta escuela llevó el nombre de San Andrés y funcionó bajo la orden de los sacerdotes franciscanos durante unos 20 años, luego de lo cual pasó a órdenes de los sacerdotes agustinos. Además de las escuelas religiosas durante la colonia existían escuelas particulares pagadas por los padres de familia y subsidiados por el cabildo.

El primer establecimiento educativo de la ciudad de Santiago de Guayaquil fundada en 1537, fue creado por los sacerdotes dominicanos, en el año 1554.

Las ciudades más importantes, además de Quito, que se encontraban dentro del territorio de la Audiencia de Quito eran Guayaquil, Cuenca, Loja, Riobamba, Ibarra, Ambato y Latacunga, estas ciudades se habían establecido, de tal forma que las viviendas se ubicaban alrededor de la iglesia mayor y de los conventos de San Francisco, Santo Domingo, de la Merced y de San Agustín. A partir del año 1570 en estas ciudades los párrocos y doctrineros de las iglesias y conventos fundaron escuelas en

las que fundamentalmente se enseñaba catequización, leer, escribir, artes y oficios.

La educación media en la colonia se desarrollaba en los colegios, también llamados en esa época mayores o máximos, y en los seminarios que en su mayoría eran considerados de nivel medio. El colegio era la institución preparatoria para ingresar a la universidad y se enseñaba filosofía, arte, retórica, teología, etc.

El primer colegio fue fundado por los sacerdotes jesuitas en 1592 en la ciudad de Quito y se lo llamó seminario de San Luis, en este establecimiento educativo se enseñaba la lengua latina, filosofía y teología, los jesuitas abrieron nuevas posibilidades de cultura, con la organización permanente de establecimientos educativos.

Luego de un siglo de creado este colegio los sacerdotes dominicanos fundaron el colegio secundario Convictorio de San Fernando el cual fue auspiciado por el convento y por algunos padres interesados en su creación los cuales proporcionaron los medios económicos necesarios para su construcción y la implementación de una biblioteca.

Después de medio siglo de enseñanza del colegio seminario de San Luis, en el que los sacerdotes jesuitas demostraron su conocimiento para

establecer colegios destinados a la instrucción de las personas, el afán de cultura de las demás ciudades y pueblos de la Audiencia de Quito, hizo que estos enviaran solicitudes a la compañía de Jesús, para que fundases establecimientos educativos en sus respectivas ciudades. De esta forma se fueron estableciendo en las ciudades de la Audiencia de Quito los colegios jesuitas.

De las demás ciudades de la audiencia la primera que se benefició con la fundación de un colegio jesuita fue la ciudad de Cuenca en el año de 1633. En 1668 se fundó el tercer colegio de jesuitas en la ciudad de Latacunga, el cual en el año 1674 pasó a funcionar como noviciado. La tercera ciudad que contó con un colegio jesuita fue Ibarra en 1685. En las ciudades de Riobamba y Guayaquil la creación de colegios jesuitas se llevó a cabo durante el primer decenio del siglo XVIII.

La creación de colegios y seminarios en las principales ciudades de la Audiencia de Quito, elevó el nivel de cultura general de los habitantes de esta zona. Las materias que se dictaban en estos establecimientos educativos eran filosofía, lógica, física, cosmología, psicología, teología dogmática y teología moral.

Las universidades coloniales tienen su origen en el modelo de la universidad española, cuya organización contenía facultades mayores y

menores. Las universidades españolas que estaban bajo la tutela de la iglesia y de la monarquía, a su vez, adoptaron los modelos de las universidades alemanas que tenían carácter investigativo, las inglesas que se basaban en la educación y autodesarrollo personal y la universidad francesa identificada con el servicio de la administración del estado.

En las facultades mayores se realizaban estudios de derecho, economía, código civil, teología y medicina, mientras que en las facultades menores se realizaban estudios de arte y filosofía entre otras.

Las universidades se clasificaban en: pontificias cuando su origen venía de bulas papales e imperiales y señoriales cuando su creación provenía de los monarcas. Las primeras universidades ecuatorianas surgieron como una extensión de los colegios y seminarios que habían fundado las órdenes religiosas y fueron las de San Fulgencio en 1603, de Santo Tomás y de San Gregorio el Magno, en la primera se enseñaba teología derecho canónico y arte, esta duró poco tiempo por la excesiva liberalidad, con la que concedían títulos de bachiller, maestro y doctor, por lo que a partir de 1786 mediante cédula real el rey Carlos III, este prohibió que el centro de estudios confiriera grados.

Como ya se mencionó anteriormente el progreso de la educación en la época de la colonia, estaba ligada a la llegada de los sacerdotes jesuitas y la propagación de establecimientos educativos bajo su organización. Estos religiosos influyeron notablemente en el fortalecimiento del sistema educativo universitario durante la colonia. En el año de 1622 el seminario de San Luis que funcionaba bajo su responsabilidad, fue elevado a la categoría de universidad, la cual se denominó Pontificia universidad de San Gregorio.

La universidad de San Gregorio fue disuelta años más tarde con el fin de que se integre todo su contingente a la de Santo Tomás, y así incorporar y reunir todas las cátedras de ambas universidades, para que luego en 1787 surja la universidad pública de Santo Tomás de Aquino. Esta universidad fue la antesala de lo que actualmente es la universidad central de Quito.

1.5 Desarrollo y educación del Ecuador durante la Gran Colombia

La Gran Colombia tuvo una duración de ocho años y fue creada en el Congreso de Angostura, realizado el 18 de Diciembre de 1822. La nueva República estuvo integrada por tres naciones: la Capitanía General de Venezuela, el Virreinato de Santa Fe de Bogota y la Real Audiencia de

Quito. Las cuales se denominaron distritos del norte, centro y sur respectivamente.

Tabla I

Cronología de la fundación de las universidades y escuelas politécnicas estatales en el Ecuador hasta 1999

Universidad	Año	Ciudad
Central del Ecuador	1826	Quito
Escuela Politécnica Nacional	1869	Quito
Estatal de Guayaquil	1897	Guayaquil
Estatal de Cuenca	1897	Cuenca
Nacional de Loja	1943	Loja
Técnica de Manabí	1952	Portoviejo
Escuela superior politécnica del litoral	1958	Guayaquil
Técnica de Ambato	1959	Ambato
Técnica de Machala	1969	Machala
Técnica de Esmeraldas	1969	Machala
Técnica de Babahoyo	1971	Babahoyo
Escuela superior politécnica del Chimborazo	1973	Riobamba
Escuela politécnica del ejército.	1977	Quito
Técnica estatal de Quevedo	1984	Quevedo
Laica Eloy Alfaro de Manabí	1985	Manta
Tecnológica Equinoccial	1986	Quito
Técnica del norte	1986	Ibarra
Estatal de Bolívar	1989	Guaranda
De Azuay	1990	Cuenca
Agraria del Ecuador	1992	Guayaquil
Técnica de Cotopaxi	1995	Latacunga
Nacional de Chimborazo	1995	Riobamba
De la península de Santa Elena	1996	Santa Elena
Escuela superior politécnica agropecuaria de Manabí.	1999	Calceta

FUENTE: Secretaría general del CONUEP.

Cada distrito se dividía en departamentos y estaba dirigido bajo la autoridad de un Intendente. El distrito del sur, que se adhirió a la Gran Colombia, luego de la batalla del Pichincha en 1824, estaba conformado por tres departamentos Ecuador, Guayaquil y Cuenca.

Durante la vida de la Gran Colombia nuestro país tuvo una cultura semejante a la del tiempo de la colonia, las diferentes ordenes religiosas asentadas en este territorio siguieron cubriendo las necesidades educativas de los habitantes.

El General Sucre se preocupó en "recomendar la creación de escuelas en el distrito del sur, este consejo fue tomado en cuenta por la ciudad de Cuenca en la que se crearon veinte establecimientos educativos"³. Años después cuando el General Sucre fue intendente del nuevo estado de Quito se preocupó por la implantación del sistema de enseñanza mutua, que consistía en enseñarse entre alumnos, este método fue de gran ayuda para los maestros que debían atender varios grados de estudio a la vez y por el poco material bibliográfico con el que contaban.

En 1826 por disposición del congreso de Cundinamarca, se crean en cada uno de los distritos de la gran Colombia las universidades centrales,

³ Castro R, 1985, Historia del Ecuador

así surge lo que es actualmente la universidad central del Ecuador con sede en Quito.

1.6 La educación durante la vida republicana del Ecuador

El deseo separatista del distrito del sur de la Gran Colombia se basaba principalmente en la autonomía económica y la situación geográfica dada la lejanía de la capital de la confederación que era la ciudad de Bogotá y la mala administración de las autoridades extranjeras. Por estos motivos se realizó el "13 de Mayo de 1830 un congreso en la ciudad de Riobamba en la que se decidió la separación del distrito del sur de la confederación bolivariana, se agradeció a Bolívar por la independencia de la nación y se encargó el gobierno interinamente al General Juan José Flores¹⁴. Como presidente de la nación el General Flores preparó la primera asamblea constituyente para el 14 de Agosto de 1830 en Riobamba, la primera declaración que se realizó en la asamblea, fue la de llamar Ecuador al antiguo distrito del Sur.

La vida republicana del Ecuador se la puede dividir para su mejor comprensión en lo referente a la educación y desarrollo en las siguientes cinco etapas: la primera incluye a la Dominación Floreana 1830-1845, la Revolución Marcista 1845-1860, la Dominación Garciana 1860-1875 y el

Progresismo 1885-1895. La segunda etapa abarca el Liberalismo Radical 1895-1912 y la Dominación Bancaria 1912-1925. La tercera etapa la Revolución Juliana 1925-1934. La cuarta etapa el Populismo Velasquista 1934-1972. La quinta etapa de Dictaduras Militares 1972-1979 y los Regímenes Democráticos 1979-2000.

1.6.1 Primera Etapa

Dominación Floreana 1830-1845. En la presidencia del General Juan José Flores se produjo un estancamiento del desarrollo de la cultura y la educación, debido a la utilización de la mayor parte del presupuesto del estado en el ejército, por lo que la educación y el desarrollo del país se mantuvieron invariables y en los mismos niveles que en la época colonial.

En su mandato presidencial, Vicente Rocafuerte impulsó la creación de las instituciones de educación pública y laica, él decía: "La instrucción de las masas afianza la libertad y destruye la esclavitud". Comenzó con la instrucción a los aborígenes. Creó el primer colegio femenino Santa María de Socorro. En 1836 creó la dirección general de estudios. Mediante su mandato como gobernador de Guayaquil fundó el colegio que lleva su nombre en el año de 1842, en el plan de estudios de dicho

⁴Dávila J, Orbe J, Gómez M, 1990, Ciencias Sociales: Historia, Geografía y Cívica.

colegio constaban las siguientes materias: castellano, inglés, francés, matemáticas, lógica, literatura, historia sagrada antigua y moderna entre otras. Existían 170 escuelas en todo el país, con 4323 alumnos y 546 alumnas. Las escuelas particulares eran 126 y las fiscales 44 de las cuales cinco eran de mujeres.

Revolución Marcista 1845-1860. Los gobiernos marcistas surgieron luego de la dominación floreana, cuando se produjo la cuarta asamblea constituyente en nuestra historia en el año de 1845. Durante el gobierno del Dr. Urbina en 1852, se promulgó la ley de libertad de enseñanza la cual revolucionó los planes y programas de estudio, los cuales promovían la facilidad del proceso educativo para la obtención de títulos y de años escolares.

En esta etapa se trató de implantar, la modalidad de enseñanza de la Universidad de Oxford, de Inglaterra, que consistía en que los estudiantes podían graduarse, luego de rendir exámenes con las calificaciones exigidas sin necesidad de asistir a clases. La demagogia reinó, en lo referente a las propuestas del apoyo a la educación, el presupuesto designado para su desarrollo fue escaso, por lo que se evidenció una carencia de infraestructura adecuada y de útiles escolares para el proceso de aprendizaje.

Dominación Garciana 1860-1875. Durante este período se destaca la figura política de García Moreno, que contribuyó notablemente al desarrollo de la educación en nuestro país. La primera acción que realizó, fue la de aumentar los salarios de los maestros urbanos y rurales de nueve a veinte y treinta pesos mensuales respectivamente. Otro dato destacable para conocer la importancia que le dio a la educación, fue el número de alumnos de nivel primario que existían en 1867 que era de 13485 y que luego de ocho años llegaron a ser 32000.

Además creó el protectorado que hoy en día se conoce como colegio técnico, trayendo maestros artesanos de norteamérica y desarrolló un plan de becas para este Colegio. Promovió la pintura y el arte en todas sus formas, fundando la escuela de bellas artes y el conservatorio de música. Otros establecimientos de instrucción que creó fueron el colegio militar, la escuela de obstetricia, un colegio normal para indígenas y el observatorio astronómico.

Un hecho negativo de su mandato fue la clausura de la universidad central, debido a que se creía que en esta se propagaba ideas liberales, lo cual estaba en total oposición a la convicción conservadora y católica de García Moreno, a pesar de esta situación dio un impulso vertiginoso al desarrollo científico-técnico del Ecuador, con la creación de la escuela politécnica nacional dirigida por jesuitas alemanes en 1869. Para este

establecimiento educativo contrató como profesores de diferentes áreas a reconocidos científicos de Europa.

El Progresismo 1885-1895. Durante esta etapa en el año de 1884, se creó el ministerio de instrucción pública, dado que desde el inicio de la república, la educación estaba adscrita a otros ministerios.

Se restableció la escuela de ciencias y se creó en Ibarra el colegio de San Alfonso. Como se puede apreciar el número de colegios en el país aumentaba considerablemente, estos existían en: Ibarra, Latacunga, Guayaquil, Cuenca, Quito, Portoviejo, Loja, Ambato, Guaranda y Riobamba.

De acuerdo a un informe emitido por el ministerio de instrucción pública en 1885 el número de alumnos de nivel medio ascendía a 40000 y si se consideraba los colegios privados, estos llegaban a 50000. Durante este período además se creó la facultad de filosofía y letras de la universidad central, se fundan establecimientos educativos femeninos religiosos en Ibarra, Otavalo y Guaranda, en Machala se creó un colegio en 1888.

Durante el mandato presidencial del Dr. Antonio Flores Jijón, este impulsó considerables mejoras en el desarrollo de la educación: Trabajó en la reforma de la ley de instrucción pública, introdujo el método

concéntrico de la enseñanza secundaria, impulsó la educación a nivel nacional, utilizando para este fin a las comunidades religiosas. Durante 1890 a 1892 se crearon 240 escuelas y se designó a los sacerdotes salesianos para que dirijan las escuelas de artes y oficios en Riobamba y la de pintura en Cuenca.

Los informes que realizaba el ministerio de instrucción indicaban que la población escolar había llegado a 74. 358 alumnos distribuidos en 1.108 escuelas en todo el país. En este período mejoró la educación pública, otros datos proporcionados por el ministerio de esa época, señalan la presencia de 234 profesores secundarios encargados de educar a 2.084 alumnos y de 5.151 alumnas, bajo la dirección de 282 profesoras. Un estudio realizado en ese tiempo por el Dr. Luis Cordero indica el recargo de las materias en los planes de estudio y decía: "Con este sistema se fatiga la memoria del estudiante por el recargo intelectual, causándoles daños prematuramente en la salud, en la edad que más necesitan robustez para el mejor desarrollo físico"

1.6.2 Segunda Etapa

Liberalismo Radical 1895-1912. Los hechos más destacables en este período fueron las modificaciones a la ley de instrucción pública, que incluye el laicismo en forma definitiva y la enseñanza primaria gratuita y

obligatoria, la creación de las escuelas nocturnas para artesanos, la escuela de bellas artes y oficios en Latacunga, el conservatorio de música en Quito, la escuela naval y la de veterinaria y se inauguró el instituto normal en Guayaquil, en Ibarra de varones y en Riobamba de señoritas, se inaugura el colegio de Portoviejo, el observatorio meteorológico de Riobamba. Se elaboraron planes de becas que consistían en enviar al exterior a estudiantes con el fin de que se especialicen en los campos necesarios para el progreso y desarrollo del país. Todos estos actos de progresos fueron impulsados en el gobierno del Gral. Eloy Alfaro.

Los efectos de la revolución liberal en la universidad, se dieron en el sentido de que la educación se volvió estatal y laica. Esta situación terminó con la dirección de las universidades por los sacerdotes y la iglesia católica.

Dominación Bancaria 1912-1925. Esta etapa de la historia ecuatoriana tuvo su origen en la revolución liberal, se caracterizó por la influencia de los bancos en la vida política del estado. En lo referente a la educación en este período, se elaboraron los planes, programas y reglamentos de estudio, de régimen escolar, de obtención de becas, de reparto del material didáctico, haciendo efectiva la gratuidad de la enseñanza. Se fundó el instituto de agronomía de Ambato. En 1917 se destinaron fondos

del ministerio de hacienda para cubrir una parte de la educación primaria. Durante 1918 el presupuesto para la instrucción primaria subió a 1'500.000 sucres. Además se dictó un segundo plan de estudios para la enseñanza secundaria. El número de estudiantes de nivel primario en 1920 era de 105.377 y el de las escuelas primarias de 1.664.

1.6.3 Tercera Etapa

Revolución Juliana 1925-1934. Entre estos años predominaron las ideas socialistas en el país, en consecuencia de esto se atiende a la educación rural, se vincula a la educación los aspectos sociales, culturales, económicos y políticos, se diversifica el diseño y la elaboración de los planes de estudio, se pretende con esto la democratización de la educación.

1.6.4 Cuarta Etapa

Populismo Velasquista 1934-1972. La política educativa en este período, tuvo como característica más señalada la secularización, que consiste en transferir bienes o funciones eclesiásticas a particulares o al estado. Esta secularización de la enseñanza, junto a otros servicios públicos, formaba parte de las necesidades de afianzar al Estado frente a la poderosa influencia que la Iglesia católica había ejercido a lo largo de todo el siglo

XIX, se constituyó el laicismo y la "nacionalización". La Iglesia y el Estado se comprometieron a colaborar en la atención al indígena, las misiones, la cultura y la educación, favoreciéndose a partir de entonces la educación en las instituciones religiosas pero la enseñanza oficial continuó siendo laica.

**Tabla II.
Cronología de la fundación de las universidades particulares que reciben subvención del estado hasta 1997**

Universidad	Año	Ciudad
Pontificia universidad católica del Ecuador	1946	Quito
Católica Santiago de Guayaquil	1962	Guayaquil
Laica Vicente Rocafuerte	1966	Guayaquil
Católica de Cuenca	1970	Cuenca
Técnica particular de Loja	1971	Loja
Tecnológica Equinoccial	1986	Quito
Del Azuay	1990	Cuenca
Politécnica salesiana	1994	Guayaquil
Escuela superior politécnica ecológica amazónica	1997	Tena

FUENTE: Secretaría general del CONUEP

El sistema educativo liberal instaurado a partir de 1895 había favorecido a las clases medias urbanas y su acceso a los colegios secundarios laicos, a las escuelas normales y a ciertas instituciones educativas militares, entre otros establecimientos mientras que en los años 40, se

atendió a la población rural, sobre todo la indígena. La Constitución de 1946 contempló la posibilidad de subvencionar la enseñanza particular (Art. 171) y en el mismo año se autorizó la creación de la universidad católica en Quito.

Los gobiernos velasquistas se caracterizaron por prescindir de toda planificación, adoptando más bien medidas de tipo asistencial de atención urgente a los requerimientos populares, dando prioridad sobre todo a la realización de obras públicas y a la extensión de la educación. La Junta Militar de 1963-66, efectuó ambas políticas, la asistencialista y la planificadora, que beneficiaron de la extensión de la educación pública.

En 1960 se creó el departamento de planificación integral de la educación en el ministerio de educación nacional. Para el decenio 1964-1973, la educación primaria presentaba hondas desigualdades entre las zonas rurales y urbanas, ofreciéndose en este nivel obligatorio sólo cuatro años de estudio en el ámbito rural, mientras que en las zonas urbanas el ciclo era de seis años y era el único que capacitaba para el acceso a la enseñanza secundaria. La política educativa favoreció sobre todo la extensión de la educación primaria en las zonas rurales, así como un considerable crecimiento de la enseñanza secundaria pública en las ramas de enseñanza general y técnica. La reforma educativa de 1964 amplió la enseñanza primaria rural a 6 años, igualándola con la urbana.

En la enseñanza secundaria la reforma de 1964 instituyó un ciclo básico y otro diversificado, este último incluyendo opciones de estudio para la formación de obreros calificados. Una ley de educación y cultura promulgada en 1977 amplió la obligatoriedad y gratuidad de la educación hasta el primer ciclo de la enseñanza secundaria, formando la denominada enseñanza básica, con 9 años de estudio.

1.6.5 Quinta Etapa

Dictaduras Militares 1972-1979. En este período la educación ha cambiado tanto en forma cuantitativa como cualitativa, hay planes programas y recursos didácticos, la formación, capacitación y mejoramiento docente son objetivos permanentes. Por otro lado la influencia de los gobiernos militares se evidencia en una falta de libertad de expresión y un fuerte control de los procesos de enseñanza.

Regímenes Democráticos 1979-2000. La nueva Constitución Política de 1979, que restauró la democracia en el país, otorgó por primera vez en la historia ecuatoriana el derecho de ciudadanía y el sufragio a los analfabetos, esta reforma no ha sido beneficiosa. Sin embargo, el voto es obligatorio solamente para los ciudadanos que sepan leer y escribir y

facultativo para los analfabetos (Art. 33 de la constitución de la república).

Los nuevos gobiernos se fijaron como una de sus principales metas en política educativa la disminución del analfabetismo, ampliando a la educación de adultos y del medio rural. La misma Constitución estableció que en el presupuesto del Estado se destinaría no menos del treinta por ciento de los ingresos corrientes del gobierno central para la educación y la erradicación del analfabetismo (Art. 71 de la constitución de la república). El programa de alfabetización "Jaime Roldós Aguilera" de 1980-1984, así como la Campaña Nacional de Alfabetización "Monseñor Leónidas Proaño", que se llevó a cabo entre los años 1989 y 1990, son las primeras evidencias de esta política, a las que hay que añadir la política decidida en favor de la educación bilingüe e intercultural dirigida a la población indígena.

En febrero de 1980 el Ministerio de Educación y Cultura promovió un debate pedagógico nacional, con el fin de sentar las bases de la política educativa de la nueva década que entonces se iniciaba. La nueva Ley de Educación de 1983 no introdujo cambios significativos en la organización de los niveles educativos, pero las estrategias de desarrollo del sistema educativo dieron prioridad a la educación preescolar. La preocupación en los 80 se centró en cuestiones relativas a la llamada calidad de la

enseñanza, en medio de las dificultades financieras que en esa década afectaron a toda América Latina. La preocupación por la calidad de la enseñanza es por una parte una reacción a alarmantes cifras como las de deserción y repetición dentro del sistema escolar, sobre todo en el nivel primario. Por otra parte, es una forma de respuesta al reto de la globalización y la competitividad internacional.

Tabla III.
Porcentaje de analfabetismo en la población de 15 años y más

Año	1950	1960	1970	1980	1990
Porcentaje	44,0	32,5	25,8	16,5	10.4

FUENTE: Nassif R, Rama G, Tedesco J, (1984), El sistema educativo en América Latina, Buenos Aires, Kapeluz

A partir de 1944, la alfabetización entró a formar parte de las prioridades en la política educativa ecuatoriana. Como se recoge en la tabla III. relativa al aumento de las tasas de escolaridad entre 1950 y 1990, el analfabetismo en la población mayor de 15 años descendió desde el 44% (1950) hasta el 16,5% en 1980. La alfabetización en el país se ha caracterizado por la integración a los centros de trabajo industrial y agropecuario, así como por su inclusión en programas de desarrollo comunitario y que abarcan a la vez programas de salud, infraestructura y

otras áreas de desarrollo. Las campañas alfabetizadoras de la década de los años 80 se preocuparon más específicamente de los indígenas. En el campo de la educación, esta cuestión obtuvo su estatuto legal a partir de 1983, año en el que la nueva ley de educación reconoció como objetivo oficial la educación bilingüe intercultural.

Tabla IV.
Cronología de la fundación de las universidades particulares que no reciben subvención del estado hasta el 2000

Universidad	Año	Ciudad
Internacional SEK	1993	Guayaquil
De especialidades espíritu santo	1993	Guayaquil
San Francisco de Quito	1995	Quito
De las Américas	1995	Quito
Escuela politécnica Javeriana del Ecuador	1995	Quito
Internacional del Ecuador	1996	Quito
Regional autónoma de los Andes	1997	Ambato
Tecnológica América Del pacífico escuela de negocios	1997	Quito
	1997	Guayaquil
Tecnológica indoamérica	1998	Ambato
Internacional Jefferson	1999	Guayaquil
Tecnológica San Antonio de Machala	1999	Machala
Casa Grande	1999	Guayaquil
Autónoma de Quito	1999	Quito
Tecnológica Israel	1999	Quito
Tecnológica empresarial de Guayaquil	2000	Guayaquil
De especialidades turísticas	2000	Quito
Cristiana latinoamericana	2000	Quito
Metropolitana	2000	Guayaquil

FUENTE Secretaría general del CONUEP

En ese mismo año se introdujo una reforma a la constitución política de 1979 en lo relativo a la educación bilingüe, definiéndola en los siguientes términos: "En los sistemas de educación que se desarrollen en las zonas de predominante población indígena, se utilizará como lengua principal de educación el quichua o la lengua de la cultura respectiva; y el castellano, como lengua de relación intercultural" (Art. 27 de la constitución de la república).

Durante los años 90, se observa un incremento en la fundación de universidades, especialmente comienza el establecimiento de universidades particulares que no reciben subvención del estado, las universidades creadas de este tipo se muestran en la tabla IV.

1.7 El sistema educativo Ecuatoriano

La educación ecuatoriana de acuerdo a la actual constitución de la república "es un derecho irrenunciables de las personas, deber inexcusable del estado... área prioritaria de la inversión pública, requisito del desarrollo nacional...". Este sistema consta de dos subsistemas el del Ministerio de educación y el Universitario. De acuerdo a lo expuesto en el libro Sistemas educativos nacionales, del Ministerio de Educación y

Cultura, el sistema educativo del Ministerio de Educación está conformado por dos subsistemas el escolarizado y el no escolarizado. El sistema escolarizado comprende la educación que se imparte en la Ley y en los reglamentos generales y especiales; abarca la educación regular, la educación compensatoria y la educación especial.

La educación Regular se desarrolla en un proceso continuo, a través de los siguientes niveles:

- Básica: integrada por la pre-primaria, primaria y ciclo básico.
- Medio: integrado por el ciclo diversificado y de especialización
- Superior sujeto a leyes especiales

Además la educación regular se somete a las disposiciones reglamentarias como el límite de edad, secuencia y duración de niveles y cursos. Los establecimientos de educación regular se denominan jardín de infantes, escuela, colegio, instituto pedagógico y instituto técnico

La educación compensatoria tiene la finalidad de establecer la igualdad de oportunidades para quienes no ingresaron a los niveles de educación regular o no los concluyeron, permite el ingreso al sistema regular en cualquier momento de acuerdo a sus necesidades y aspiraciones. La educación compensatoria comprende:

- Nivel Primario compensatorio
- Ciclo básico compensatorio
- Ciclo diversificado compensatorio y
- Formación y capacitación artesanal

La educación especial atiende a personas excepcionales que por diversas causas no pueden adaptarse a la educación regular. La educación no escolarizada favorece la realización de estudios fuera de las instituciones educativas, sin requisito de currículo académico.

Las instituciones educativas, tienen como misión la formación humana y la promoción cultural, y están destinadas a cumplir con los fines estipulados en la Ley y sus Reglamentos. Las instituciones educativas se clasifican de la siguiente forma:

Por el financiamiento

- Oficiales: fiscales, municipales y de otras instituciones públicas
- Particulares: pertenecen a personas naturales o jurídicas de derecho privado pueden ser laicos o confesionales.
- Otros: los que cuentan con financiamiento parcial de entidades públicas y privadas

Por las jornadas de trabajo

- Matutinos
- Vespertinos
- Nocturnos
- De doble jornada

Por el alumnado

- Masculinos
- Femeninos
- Mixtos

Por la ubicación geográfica

- Urbanos
- Rurales

Fuente: Ministerio de Educación y Cultura

1.8 Educación Rural en el Ecuador

Esta área de educación se encuentra bajo la responsabilidad del departamento de educación rural, adscrito a la dirección nacional de educación regular y especial del ministerio de educación y cultura. Fue creada con el objetivo fundamental de aplicar una política educacional acorde con la problemática y las necesidades del campesino

ecuatoriano. La educación rural tiene tres campos de acción que son la educación comunitaria, la nuclearización y los proyectos comunitarios.

La educación comunitaria esta orientada al desarrollo educativo de la población rural en el marco del avance tecnológico y científico mundial. La nuclearización es un programa educativo que se fundamenta en la participación de todos los componentes de la realidad campesina. Los proyectos comunitarios se generan con el apoyo de la labor de los otros componentes de la educación rural con el propósito de cumplir con los objetivos de la comunidad.

1.9 Datos estadísticos sobre la educación en el Ecuador

La falta de presupuesto del estado para la educación constituye un factor importante en la calidad de esta. De acuerdo a la Constitución de la Republica vigente el 30% del presupuesto debe ser destinado a la educación. Sin embargo esta ha ido decreciendo como se observa en la tabla III.

Este factor ha precipitado la crisis de la educación, la que se ve reflejada en la paralización de la construcción de instituciones públicas, reparación y mantenimiento de los establecimientos existentes, remuneración de los maestros, etc. Paralelamente a esto la tendencia mundial ha sido elevar la inversión en la educación, en vista de la creciente demanda de mano

Tabla V.
Porcentaje del presupuesto del estado destinado para la educación

Año	Porcentaje	Monto \$
1986	29,9	55912
1990	16,5	223616
1991	19,7	357202
1992	20,1	605077
1993	19,6	860200
1994	18	1114500
1995	16,2	1316601
1997	12,8	2445500
1998	14,2	3345000
1999	9,2	3558500

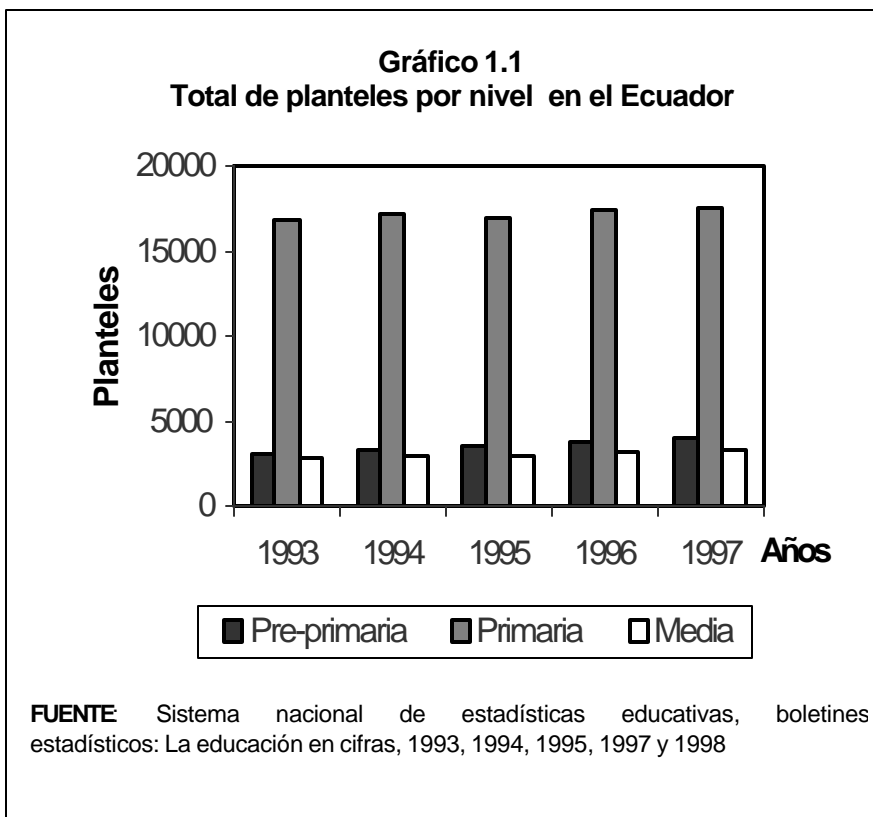
FUENTE Presupuesto General del Estado

de obra calificada, mientras tanto el estado ecuatoriano la reduce, apenas destina el 2,3% del PIB a este sector, el cual representa uno de los niveles más bajos de América Latina. De acuerdo a proyecciones obtenidas del informe social del Ecuador para que a finales del 2010 se pueda mejorar el nivel educativo de la nación por lo menos hasta el

noveno año de educación básica se necesitaría elevar la inversión en educación hasta el 8.4% del PIB. De igual forma en este informe se señala: "En los últimos 25 años podría explicarse la pobreza semiestructural de América Latina por la insuficiente educación que recibieron los jóvenes que ingresaron al mercado de trabajo...", pues hay una fuerte correlación entre el grado de educación y el nivel de desarrollo de un país; "El aumento del ingreso tiende a ser mayor ante mejoramientos del nivel de la educación..." según el informe del ILDIS.

A continuación se realiza un estudio de los datos obtenidos de los boletines estadísticos: La educación en cifras del sistema nacional de los años 1993, 1994, 1995, 1996, 1997 y 1998.

El número de planteles por nivel se puede apreciar en el gráfico 1.1, en promedio desde el año 1993 hasta 1997 la variación porcentual del número de planteles pre-primarios de año a año es del 7%, la de la primaria el 1.1% y la de nivel medio de 3.18%, de acuerdo a los valores que se muestran en la tabla IV.



En promedio los porcentajes de los planteles de los niveles pre primario, primario y medio son respectivamente 14.8%, 72.4% y 12.7% con respecto al total de planteles educativos en el país, de acuerdo a los datos que se muestran en la tabla VI de los años 1993 a 1997. Al comparar estos porcentajes respectivamente con los de profesores y alumnos se observa que a pesar de tener el más bajo porcentaje de establecimientos, el nivel medio tiene un número de profesores similar al

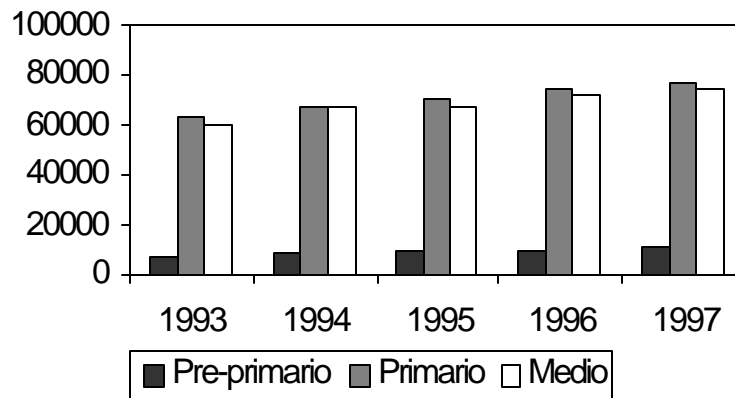
del nivel primario, el cual tiene el mayor porcentaje de planteles educativos.

Tabla VI.
Total de planteles por nivel en el Ecuador

Año	Pre-primaria	Primaria	Media
1993	3060	16825	2868
1994	3335	17194	2909
1995	3487	16974	2976
1996	3723	17367	3144
1997	4009	17569	3249

FUENTE Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Gráfico 1.2
Total de profesores por nivel en el Ecuador



FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

El número de profesores para los niveles primario y medio es similar considerando la diferencia, que existe entre el número de planteles para estos niveles. La variación porcentual promedio para los niveles pre-primario, primario y medio es respectivamente 11.92%, 4.74% y 5.97% del número de profesores, como se muestra en el gráfico 1.2.

Tabla VII.
Total de profesores por nivel en el Ecuador

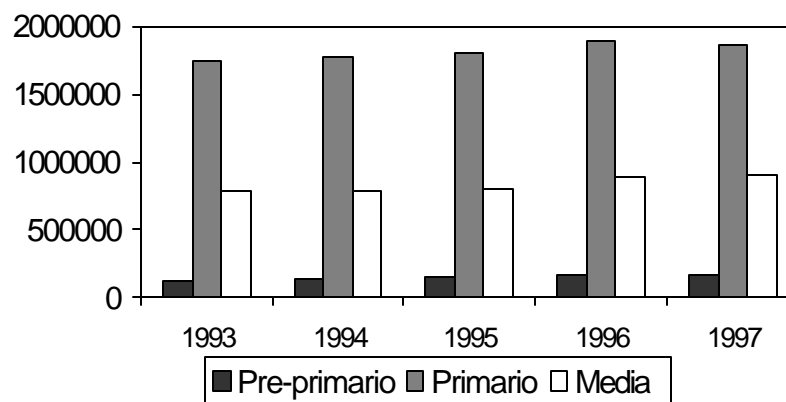
Año	Preprimaria	Primaria	Media
1993	7020	63708	59449
1994	8246	67446	67446
1995	9296	70162	67414
1996	9980	74601	72099
1997	10992	76642	74639

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Los porcentajes promedios en los años 1993 al 1997 del número de profesores por cada nivel con respecto al total son para el pre primario 5.2%, el primario 64.9% y para el medio 29.8%. Los valores que corresponden a estos porcentajes se muestran en la tabla VII.

El número de alumnos de nivel primario es considerablemente mayor que los de pre-primario y medio, como se lo muestra en el gráfico 1.3, en los tres niveles se puede apreciar un aumento de estudiantes, en promedio el porcentaje de aumento para los tres niveles son 7.1%, 1.8% y 3.9%.

Gráfico 1.3
Total de alumnos por nivel en el Ecuador



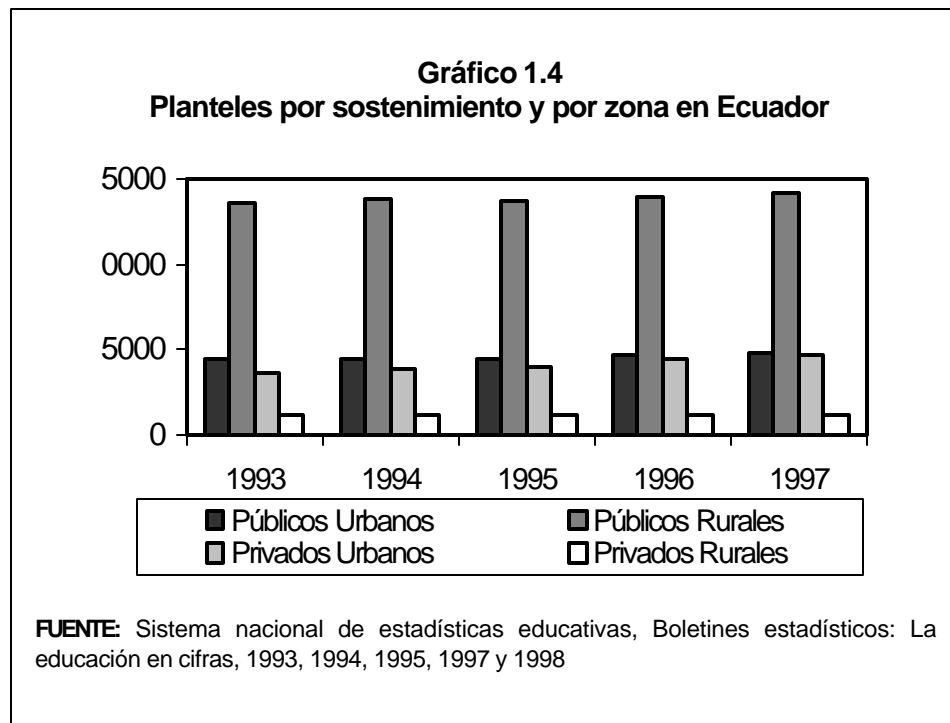
FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Tabla VIII.
Total de alumnos por nivel en el Ecuador

Año	Preprimaria	Primaria	Media
1993	127355	1742984	785522
1994	136158	1777304	787290
1995	145858	1812255	806096
1996	156772	1888172	895711
1997	167582	1873349	911572

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

En lo que se refiere al porcentaje del número de alumnos por nivel con respecto al total del país en promedio para los años de 1993 a 1997 se tiene en el nivel pre primario 5.2%, en el nivel primario 64.9% y en el medio 29.8%, estos valores se los obtiene de los datos que se muestran en la tabla VIII.



La clasificación del tipo de planteles por sostenimiento se refiere a: los establecimientos educativos que son financiados por el estado y los que son financiados por fondos privados, además se pueden clasificar por la zona en que están ubicados como rurales y urbanos. En el gráfico 1.4 se puede apreciar claramente que en todo el país los colegios públicos

rurales son desde el año 1993 a 1997 aproximadamente 15000 con un porcentaje promedio de crecimiento del 1%, mientras que los demás durante este mismo período no sobrepasan los 5000 establecimientos.

Tabla IX.
Planteles por sostenimiento y por zona en Ecuador

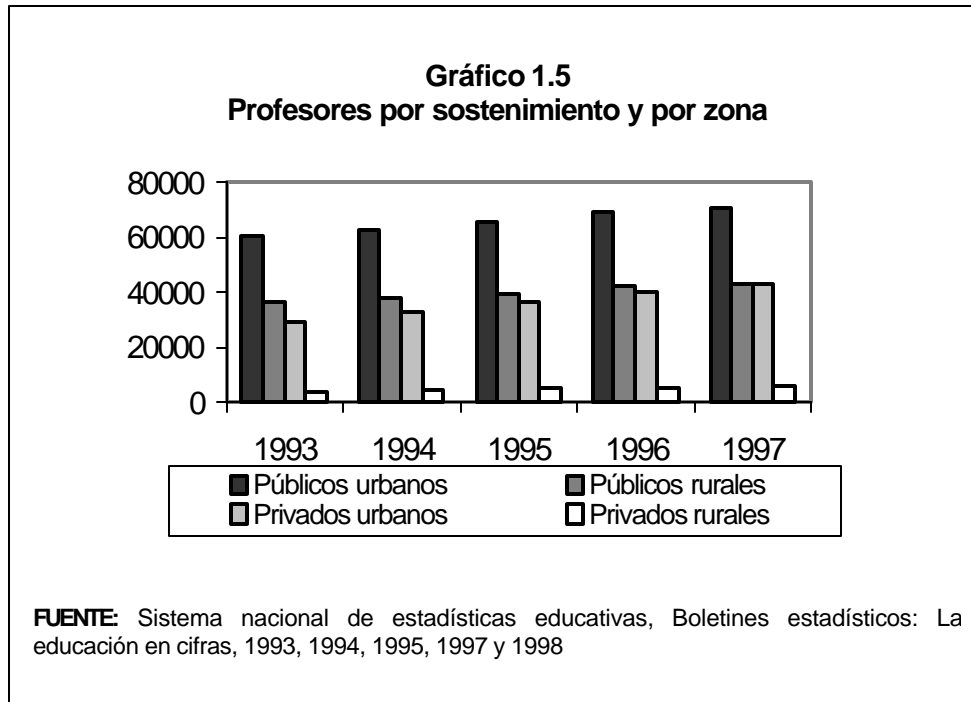
Año	Públicos		Privados	
	Urbanos	Rurales	Urbanos	Rurales
1993	4449	13548	3625	1151
1994	4482	13837	3867	1252
1995	4497	13659	4013	1278
1996	4682	13903	4429	1220
1997	4819	14141	4727	1204

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

En la tabla IX se puede apreciar los valores totales de los tipos de planteles: públicos urbanos y rurales, privados urbanos y rurales. El porcentaje promedio de estos tipos de establecimientos educativos con respecto al total en los años que se muestran en la tabla son respectivamente 19.3%, 58.1%, 17.3% y 5.1%.

El número de profesores por sostenimiento y por zona como se puede apreciar en el gráfico 1.5 es mayor para el sector público urbano con un crecimiento promedio entre los años 1993 a 1997 del 3.9%. Al comparar el número de establecimientos educativos con el de profesores por

sostenimiento y por zona de las tablas IX y X se obtiene que existen en promedio 14.33 y 2.8 profesores por cada plantel público urbano y rural, y 8.7 y 4.1 profesores por cada plantel privado urbano y rural.



Estos resultados muestran que a pesar de que el número de establecimientos públicos rurales en todo el país es aproximadamente 3.01 veces mayor que los públicos urbanos y la relación del número de profesores en estos tipos de planteles es de 1.65 veces más profesores públicos urbanos que rurales, la tasa de 2.8 de profesores por planteles públicos urbanos es la más baja.

Tabla X.
Profesores por sostenimiento y por zona en Ecuador

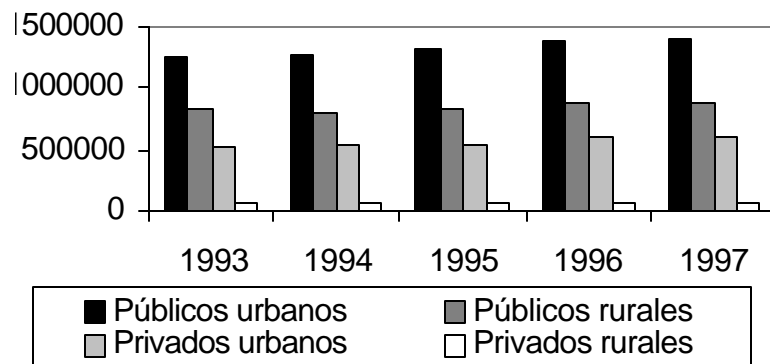
Año	Públicos		Privados	
	Urbanos	Rurales	Urbanos	Rurales
1993	60430	36651	29360	3736
1994	63143	37689	33043	4643
1995	65782	39634	36342	5114
1996	69113	42158	39950	5459
1997	70435	42914	42820	6104

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

El porcentaje promedio del número de estudiantes con respecto al total del país de los planteles públicos urbanos es de 47.3%, seguido por los públicos rurales con 30.1%, los privados urbanos con 19.9% y los privados rurales con apenas 2.5%. En el gráfico 1.6 podemos apreciar esta distribución.

A partir de los datos que se muestran en las tablas X y XI se obtienen las tasas del número de estudiantes promedio por cada profesor en los tipos de establecimientos que se están analizando, los resultados de este análisis son: que existen 20.1, 21.2, 15.5, 14.5 estudiantes por cada profesor en los colegios públicos urbanos, públicos rurales, privados urbanos y privados rurales respectivamente.

Gráfico 1.6
Alumnos por sostenimiento y por zona en Ecuador



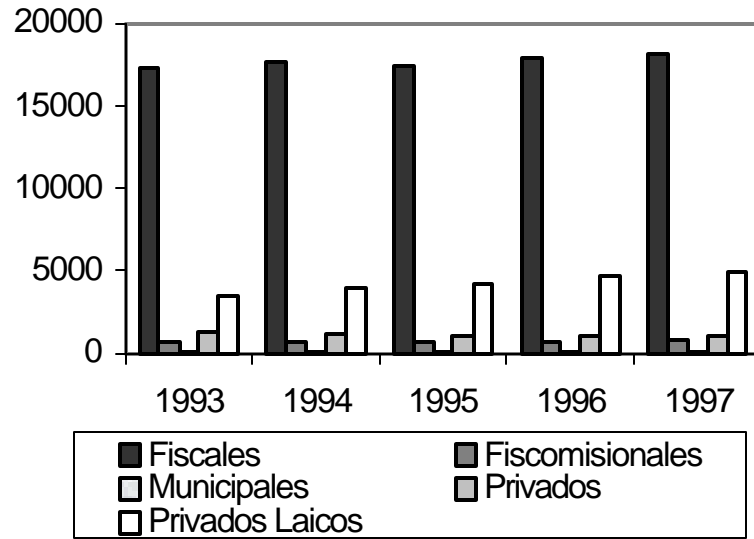
FUENTE Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Tabla XI.
Alumnos por sostenimiento y por zona en Ecuador

Año	Públicos		Privados	
	Urbanos	Rurales	Urbanos	Rurales
1993	1245247	824545	513742	65095
1994	1266940	816277	546869	70566
1995	1311868	834766	544394	73181
1996	1397456	874225	595373	73601
1997	1407951	865855	593425	75278

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Gráfico 1.7
Planteles por sostenimiento en Ecuador



FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Una clasificación más detallada de los tipos de planteles por sostenimiento se muestra en el gráfico 1.7, en el que se aprecia una marcada diferencia de los colegios fiscales sobre los demás del país los cuales en promedio representan 73.98%. Los colegios fiscomisionales y municipales que en general son considerados dentro de la clasificación de fiscales representan conjuntamente en promedio 3.05% del total de planteles. Por otro lado el porcentaje de crecimiento promedio que han tenido durante los años 1993 a 1997 los colegios de tipo privado son:

para los privados religiosos decrecimiento del -5.4%, y para los privados laicos el crecimiento es del 8.9%.

Tabla XII
Planteles por sostenimiento en Ecuador

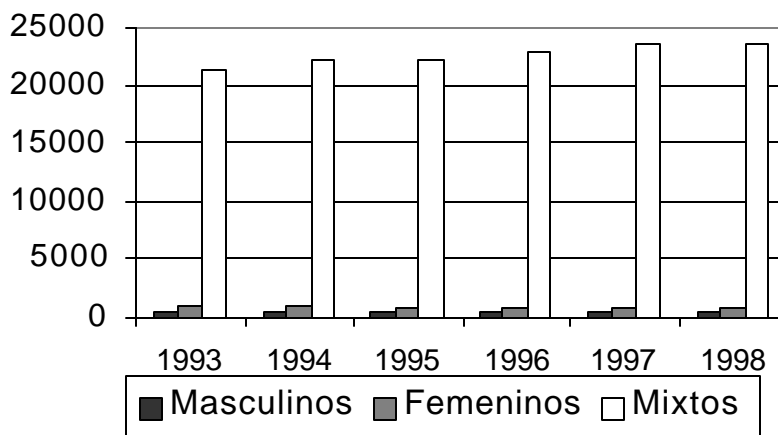
Año	Fiscal	Fiscomisional	Municipal	Privado Religioso	Privado Laico
1993	17278	643	56	1300	3476
1994	17627	634	58	1217	3902
1995	17446	646	54	1065	4226
1996	17827	701	57	1033	4616
1997	18162	733	66	1037	4883

FUENTE Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

En la tabla XII se puede observar los valores que corresponden al número de colegios clasificados por sostenimiento y de acuerdo al análisis hecho anteriormente se puede comprobar a simple vista el decrecimiento de los colegios privados religiosos, además aunque el crecimiento porcentual de los colegios privados laicos es mayor que el de los fiscales la cantidad es similar.

La clasificación de planteles educativos por género se muestra en el gráfico 1.8, se puede apreciar claramente que los establecimientos de tipo mixtos que representan en promedio el 94.6% superan ampliamente a los masculinos y femeninos, con un 2% y 3.3% respectivamente.

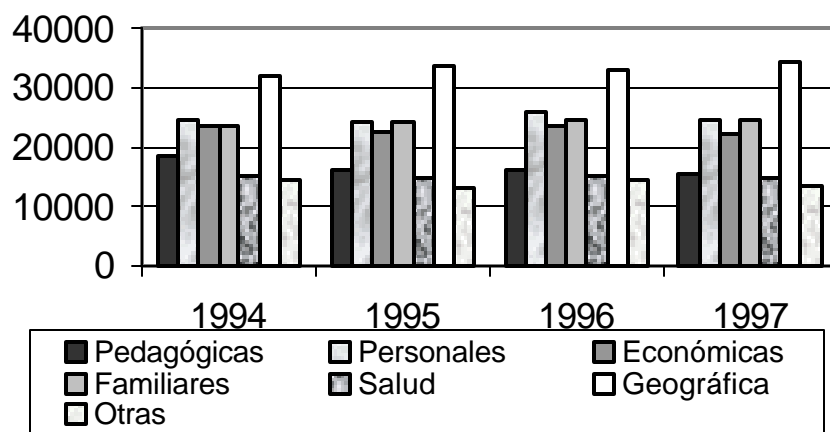
Gráfico 1.8
Planteles por género en Ecuador



FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Las causas de deserción de los estudiantes de los planteles educativos durante su curso escolar se han agrupado como se muestra en el gráfico 1.9 en siete causas, al nivel de todo el país la causa de deserción más común es la geográfica que en promedio desde 1994 a 1997 es 22.1%, seguida por las personales, económicas y familiares con 16.5%, 15.25 y 16.1% respectivamente. Las causas pedagógicas, salud y otras con porcentajes de 11%, 9.9% y 9.1% respectivamente representan tienen una menor incidencia en la deserción estudiantil.

Gráfico 1.9
Causas de deserción en Ecuador



FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Tabla XIII
Causas de deserción en Ecuador

Año	Pedagógicas	Personales	Económicas	Familiares	Salud	Geográfica	Otras	Total
1994	18499	24662	23403	23634	15032	31845	14158	151233
1995	16007	24228	22542	24209	14731	33635	12963	148315
1996	16192	25872	23437	24435	15098	32899	14228	152161
1997	15417	24273	22267	24273	14852	34423	13372	147862

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

En base a los valores de las tablas XIII y XIV se puede calcular los porcentajes promedio del total de estudiantes que desertan de sus cursos a nivel de todo el país el cual es del 5.28%, para la principal

causa la geográfica se obtiene que 1.17 estudiantes de cada 100 desertan del curso.

Tabla XIV
Alumnos por zona en todo el país

Años	Urbanos	Rurales	Total
1994	1813809	886843	2700652
1995	1856262	907947	2764209
1996	1992829	947826	2940655
1997	2001376	941133	2942509
Total	7664276	3683749	11348025
Promedio	1916069	920937,25	2837006,25

FUENTE Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras. 1993. 1994. 1995. 1997 v 1998

En la zona urbana de todo el país como se puede apreciar en el gráfico 1.10 las mayores causas de deserción de los estudiantes son las geográficas y las personales las que en promedio representan el 18.61% y el 17.65% respectivamente.

El porcentaje del total de estudiantes de las zonas urbanas de todo el país que desertan de su curso escolar se lo obtiene basándose en los valores de las tablas XIV y XV el cual es del 5.23% y para las principales causas de deserción como las de tipo geográfico, personales y económicas son del 0.97%, 0.92% y 0.85% respectivamente; además el total de las causas de deserción en la zona urbana con respecto a todo

el país es 353% lo que significa que de cada 100 estudiantes en todo el país existen 3.53 que desertan en la zona urbana.

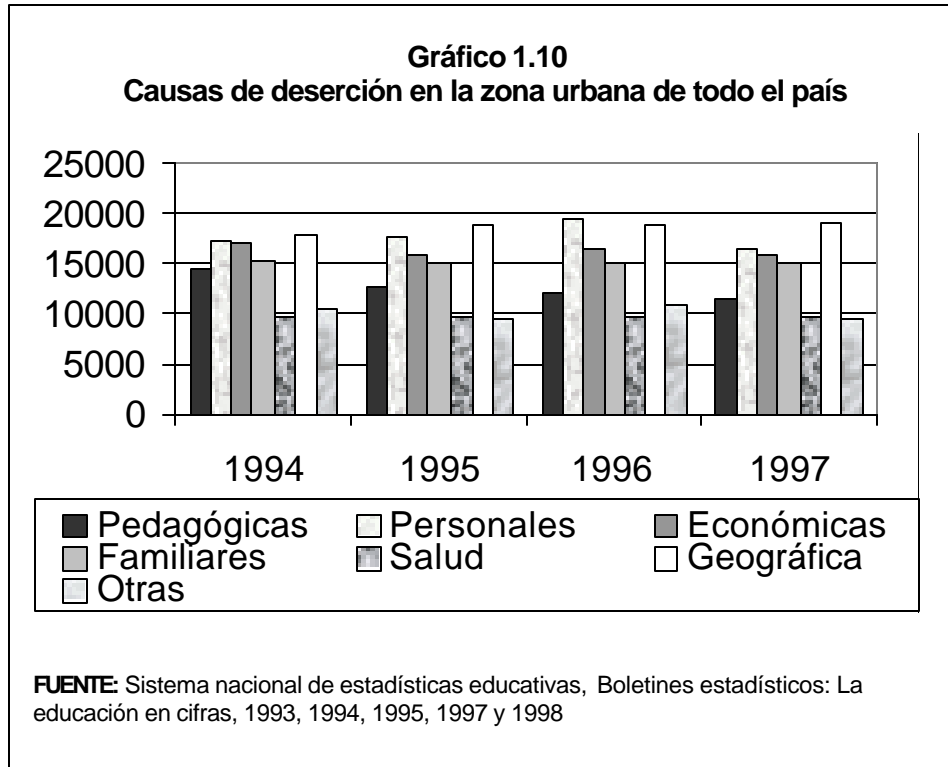
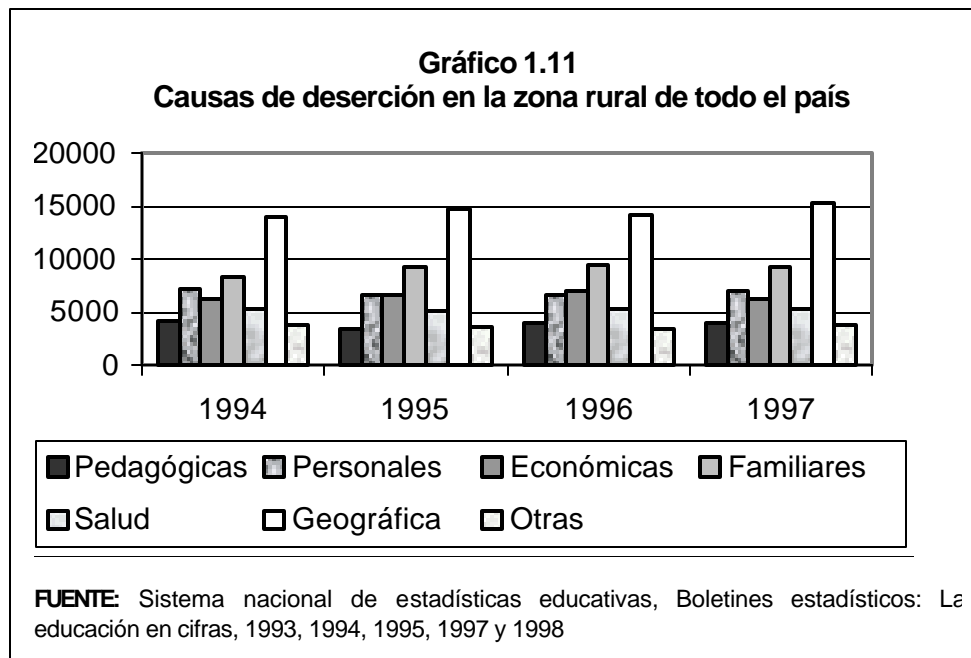


Tabla XIV.
Causas de deserción en la zona urbana de todo el país

Año	Pedagógicas	Personales	Económicas	Familiares	Salud	Geográfica	Otras	Total
1994	14313	17484	17103	15348	9773	17909	10453	102383
1995	12614	17705	15934	15013	9645	18855	9366	99132
1996	12180	19286	16424	15002	9820	18716	10746	102194
1997	11555	16342	15945	15096	9630	19077	9534	97179

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

En la deserción estudiantil de la zona rural se destaca como principal causa la geográfica, como se puede observar en el gráfico 1.11, que en promedio es del 29.31%, seguida por las familiares con un 18.15%. Comparando los valores de la tabla XIV y XVI podemos obtener los porcentajes de deserción de la zona rural con respecto al total de alumnos en esta zona y a todo el país estos son respectivamente el 5.39% y el 1.75%.



La no promoción de los estudiantes matriculados en un curso al siguiente, es una parte muy importante dentro del análisis de la educación del país, concretamente para el desarrollo de esta tesis se estudiarán los datos que corresponden a la no promoción de los

estudiantes de nivel medio de todo el país y clasificados por zona. Por el contrario de las causas de deserción de los estudiantes en las que se

Tabla XVI
Causas de deserción en la zona rural de todo el país

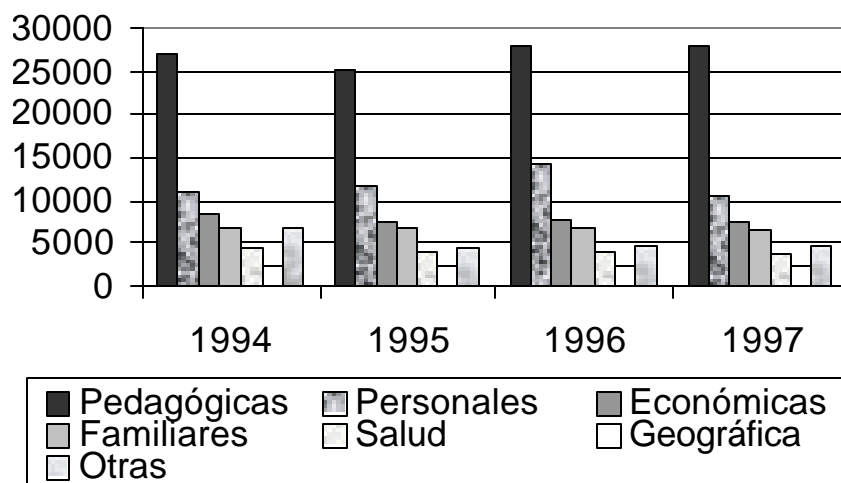
Año	Pedagógicas	Personales	Económicas	Familiares	Salud	Geográfica	Otras	Total
1994	4186	7178	6300	8286	5259	13936	3705	48850
1995	3393	6523	6608	9196	5086	14780	3597	49183
1996	4012	6586	7013	9413	5278	14183	3482	49967
1997	3862	6916	6322	9177	5222	15346	3838	50683

FUENTE tabla elaborada por la autora, en base a los datos presentados por Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

destacan principalmente las de tipo geográfico y personales en las causas de no promoción en todo el país como se puede observar en el gráfico 1.12 sobresalen las causas pedagógicas y las personales que representan en promedio con respecto al total de casos de no promoción el 41.4% y 18.2% respectivamente.

El total de alumnos de nivel medio no promovidos al año siguiente de estudio con respecto al total de los alumnos en todo el país los cuales se muestran en la tabla XII representa el 2.31%, este porcentaje contiene a las principales causas de no promoción las causas pedagógicas y las personales con .96% y 0.42% cada una.

Gráfico 1.12
Causas de no promoción en el nivel medio en todo el país



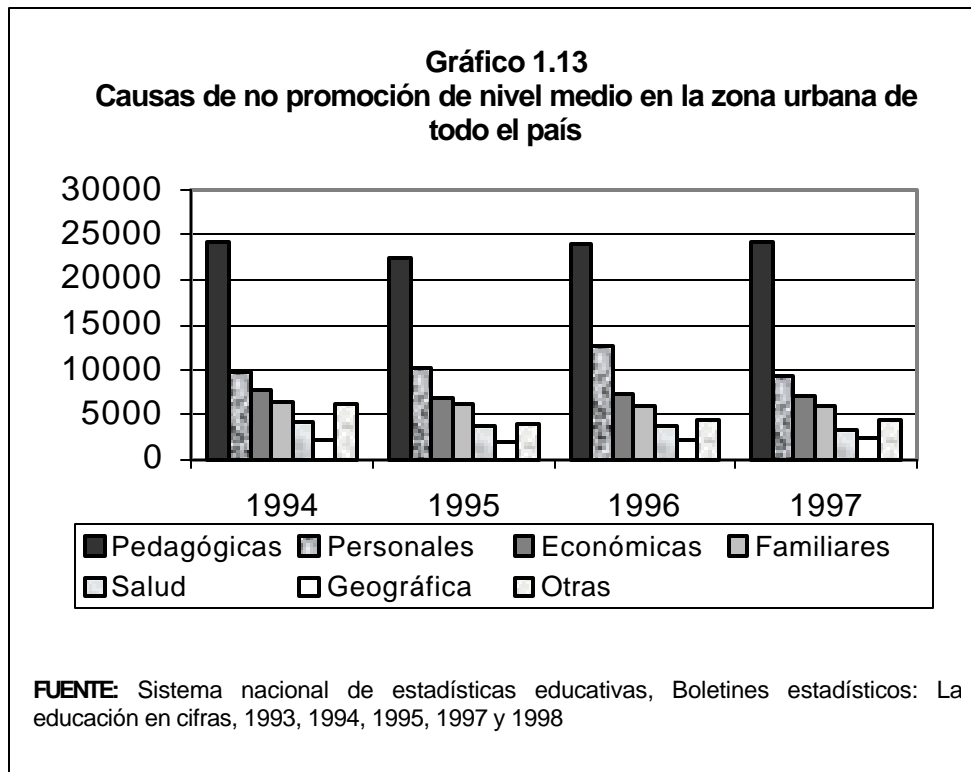
FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Tabla XVII
Causas de no promoción de nivel medio en todo el país

Año	Pedagógicas	Personales	Económicas	Familiares	Salud	Geográfica	Otras	Total
1994	27071	11065	8518	6967	4490	2503	6664	67278
1995	25396	11544	7555	6899	3962	2289	4452	62097
1996	27923	14417	7939	6662	4134	2515	4852	68442
1997	28081	10709	7659	6525	3699	2486	4818	63977

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Con respecto a la no promoción en la zona urbana se puede apreciar en el gráfico 1.13 que se mantiene la tendencia de las principales causas como a nivel de todo el país; del total de las causas de no promoción del nivel medio de la zona urbana las de tipo pedagógico representan en promedio para los años que se consideran en la tabla XVII el 40.5%, para las de tipo personal el 18.03% y las económicas el 12.4%.



Con respecto al número total de estudiantes de nivel medio en el país, el total de alumnos no promovidos en la zona urbana en este mismo nivel representa el 6.94%, dentro de este porcentaje las principales causas de no promoción tienen una participación de: 2.83% las pedagógicas, 1.25%

las personales y 0.86% las económicas, para obtener estos resultados comparamos los valores de la tabla VIII y los de la tabla XVIII. De igual modo se puede comparar cuanto representa el total de alumnos no promovidos de nivel medio de la zona urbana de todo el país con el total de estudiantes matriculados en esta zona en todo el país que en promedio sería el 3.03%.

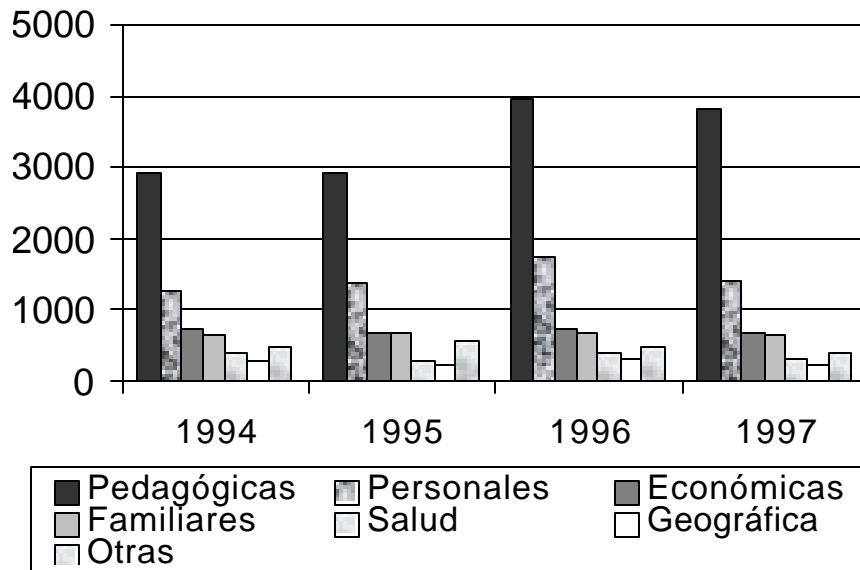
Tabla XVIII
Causas de no promoción de nivel medio en la zona urbana de todo el país

Año	Pedagógicas	Personales	Económicas	Familiares	Salud	Geográfica	Otras	Total
1994	24161	9809	7784	6327	4083	2226	6181	60571
1995	22480	10144	6863	6191	3678	2027	3901	55284
1996	23966	12683	7194	5958	3731	2198	4380	60110
1997	24255	9303	6983	5856	3360	2257	4409	56423

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Las causas de no promoción de los estudiantes de la zona rural se puede observar en el gráfico 1.14 y los valores en la tabla XIX, en esta zona las principales causas son : las pedagógicas y personales que representan el 46.2% y 19.7%, y con respecto al total de estudiantes matriculados en esta zona las causas e no promoción representan el 0.79%, por otro lado la relación de estas causas con respecto al total de los estudiantes matriculados en el nivel medio es 0.87%.

Gráfico 1.14
Causas de no promoción de nivel medio en la zona rural de todo el país



FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Tabla XIX
Causas de no promoción de nivel medio en la zona rural en Ecuador

Año	Pedagógicas	Personales	Económicas	Familiares	Salud	Geográfica	Otras	Total
1994	2910	1256	734	640	407	277	483	6707
1995	2916	1400	692	708	284	262	551	6813
1996	3957	1734	745	704	403	317	472	8332
1997	3826	1406	676	669	339	229	409	7554

FUENTE: Sistema nacional de estadísticas educativas, Boletines estadísticos: La educación en cifras, 1993, 1994, 1995, 1997 y 1998

Capítulo II

2. Diseño del cuestionario y codificación de variables.

2.1 Introducción

En base a la reforma curricular vigente desde el mes de abril de 1996, elaborada por el Ministerio de Educación y Cultura, la cual contempla los programas de estudio correspondiente a los diferentes niveles de educación, se elaboraron dos cuestionarios, para las asignaturas de matemáticas y lenguaje, los que contienen preguntas formuladas para medir el conocimiento de lo aprendido, desde el primero hasta el noveno año de estudio, de los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil.

2.2 Definiciones

Universo.- Al conjunto bien definido de entes que posee características medibles, las cuales pueden ser cuantitativas o cualitativas se denomina universo. Las características cuantitativas son aquellas que pueden ser expresadas en una escala numérica, mientras que las cualitativas son las que definen si el ente posee o no cierto atributo y a que nivel.

Población.- Dado un espacio muestral (Ω, ξ) , variable aleatoria es la función X que a cada elemento $\omega \in \Omega$, le asigna un valor $X(\omega) \in \mathfrak{R}$ un número real. Si el conjunto de llegada de la función X es finito o infinito contable la variable aleatoria es discreta; Si el conjunto de llegada es un intervalo de números reales o la unión de intervalos entonces la variable aleatoria es continua.

Muestra.- Al conjunto X_1, X_2, \dots, X_n , de tamaño n tomado de una población X de tamaño N se denomina muestra, donde $n \leq N$.

Censo.- Etimológicamente proviene del latín *censum* que significa contar. Es el proceso mediante el cual se mide las características de interés de todas las unidades de investigación que conforman la población, para analizarlas.

2.3 Marco censal

En base a los conceptos definidos en la sección 2.2, en el presente estudio el universo está conformado por todos los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil. En la tabla XX se puede observar los colegios ubicados en la zona rural de Guayaquil y el total de alumnos matriculados en décimo año en estos colegios.

Tabla XX
Ubicación y número de alumnos de los colegios fiscales rurales del cantón Guayaquil.

Colegio	Jornada de trabajo	Estudiantes matriculados en 10^{mo} año	Parroquia	Dirección
Elías Severo Bohórquez	Vespertino	13	Chongón	Recinto P. de Sabana Grande
Luis Fernando Vivero	Matutino	84	Posorja	Km 1 1/2 vía a Data Posorja
Pablo Weber Cubillo	Matutino	35	Juan Gómez Rendón	Km 67 vía a la costa
Primero de Junio	Nocturno	17	Tenguel	Calle L. Febres Cordero
Primero de Junio	Matutino	34	Tenguel	Calle L. Febres Cordero
Puná	Matutino	25	Puná	Barrio lindo

FUENTE: Registros de la dirección provincial de educación del Guayas

El universo de estudio es de tamaño $N= 208$ estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil. Las características de interés que se desean medir son edad, sexo, las correspondientes a la medición del conocimiento en lenguaje y las de la medición del conocimiento en matemáticas de las unidades de investigación definidas en el universo.

Gráfico 2.1
Ubicación geográfica de las parroquias de los colegios fiscales rurales del cantón Guayaquil



Para este estudio en particular resultó conveniente realizar un censo, debido a que el tamaño de la población es pequeño y la posibilidad de entrevistar a todos los estudiantes fue factible, considerando además que el número de colegios en los que estaban repartidos los estudiantes es seis.

La ubicación de los colegios fiscales en las parroquias rurales del cantón Guayaquil se puede observar en el gráfico 2.1. En cada una de las parroquias existe un colegio fiscal, a excepción de la parroquia Tenguel en que se encuentran ubicados dos colegios que funcionan con jornadas matutina y nocturna.

Tabla XXI
Clasificación por género del alumnado de los colegios rurales del cantón Guayaquil

Colegio	Hombres	Mujeres	Total de estudiantes en 10^{mo} año	Total de estudiantes del colegio
Elías Severo Bohórquez	9	4	13	64
Luis Fernando Vivero	42	42	84	388
Pablo Weber Cubillo	24	11	35	189
Primero de Junio	8	9	17	59
Primero de Junio	16	18	34	216
Puná	12	13	25	138
TOTAL	111	97	208	1054

FUENTE Registros de la dirección provincial de educación del Guayas.

El número de estudiantes de sexo masculino y femenino que se encuentran matriculados es de 111 y 97 respectivamente, en la tabla XXI se puede observar esta clasificación por cada colegio, además del total de estudiantes matriculados en el colegio.

El personal que labora en los planteles a investigar se clasifica por los cargos que desempeñan en directivos, profesores, administradores y de servicio, además como se puede apreciar en la tabla XXII se han clasificado por sexo y por colegio.

Tabla XXII
Clasificación por actividad del personal que labora en los colegios
fiscales rurales del cantón Guayaquil

Colegio	Direc.		Prof.		Adm.		Servicio		Total	
	H	M	H	M	H	M	H	M	H	M
Elías Severo Bohórquez	1	0	4	2	0	1	1	0	6	3
Luis Fernando Vivero	1	0	8	6	2	0	2	0	13	6
Pablo Weber Cubillo	3	0	15	4	1	2	5	0	24	6
Primero de Junio	1	0	4	9	1	1	0	0	6	10
Primero de Junio	2	0	8	10	1	1	2	0	13	11
Puná	0	1	6	4	0	1	1	0	7	6
TOTAL	8	1	45	35	5	6	11	0	69	42

FUENTE Registros de la dirección provincial de educación del Guayas.

2.4 Diseño del cuestionario

Para la investigación que se realizará en esta tesis, los cuestionarios que se elaboraron, son similares a los exámenes tomados en los planteles educativos, pero con la diferencia de que estos evaluarán el **conocimiento** que deben tener los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil, de lo aprendido hasta ese año en las asignaturas de matemáticas y lenguaje.

2.4.1 Temas y subtemas para la evaluación

El primer paso que se dio para la elaboración de los cuestionarios, fue el de agrupar los temas y subtemas del contenido de las asignaturas matemáticas y lenguaje, desde primero hasta noveno año de educación básica. La elaboración del cuestionario de matemáticas, se basó en los siguientes temas: operaciones aritméticas, sistema numérico, sistema de funciones, geometría, sistemas de medida y estadística y probabilidad, los subtemas correspondientes a estas áreas se muestran en la tabla XIX. En lo que respecta al cuestionario de lenguaje, se elaboraron quince preguntas para evaluar los conocimientos de los estudiantes en: morfosintaxis, fonología, vocabulario, lectura comprensiva, ortografía,

Tabla XXIII
Temas y subtemas para la evaluación de matemáticas

Temas	Subtemas	Contenido
Operaciones aritméticas	Suma	Números enteros y fracciones
	Resta	Números enteros y fracciones
	Multiplicación	Números enteros y fracciones
	División	Números enteros y fracciones
Sistema numérico	Características	Diferenciación de números naturales, enteros, racionales, irracionales y reales
	Orden	Valor absoluto Relaciones de orden
	Operaciones	Potenciación y radicación
	Divisibilidad	Múltiplos y divisores
	Proporcionalidad	Razones y proporciones Proporcionalidad directa e inversa Regla de tres simple
Sistema de funciones	Lógica	Proposiciones simples Estado de verdad
	Conjuntos	Noción de conjuntos y elementos Representación Subconjuntos Operaciones de conjuntos
	Funciones	Nociones básicas Evaluación de funciones Factorización Ecuación lineal con una incógnita
Geometría y sistema de métrico	Polígonos	Perímetro y áreas
	Círculo	Área del círculo
	Triángulo rectángulo	Teorema de Pitágoras
	Sistema métrico	Conversión de unidades
Sistema de estadística y probabilidad	Estadística	Medida de tendencia central: la media
	Probabilidad	Noción de probabilidad

FUENTE: Marcillo S, 1998, Programa curricular institucional, Quito - Ecuador, Ministerio de Educación y Cultura.

semántica; la tabla XXIV muestra los contenidos de cada tema seleccionado para la evaluación.

Tabla XXIV
Temas y subtemas para la evaluación de lenguaje

Temas	Subtemas	Contenido
Morfosintaxis	Oración	Noción general Concordancia Tipo de oraciones Oraciones simples
	Forma y función de las palabras en la oración	Noción básica Clasificación semántica de sustantivos, adjetivos, verbos y artículos Concordancia
	Verbo	Noción básica Concordancia de tiempo y persona
Fonología	Acento de las palabras	Identificación de palabras agudas, graves y esdrújulas
	Unión de vocales en las palabras	Identificación de hiatos, diptongos y triptongos.
Vocabulario	Significado de palabras	Definir el significado de las palabras
	Antónimos Sinónimos	Identificación de antónimos Identificación de sinónimos
Lectura	Lectura comprensiva	Preguntas en base a la lectura
Ortografía	Uso de la tilde	En palabras agudas, graves, esdrújulas y sobresdrújulas.

FUENTE: Marcillo S,1998, Programa curricular institucional, Quito - Ecuador, Mnisterio de Educación y Cultura.

Una vez que se establecieron los temas y subtemas a evaluar, se procedió a elaborar cuestionarios de prueba, con preguntas que estén

relacionadas a los temas seleccionados. Un punto muy importante, que se debió considerar para la formulación de las preguntas, fue la duración que debía tener cada prueba, la cual no podía exceder los 60 minutos. Para la calificación de las pruebas, se consideró ponderar de acuerdo el grado de dificultad y de importancia, el puntaje de cada pregunta, tomando en cuenta que el puntaje total de cada cuestionario es de cien puntos.

Los cuestionarios elaborados para la evaluación de los estudiantes en matemáticas y lenguaje, basándose en lo descrito anteriormente se muestran en el anexo 1.

2.5 Definición de variables

La definición de las variables de estudio, es una descripción de lo que se desea medir; en esta definición se establece el nombre de cada variable y la codificación de los valores que puede tomar la variable, los códigos se utilizarán en la etapa del análisis estadístico.

2.5.1 Variables generales

Las variables denominadas generales utilizadas en este análisis son: la edad y el sexo del estudiante.

Variable 1: Edad

Con esta variable se identifica la edad de los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil entrevistados el 12 de Diciembre del 2000 para conocer si corresponde, con la que se espera que tengan al estar cursando el décimo año de educación básica, o si existe alguna relación entre la edad con la nota obtenida.

Codificación:

Es de tipo numérica

Variable 2: Sexo

Por medio de esta variable se desea identificar el sexo de los estudiantes evaluados para poder obtener proporciones de resultados de acuerdo a su género.

Sexo	Codificación
Femenino	0
Masculino	1

2.5.2 Variables de la prueba de matemáticas

Para la prueba de matemáticas las variables definidas son 31, de las cuales 30 conciernen específicamente a cada uno de los temas evaluados y una que es la calificación total del estudiante en esta prueba.

Variable 3: Suma de números enteros

Esta variable mide el conocimiento de los estudiantes para efectuar una suma de números enteros.

Codificación:

Realizó la suma incorrectamente:	0
Realizó correctamente sólo la suma de unidades:	1
Realizó correctamente la suma de unidades y decenas:	2
Realizó correctamente la suma de unidades, decenas y centenas:	3
Realizó correctamente la suma:	4

Variable 4: Suma de números racionales

Se utiliza esta variable para evaluar si los estudiantes saben efectuar una suma de números racionales.

Codificación:

Incorrecta resolución de la suma de números racionales:	0
---	---

Correcta resolución de la suma de números racionales: 1

Variable 5: Resta de números enteros

Mediante esta variable se determina el conocimiento de los estudiantes para efectuar una resta de números enteros.

Codificación:

Realizó la resta incorrectamente:	0
Realizó correctamente sólo la resta de unidades:	1
Realizó correctamente la resta de unidades y decenas:	2
Realizó correctamente la resta de unidades, decenas y centenas:	3
Realizó correctamente la resta:	4

Variable 6: Resta de números racionales

Para evaluar si los estudiantes saben efectuar una operación de resta de números racionales, se utiliza esta variable.

Codificación:

Incorrecta resolución de la resta de números racionales:	0
Correcta resolución de la resta de números racionales:	1

Variable 7: Multiplicación de números enteros

La determinación del conocimiento de los estudiantes para efectuar una multiplicación de números enteros, se la realiza con esta variable.

Codificación:

Realizó incorrectamente la multiplicación:	0
Realizó correctamente la multiplicación, sólo por la primera cifra:	1
Realizó correctamente la multiplicación, sólo por la segunda cifra:	2
Realizó correctamente toda la multiplicación:	3

Variable 8: Multiplicación de números racionales

Por medio de esta variable se evalúa si los estudiantes saben efectuar una multiplicación de números racionales.

Codificación:

Realizó incorrectamente la multiplicación de números racionales:	0
Realizó correctamente la multiplicación de números racionales:	1

Variable 9: División de números enteros

Se emplea esta variable para reconocer si los estudiantes saben efectuar una multiplicación de números enteros.

Codificación:

Realizó incorrectamente la división:	0
Realizó correctamente la división:	1

Variable 10: División de números racionales

La definición de esta variable es para obtener una medida del conocimiento de los estudiantes para realizar una multiplicación de números racionales.

Codificación:

No realizó correctamente la división de números racionales: 0

Realizó correctamente la división de números racionales: 1

Variable 11: Definiciones de conjunto de números

Esta variable mide el grado de conocimiento que tienen los estudiantes sobre la clasificación de los números.

Codificación:

Número de definiciones correctas: 0, 1, 2, 3, 4.

Variable 12: Ejercicio de valor absoluto

Se utiliza esta variable para evaluar si los alumnos saben determinar el conjunto solución de una expresión que contenga el valor absoluto de números reales.

Codificación:

Sin respuesta: 0

Respuesta incorrecta: 1

Descomponen el valor absoluto pero no resuelven correctamente 2

Variable 13: Número de relaciones de orden correctamente colocados

Mediante esta variable se determina el conocimiento de los estudiantes con respecto a establecer relaciones de orden de números reales.

Codificación:

Número de relaciones de orden correctas: 0, 1, 2, 3, 4.

Variable 14: Ejercicios de potenciación y radicación correctos

Para evaluar el conocimiento de los estudiantes acerca de las propiedades de potenciación y radicación de números reales, se utiliza esta variable.

Codificación:

Número de ejercicios correctos: 0, 1, 2, 3, 4.

Variable 15: Ejercicios de divisibilidad correctos

La determinación del conocimiento de los estudiantes para resolver ejercicios de divisibilidad, se la realiza utilizando esta variable.

Codificación:

Número de respuestas correctas: 0, 1, 2, 3.

Variable 16: Planteamiento y resolución de una ecuación lineal con una incógnita

Por medio de esta variable se analiza si los estudiantes tienen nociones para plantear y resolver una ecuación lineal con una incógnita.

Codificación:

Sin respuesta:	0
Problema no planteado y mal resuelto:	1
Problema no planteado y bien resuelto:	2
Problema mal planteado y mal resuelto:	3
Problema mal planteado y bien resuelto:	4
Problema bien planteado y mal resuelto:	5
Problema bien planteado y bien resuelto:	6

Variable 17: Planteamiento y resolución de un problema de interés simple

Esta variable mide el conocimiento de los estudiantes para plantear y resolver un problema de interés simple.

Codificación:

Sin respuesta:	0
Problema no planteado y mal resuelto:	1
Problema no planteado y bien resuelto:	2
Problema mal planteado y mal resuelto:	3

Problema mal planteado y bien resuelto:	4
Problema bien planteado y mal resuelto:	5
Problema bien planteado y bien resuelto:	6

Variable 18: Planteamiento y resolución de un ejercicio de regla de tres

La presente variable se utiliza para determinar si los estudiantes tienen nociones para plantear y resolver un ejercicio de regla de tres.

Codificación:

Sin respuesta:	0
Problema no planteado y mal resuelto:	1
Problema no planteado y bien resuelto:	2
Problema mal planteado y mal resuelto:	3
Problema mal planteado y bien resuelto:	4
Problema bien planteado y mal resuelto:	5
Problema bien planteado y bien resuelto:	6

Variable 19: Ejercicios de conversiones del sistema métrico

Mediante esta variable se mide el conocimiento que tienen los estudiantes para realizar conversiones del sistema métrico.

Codificación:

Número de conversiones correctas: 0, 1, 2, 3.

Variable 20: Propiedades de conjuntos

Por medio de esta variable se analiza el nivel de conocimiento de los estudiantes para aplicar las propiedades de los conjuntos.

Codificación:

Número de propiedades correctamente halladas: 0, 1, 2, 3, 4.

Variable 21: Lógica matemática

Esta variable se utiliza para obtener el nivel de conocimiento que tienen los estudiantes acerca de lógica matemática.

Codificación:

Número de proposiciones lógicas correctas: 0, 1, 2, 3.

Variable 22: Evaluación de una función lineal

La utilización de esta variable es para medir el nivel de conocimiento que tienen los estudiantes para evaluar una función lineal.

Codificación:

Número de funciones correctamente evaluadas: 0, 1, 2, 3, 4.

Variable 23: Perímetro del cuadrado

Mediante esta variable se determina el conocimiento que tienen los estudiantes sobre el cálculo del perímetro y la determinación del lado de un cuadrado.

Codificación:

Sin respuesta:	0
Problema no planteado y mal resuelto:	1
Problema no planteado y bien resuelto:	2
Problema mal planteado y mal resuelto:	3
Problema mal planteado y bien resuelto:	4
Problema bien planteado y mal resuelto:	5
Problema bien planteado y bien resuelto:	6

Variable 24: Área de Triángulos

Para evaluar si los estudiantes tienen nociones para calcular el área de un triángulo, se utiliza esta variable.

Codificación:

Sin respuesta:	0
Problema no planteado y mal resuelto:	1
Problema no planteado y bien resuelto:	2
Problema mal planteado y mal resuelto:	3
Problema mal planteado y bien resuelto:	4
Problema bien planteado y mal resuelto:	5
Problema bien planteado y bien resuelto:	6

Variable 25: Área del círculo

La determinación del nivel de conocimiento de los estudiantes para calcular el área de un círculo, se la realiza utilizando esta variable.

Codificación:

Sin respuesta:	0
Problema no planteado y mal resuelto:	1
Problema no planteado y bien resuelto:	2
Problema mal planteado y mal resuelto:	3
Problema mal planteado y bien resuelto:	4
Problema bien planteado y mal resuelto:	5
Problema bien planteado y bien resuelto:	6

Variable 26: Utilización del teorema de Pitágoras

Para evaluar el nivel de conocimiento que tienen los estudiantes para aplicar el teorema de Pitágoras, se utiliza esta variable.

Codificación:

Sin respuesta:	0
Problema no planteado y mal resuelto:	1
Problema no planteado y bien resuelto:	2
Problema mal planteado y mal resuelto:	3
Problema mal planteado y bien resuelto:	4
Problema bien planteado y mal resuelto:	5

Problema bien planteado y bien resuelto: 6

Variable 27: Aplicación del trinomio cuadrado perfecto y de la diferencia de cuadrados perfectos

Esta variable mide la noción que tienen los estudiantes para aplicar los casos de factorización trinomio cuadrado perfecto y diferencia de cuadrados perfectos.

Codificación:

Mal factorizadas las dos expresiones:	0
Aplicó bien el trinomio pero no la diferencia:	1
Aplicó bien la diferencia pero no el trinomio:	2
Aplicó bien los dos casos:	3

Variable 28: Aplicación del trinomio de la forma $x^2 + bx + c$.

Mediante esta variable se determina el nivel de conocimiento que tienen los estudiantes para factorizar un trinomio de la forma $x^2 + bx + c$.

Codificación:

Mal factorizada la expresión:	0
Aplicó bien el caso:	1

Variable 29: Ecuación lineal con una incógnita

Para medir el conocimiento de los estudiantes para resolver una ecuación lineal con una incógnita, se utiliza esta variable.

Codificación:

Ecuación mal resuelta: 0

Ecuación bien resuelta: 1

Variable 30: Sistema de dos ecuaciones lineales con dos incógnitas

La determinación del conocimiento de los estudiantes para resolver correctamente un sistema de dos ecuaciones lineales con dos incógnitas, se la realiza utilizando esta variable.

Codificación:

Sin respuesta: 0

Problema no planteado y mal resuelto: 1

Problema no planteado y bien resuelto: 2

Problema mal planteado y mal resuelto: 3

Problema mal planteado y bien resuelto: 4

Problema bien planteado y mal resuelto: 5

Problema bien planteado y bien resuelto: 6

Variable 31: Ejercicio de probabilidad

Por medio de esta variable se analiza el nivel de conocimiento de los estudiantes en principios básicos de probabilidad.

Codificación:

Sin respuesta: 0

Mal respondido: 1

Bien respondido: 2

Variable 32: Ejercicio de Estadística

Por medio de esta variable se pretende medir el nivel de conocimiento básico en estadística, de los estudiantes.

Codificación:

Sin respuesta: 0

Mal respondido: 1

Variable 33: Calificación de matemáticas

Con esta variable se calcula la nota del estudiante en la asignatura matemáticas, la cual se basa en una ponderación del puntaje de cada tema de acuerdo a su importancia y dificultad.

Codificación:

Es de tipo numérica

2.5.3 Variables de la prueba de lenguaje

Las variables de estudio para la prueba de lenguaje son 26 que conciernen a cada uno de los temas evaluados en el cuestionario, más una variable que representa la calificación total del estudiante en esta prueba. La codificación que se realizó en las variables de la prueba de lenguaje, está basada en el número de respuestas correctas, que para cada pregunta del cuestionario tuvieron los estudiantes, así, por ejemplo la codificación *malo* significa que el estudiante no contestó correctamente la pregunta, las codificaciones *regular*, *bueno*, *muy bueno* y *excelente* se obtienen clasificando en orden ascendente el número de respuestas correctas, siendo excelente el mayor número de respuestas correctas.

Variable 34: Número de diptongos identificados correctamente.

Esta variable se utilizará para medir el conocimiento que tienen los estudiantes para identificar las palabras que contienen diptongos.

Número de respuestas correctas	Codificación
0	Malo
1-2	Regular
3-5	Bueno
6-8	Muy bueno
9-10	Excelente

Variable 35: Número de triptongos identificados correctamente.

Por medio de esta variable se determina el nivel de conocimiento que tienen los estudiantes para identificar las palabras que contienen triptongos.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2	Bueno
3-4	Muy bueno
5	Excelente

Variable 36: Número de hiatos identificados correctamente.

Se utiliza esta variable para determinar el nivel del conocimiento que tienen los estudiantes para identificar las palabras que contienen hiatos.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2-3	Bueno
4-5	Muy bueno
6-7	Excelente

Variable 37: Número de palabras con acento agudo identificadas correctamente.

Para evaluar el conocimiento que tienen los estudiantes para identificar las palabras que tienen acento agudo, se utiliza esta variable.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2	Bueno
3-4	Muy bueno
5	Excelente

Variable 38: Número de palabras con acento grave identificadas correctamente.

La determinación del conocimiento de los estudiantes para identificar las palabras que tienen acento grave, se la realiza utilizando esta variable.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2-3	Bueno
4-5	Muy bueno
6-7	Excelente

Variable 39: Número de palabras con acento esdrújulo identificadas correctamente.

Con esta variable se mide el conocimiento que tienen los estudiantes para identificar las palabras que tienen acento esdrújulo.

Número de respuestas correctas	Codificación
0	Malo
1	Bueno
2	Excelente

Variable 40: Número de palabras con acento sobresdrújulo identificadas correctamente.

Se utiliza esta variable para analizar el nivel de conocimiento que tienen los estudiantes para identificar las palabras que tienen acento sobresdrújulo.

Número de respuestas correctas	Codificación
0	Malo
1	Bueno

Variable 41: Número de oraciones con correcta semántica.

Para medir el conocimiento que tienen los estudiantes para completar correctamente la semántica de las oraciones, se utiliza esta variable.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2-3	Bueno
4-5	Muy bueno
6-7	Excelente

Variable 42: Número de sinónimos correctos.

Se utiliza esta variable para evaluar el conocimiento que tienen los estudiantes para identificar los sinónimos de las palabras.

Número de respuestas correctas	Codificación
0	Malo
1-2	Regular
3-4	Bueno
5-6	Muy bueno
7-8	Excelente

Variable 43: Número de antónimos correctos.

El conocimiento que tienen los estudiantes para identificar los antónimos de las palabras se evalúa por medio de esta variable.

Número de respuestas correctas	Codificación
0	Malo
1-2	Regular
3-4	Bueno
5-6	Muy bueno
7-8	Excelente

Variable 44: Número de palabras aumentativas clasificadas correctamente

Para determinar el conocimiento que tienen los estudiantes para identificar las palabras clasificadas como aumentativas, se utiliza esta variable.

Número de respuestas correctas	Codificación
0	Malo
1	Bueno
2	Excelente

Variable 45: Número de palabras diminutivas clasificadas correctamente

Con esta variable se mide el nivel de conocimiento que tienen los estudiantes para identificar las palabras clasificadas como diminutivas.

Número de respuestas correctas	Codificación
0	Malo
1	Bueno
2	Excelente

Variable 46: Número de palabras despectivas clasificadas correctamente

Se utiliza esta variable para evaluar el conocimiento que tienen los estudiantes, para identificar las palabras clasificadas como despectivas.

Número de respuesta correctas	Codificación
0	Malo
1	Bueno

Variable 47: Número de palabras simples clasificadas correctamente

Por medio de esta variable se evalúa el conocimiento que tienen los estudiantes para identificar las palabras clasificadas como simples nombres.

Número de respuestas correctas	Codificación
0	Malo
1	Bueno
2-3	Excelente

Variable 48: Número de palabras definidas correctamente

Para determinar el nivel del conocimiento del vocabulario que tienen los estudiantes se utilizará esta variable

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2	Bueno
3	Muy bueno
4	Excelente

Variable 49: Número de verbos correctamente conjugados

Esta variable se utilizará para medir el conocimiento que tienen los estudiantes para conjugar los verbos correctamente.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2-3	Bueno
4-5	Muy bueno
6	Excelente

Variable 50: Número de sujeto de la oración correctamente identificado

Se utilizará esta variable para evaluar el conocimiento que tienen los estudiantes para identificar el sujeto de las oraciones.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2	Bueno
3	Muy bueno
4	Excelente

Variable 51: Número de predicado de la oración correctamente identificado

Para analizar el nivel de conocimiento que tienen los estudiantes para identificar el predicado de las oraciones, se utilizará esta variable.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2	Bueno
3	Muy bueno
4	Excelente

Variable 52: Número de sustantivo de la oración correctamente identificado

Por medio de esta variable se evalúa el conocimiento que tienen los estudiantes para identificar el sustantivo de las oraciones.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2	Bueno
3-4	Muy bueno
5	Excelente

Variable 53: Número de verbo de la oración correctamente identificado

Se utiliza esta variable para medir el conocimiento que tienen los estudiantes para identificar el verbo de las oraciones.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2	Bueno
3	Muy bueno
4	Excelente

Variable 54: Número de frases correctamente identificadas

Se utiliza esta variable para determinar el nivel de conocimiento que tienen los estudiantes para identificar las frases.

Número de respuestas correctas	Codificación
0	Malo
1	Bueno
2	Excelente

Variable 55: Número de oraciones correctamente identificadas

Para medir el conocimiento que tienen los estudiantes, para identificar las oraciones se utiliza esta variable.

Número de respuestas correctas	Codificación
0	Malo
1	Bueno
2-3	Excelente

Variable 56: Número de oraciones clasificadas correctamente

Por medio de esta variable se evalúa el conocimiento que tienen los estudiantes para clasificar las oraciones admirativas, interrogativas, enunciativas, exhortativas, exclamativas, negativas, desiderativas, y afirmativas.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2-3	Bueno
4-5	Muy bueno
6-7	Excelente

Variable 57: Número de frases escritas con correcta sintaxis

Mediante esta variable se determinará el nivel de conocimiento que tienen los estudiantes para corregir la estructura de la oración.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2-3	Bueno
4-5	Muy bueno
6	Excelente

Variable 58: Número de acentos tildados correctamente en la redacción

Se utiliza esta variable para medir el conocimiento que tienen los estudiantes, para tildar correctamente las sílabas, de acuerdo al acento que tienen las palabras

Número de respuestas correctas	Codificación
0	Malo
1-6 7-11	Regular Bueno
12-17	Muy bueno
18-21	Excelente

Variable 59: Número de respuestas correctas

Esta variable se utiliza para evaluar la capacidad que tienen los estudiantes para comprender una lectura, mediante las respuestas que realizan a preguntas formuladas sobre la lectura.

Número de respuestas correctas	Codificación
0	Malo
1	Regular
2-3	Bueno
4-5	Muy bueno
6	Excelente

Variable 60: Calificación de lenguaje

Por medio de esta variable se determina la nota del estudiante en la asignatura lenguaje, esta nota se basa en una ponderación del puntaje de cada tema de acuerdo a su importancia y dificultad.

Codificación:

Es de tipo numérica

Capítulo III

3. Análisis Estadístico Univariado

3.1 Introducción

En el presente capítulo se mostrarán los resultados obtenidos en el tratamiento estadístico de los datos de cada una de las variables aleatorias, descritas en el capítulo anterior. El objetivo de este análisis es determinar, el nivel de conocimiento de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil, en los temas evaluados correspondientes a matemáticas y lenguaje.

3.2 Análisis univariado de las variables generales

En el capítulo dos se definió como variables generales a la edad y el sexo, las cuales proveen información acerca del estudiante. La importancia del análisis de estas variables es para determinar su relación

con el conocimiento de los estudiantes en las asignaturas lenguaje y matemáticas.

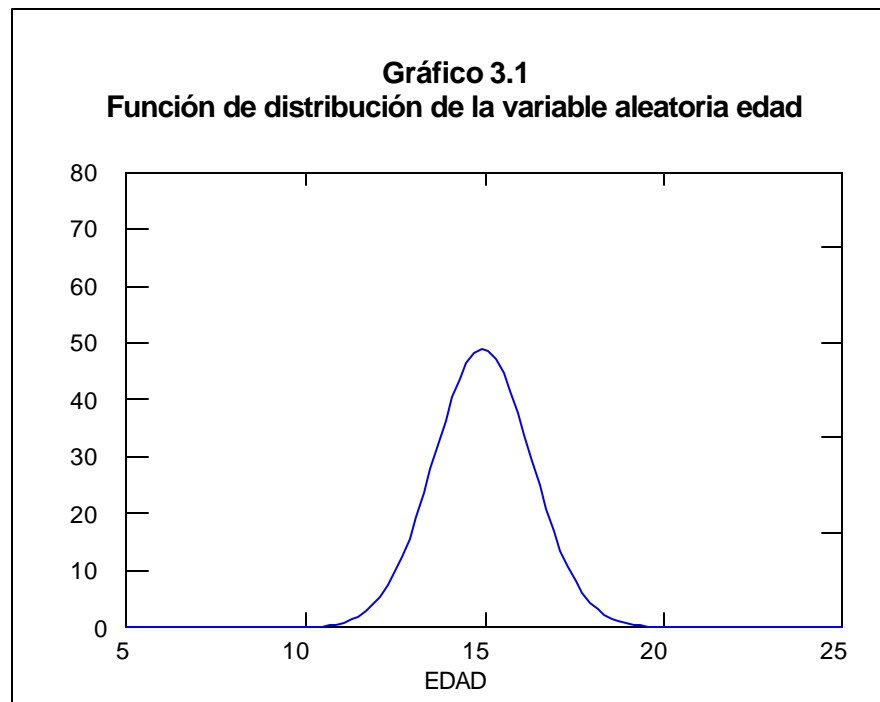
3.2.1 Variable aleatoria: Edad

Como se puede observar en la tabla XXV, el resultado de la media poblacional de la edad de los estudiantes es de 14.93 años y la desviación estándar, que representa la variación absoluta con respecto a la media, es 1.363, este valor es bajo, si se considera que el coeficiente de variación porcentual de la desviación estándar con respecto a al media es del 9.1%.

Media	14,93	Mínimo	10
Mediana	15	Máximo	22
Desviación estándar	1,363	Sesgo	1,046
Varianza	1,857	Kurtosis	4,825
Error estándar	0,105	Primer cuartil	14
Rango	12	Tercer cuartil	15
Moda	14	Rango intercuartil	1
Suma	2484	Coefficiente de variación	0.091

El rango obtenido, indica que existen 12 años de diferencia entre la edad máxima (22 años) y la edad mínima (10 años). Este resultado pudo

afectar el cálculo de la media, pues existen valores extremos muy grandes y muy pequeños, pero el resultado obtenido de la mediana (15 años), indica que no existe mayor diferencia, del valor alrededor del cual se agrupan las observaciones.



El coeficiente de sesgo de esta variable es positivo (1.046), e indica que la distribución está ligeramente sesgada hacia la derecha. Esta situación se puede apreciar mejor en el gráfico 3.1. En lo que respecta al coeficiente de kurtosis (4.825) se determinó que la distribución de la variable edad es leptocúrtica, es decir que su forma es más empinada que la distribución normal.

Para determinar si la variable edad está distribuida normalmente con una media 14.93 y varianza 1.857, se realizó el siguiente contraste de hipótesis:

$$H_0: X \sim N(14.93, 1.857)$$

Vs.

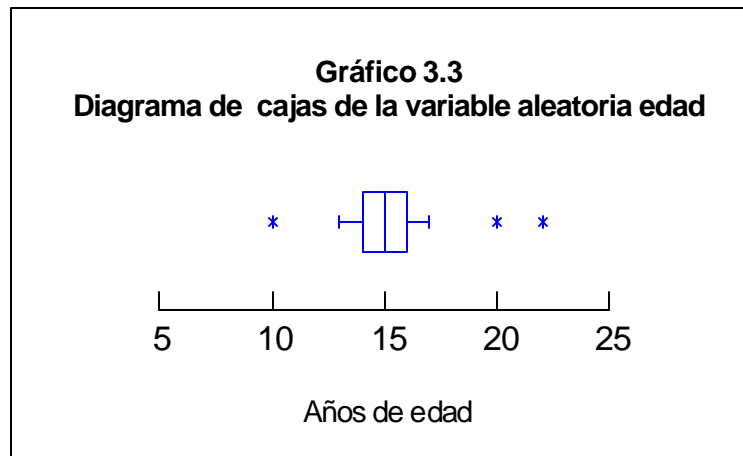
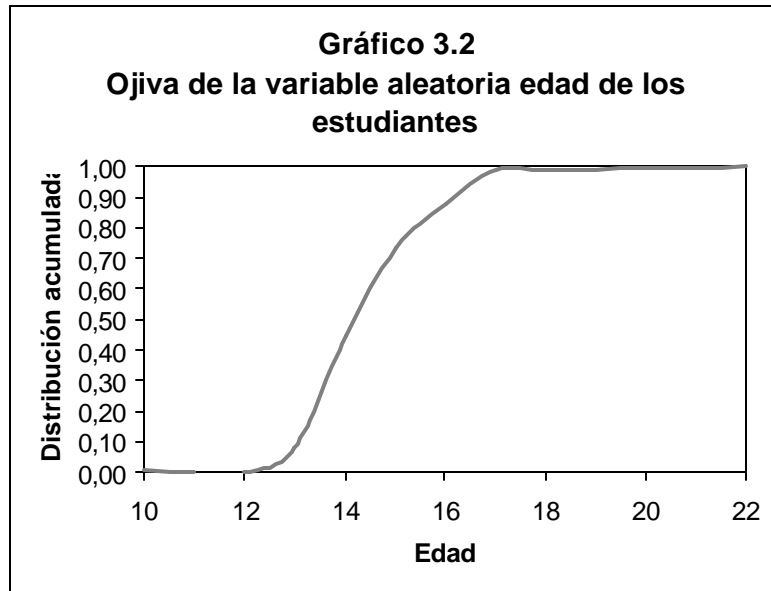
$$H_1: \neg H_0$$

El valor p obtenido al realizar la prueba fue 0.0002, por lo tanto existe evidencia estadística para rechazar la hipótesis nula H_0 , en favor de H_1 .

Tabla XXVI
Frecuencias de la variable aleatoria edad de los estudiantes

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
10	1	0,01	1	0,01
13	12	0,07	13	0,08
14	61	0,37	74	0,44
15	48	0,29	122	0,73
16	24	0,14	146	0,87
17	19	0,11	165	0,99
20	1	0,01	166	0,99
22	1	0,01	167	1,00

.La ojiva de la variable aleatoria edad se puede observar en el gráfico 3.2.



En el gráfico 3.3 muestra el diagrama de cajas de la variable aleatoria edad de los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil.

3.2.2 Variable aleatoria sexo

La variable aleatoria sexo tiene dos resultados posibles 0 (si es de sexo femenino) o 1 (si es de sexo masculino). La probabilidad de obtener éxito, es decir 0, es $p = 0.51$ la probabilidad de fracaso que corresponde al valor 1, es $q = 1 - p = 0.49$, entonces X es una variable aleatoria Bernulli, tal que

$$P(X=x) = \binom{1}{x} 0.51^x (0.49)^{1-x}$$
$$x = 0, 1$$

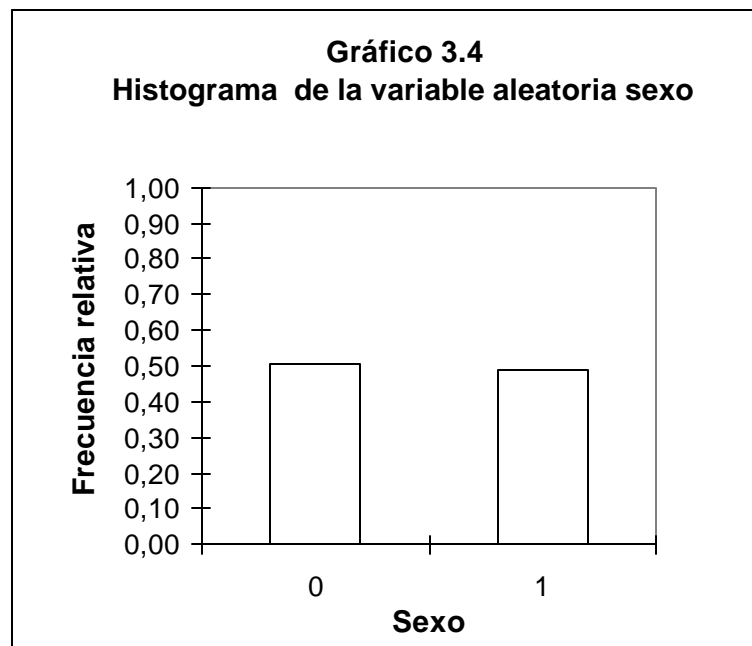
Tabla XXVII
Parámetros poblacionales de la variable aleatoria sexo

Media	0,509	Mínimo	0
Mediana	1	Máximo	1
Desviación estándar	0,504	Sesgo	0,036
Varianza	0,254	Kurtosis	-2,023
Error estándar	0,039	Primer cuartil	0
Rango	1	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	82	Coefficiente de variación	0.99

Como se puede observar en la tabla XXVII, el valor de la moda es 0, este valor corresponde al código de estudiantes de sexo femenino, aproximadamente 51 de cada 100 estudiantes entrevistados son mujeres. Si

se observa el valor de la suma (β_2) este indica el número de estudiantes de sexo masculino que fueron evaluados.

El coeficiente de asimetría de esta distribución es positivo (0.036) la distribución esta ligeramente sesgada hacia la derecha. El coeficiente de kurtosis (-2.023) que también es negativo indica que la distribución es platicúrtica, es decir que la curva es más aplanada que la distribución normal.



Para poder apreciar mejor la distribución de probabilidades de los datos de la variable aleatoria sexo, en el gráfico 3.4 se presenta el histograma de frecuencia relativa.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.51 + 0.49e^t$$

3.3 Análisis univariado de las variables aleatorias correspondientes a la prueba de matemáticas

El análisis univariado que se realiza a continuación, corresponde a las variables aleatorias que se utilizaron, para evaluar el conocimiento de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil, en matemáticas. En esta sección se analizan un total de 31 variables relacionadas a esta evaluación.

3.3.1 Variable aleatoria suma de números enteros

En la tabla XXVIII se puede observar los valores de los parámetros de esta variable, el resultado de la moda es 4, es decir que el valor que más se repite corresponde a la suma correctamente realizada, esta situación se confirma dado que por cada 100 estudiantes 81 realizaron correctamente la suma. La media es 3.545 que también es un valor cercano a cuatro y la mediana (4) es otra medida que afirma el hecho de que los estudiantes saben realizar sumas de números enteros correctamente.

Una medida de la dispersión que tienen las observaciones, es la desviación estándar que es 1.057, este valor es ligeramente bajo, dado

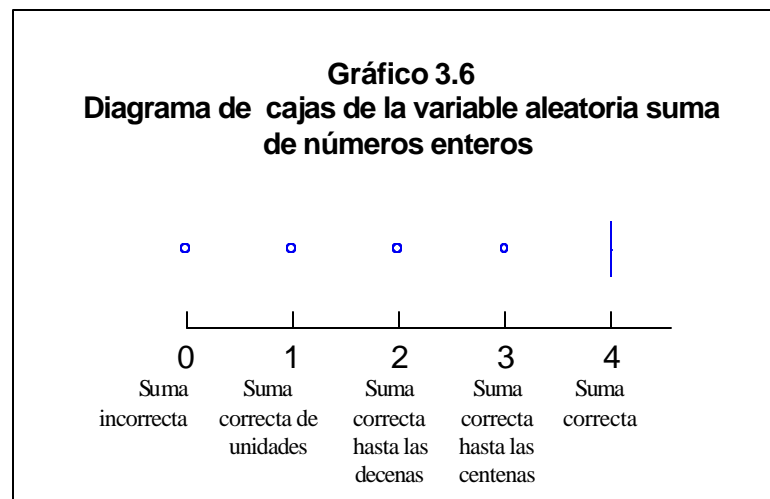
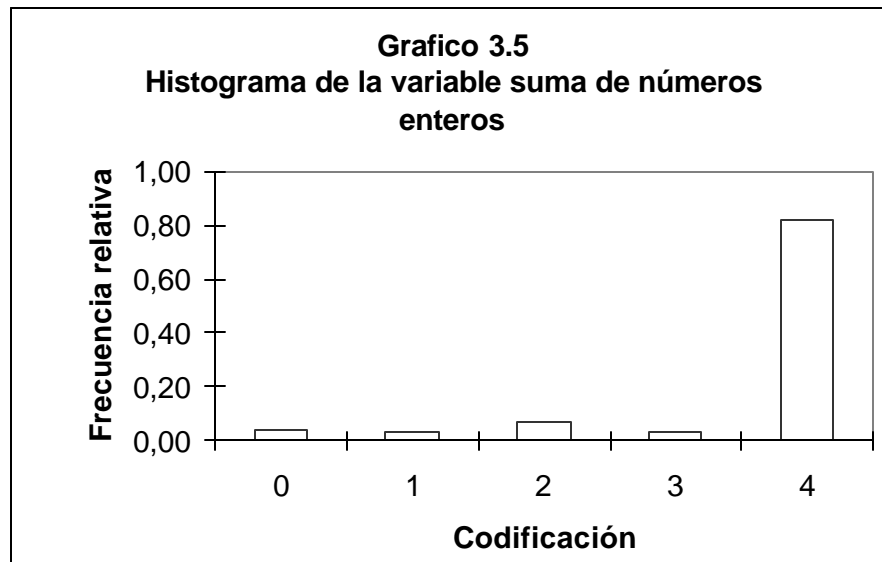
que representa una variación relativa porcentual con respecto a la media del 29%.

Media	3,545	Mínimo	0
Mediana	4	Máximo	4
Desviación estándar	1,057	Sesgo	-2,289
Varianza	1,1172	Kurtosis	4,113
Error estándar	0,082	Primer cuartil	4
Rango	4	Tercer cuartil	4
Moda	4	Rango intercuartil	0
Suma	592	Coefficiente de variación	0.298

En el gráfico 3.5 se puede apreciar que la distribución está sesgada hacia la izquierda, el coeficiente asimetría es negativo (-2.289), lo que indica que la pregunta es fácil, pues existe una mayor proporción de observaciones que toman los valores mayores de la variable. Por otro lado el coeficiente de kurtosis (4.113), indica que la distribución es leptocúrtica, es decir que tiene un pico alto o es más picuda que la distribución normal.

Otros resultados obtenidos son que, el 4% de los estudiantes entrevistados no saben sumar números enteros, el 4% saben sumar

solamente las unidades, el 7% saben sumar hasta las decenas y el 4% saben sumar hasta las centenas.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.04 + 0.04e^t + 0.07e^{2t} + 0.04e^{3t} + 0.81e^{4t}$$

3.3.2 Variable aleatoria suma de números racionales

En esta variable aleatoria existen dos resultados posibles 0 (éxito) si la respuesta es incorrecta y 1 (fracaso) si la respuesta es correcta, la probabilidad de obtener éxito es $p = 0.84$ la probabilidad de fracaso es $q = 1 - p = 0.16$, entonces X es una variable aleatoria Bernulli.

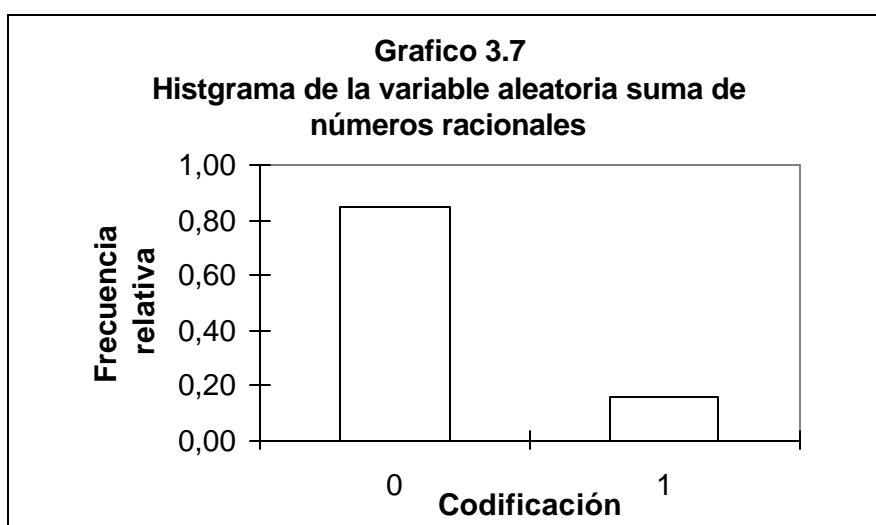
$$P(X=x) = \binom{1}{x} 0.84^x (0.16)^{1-x}$$

$$x = 0, 1$$

Los parámetros de esta variable aleatoria son mostrados en la tabla XXIX. De estos resultados se puede destacar el valor de la moda que es 0, es decir que es el valor que más se repite, esto indica que los estudiantes no saben sumar números racionales, pues aproximadamente por cada 100 estudiantes 84 están en ésta situación.

Media	0,156	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,364	Sesgo	1,917
Varianza	0,132	Kurtosis	1,693
Error estándar	0,028	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	26	Coficiente de variación	2.333

La dispersión de los datos, de acuerdo al coeficiente de variación porcentual, representa el 233.3% con respecto a la magnitud de la media, sin embargo se puede observar en el gráfico 3.7 que las observaciones están agrupadas alrededor de cero. El coeficiente de sesgo obtenido es positivo (1.917), este indica que la distribución está sesgada hacia la derecha, por lo tanto la pregunta es difícil. El coeficiente de kurtosis (1.693) significa que la distribución es platicúrtica, es decir más achatada que la distribución normal.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.84 + 0.16e^t$$

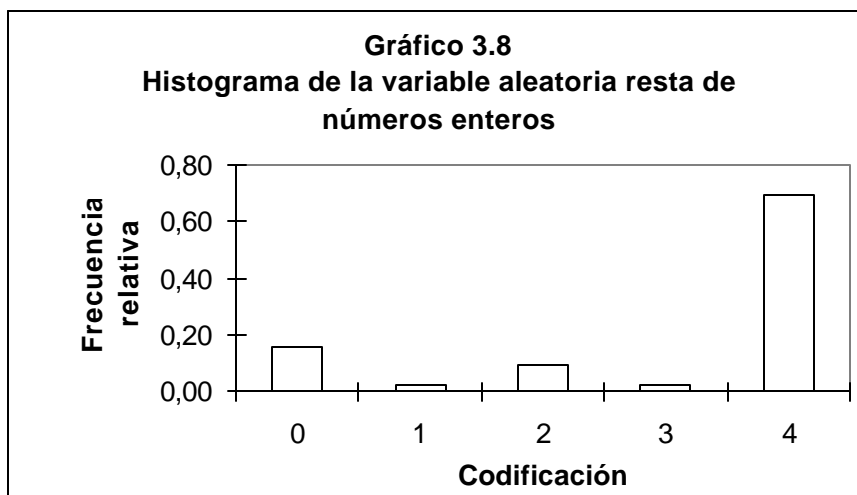
3.3.3 Variable aleatoria resta de números enteros

En la tabla XXX se puede observar los valores de los parámetros que definen ésta población. El valor de la media 3.066 indica que alrededor de éste valor, se agrupan las observaciones; la moda es 4, esto quiere decir que el valor que más se repite corresponde a los estudiantes que realizaron la resta de números enteros correctamente. La desviación estándar que es 1.533, y al coeficiente de variación es 0.5, estos valores indican que, la dispersión relativa de las observaciones es del 50% con respecto a la media.

Media	3,066	Mínimo	0
Mediana	4	Máximo	4
Desviación estándar	1,533	Sesgo	-1,248
Varianza	2,3501	Kurtosis	-0,154
Error estándar	0,119	Primer cuartil	2
Rango	4	Tercer cuartil	4
Moda	4	Rango intercuartil	2
Suma	512	Coefficiente de variación	0.5

La asimetría de esta distribución está dada por el coeficiente de sesgo que es negativo(-1.248), por lo tanto la distribución está sesgada hacia la izquierda, esta situación se puede apreciar más claramente en el gráfico

3.8, donde se puede observar que la mayoría de las observaciones toman el valor de 4, aproximadamente 60 de cada 100 estudiantes saben restar números enteros correctamente.



Los códigos utilizados en esta variable son 0 si no sabe restar, 1 si sabe restar unidades, 2 si sabe restar unidades y decenas, 3 si sabe restar unidades decenas y centenas, y 4 si efectúa correctamente toda la resta.

El coeficiente de kurtosis calculado es -0.154, el cual comparado con el coeficiente de kurtosis de la normal que es tres, nos indica que la distribución de la variable estudiada en esta sección es platicúrtica, es decir más achatada que la distribución normal.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.16 + 0.02e^t + 0.10e^{2t} + 0.02e^{3t} + 0.69e^{4t}$$

3.3.4 Variable aleatoria resta de fracciones

En esta variable aleatoria existen dos resultados posibles 0 (éxito) si la respuesta es incorrecta y 1 (fracaso) si la respuesta es correcta, la probabilidad de obtener éxito es $p = 0.94$ la probabilidad de fracaso es $q = 1 - p = 0.06$, entonces X es una variable aleatoria Bernulli.

$$P(X=x) = \binom{1}{x} 0.94^x (0.06)^{1-x}$$

$$x = 0, 1$$

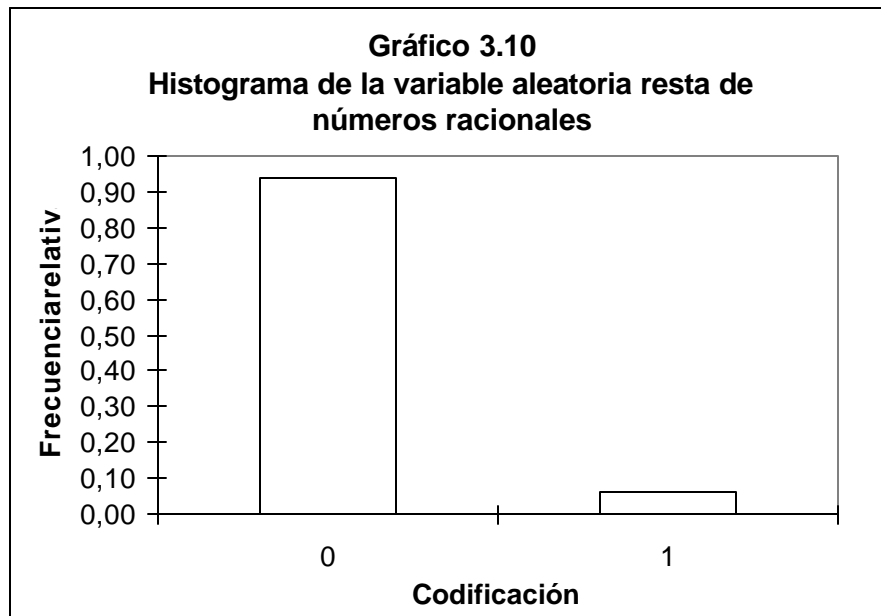
Del conjunto de parámetros obtenidos para esta variable, los cuales se muestran en la tabla XXXII, se puede resaltar que las medidas de tendencia central, se encuentran todas alrededor de un mismo valor, pues la media (0.06), la mediana (0) y la moda(0) son o están situadas alrededor del valor cero, al analizar las frecuencias de esta variable se encontró que por cada 100 estudiantes 94 no saben restar fracciones correctamente.

Tabla XXXII
Parámetros poblacionales de la variable aleatoria resta de fracciones

Media	0,06	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,238	Sesgo	3,744
Varianza	0,0566	Kurtosis	12,16
Error estándar	0,018	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	10	Coefficiente de variación	

Si se analiza el coeficiente de sesgo de esta variable aleatoria, cuyo valor es 3.744 este señala que la distribución esta muy sesgada hacia la derecha, por lo tanto la pregunta es muy difícil, esta situación se puede apreciar mejor en el gráfico 3.10.

Otro parámetro que se puede apreciar mejor gráficamente es el coeficiente de kurtosis que es 12.16, el cual es muy alto e indica que la distribución de esta variable aleatoria es leptocúrtica, es decir tiene un pico alto o es más apuntada que la distribución normal.



La frecuencia relativa correspondiente al código 1, indica que solamente 6 de cada 100 estudiantes evaluados, realizaron la resta correctamente.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0,94 + 0,06e^t$$

3.3.5 Variable aleatoria multiplicación de enteros

En lo que respecta a las medidas de tendencia central, los valores obtenidos de la media, mediana y moda son respectivamente 1.689, 2 y 3, con estos resultados se puede afirmar que las observaciones no están localizadas alrededor de un solo punto. La desviación estándar de la variable (1.33) es muy alta, si se considera que ésta, representa una

variación relativa del 78.7 % de las observaciones con respecto a la media.

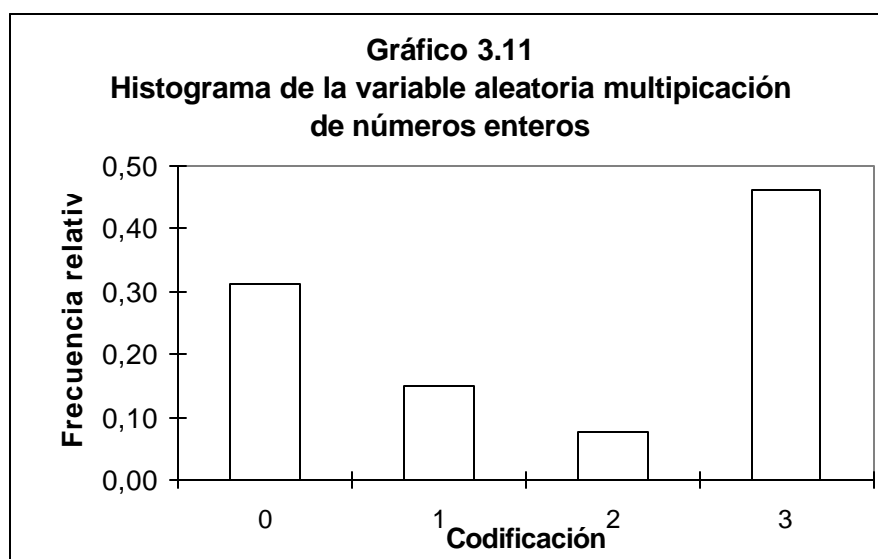
Los valores de los parámetros que definen esta población se muestran en la tabla XXXIII. Entre estos están, el valor mínimo (0) para el cual se determinó 31 de cada 100 estudiantes corresponden a este valor, es decir no saben multiplicar enteros y el valor máximo (1) para el que se obtuvo que 46 de cada 100 estudiantes si saben multiplicar enteros.

Tabla XXXIII
Parámetros poblacionales de la variable aleatoria
multiplicación de enteros

Media	1,689	Mínimo	0
Mediana	2	Máximo	3
Desviación estándar	1,33	Sesgo	-0,219
Varianza	1,768	Kurtosis	-1,747
Error estándar	0,103	Primer cuartil	0
Rango	3	Tercer cuartil	3
Moda	3	Rango intercuartil	3
Suma	282	Coefficiente de variación	0.787

El coeficiente de asimetría de esta variable es negativo (-0.219), por lo tanto la distribución está sesgada hacia la izquierda, por lo tanto la pregunta es fácil, sin embargo se puede observar en el gráfico 3.11 que una gran proporción de las observaciones toman valores de cero.

El resultado obtenido del coeficiente de kurtosis es -1.747 , como es menor a 3, que es el coeficiente de kurtosis de la normal, este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.



La codificación utilizada es 0 si no sabe multiplicar, 1 si multiplicó correctamente solo la primera cifra, 2 si multiplicó correctamente solo la segunda cifra y 3 si realizó correctamente toda la multiplicación.

En la tabla XXXIV se muestran la frecuencias de la variable aleatoria multiplicación de enteros. Además de los resultados comentados anteriormente sobre esta variable en base a los datos de esta tabla se puede decir que por cada 100 estudiantes entrevistados, 15 multiplicaron correctamente solo la primera cifra y 8 solo la segunda cifra.

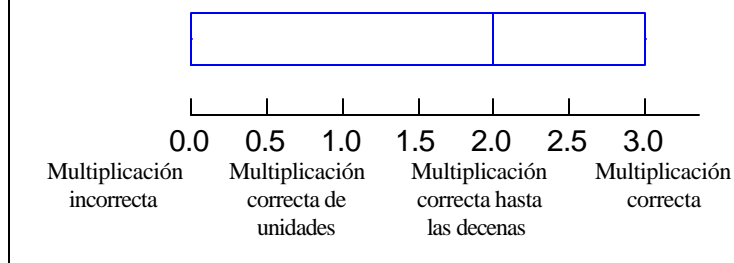
Tabla XXXIV
Frecuencias de la variable aleatoria multiplicación de números enteros

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	52	0,31	52	0.31
1	25	0,15	77	0.46
2	13	0,08	90	0.53
3	77	0,46	167	1.00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.31 + 0.15e^t + 0.08e^{2t} + 0.46e^{3t}$$

Gráfico 3.12
Diagrama de cajas de la variable aleatoria multiplicación de números enteros



3.3.6 Variable aleatoria multiplicación de fracciones

En esta variable aleatoria existen dos resultados posibles 0 (éxito) si la respuesta es incorrecta y 1 (fracaso) si la respuesta es correcta, la

probabilidad de obtener éxito es $p = 0.99$ la probabilidad de fracaso es $q = 1-p = 0.01$, entonces X es una variable aleatoria Bernulli.

$$P(X=x) = \binom{1}{x} 0.99^x (0.01)^{1-x}$$

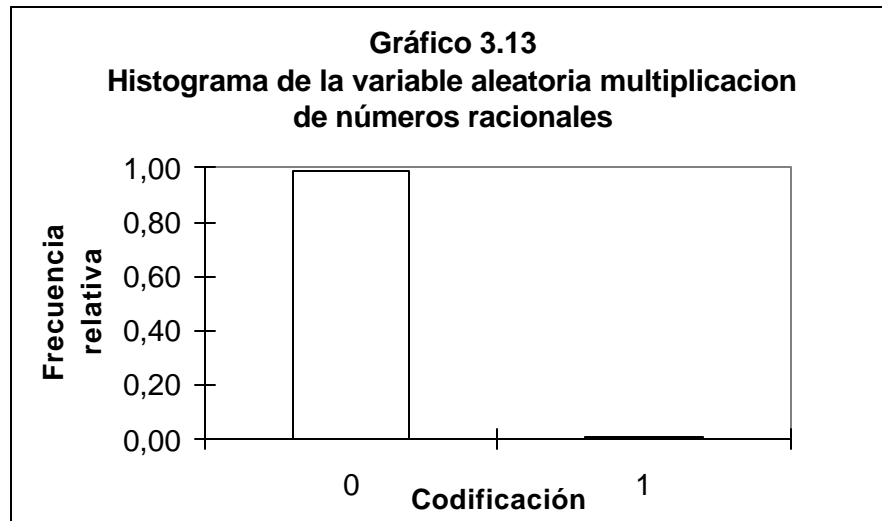
$$x = 0, 1$$

Tabla XXXV
Parámetros poblacionales de la variable aleatoria
multiplicación de números racionales

Media	0,012	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,109	Sesgo	9,054
Varianza	0,0119	Kurtosis	80,95
Error estándar	0,008	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	2	Coefficiente de variación	9.083

Como se puede observar en la tabla de parámetros XXXV, todas las medidas de tendencia central están alrededor de cero, la media es 0.012, la mediana es 0 y la moda también es cero, estos resultados indican que existe una gran proporción de observaciones que toman el valor de cero, más concretamente por cada 100 estudiantes entrevistados 99 no saben multiplicar número fraccionarios.

El valor del coeficiente de sesgo es muy alto (9.054), este indica que la variable de estudio esta marcadamente sesgada hacia la derecha, por lo tanto esta pregunta es muy difícil. Esta situación se puede observar en el gráfico 3.13.



El coeficiente de kurtosis obtenido para esta variable es 80.95, este valor es extremadamente alto, lo que indica que la picudez de esta distribución es mucho mayor que la de la de la curva normal, es decir que tiene forma leptocúrtica.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.99 + 0.01e^t$$

3.3.7 Variable aleatoria división de enteros

En esta variable aleatoria existen dos resultados posibles 0 si la respuesta es incorrecta (éxito) y 1 si la respuesta es correcta (fracaso), la probabilidad de obtener éxito es $p = 0.70$ la probabilidad de fracaso es $q = 1-p = 0.30$, entonces X es una variable aleatoria Bernulli.

$$P(X=x) = \binom{1}{x} 0.70^x (0.30)^{1-x}$$

$$x = 0, 1$$

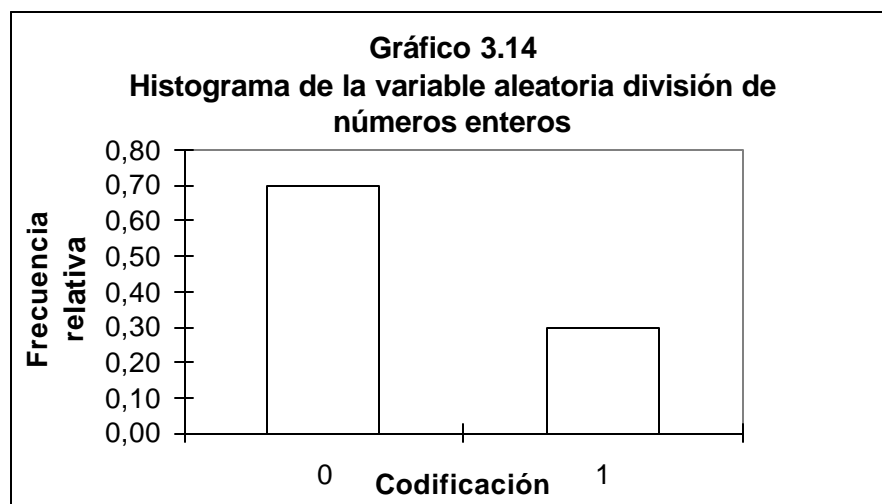
Tabla XXXVI
Parámetros poblacionales de la variable aleatoria división de números enteros

Media	0,299	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,459	Sesgo	0,884
Varianza	0,2107	Kurtosis	-1,234
Error estándar	0,036	Primer cuartil	0
Rango	1	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	50	Coefficiente de variación	1.535

Como se puede observar en la tabla de parámetros XXXVI las medidas de tendencia central la media, mediana y moda son 0.299, 0 y 0 respectivamente, esta situación indica que existe un gran número de observaciones que se agrupan alrededor del valor cero, es decir que la

mayor parte de los estudiantes entrevistados no sabe dividir números enteros.

En el análisis de esta variable se obtuvo que 70 de cada 100 estudiantes entrevistados no saben dividir números enteros. El valor del coeficiente de asimetría (0.884) demuestra que la distribución está sesgada hacia la derecha, por lo tanto la pregunta es difícil, esta situación se puede observar más claramente en el gráfico 3.14. Otra medida que puede ser apreciada mejor gráficamente es el coeficiente de kurtosis que para esta variable es -1.234, este valor indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.70 + 0.30e^t$$

3.3.8 Variable aleatoria división de fracciones

En esta variable aleatoria existen dos resultados posibles 0 (éxito) si la respuesta es incorrecta y 1 (fracaso) si la respuesta es correcta, la probabilidad de obtener éxito es $p = 0.96$ la probabilidad de fracaso es $q = 1 - p = 0.04$, entonces X es una variable aleatoria Bernulli.

$$P(X=x) = \binom{1}{x} 0.96^x (0.04)^{1-x}$$

$$x = 0, 1$$

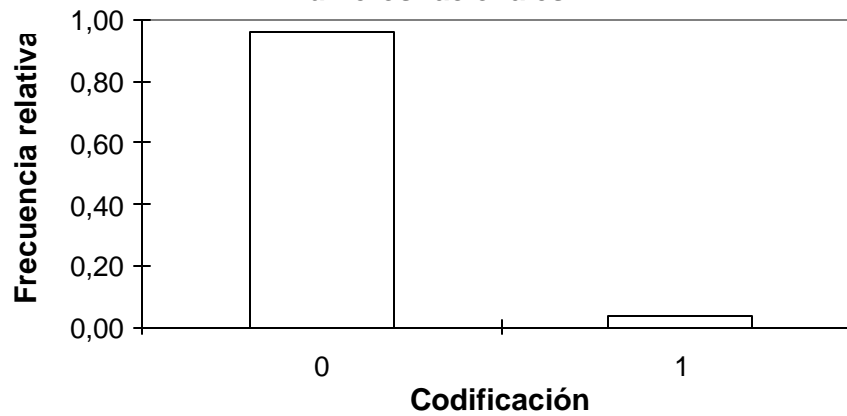
Al analizar los resultados de la tabla XXXVII, las medidas de tendencia central de esta variable la media (0.03), la mediana (0) y la moda (0), se puede asegurar que existe un gran número de estudiantes que no saben dividir fracciones.

Del análisis del coeficiente de asimetría de esta variable se puede entender más claramente si se observa el gráfico 3.15, pues como del sesgo es 5.567 este indica que la distribución esta sesgada hacia la derecha. Por otro lado el coeficiente de kurtosis es muy alto 29.34 este indica que la distribución es leptocúrtica es decir, más picuda o elevada que la de la normal.

Tabla XXXVII
Parámetros poblacionales de la variable aleatoria
división de números racionales

Media	0,03	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,171	Sesgo	5,567
Varianza	0,029	Kurtosis	29,34
Error estándar	0,013	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	5	Coefficiente de variación	5.7

Gráfico 3.15
Histograma de la variable aleatoria división de
números racionales



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.96 + 0.04e^t$$

3.3.9 Variable aleatoria definiciones de números

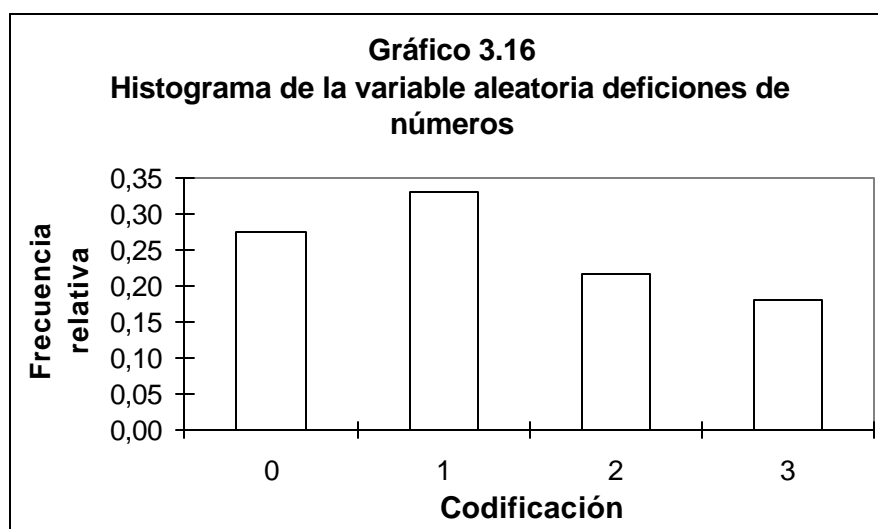
El máximo número de respuestas posibles en esta variable es cuatro sin embargo el valor máximo de las observaciones es tres, esto significa que ningún estudiante definió correctamente todos conceptos.

Tabla XXXVIII
Parámetros poblacionales de la variable aleatoria
definiciones de números

Media	1,299	Mínimo	0
Mediana	1	Máximo	3
Desviación estándar	1,061	Sesgo	0,293
Varianza	1,125	Kurtosis	-1,13
Error estándar	0,082	Primer cuartil	0
Rango	3	Tercer cuartil	2
Moda	1	Rango intercuartil	2
Suma	217	Coficiente de variación	0.816

Los parámetros que describen esta población se muestran en la tabla XXXVIII. Las medidas de tendencia central media (1.299), mediana (1) y moda (1) indican que alrededor de estos valores se agrupan los datos. Analizando las observaciones se obtuvo como resultado que por cada 100 estudiantes 33 respondieron bien solo un concepto de números. La desviación estándar (1.061), es muy alta, representa el 81.6% de variación de las observaciones con respecto a la media.

El coeficiente de asimetría es positivo (0.293) por lo tanto la distribución está sesgada hacia la derecha, lo que significa que la pregunta es medianamente difícil; en lo que respecta al coeficiente de kurtosis (-1.13) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. El histograma de frecuencia absoluta de esta variable es mostrado en el gráfico 3.16.



En la tabla XXXIX se puede apreciar las frecuencias de la variable analizada en esta sección, a partir de los valores de la frecuencia relativa se puede decir que de cada 100 estudiantes evaluados 28 no tienen conocimiento sobre definiciones de números, mientras que 18 de 100 si lo tienen.

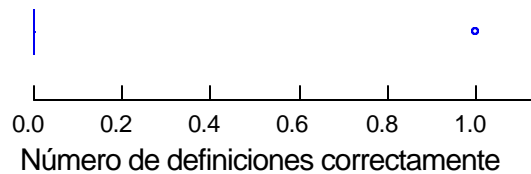
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.28 + 0.33e^t + 0.22e^{2t} + 0.18e^{3t}$$

Tabla XXXIX
Frecuencias de la variable aleatoria definiciones de números

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	46	0,28	46	0,28
1	55	0,33	101	0,61
2	36	0,21	137	0,82
3	30	0,18	167	1,00

Grafico 3.17
Diagrama de cajas de la variable aleatoria definiciones de números



3.3.10 Variable aleatoria desigualdad

Como se puede observar en la tabla XL, los valores de tendencia central, la media, la mediana y la moda tienen valores de cero o muy cercano a cero, esto quiere decir que las observaciones se agrupan alrededor de este valor; de los estudiantes entrevistados se obtuvo que 92 de cada 100 estudiantes no saben resolver este tipo de ejercicios. La desviación estándar es 3.74, este valor es muy alto si se considera el coeficiente de

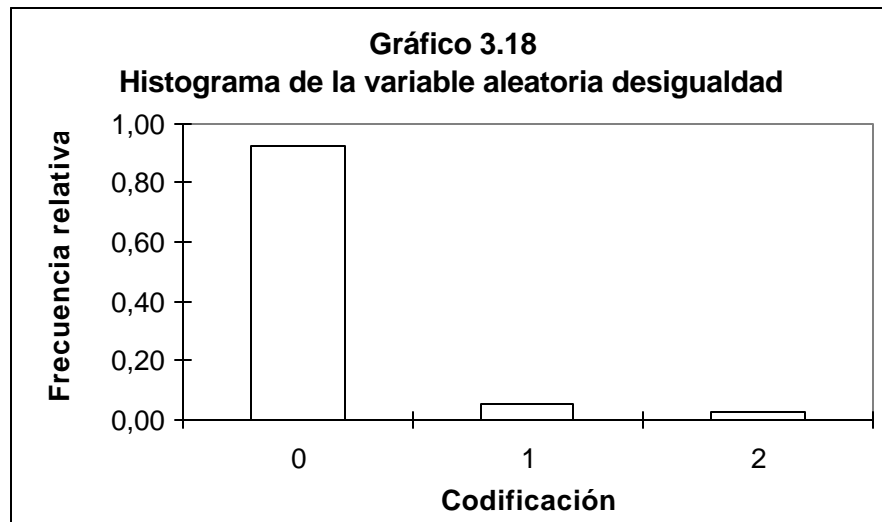
variación, es decir la variación relativa con respecto a la media la cual es del 366.6%.

Media	0,102	Mínimo	0
Mediana	0	Máximo	2
Desviación estándar	0,374	Sesgo	3,917
Varianza	0,139	Kurtosis	15,31
Error estándar	0,029	Primer cuartil	0
Rango	2	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	17	Coefficiente de variación	3.666

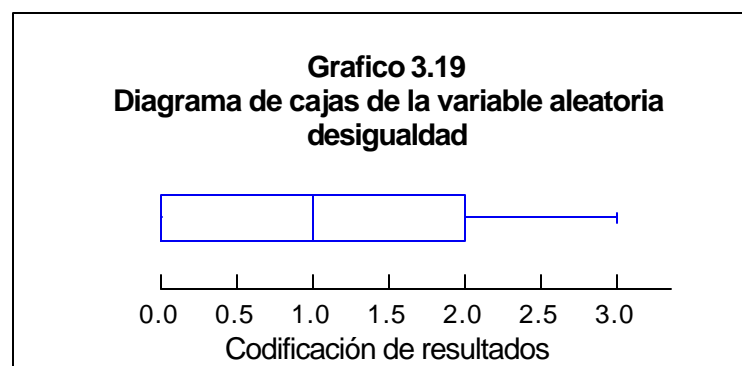
En el gráfico 3.18 se muestra el histograma de frecuencia relativa de esta variable aleatoria, en el cual se puede apreciar que la distribución está sesgada hacia la derecha, dado que el coeficiente de asimetría es positivo (3.917) por lo tanto la media poblacional es mayor que la mediana; en lo que respecta al coeficiente de kurtosis (15.31) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o que es más apuntada que la distribución normal.

Los resultados posibles en esta variable son 0 (sin respuesta), 1 (respuesta incorrecta), 2 (solo planteamiento) y 3 (planteamiento y

respuesta).El máximo valor es 2, lo que significa que ningún estudiante pudo resolver correctamente la desigualdad.



En la tabla XLI se muestran las frecuencias de la variable aleatoria analizada, de las que se obtuvo como resultado que por cada 100 estudiantes entrevistados solamente 2 saben descomponer la desigualdad pero no resolverla, y que 6 de cada 100 resolvieron la desigualdad sin descomponerla.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.92 + 0.05e^t + 0.02e^{2t}$$

Tabla XLI
Frecuencias de la variable aleatoria desigualdad

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	154	0,92	154	0,92
1	9	0,06	163	0,98
2	4	0,02	167	1,00

3.3.11 Variable aleatoria relación de orden

El número máximo posible de relaciones de orden, identificadas correctamente es 5. En la tabla XLII se muestran los resultados de los parámetros poblacionales, en la cual se pueden observar los valores de las medidas de tendencia central la media (1.575), la mediana (0) y la moda (0) los cuales indican alrededor de estos valores se encuentran las observaciones.

El primer cuartil que es 0 indica que al menos el 25% de las observaciones toman ese valor, mientras que el tercer cuartil que es 2 indica que el 75% de las observaciones están bajo ese valor, este hecho

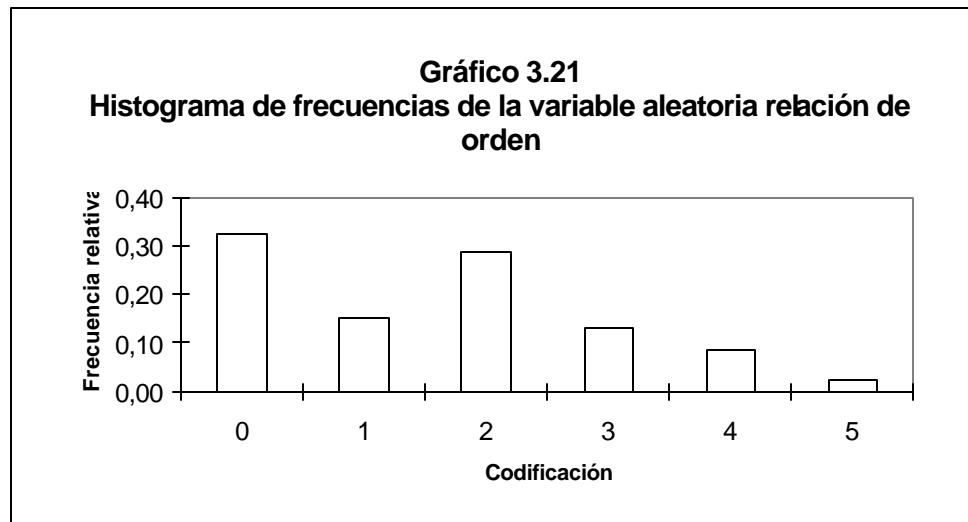
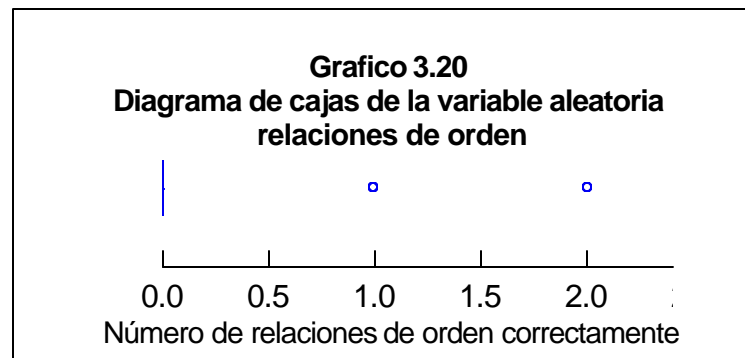
demuestra que los estudiantes en su mayoría no saben identificar las relaciones de orden. La desviación estándar de esta variable es 1.399, es alta porque representa una variación relativa del 88.8% de las observaciones con respecto a la media poblacional, de acuerdo con el coeficiente de variación.

Tabla XLII
Parámetros poblacionales de la variable aleatoria
relación de orden

Media	1,575	Mínimo	0
Mediana	2	Máximo	5
Desviación estándar	1,399	Sesgo	0,473
Varianza	1,957	Kurtosis	-0,67
Error estándar	0,108	Primer cuartil	0
Rango	5	Tercer cuartil	2
Moda	0	Rango intercuartil	2
Suma	263	Coefficiente de variación	0.888

En el gráfico 3.21 se puede apreciar que la distribución está sesgada hacia la derecha, dado que el coeficiente de sesgo es positivo(0.473) la distribución está sesgada hacia la izquierda; en lo que respecta al coeficiente de kurtosis (-0.67) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. De los resultados obtenidos de las frecuencias de esta variable aleatoria mostrados en la tabla XLIII se obtuvo que solo 2 de cada 100 estudiantes entrevistados respondieron correctamente todas las relaciones de orden,

sin embargo 32 de cada 100 no tienen conocimiento sobre relaciones de orden.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.32 + 0.15e^t + 0.29e^{2t} + 0.13e^{3t} + 0.08e^{4t} + 0.02e^{5t}$$

Tabla XLIII
Frecuencias de la variable aleatoria relación de orden

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	54	0,32	54	0,32
1	25	0,15	79	0,47
2	48	0,29	127	0,76
3	22	0,13	149	0,89
4	14	0,08	163	0,98
5	4	0,02	167	1,00

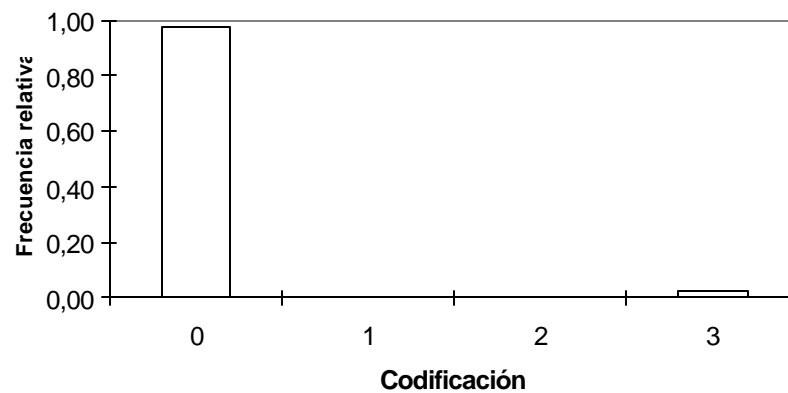
3.312 Variable aleatoria potenciación

Los parámetros de esta población son mostrados en la tabla XLIV, las medidas de tendencia central media, mediana, y moda están alrededor de cero, lo que indica que los estudiantes en una gran proporción no saben resolver ejercicios de potenciación y radicación, del análisis de las observaciones se obtuvo que 98 de cada 100 estudiantes entrevistados está en esta situación. El valor máximo calculado de las observaciones es 3, sin embargo el número máximo de respuestas correctas es 4, esto significa que ningún estudiante pudo responder correctamente todos los ejercicios. El primer y el tercer cuartil son iguales a cero lo que indica que por lo menos el 75% de las observaciones toman este valor.

Tabla XLIV
Parámetros poblacionales de la variable aleatoria
potenciación

Media	0,072	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	0,46	Sesgo	6,284
Varianza	0,211	Kurtosis	37,94
Error estándar	0,036	Primer cuartil	0
Rango	3	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	12	Coeficiente de variación	6.388

Gráfico 3.22
Histograma de frecuencias de la variable aleatoria potenciación



La desviación estándar de esta variable es alta, representa el 638.8% de variación de las observaciones con respecto a la media, debido a que existen observaciones que toman valores muy altos.

Tabla XLV
Frecuencias de la variable aleatoria potenciación

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	163	0,98	163	0,98
1	0	0,00	163	0,98
2	0	0,00	163	0,98
3	4	0,02	167	1,00
4	0	0,00	167	1,00

El coeficiente de asimetría es positivo (6.284) por lo tanto la distribución está sesgada hacia la derecha, lo que indica que esta pregunta es muy difícil lo cual se puede observar en el gráfico 3.22; en lo que respecta al coeficiente de kurtosis (37.94) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o que es más apuntada que la distribución normal. En la tabla XLV se muestran las frecuencias de esta variable aleatoria, se puede apreciar en la frecuencia relativa que por cada 100 estudiantes entrevistados solo 2 respondieron correctamente tres de cuatro ejercicios.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.98 + 0.02e^t$$

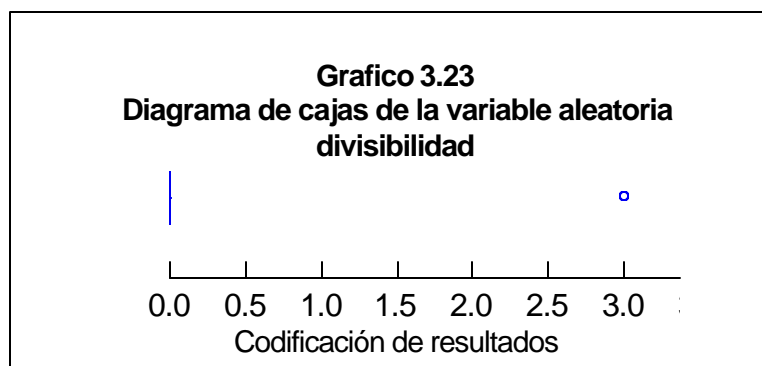
3.3.13 Variable aleatoria divisibilidad

Las respuestas posibles son 0, 1, 2 y 3. En la tabla XLVI se muestra los valores de los parámetros de esta variable, obtenidos al analizar las observaciones. El valor de la media (1.593) como medida de tendencia central se ve afectado por los valores extremos que toman las observaciones, mientras que la mediana y la moda que son iguales a 3, indican sobre que valor se encuentra el 50% de las observaciones y el valor más se repite, respectivamente. Los estudiantes que obtuvieron esta cantidad de respuestas son 51 por cada 100.

Media	1,593	Mínimo	0
Mediana	3	Máximo	3
Desviación estándar	1,477	Sesgo	-0,121
Varianza	2,181	Kurtosis	-1,979
Error estándar	0,114	Primer cuartil	0
Rango	3	Tercer cuartil	3
Moda	3	Rango intercuartil	3
Suma	266	Coficiente de variación	0.927

El valor mínimo es igual al primer cuartil, es decir 0, para esta variable en particular como no existen valores posibles menores a cero esto significa que por lo menos el 25% de las observaciones toman este valor. La

desviación estándar es 1.477 la cual es considerada alta ya que representa el 92.7% de la variación de las observaciones con respecto a la media.



El coeficiente de asimetría es negativo (-0.121) lo que significa que la distribución está sesgada hacia la izquierda; en lo que respecta al coeficiente de kurtosis (-1.979) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o es más apuntada que la distribución normal.

De la tabla XLVII en la que se muestran las frecuencias de la variable analizada en esta sección, se obtiene que por cada 100 estudiantes 45 no tienen conocimiento sobre divisibilidad.

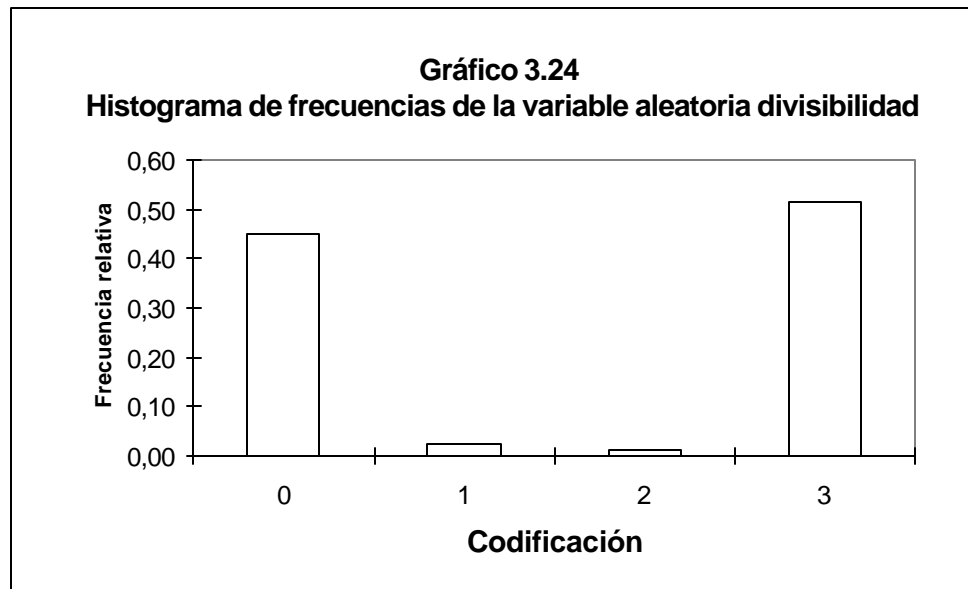


Tabla XLVII
Frecuencias de la variable aleatoria divisibilidad

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	75	0,45	75	0,45
1	4	0,02	79	0,47
2	2	0,01	81	0,49
3	86	0,51	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.45 + 0.02e^t + 0.01e^{2t} + 0.51e^{3t}$$

3.3.14 Variable aleatoria ejercicio de proporcionalidad

La dispersión de los datos es alta, pues la desviación estándar (0.526) representa el 204% del valor de la media. Como se puede observar en la tabla XLVIII todas las medidas de tendencia central están alrededor de 0, es decir que existe una concentración de las observaciones alrededor de este valor. Del análisis de las observaciones se obtuvo que 78 de cada 100 estudiantes no respondieron el ejercicio.

Media	0,257	Mínimo	0
Mediana	0	Máximo	2
Desviación estándar	0,526	Sesgo	1,955
Varianza	0,276	Kurtosis	2,977
Error estándar	0,041	Primer cuartil	0
Rango	2	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	43	Coeficiente de variación	2.046

El máximo de esta variable aleatoria es 2, lo que indica que ningún estudiante planteo y resolvió correctamente el ejercicio. Por otro lado el mínimo es igual al primer cuartil que es 0, es decir que por lo menos el 75% de las observaciones toman este valor.

Gráfico 3.25
Histograma de frecuencias de la variable aleatoria ejercicio de proporcionalidad

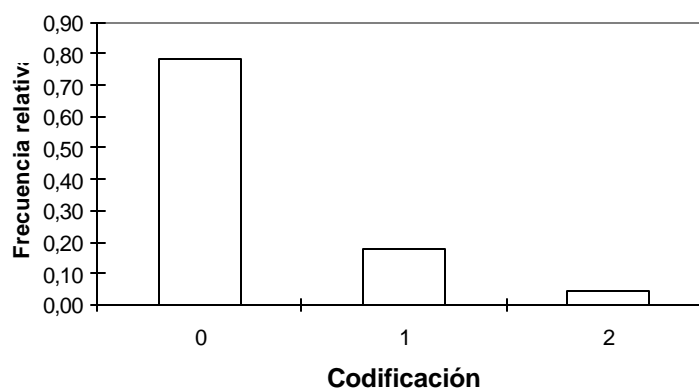
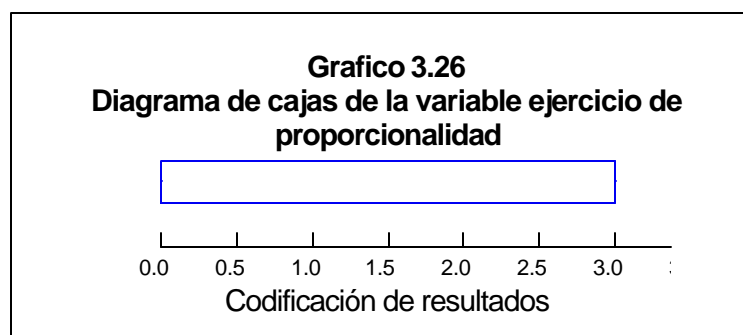


Tabla II
Frecuencias de la variable aleatoria ejercicio de proporcionalidad

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	131	0,78	131	0,78
1	29	0,17	160	0,96
2	7	0,04	167	1,00
3	0	0,00	167	1,00
4	0	0,00	167	1,00
5	0	0,00	167	1,00
6	0	0,00	167	1,00

En el gráfico 3.25 se puede observar que la distribución está sesgada hacia la derecha como lo indica el coeficiente de asimetría que es positivo (1.955), esto significa que es una pregunta difícil, además es

ligeramente platicúrtica, es decir más achatada que la distribución normal, pues el coeficiente de kurtosis es 2.977. Otro resultado obtenido es que por cada 100 estudiantes entrevistados solo 4 respondieron el problema correctamente sin plantearlo. En la tabla IL se muestran las frecuencias de la variable analizada en esta sección.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.78 + 0.17e^t + 0.04e^{2t}$$

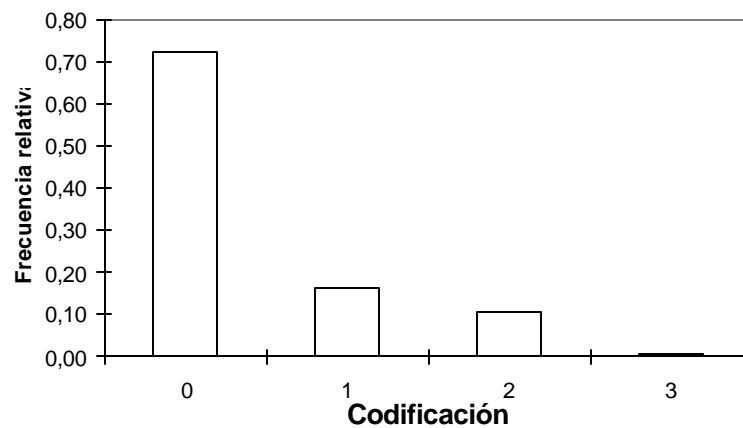
3.3.15 Variable aleatoria ejercicio de interés simple

De los valores de las medidas de tendencia media (0.395), mediana (0) y moda (0), se puede decir que las observaciones están localizadas alrededor del punto cero. La dispersión de las observaciones de esta variable es muy alta pues la desviación estándar 0.702, representa el 177.7% de variación con respecto a al media.

Tabla L
Parámetros poblacionales de la variable aleatoria
ejercicio de interés simple

Media	0,395	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	0,702	Sesgo	1,593
Varianza	0,492	Kurtosis	1,356
Error estándar	0,054	Primer cuartil	0
Rango	3	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	66	Coefficiente de variación	1.777

Gráfico 3.27
Histograma de frecuencias de la variable aleatoria interés simple



El valor máximo de esta variable es 3, este valor corresponde a los estudiantes que plantearon mal el problema pero la respuesta fue correcta. Ningún estudiante planteó y resolvió bien el problema. En la tabla L se muestran los parámetros de esta variable aleatoria. El

coeficiente de asimetría de esta variable es positivo(1.597) por lo tanto la distribución está sesgada hacia la derecha como se puede apreciar en el gráfico 3.27; en lo que respecta al coeficiente de kurtosis (1.356) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

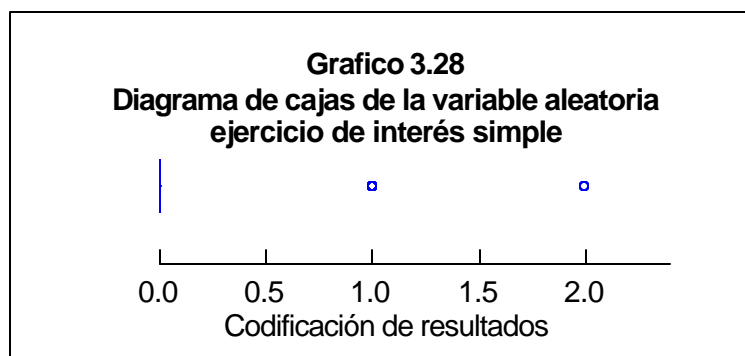
Tabla LI
Frecuencias de la variable aleatoria ejercicio de interés simple

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	121	0,72	121	0,72
1	27	0,16	148	0,89
2	18	0,11	166	0,99
3	1	0,01	167	1,00
4	0	0,00	167	1,00
5	0	0,00	167	1,00
6	0	0,00	167	1,00

En la tabla LI se muestran las frecuencias de esta variable aleatoria, con los siguientes resultados, de cada 100 estudiantes entrevistados 1 planteó y resolvió mal el problema, 11 no plantearon el problema pero lo resolvieron bien y 16 no plantearon el problema y lo resolvieron mal.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.72 + 0.16e^t + 0.11e^{2t} + 0.01e^{3t}$$



3.3.16 Variable aleatoria ejercicio de regla de tres

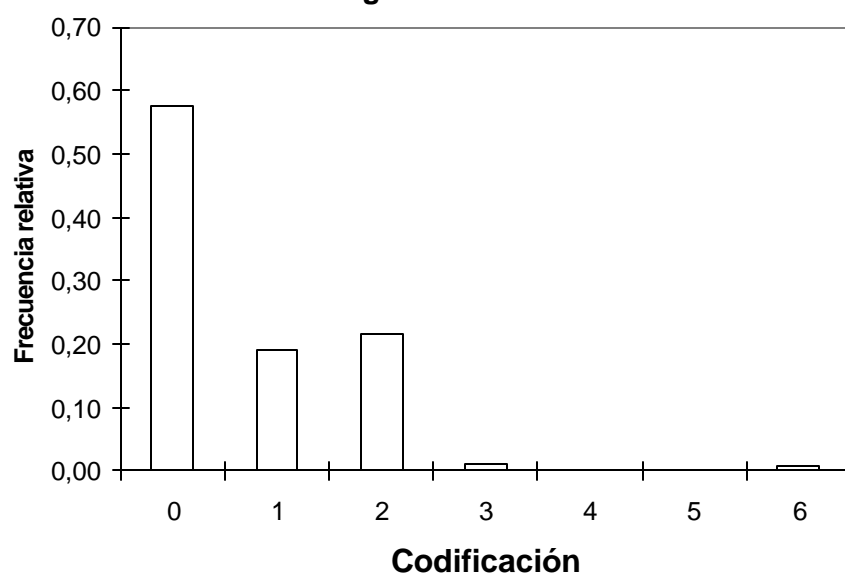
Observando los resultados de tabla LII el valor del máximo obtenido indica que al menos un estudiante planteo y resolvió bien el ejercicio de proporcionalidad, las medidas de tendencia central indican que las observaciones se agrupan hacia el valor 0. El valor del tercer cuartil que es 1 significa que por lo menos el 75% de las observaciones son iguales o menores a este valor, es decir que más de la mitad de los estudiantes entrevistados no saben resolver este ejercicio. Por otro lado la desviación estándar representa el 136.5% de variación de las observaciones con respecto a la media.

El coeficiente de sesgo es positivo (1.59) lo que indica que la distribución está sesgada hacia la derecha; en lo que respecta al coeficiente de kurtosis (4.466) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o que es más apuntada que la distribución normal, esto se puede apreciar si se observa el gráfico 3.29.

Tabla LII
Parámetros poblacionales de la variable aleatoria
ejercicio de regla de tres

Media	0,695	Mínimo	0
Mediana	0	Máximo	6
Desviación estándar	0,949	Sesgo	1,59
Varianza	0,900	Kurtosis	4,468
Error estándar	0,073	Primer cuartil	0
Rango	6	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	116	Coefficiente de variación	1.365

Gráfico 3.29
Histograma de frecuencias de la variable aleatoria ejercicio de
regla de tres



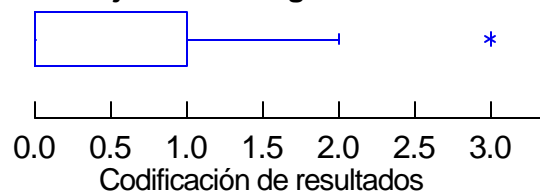
En la tabla LIII se muestran las frecuencias de esta variable aleatoria, los resultados que se obtuvieron fueron que por cada 100 estudiantes entrevistados 57 no contestaron, 20 no plantearon bien el problema y no

lo respondieron, 21 no lo plantearon pero si lo resolvieron y 1 lo planteó bien pero no lo resolvió.

Tabla LIII
Frecuencias de la variable aleatoria ejercicio de regla de tres

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	96	0,57	96	0,57
1	32	0,20	128	0,77
2	36	0,21	164	0,98
3	2	0,01	166	0,99
4	0	0,00	166	0,99
5	0	0,00	166	0,99
6	1	0,01	167	1,00

Grafico 3.30
Diagrama de cajas de la variable aleatoria ejercicio de regla de tres



La función generadora de momentos de esta variable aleatoria es:

$$M(t) = 0.57 + 0.20e^t + 0.21e^{2t} + 0.01e^{3t} + 0.01e^{4t}$$

3.3.17 Variable aleatoria sistema métrico

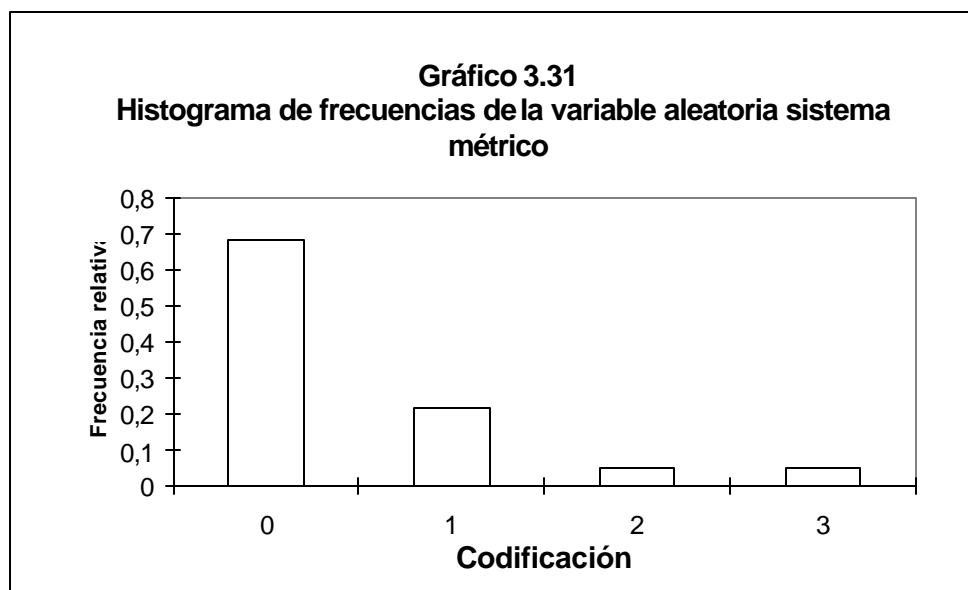
Como se puede observar en la tabla LIV, las medidas de tendencia central están alrededor del valor 0, lo que significa que hacia este valor se agrupan las observaciones. La desviación estándar representa el 173.3% de la variación de las observaciones con respecto a la media, lo cual es un porcentaje muy alto. El valor del tercer cuartil es 1 lo cual indica que al menos el 75% de las observaciones están por debajo de este valor.

Tabla LIV
Parámetros poblacionales de la variable aleatoria
sistema métrico

Media	0,473	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	0,82	Sesgo	1,845
Varianza	0,672	Kurtosis	2,731
Error estándar	0,063	Primer cuartil	0
Rango	3	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	77	Coeficiente varianza	1.733

Como se puede apreciar en el grafico 3.31, la distribución de esta variable está sesgada hacia la derecha pues el coeficiente de asimetría es positivo (1.845), por lo tanto la pregunta es medianamente difícil; en lo que respecta al coeficiente de kurtosis (2.731) este indica que la

distribución es ligeramente platicúrtica, es decir más achatada que la distribución normal.



Los resultados obtenidos de la frecuencia relativa mostrados en la tabla LV son que por cada 100 estudiantes entrevistados 68 no contestaron, 21 no lo plantearon y lo resolvieron mal, 4 no lo plantearon pero si lo respondieron bien y 5 lo plantearon y resolvieron mal.

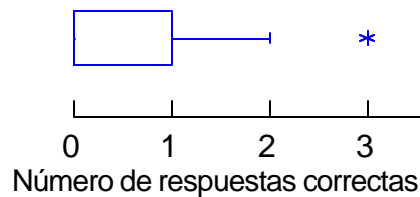
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.68 + 0.22e^t + 0.05e^{2t} + 0.05e^{3t}$$

Tabla LV
Frecuencias de la variable aleatoria sistema métrico

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	114	0,68	114	0,68
1	36	0,21	150	0,90
2	8	0,04	158	0,95
3	9	0,05	167	1,00

Grafico 3.32
Diagrama de cajas de la variable aleatoria conversiones del sistema métrico



3.3.18 Variable aleatoria operaciones de conjuntos

En la tabla LVI el resultado se muestra los valores de los parámetros obtenidos para esta variable aleatoria. Las medidas de tendencia central indican que las observaciones se agrupan alrededor del valor 0, es decir a cero respuestas correctas. El valor del tercer cuartil que es 1 indica que al menos el 75% de las observaciones son iguales o menores a 1.

La dispersión de las observaciones de esta población, se puede analizar con el valor de la desviación estándar es muy alta (0.964), la cual representa el 180.8% de la media de esta variable.

Tabla LVI
Parámetros poblacionales de la variable aleatoria
operaciones de conjuntos

Media	0,533	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	0,964	Sesgo	1,71
Varianza	0,929	Kurtosis	1,902
Error estándar	0,07	Primer cuartil	0
Rango	3	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	89	Coefficiente de variación	1.808

El coeficiente de asimetría es positivo (1.71) la distribución está sesgada hacia la derecha lo cual se puede observar en el gráfico 3.33 del histograma de frecuencia relativa; en lo que respecta al coeficiente de kurtosis (1.902) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. El máximo de esta variable es 3 lo que indica que ningún estudiantes respondió correctamente todos los ejercicios propuestos, los resultados que se obtuvieron al calcular las

frecuencias de esta variable son que por cada 100 estudiantes entrevistados 67 no respondieron, 20 realizaron 1 conversión, 5 realizaron 2 conversión y 8 realizaron conversión 3. E la tabla LVII se muestran las frecuencias de esta variable.

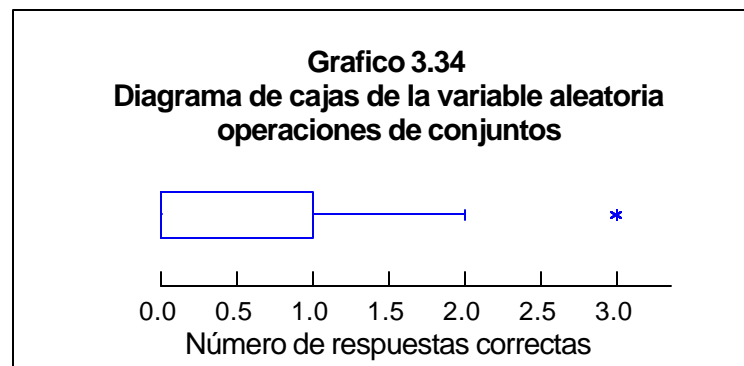
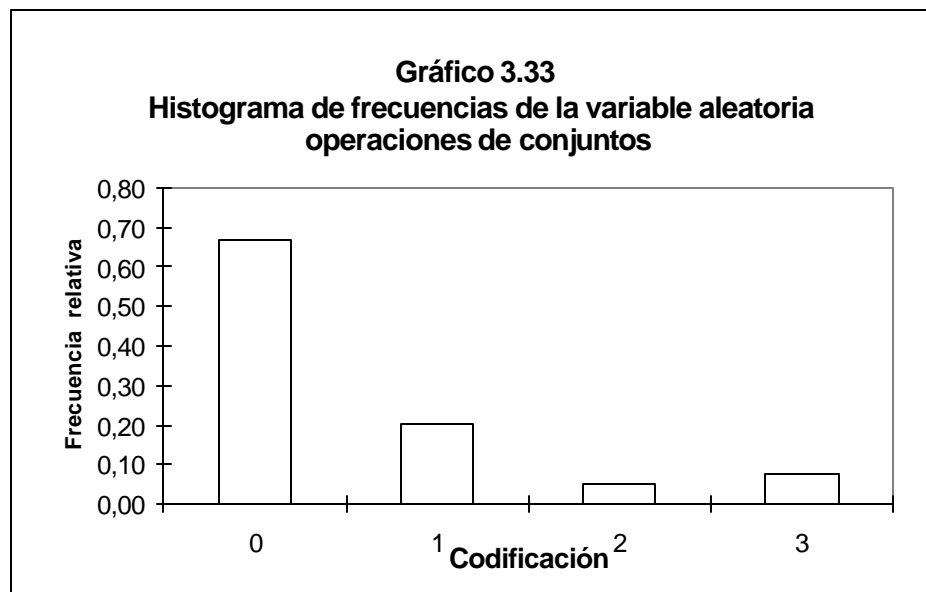


Tabla LVII
Frecuencias de la variable aleatoria operaciones de conjuntos

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	112	0,67	112	0,67
1	34	0,20	146	0,87
2	8	0,05	154	0,92
3	13	0,08	167	1,00
4	0	0,00	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.67 + 0.20e^t + 0.050e^{2t} + 0.08e^{3t}$$

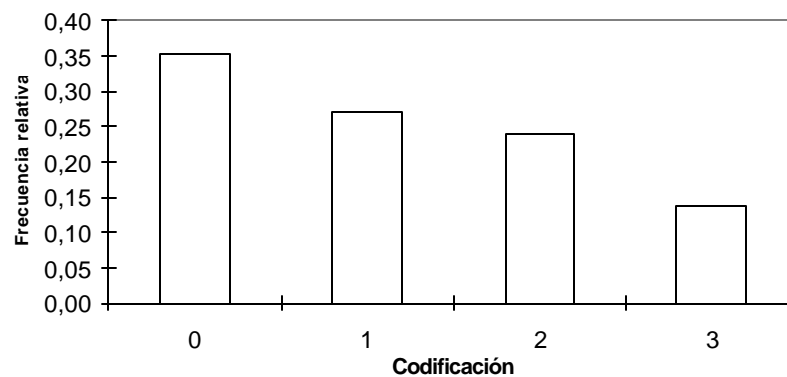
3.3.19 Variable aleatoria proposiciones lógicas

Esta variable representa el número de proposiciones lógicas identificadas correctamente. Como se puede observar en la tabla LVIII, el resultado de las medidas de tendencia central la media es 1.162 y la mediana y la moda son 0, esto indica que las observaciones se agrupan alrededor de estos valores. Por otro lado la dispersión de las observaciones es muy alta, pues la desviación estándar 1.06 representa el 91.2% de la media de esta variable.

Tabla LVIII
Parámetros poblacionales de la variable aleatoria
proposiciones lógicas

Media	1,162	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	1,06	Sesgo	0,377
Varianza	1,1236	Kurtosis	-1,128
Error estándar	0,082	Primer cuartil	0
Rango	3	Tercer cuartil	2
Moda	0	Rango intercuartil	2
Suma	194	Coefficiente de variación	0.912

Gráfico 3.35
Histograma de frecuencias de la variable aleatoria
proposiciones lógicas



El valor del tercer cuartil que es 2 indica que al menos el 75% de los datos son menores o iguales a este valor. El coeficiente de asimetría indica que la distribución está sesgada hacia la derecha, pues el valor

del sesgo es 0.377; en lo que respecta al coeficiente de kurtosis (-1.128) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

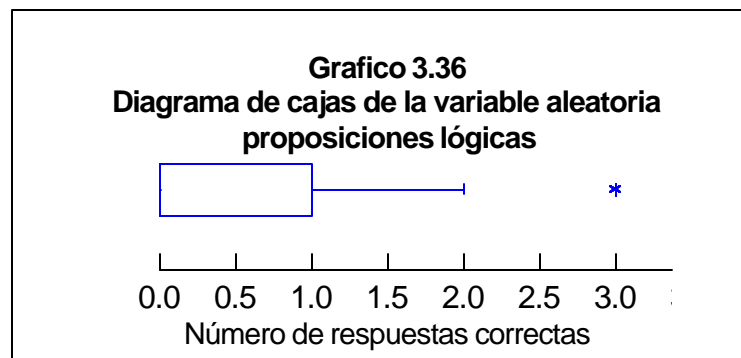
En base a los resultados de las frecuencias de las observaciones analizadas los cuales se muestran en la tabla LIX, se obtuvo que por cada 100 estudiantes entrevistados 35 no respondieron, 27 respondieron 1 ejercicio, 24 respondieron 2 ejercicio y 14 respondieron correctamente todos los ejercicios.

Tabla LIX
Frecuencias de la variable aleatoria proposiciones lógicas

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	59	0,35	59	0,35
1	45	0,27	104	0,62
2	40	0,24	144	0,86
3	23	0,14	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.35 + 0.27e^t + 0.24e^{2t} + 0.14e^{3t}$$



3.3.20 Variable aleatoria evaluación de funciones

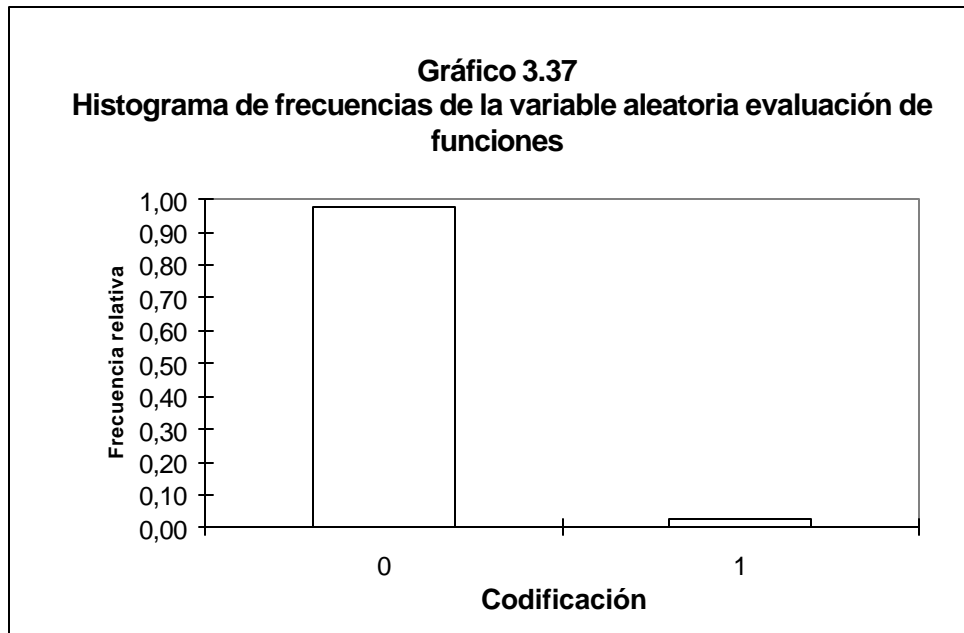
El número máximo de evaluaciones correctas de esta variable es 4. Los parámetros que definen esta variables son mostrados la tabla LX de estos valores las medidas de tendencia central la media, la mediana y la moda toman valores cercanos a 0, lo que indica que hacia este valor se localizan las observaciones.

La dispersión de esta variable es muy alta, pues la desviación estándar es 6.375 veces más mayor a la media obtenida. Por otro lado el máximo obtenido es 1 lo que indica que a lo sumo los estudiantes evaluaron bien una de las cuatro funciones propuestas.

Tabla LX
Parámetros poblacionales de la variable aleatoria
evaluación de funciones

Media	0,024	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,153	Sesgo	6,284
Varianza	0,0234	Kurtosis	37,94
Error estándar	0,012	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	4	Coefficiente de variación	6.375

Los valores del primer y tercer cuartil son iguales a 0, esto significa que al menos el 75% de las observaciones, toman este valor, es decir que por lo menos esta proporción de estudiantes no realizaron correctamente ninguna evaluación. La simetría que tiene la distribución de esta variable aleatoria está dada por el coeficiente de sesgo que es positivo 6.284, lo que significa que la distribución está sesgada hacia la derecha; en lo que respecta al coeficiente de kurtosis (37.94) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o que es más apuntada que la distribución normal. Estas características se las puede apreciar en el gráfico 3.37.



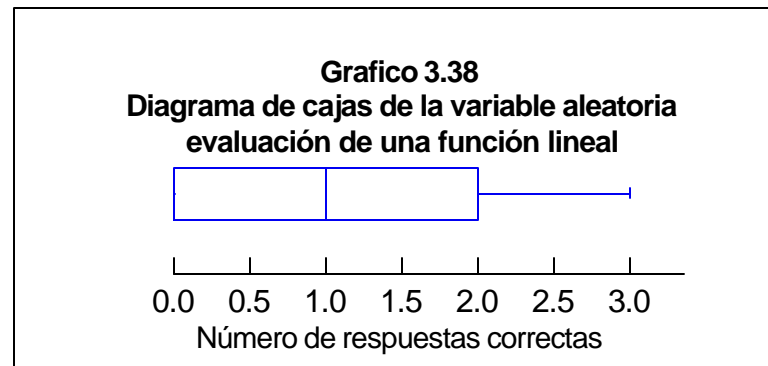
En la tabla LVIII se muestran las frecuencias de la variable analizada en esta sección, de la cuales se obtuvo que por cada 100 estudiantes 98 no saben evaluar funciones y 2 evaluaron correctamente solo una de cuatro funciones.

Tabla LXI
Frecuencias de la variable aleatoria evaluación de funciones

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	163	0,98	163	0,98
1	4	0,02	167	1,00
2	0	0,00	167	1,00
3	0	0,00	167	1,00
4	0	0,00	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.98 + 0.02e^t$$



3.3.21 Variable aleatoria perímetro del cuadrado

Como se puede observar en la tabla LXII se muestran los parámetros de esta variable aleatoria. El máximo obtenido es 3 lo que significa que ningún alumno planteó y respondió correctamente el ejercicio. Por otro lado el valor de los tres cuartiles es cero lo que indica que al menos el 75% de las observaciones son menores o iguales a este valor. La media es 0.192, este valor indica que las observaciones se agrupan alrededor de cero. La dispersión de las observaciones es muy alta pues la desviación estándar que es 0.59 es 3.072 veces mayor que la media.

Tabla LXII
Parámetros poblacionales de la variable aleatoria
perímetro del cuadrado

Media	0,192	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	0,59	Sesgo	2,949
Varianza	0,3481	Kurtosis	7,364
Error estándar	0,046	Primer cuartil	0
Rango	3	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	32	Coefficiente de variación	3.072

Como se puede observar en el gráfico 3.39 la distribución de esta variable aleatoria esta sesgada hacia la derecha, debido a que el coeficiente de sesgo es positivo (2.949), por lo tanto es una pregunta difícil; en lo que respecta al coeficiente de kurtosis (7.364) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o es más apuntada que la distribución normal.

En la tabla LXIII se muestran las frecuencias absolutas y relativas; acumuladas y acumuladas relativas de esta variable aleatoria, de las cuales se obtuvo que por cada 100 estudiantes entrevistados 90 no respondieron, 2 no plantearon y resolvieron mal el ejercicio, 7 no plantearon pero si resolvieron el ejercicio correctamente y 1 planteo y resolvió el ejercicio mal.

Gráfico 3.39
Histograma de frecuencias de la variable aleatoria perímetro del cuadrado

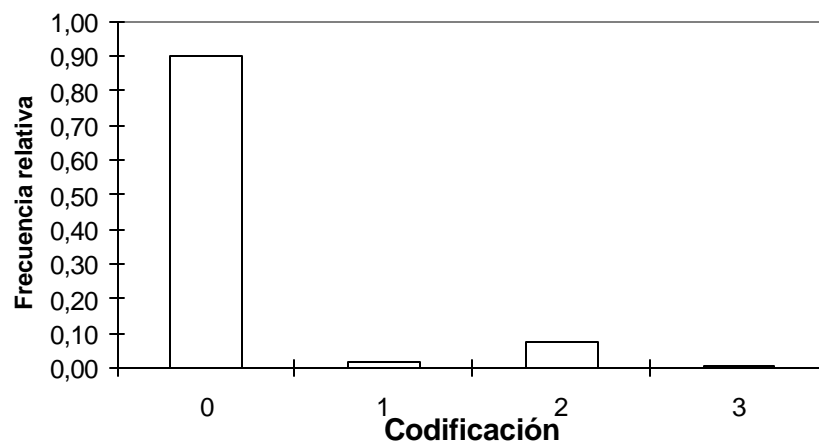
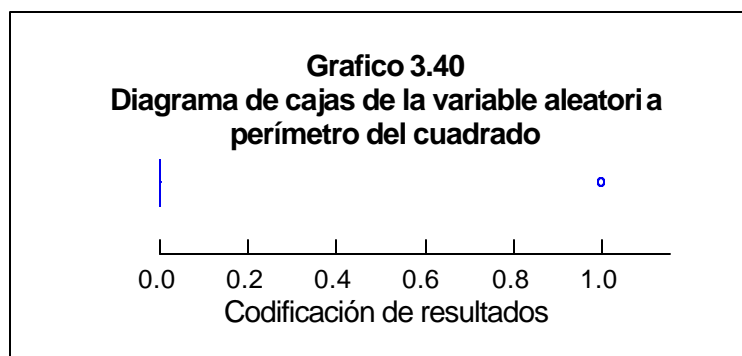


Tabla LXIII
Frecuencias de la variable aleatoria perímetro del cuadrado

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	150	0,90	150	0,90
1	3	0,02	79	0,92
2	13	0,07	81	0,99
3	1	0,01	167	1,00
4	0	0,00	167	1,00
5	0	0,00	167	1,00
6	0	0,00	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.90 + 0.02e^t + 0.07e^{2t} + 0.01e^{3t}$$



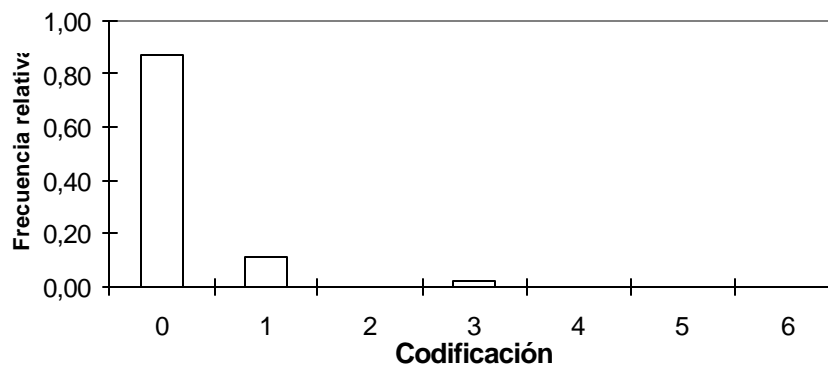
3.3.22 Variable aleatoria área del triángulo

Como se puede observar en la tabla LXIV, todos los valores de tendencia central están alrededor de cero, valor que corresponde a los estudiantes que no respondieron: además el valor del tercer cuartil que también es cero indica que al menos el 75% de los valores observados son iguales a cero. La dispersión de las observaciones está dada por la mediada de la desviación estándar 0.541, la cual es alta ya que representa 3.005 veces más que la media de esta variable aleatoria. El coeficiente de asimetría es positivo(3.803) por lo tanto la distribución está sesgada hacia la derecha, lo que significa que existe una mayor proporción de observaciones hacia los valores menores que puede tomar la variable; en lo que respecta al coeficiente de kurtosis (15.984) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o es más apuntada que la distribución normal.

Tabla LXIV
Parámetros poblacionales de la variable aleatoria área del triángulo

Media	0,18	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	0,541	Sesgo	3,803
Varianza	0,2927	Kurtosis	15,984
Error estándar	0,042	Primer cuartil	0
Rango	3	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	30	Coefficiente de variación	3.005

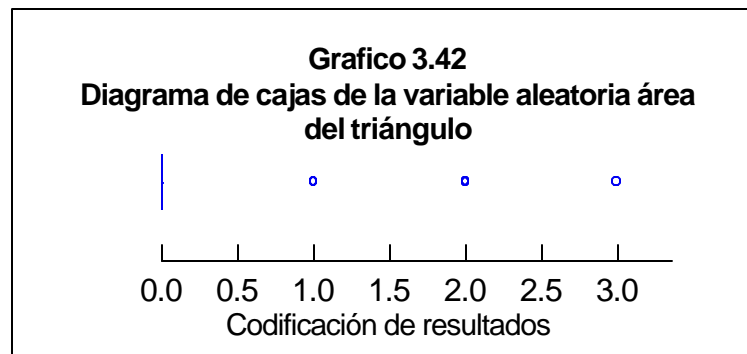
Gráfico 3.41
Histograma de frecuencias de la variable aleatoria área del triángulo



Los resultados obtenidos a partir de las frecuencias de esta variable son que por cada 100 estudiantes entrevistados 87 no saben calcular el área del triángulo, 11 no plantearon el problema y lo resolvieron mal y 2

plantearon y resolvieron mal el problema. La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.87 + 0.11e^t + 0.02e^{3t}$$



3.3.23 Variable aleatoria área del círculo

Al analizar esta variable se presenta una situación parecida a la utilizada para medir el conocimiento de los estudiantes de décimo año para medir el área del triángulo, pues las medidas de tendencia central también se encuentran alrededor del valor cero, las cuales pueden observarse en la tabla LXV. Además el valor del tercer cuartil también es igual a 0, esto significa que al menos el 75% de las observaciones toman este valor.

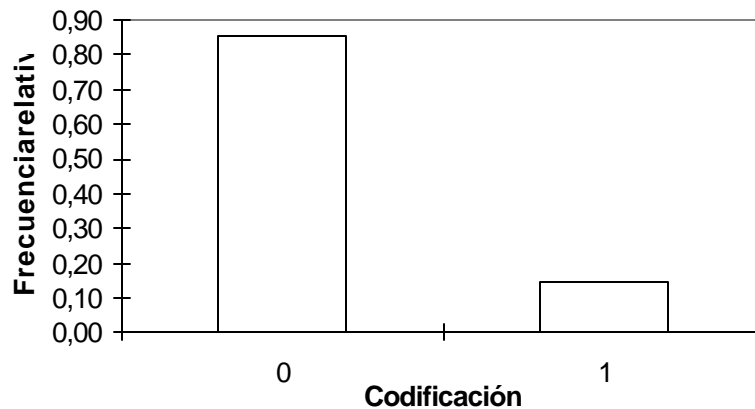
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.86 + 0.14e^t$$

Tabla LXVI
Parámetros poblacionales de la variable aleatoria área del círculo

Media	0,144	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,352	Sesgo	2,05
Varianza	0,1239	kurtosis	2,228
Error estándar	0,027	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	24	Coefficiente de variación	2.444

Gráfico 3.43
Histograma de frecuencias de la variable aleatoria área del círculo



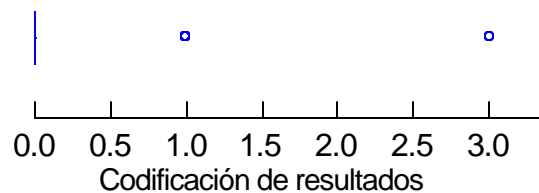
La dispersión de esta variable también es muy alta, pues la desviación estándar que es 0.352, es 2.444 veces mayor a la media de las observaciones. El coeficiente de asimetría es positivo(2.05) por lo tanto la distribución está sesgada hacia la derecha, es decir que la mayor

proporción de los datos toman valores pequeños de la variable; en lo que respecta al coeficiente de kurtosis (2.228) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. En la tabla LXVI se muestran las frecuencias de la variable aleatoria área del círculo, de las que se obtuvo que por cada 100 estudiantes entrevistados 86 no respondieron el ejercicio y 14 no lo plantearon y lo resolvieron mal.

Tabla LXVI
Frecuencias de la variable aleatoria área del círculo

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	143	0,86	143	0,86
1	24	0,14	167	1,00
2	0	0,00	167	1,00
3	0	0,00	167	1,00
4	0	0,00	167	1,00
5	0	0,00	167	1,00
6	0	0,00	167	1,00

Grafico 3.44
Diagrama de cajas de la variable aleatoria área del círculo



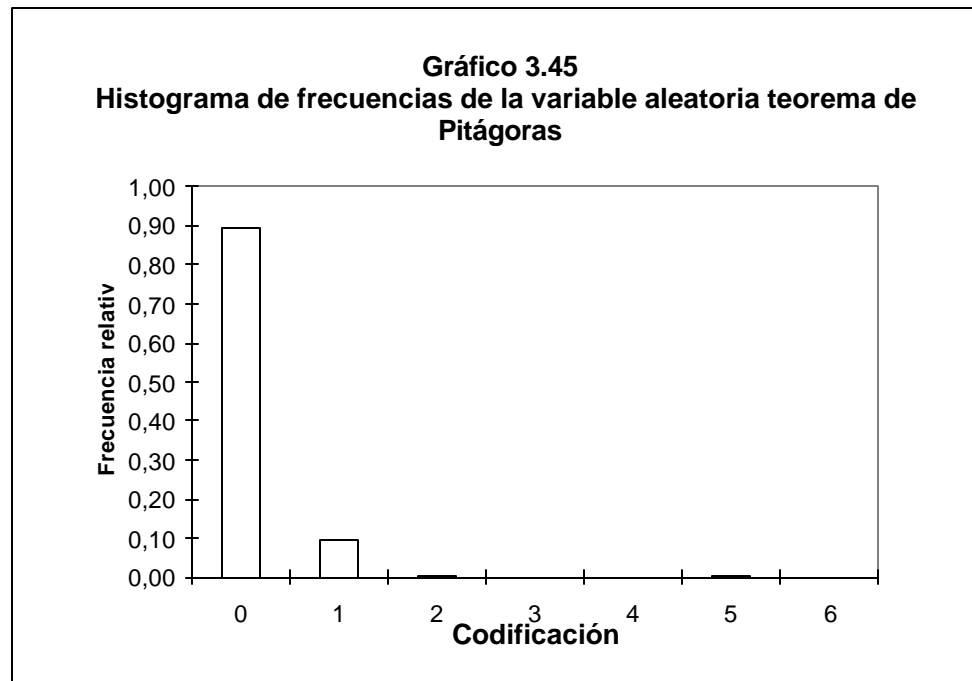
3.3.24 Variable aleatoria teorema de pitágoras

En la tabla LXVII se muestran los valores de los parámetros de esta variable, de los cuales se puede observar el máximo obtenido es 5, este valor corresponde a los estudiantes que plantearon bien el problema pero lo resolvieron mal. Las medidas de tendencia central siguen estando alrededor de 0, lo que indica que alrededor de este valor se agrupan las observaciones. El valor del tercer cuartil que es cero, indica que el 75% de las observaciones son iguales o mayores a cero.

Media	0,138	Mínimo	0
Mediana	0	Máximo	5
Desviación estándar	0,502	Sesgo	6,387
Varianza	0,252	Kurtosis	53,975
Error estándar	0,039	Primer cuartil	0
Rango	5	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	23	Coefficiente de variación	3.537

El coeficiente de asimetría de esta variable aleatoria es positivo (6.387) por lo tanto la distribución está sesgada hacia la derecha, lo cual puede observarse en el gráfico 3.45; en lo que respecta al coeficiente de kurtosis (53.975) este indica que la distribución es leptocúrtica, es decir

que tiene un pico muy alto o que es más apuntada que la distribución normal.



De los resultados de las frecuencias de la variable teorema de Pitágoras mostrados en la tabla LXVIII se obtuvo que por cada 100 estudiantes 89 no saben resolver un ejercicio de este tipo, 9 no lo plantearon pero lo resolvieron mal, 1 no lo planteó y lo resolvió correctamente y 1 planteo bien el ejercicio pero no halló la respuesta.

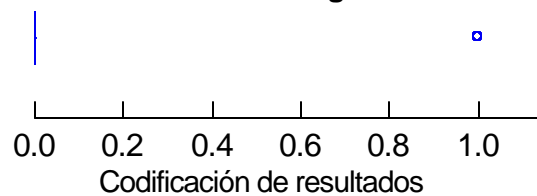
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.89 + 0.09e^t + 0.01e^{2t} + 0.01e^{5t}$$

Tabla LXVIII
Frecuencias de la variable aleatoria teorema de pitágoras

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	149	0,89	149	0,89
1	16	0,09	165	0,98
2	1	0,01	166	0,99
3	0	0,00	166	1,00
4	0	0,00	166	1,00
5	1	0,01	167	1,00
6	0	0,00	167	1,00

Grafico 3.46
Diagrama de cajas de la variable aleatoria teorema de Pitágoras



3.3.25 Variable aleatoria factorización de dos polinomios

Como se puede observar en la tabla LXIX, el valor de la media es 0.569, la cual indica que hacia este valor se agrupan las observaciones, la mediana que es cero indica que al menos el 50% de las observaciones

son menores o iguales a cero y la moda que también es cero el resultado que más se repite.

La dispersión de las observaciones es muy alta, pues la desviación estándar que es 1.073, representa el 188.5% del valor de la media de esta variable aleatoria.

Tabla LXIX
Parámetros poblacionales de la variable aleatoria
factorización de dos polinomios

Media	0,569	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	1,073	Sesgo	1,583
Varianza	1,1513	Kurtosis	0,814
Error estándar	0,083	Primer cuartil	0
Rango	3	Tercer cuartil	3
Moda	0	Rango intercuartil	3
Suma	95	Coeficiente de variación	1.885

El coeficiente de asimetría es positivo(1.583) lo que significa que la distribución está sesgada hacia la derecha, por lo tanto la pregunta es difícil; en lo que respecta al coeficiente de kurtosis (0.814) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. Estas características se pueden apreciar en el gráfico 3.47.

En la tabla LXX se muestran las frecuencias de la variable aleatoria factorización de dos polinomios de las cuales se obtuvo que por cada 100 estudiantes entrevistados 75 no respondieron, 7 resolvieron bien el trinomio pero no la diferencia, 5 resolvieron bien la diferencia pero no el trinomio y 13 resolvieron correctamente ambos casos.

Gráfico 3.47
Histograma de frecuencias de la variable aleatoria factorización de dos polinomios

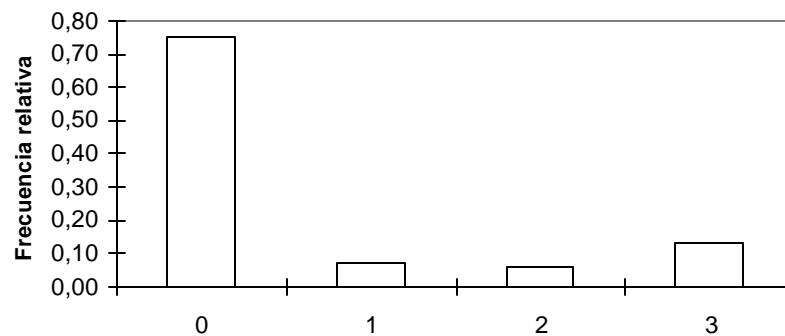
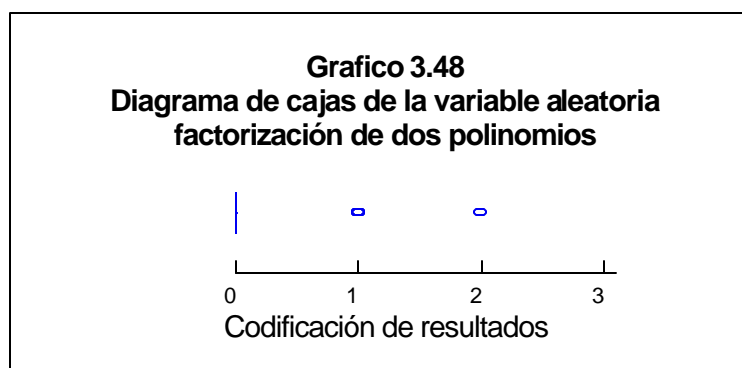


Tabla LXX
Frecuencias de la variable aleatoria factorización de dos polinomios

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	125	0,75	125	0,75
1	11	0,07	136	0,81
2	9	0,05	145	0,87
3	22	0,13	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.75 + 0.07e^t + 0.05e^{2t} + 0.13e^{3t}$$



3.3.26 Variable aleatoria factorización de un polinomio de la forma $ax^2 + ax + c$

Los parámetros de esta variable aleatoria son mostrados en la tabla LXXI. Las medidas de tendencia central que son la media, la mediana y la moda toman valores cercanos a cero, lo que significa que hacia este valor se encuentran localizadas las observaciones. Los resultados obtenidos al analizar esta variable son que por cada 100 estudiantes 81 no saben aplicar este caso y solamente 19 si lo aplicaron bien. En esta variable aleatoria existen dos resultados posibles 0 (éxito) si la respuesta es incorrecta y 1(fracaso) si la respuesta es correcta, la probabilidad de

obtener éxito es $p = 0.81$ la probabilidad de fracaso es $q = 1-p = 0.19$, entonces X es una variable aleatoria Bernulli.

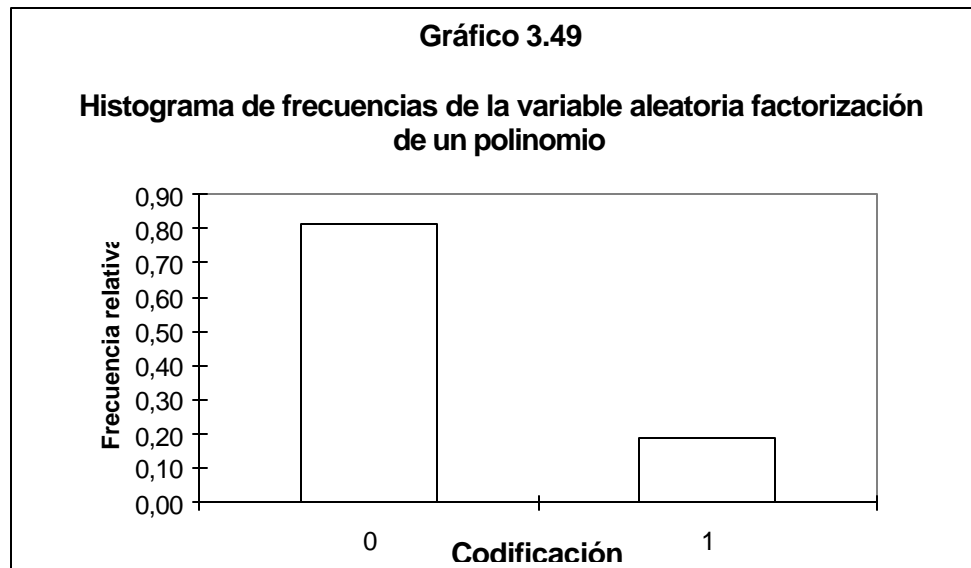
$$P(X=x) = \binom{1}{x} 0.81^x (0.19)^{1-x}$$

$$x = 0, 1$$

Tabla LXXI
Parámetros poblacionales de la variable aleatoria
factorización de un polinomio

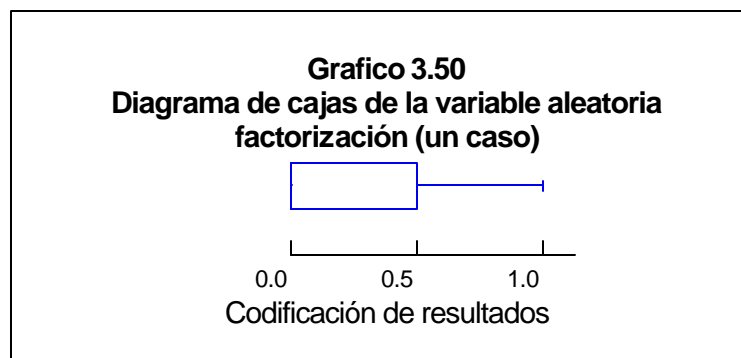
Media	0,186	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,39	Sesgo	1,632
Varianza	0,152	Kurtosis	0,671
Error estándar	0,03	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	31	Coefficiente de variación	2.096

La dispersión de los datos es muy alta, pues la desviación estándar que es 0.39 es 2.096 veces mayor que la media de esta variable aleatoria. El coeficiente de asimetría es positivo (1.632), por lo tanto la distribución está sesgada hacia la derecha, esta característica de la variable se puede observar mejor en el gráfico 3.49; en lo que respecta al coeficiente de kurtosis (0.671) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.18 + 0.19e^t$$



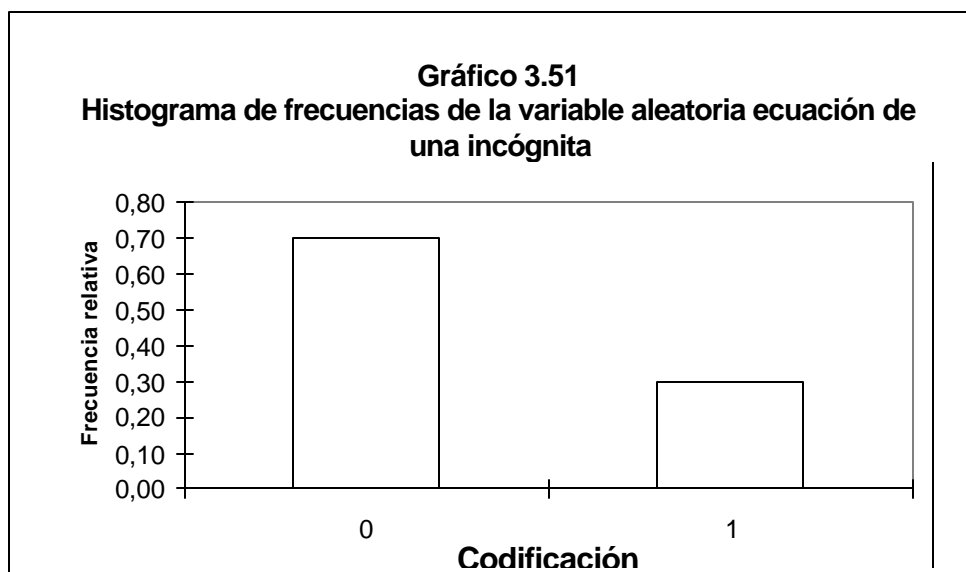
3.3.27 Variable aleatoria ecuación con una incógnita

Las respuestas posibles son de esta variable son sin respuesta 0, respuesta incorrecta 1 y respuesta correcta 2. Como se puede observar en la tabla LXXII el resultado, de las medidas de tendencia central la media que es 0.299 indica que alrededor de este valor se localizan las observaciones mientras que la mediana y la moda que son iguales a cero indican que por lo menos el 50% de las observaciones toman este valor y que es el valor que más se repite respectivamente.

Media	0,299	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,459	Sesgo	0,884
Varianza	0,2107	Kurtosis	-1,234
Error estándar	0,036	Primer cuartil	0
Rango	1	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	50	Coficiente de variación	1.535

La desviación estándar, cuyo valor es 0.459, este valor es muy alto si se considera que representa el 153.5% con respecto a al media. El coeficiente de asimetría es positivo (0.884) lo que significa que la distribución está sesgada hacia la derecha, es decir la pregunta es difícil.

Por otro lado el coeficiente de kurtosis es -1.234 este valor indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

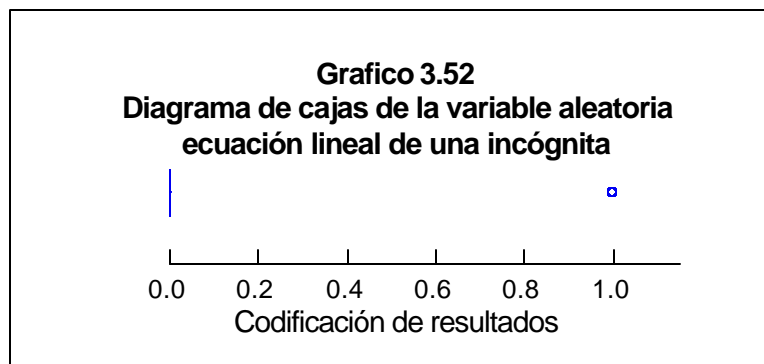


En esta variable aleatoria existen dos resultados posibles 0 (éxito) si la respuesta es incorrecta y 1 (fracaso) si la respuesta es correcta, la probabilidad de obtener éxito es $p = 0.70$ la probabilidad de fracaso es $q = 1-p = 0.30$, entonces X es una variable aleatoria Bernulli.

$$P(X=x) = \binom{1}{x} 0.70^x (0.30)^{1-x} \quad x=0, 1$$

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.70 + 0.30e^t$$



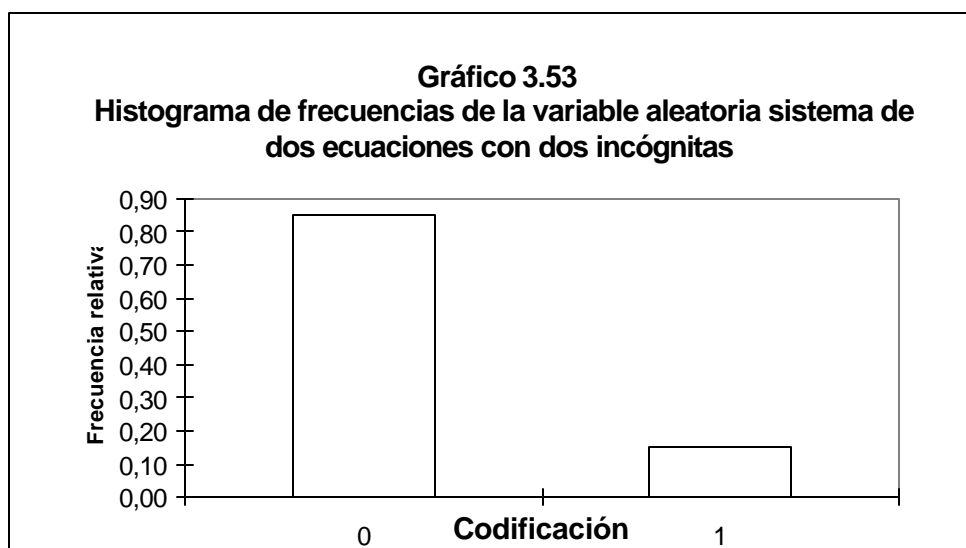
3.3.28 Variable aleatoria sistema de dos ecuaciones con dos incógnitas

En la tabla LXXIII se muestran los valores de los parámetros poblacionales de esta variable aleatoria. Las medidas de tendencia central la media (0.15), la mediana y la moda que son iguales a cero, indican que alrededor del valor cero se agrupan las observaciones. Los valores del primer y tercer cuartil también son iguales a cero, este hecho confirma que por lo menos el 75% de las observaciones toman este valor.

Tabla LXXIII
Parámetros poblacionales de la variable aleatoria
sistema de dos ecuaciones con dos incógnitas

Media	0,15	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,358	Sesgo	1,982
Varianza	0,1282	Kurtosis	1,95
Error estándar	0,028	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	25	Coefficiente de variación	2.386

La dispersión de las observaciones es alta pues la desviación estándar (0.358), es 2.86 veces mayor que la media, aunque el rango de la variable es 1, siendo el máximo valor posible 6 que significa que el estudiante planteó y resolvió bien la ecuación.



La simetría de la distribución de la variable aleatoria está dada por el coeficiente de sesgo que es positivo 1.982, por lo tanto la distribución está sesgada hacia la izquierda; en lo que respecta al coeficiente de kurtosis (1.95) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. Estas características pueden ser apreciadas en el gráfico 3.53.

En la tabla LXXIV se muestran las frecuencias absolutas y relativas; acumuladas y acumuladas relativas, de las cuales se obtuvo los siguientes resultados, por cada 100 estudiantes entrevistados 85 no respondieron y 15 no lo plantearon y lo respondieron mal.

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	142	0,85	142	0,85
1	25	0,15	167	1,00
2	0	0	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.85 + 0.15e^t$$

3.3.29 Variable aleatoria probabilidad

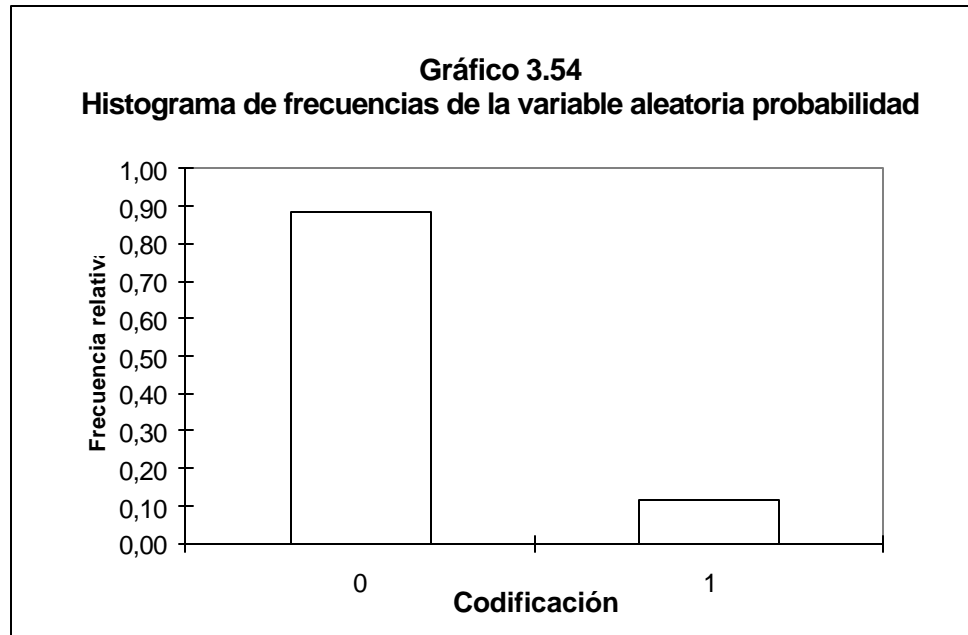
En esta variable hay tres posibilidades que son sin respuesta (0), respuesta incorrecta (1) y respuesta correcta (2). Como se puede observar en la tabla LXXV, los resultados de las medidas de tendencia central que son la media, la mediana y la moda todos indican que las

observaciones se agrupan hacia el valor de cero. El valor máximo hallado es 1 lo que significa que ningún estudiante resolvió correctamente el ejercicio. La suma calculada es 19, este valor indica para esta variable aleatoria el número de estudiantes que respondieron incorrectamente.

Media	0,114	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,318	Sesgo	2,455
Varianza	0,101	Kurtosis	4,075
Error estándar	0,025	Primer cuartil	0
Rango	1	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	19	Coefficiente de variación	2.789

El coeficiente de asimetría de la variable aleatoria es positivo(2.455), por lo tanto la distribución está sesgada hacia la derecha, es decir que la mayor proporción de las observaciones toma el menor valor de la variable; en lo que respecta al coeficiente de kurtosis (4.075) este indica que la distribución es leptocúrtica, es decir que tiene un pico muy alto o que es más apuntada la distribución normal. La función generadora de momentos de esta variable de estudio es:

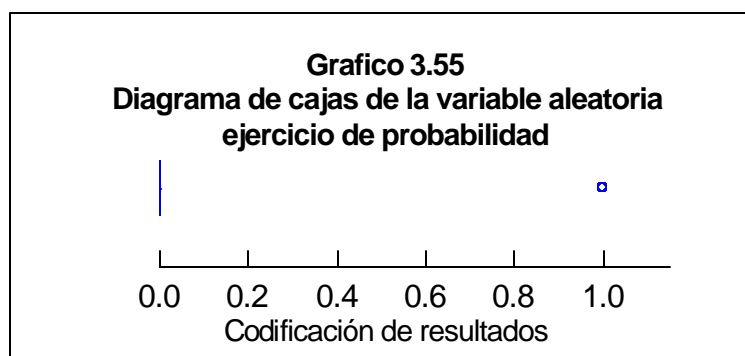
$$M(t) = 0.89 + 0.11e^t$$



En la tabla LXXVI se muestran las frecuencias calculadas de esta variable aleatoria, de las cuales se obtuvo como resultado que por cada 100 estudiantes entrevistados 89 no respondieron y los restantes respondieron incorrectamente

Tabla LXXVI
Frecuencias de la variable aleatoria probabilidad

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	148	0,89	148	0,89
1	19	0,11	167	1,00
2	0	0	167	1,00



3.3.30 Variable aleatoria estadística

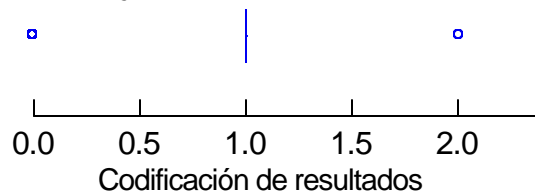
Las respuestas posibles en esta variable aleatoria son sin respuesta 0, respuesta correcta 1 y respuesta incorrecta 2. En la tabla LXXVII se muestra los resultados de los parámetros de la variable aleatoria analizada en esta sección. De estos resultados las medidas de tendencia central toman valores cercanos o iguales a uno, lo cual indica que las observaciones se localizan alrededor de este valor, además el primer y tercer cuartil también son iguales a uno, esto significa que por lo menos el 75% de las observaciones son menores o iguales a uno. La dispersión de los datos se mide en base a la desviación estándar, al rango y al varianza, el valor de la desviación que es 0.46, indica que existe una dispersión más o menos alta, pues esta representa el 57.7% de la media.

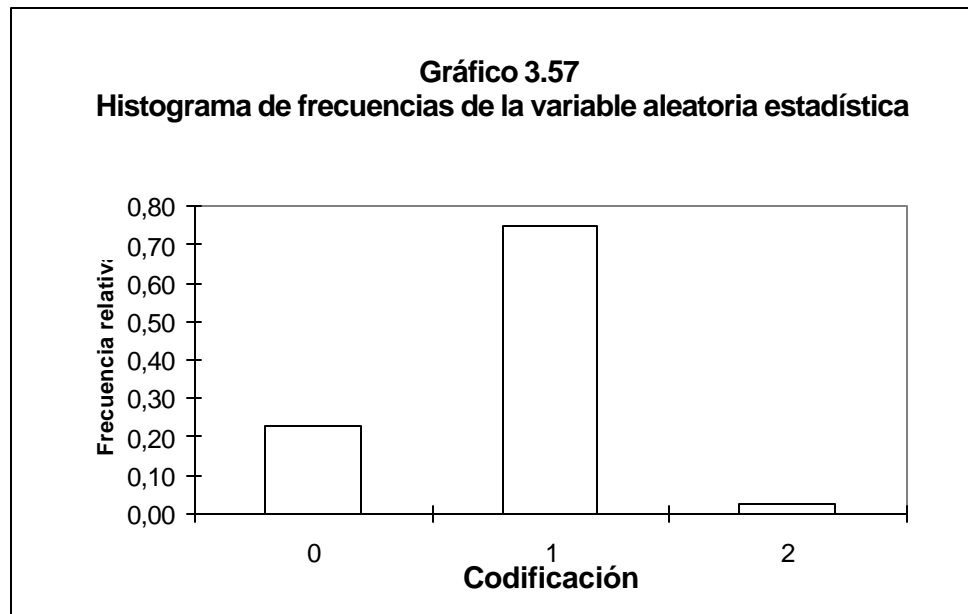
Tabla LXXVII
Parámetros poblacionales de la variable aleatoria estadística

Media	0,796	Mínimo	0
Mediana	1	Máximo	2
Desviación estándar	0,46	Sesgo	-0,701
Varianza	0,211	Kurtosis	0,287
Error estándar	0,036	Primer cuartil	1
Rango	2	Tercer cuartil	1
Moda	1	Rango intercuartil	0
Suma	133	Coefficiente de variación	0.577

El coeficiente de asimetría es negativo (-0.701), por lo tanto la distribución está sesgada hacia la izquierda; en lo que respecta al coeficiente de kurtosis (0.287) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. Estas características se muestran en el gráfico 3.57.

Gráfico 3.56
Diagrama de cajas de la variable aleatoria ejercicio de estadística





En la tabla LXXVIII se muestran las frecuencias absolutas y relativas; acumuladas y acumuladas relativas, de estos resultados se obtuvo que por cada 100 estudiantes entrevistados 23 no respondieron, 75 respondieron mal y solo 4 respondieron correctamente.

Tabla LXXVIII
Frecuencias de la variable aleatoria estadística

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	38	0,23	38	0,23
1	125	0,75	163	0,98
2	4	0,02	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.23 + 0.75e^t + 0.02e^{2t}$$

3.3.31 Variable aleatoria calificación de matemáticas

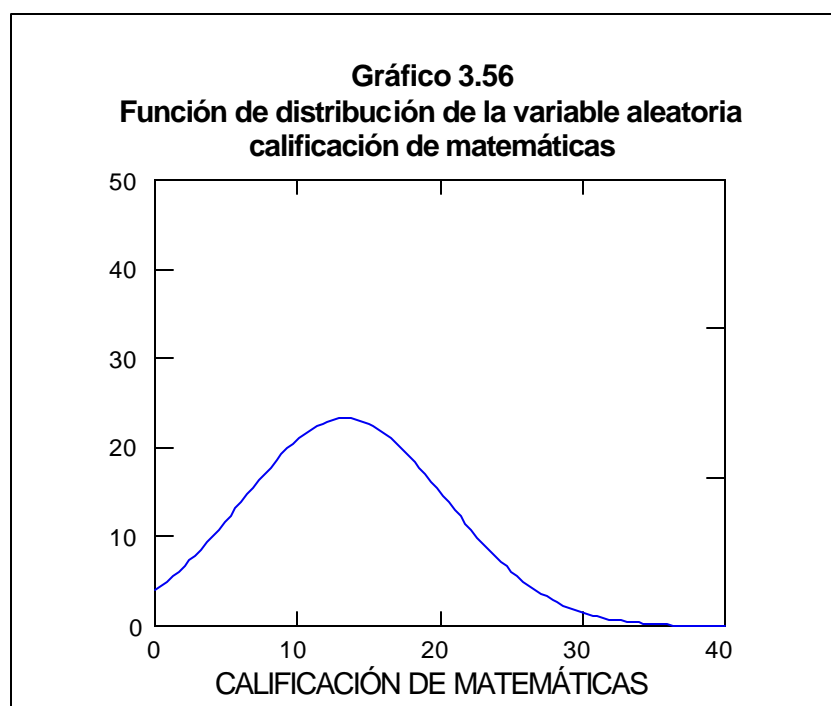
El máximo resultado posible de esta variable aleatoria es 100, sin embargo como se puede apreciar en la tabla 123, el máximo, esto indica. Como se puede observar en la tabla LXXIX, la media es 13.311, este resultado es la nota promedio sobre 100 que obtuvieron los estudiantes, esto refleja que el conocimiento de la asignatura matemáticas es muy bajo.

Media	13,311	Mínimo	0
Mediana	13	Máximo	32
Desviación estándar	7,154	Sesgo	0,358
Varianza	51,18	kurtosis	-0,415
Error estándar	0,554	Primer cuartil	8
Rango	32	Tercer cuartil	18
Moda	12	Rango intercuartil	10
Suma	2219	Coefficiente de variación	0.537

La dispersión de las observaciones, es medida a través del rango, de la varianza y de la desviación estándar, el valor de esta última medida es

considerado alto ya que representa el 53.7% del valor de la media de la variable aleatoria.

El coeficiente de sesgo es positivo (0.358), este indica que la distribución está sesgada hacia la derecha; en lo que respecta al coeficiente de kurtosis (-0.415) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal, estas características se pueden observar en el gráfico 3.56.



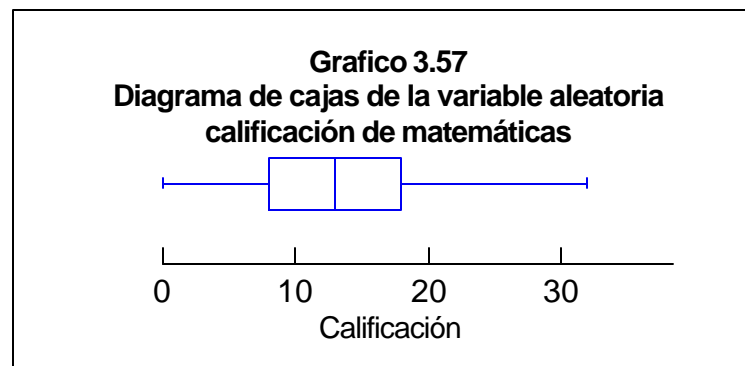
Para determinar si la variable calificación de matemáticas está distribuida normalmente con una media 13.311 y varianza 51.18, se realizó el siguiente contraste de hipótesis:

$$H_0: X \sim N(13.311, 51.18)$$

Vs.

$$H_1: \neg H_0$$

El valor p obtenido al realizar la prueba Kolmogorov-Smirnov fue 0.140 y la máxima diferencia 0.089, por lo tanto existe evidencia estadística para aceptar la hipótesis nula H_0 .



3.4 Análisis univariado de las variables de lenguaje

En esta sección se realizará el tratamiento estadístico univariado, de las variables aleatorias concernientes a la prueba de lenguaje. Las variables aleatorias correspondientes a esta prueba son veintiséis, las cuales están definidas en la sección 2.6.

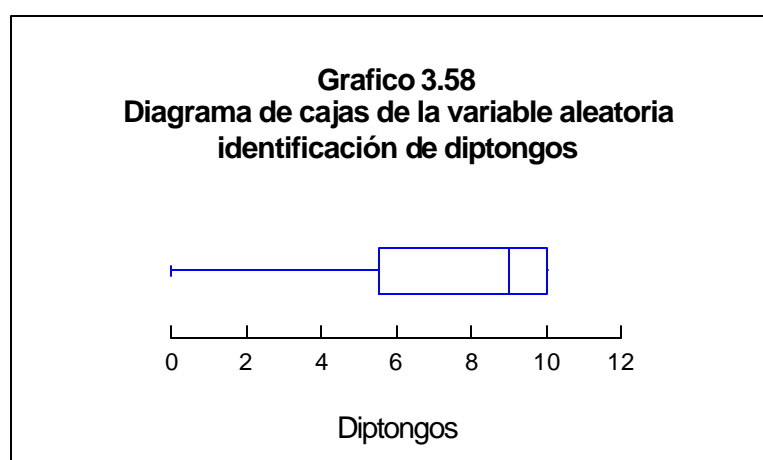
3.4.1 Variable aleatoria identificación de diptongos

En la tabla de parámetros LXXX se puede observar, que el valor de la moda es 10, este valor corresponde al máxima número de respuestas correctas que puede alcanzar el estudiante en esta pregunta, aproximadamente 32 de cada 100 estudiantes alcanzan la máxima calificación en esta variable y más de la mitad, 52 de cada 100 tienen una calificación de excelente.

Tabla LXXX
Parámetros poblacionales de la variable aleatoria
identificación de diptongos

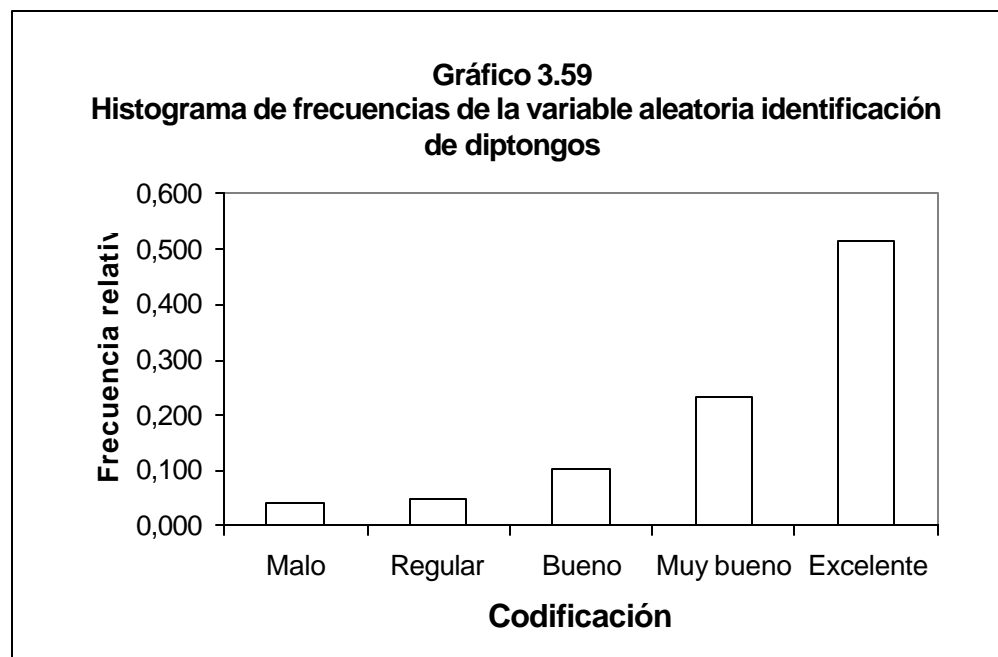
Media	7,204	Mínimo	0
Mediana	9	Máximo	10
Desviación estándar	3,299	Sesgo	-1,122
Varianza	10,883	Kurtosis	0,074
Error estándar	0,25	Primer cuartil	5,5
Rango	10	Tercer cuartil	10
Moda	10	Rango intercuartil	4,5
Suma	1203	Coficiente de variación	0,457

La media de esta variable es 7.204, este valor está incluido en la codificación muy bueno, que corresponde al 23% de los estudiantes evaluados. La variación estándar es alta ya que representa el 45.7% del valor de la media, este resultado se obtiene a partir del coeficiente de variación que es 0.457.



El coeficiente de sesgo es negativo (-1.122), por lo tanto indica que la distribución está sesgada hacia la izquierda, esto significa que la pregunta es muy fácil, esta situación se puede observar en el gráfico 3.59. En lo que respecta al coeficiente de kurtosis (0.074), este indica que la forma de la distribución es platicúrtica, es decir más achatada que la distribución normal.

Para apreciar de mejor manera la distribución de probabilidades de los datos de la variable aleatoria identificación de diptongos, en la tabla LXXXI se muestran las frecuencias de los valores que toma esta variable. Se puede observar que 10 de cada 100 estudiantes no saben identificar palabras que contienen diptongos, este valor corresponde a la codificación malo.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.10 + 0.02e^t + 0.02e^{2t} + 0.04e^{3t} + 0.07e^{4t} + 0.09e^{6t} + 0.03e^{7t} + 0.11e^{8t} + 0.20e^{9t} + 0.32e^{10t}$$

Tabla LXXXI
Frecuencias de la variable aleatoria identificación de
diptongos

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	17	0,10	17	0,10
1	4	0,02	21	0,13
2	4	0,02	25	0,15
3	6	0,04	31	0,19
4	11	0,07	42	0,25
5	0	0,00	42	0,25
6	15	0,09	57	0,34
7	5	0,03	62	0,37
8	19	0,11	81	0,49
9	33	0,20	114	0,68
10	53	0,32	167	1,00

3.4.2 Variable aleatoria identificación de triptongos

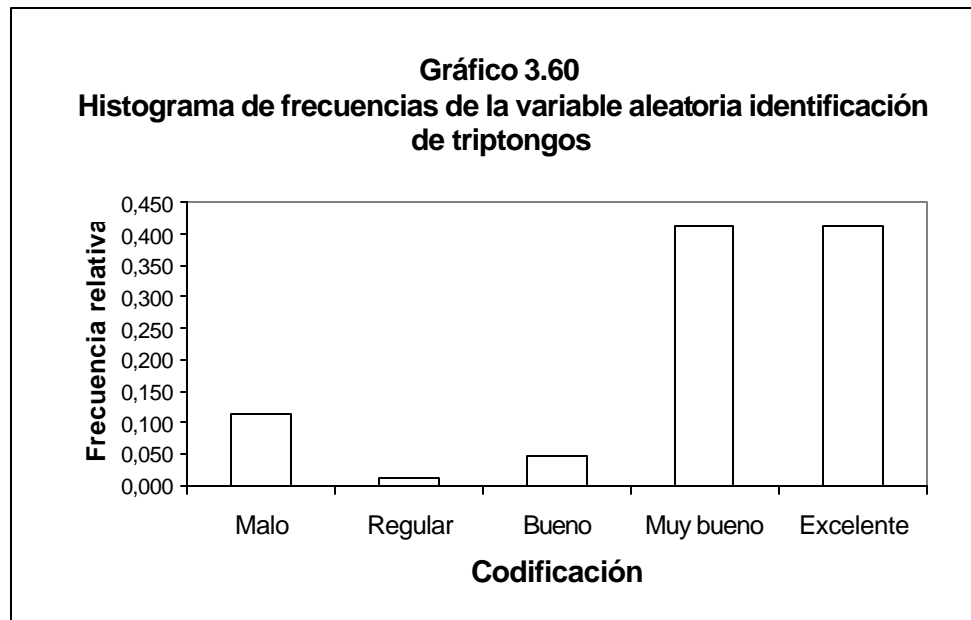
En la tabla de parámetros LXXXII se puede observar que la moda de esta variable es 5, el cual es el número máximo de respuestas correctas que puede tener un estudiante en esta variable, esto nos indica que existe una mayor proporción de estudiantes (41 de cada 100 estudiantes), que saben identificar correctamente triptongos, esta proporción corresponde a la calificación excelente.

El valor de la media de esta variable es 3.659 el cual se incluye en la calificación muy bueno, de los estudiantes evaluados 42 de cada 100 obtienen esta calificación. La desviación estándar 1.612 se considera alta debido a que representa el 44% del valor de la media, este valor se obtiene a partir del coeficiente de variación (0.440).

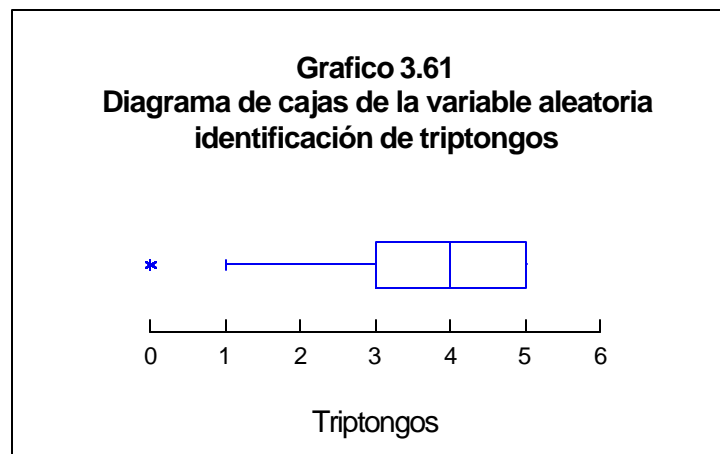
Tabla LXXXII
Parámetros poblacionales de la variable aleatoria
identificación de triptongos

Media	3,659	Mínimo	0
Mediana	4	Máximo	5
Desviación estándar	1,612	Sesgo	-1,228
Varianza	2,598	Kurtosis	0,451
Error estándar	0,125	Primer cuartil	3
Rango	7	Tercer cuartil	5
Moda	5	Rango intercuartil	2
Suma	611	Coefficiente de variación	0.440

Se puede observar en el gráfico 3.60, que el histograma de frecuencias de esta variable, está sesgado hacia la izquierda, por lo tanto la pregunta es considerada fácil, esta situación se presenta debido a que el coeficiente de asimetría es negativo (-1.228. Por otro lado el coeficiente de kurtosis (0.451), indica que la distribución es platicúrtica, es decir que la forma de su curva es más achatada que la distribución normal.



De la tabla LXXXIII podemos concluir que 11 de cada 100 estudiantes no saben identificar triptongos, esta proporción corresponde a la calificación malo.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.011 + 0.01e^t + 0.05e^{2t} + 0.17e^{3t} + 0.25e^{4t} + 0.41e^{5t}$$

Tabla LXXXIII
Frecuencias de la variable aleatoria identificación
de triptongos

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	19	0,11	19	0,11
1	2	0,01	21	0,13
2	8	0,05	29	0,17
3	28	0,17	57	0,34
4	41	0,25	98	0,59
5	69	0,41	167	1,00

3.4.3 Variable aleatoria identificación de hiatos

Los parámetros que definen esta población se muestran en la tabla LXXXIV. La moda de esta variable aleatoria es 0, es decir que la observación que más se repite, corresponde a los estudiantes que no saben identificar palabras que contienen hiato, 32 de cada 100 estudiantes están en esta situación. El valor de la media (2.132) está incluido en la calificación bueno y se determinó que 44 de cada 100

estudiantes obtienen esta calificación. Por otro lado la desviación estándar calculada es 1.987, la cual es considerada alta, en base al coeficiente de variación que indica, que ésta, representa el 93.1% de la media.

Tabla LXXXIV
Parámetros poblacionales de la variable aleatoria
identificación de hiatos

Media	2,132	Mínimo	0
Mediana	2	Máximo	7
Desviación estándar	1,987	Sesgo	0,067
Varianza	3,948	kurtosis	-0,043
Error estándar	0,154	Primer cuartil	0
Rango	5	Tercer cuartil	3
Moda	0	Rango intercuartil	3
Suma	356	Coefficiente de variación	0.931

En el gráfico 3.62 se puede observar que la distribución de esta variable aleatoria, está sesgada positivamente, dado que el coeficiente de sesgo es 0.067, por lo que se considera que es una pregunta difícil. El coeficiente de kurtosis (-0.043), indica que la distribución tiene forma platicúrtica, es decir que es más aplanada que la de la normal.

En la tabla LXXXV se muestran las frecuencias de las observaciones de esta variable aleatoria. La calificación excelente corresponde a los

valores 6 y 7 de la variable, y solamente 8 de cada 100 estudiantes obtuvieron esta calificación.

Gráfico 3.62
Histograma de frecuencias de la variable aleatoria identificación de hiatos

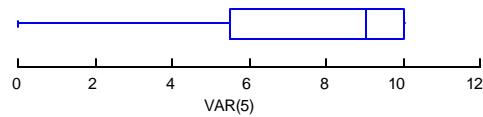
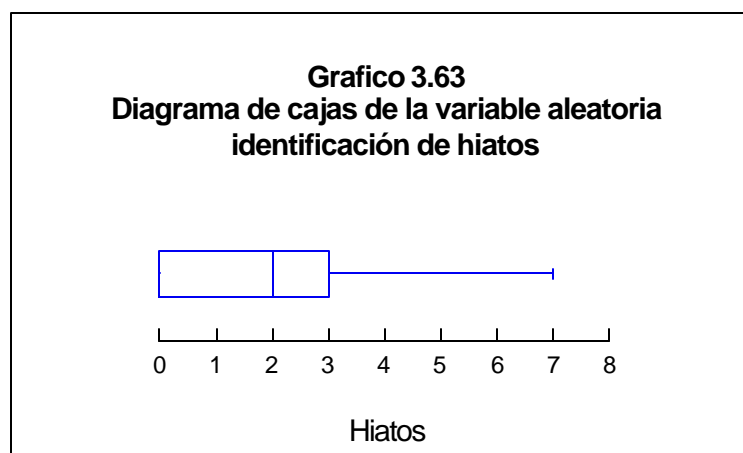


Tabla LXXXV
Tabla de frecuencias de la variable aleatoria identificación de hiatos

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	54	0,32	54	0,32
1	7	0,04	61	0,37
2	46	0,28	107	0,64
3	26	0,16	133	0,80
4	12	0,07	145	0,87
5	9	0,05	154	0,92
6	5	0,03	159	0,95
7	8	0,05	167	1,00



3.4.4 Variable aleatoria identificación de palabras agudas

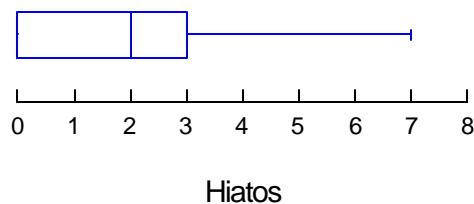
Los resultados de los parámetros poblacionales obtenidos se pueden observar en la tabla LXXXVI. Un valor representativo de esta población es la moda cuyo valor es 0, esto indica que existe una gran proporción de estudiantes que no saben identificar palabras agudas, concretamente 28 de cada 100 estudiantes están en esta situación, la cual corresponde al calificación malo. El resultado obtenido de la media es 2.024, el cual corresponde a la calificación bueno, aproximadamente 17 de cada 100 estudiantes obtuvieron esta calificación. El valor de la desviación estándar (1.632), es considerado alto, debido a que representa el 80.6% del valor de la media poblacional, este valor se obtiene a partir del coeficiente de variación.

Tabla LXXXVI
Parámetros poblacionales de la variable aleatoria
identificación de palabras agudas

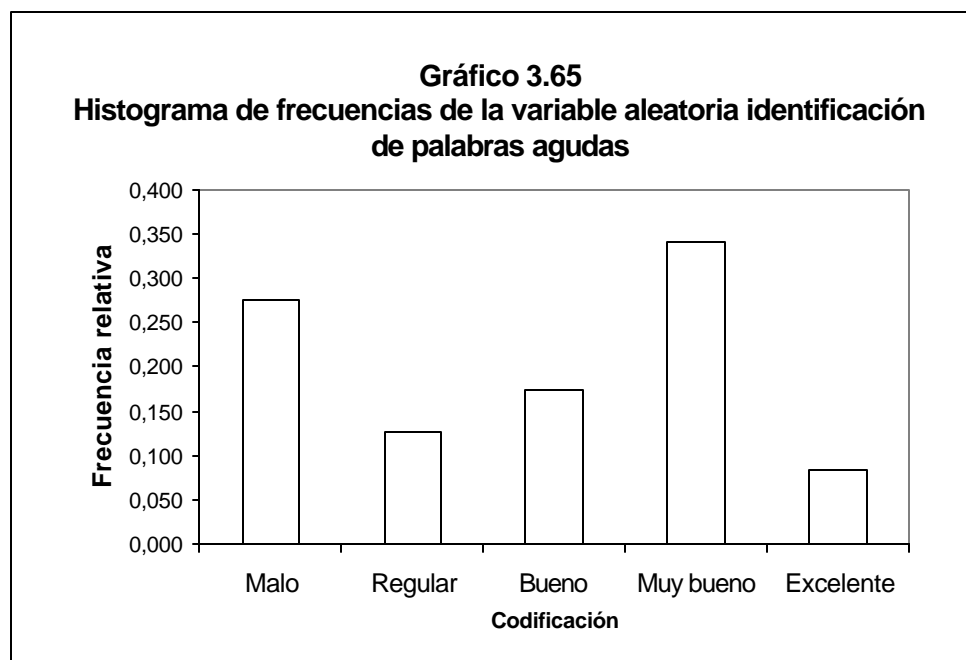
Media	2,024	Mínimo	0
Mediana	2	Máximo	5
Desviación estándar	1,632	Sesgo	0,197
Varianza	2,663	kurtosis	-1,113
Error estándar	0,126	Primer cuartil	0
Rango	5	Tercer cuartil	3
Moda	0	Rango intercuartil	3
Suma	338	Coefficiente de variación	0.806

El valor del tercer cuartil es 3, esto significa que el 75% de las observaciones son iguales o menores que este valor, mediante el análisis de las observaciones se obtuvo que 81 de cada 100 estudiantes, identificaron correctamente tres o menos palabras con acento agudo.

Grafico 3.64
Diagrama de cajas de la variable aleatoria
identificación de palabras agudas



El coeficiente de sesgo de esta variable aleatoria es 0.197, por lo que la distribución está sesgada hacia la derecha, esto significa que la pregunta es difícil. Esta situación se puede apreciar más claramente en el gráfico 3.65.



El coeficiente de kurtosis cuyo valor es -1.113, indica que la distribución es platicúrtica, lo cual significa que tiene una forma más achatada que la de distribución normal. De las frecuencias de la variable aleatoria analizada los cuales se muestran en la tabla LXXXVI, se obtuvo que 34 de cada 100 estudiantes entrevistados obtuvieron la calificación muy bueno y que 8 de cada 100 la calificación excelente.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.28 + 0.13e^t + 0.17e^{2t} + 0.23e^{3t} + 0.11e^{4t} + 0.08e^{5t}$$

Tabla LXXXVII
Frecuencias de la variable aleatoria identificación
de palabras agudas

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	46	0,28	46	0,28
1	21	0,13	67	0,40
2	29	0,17	96	0,57
3	39	0,23	135	0,81
4	18	0,11	153	0,92
5	14	0,08	167	1,00

3.4.5 Variable aleatoria identificación de palabras graves

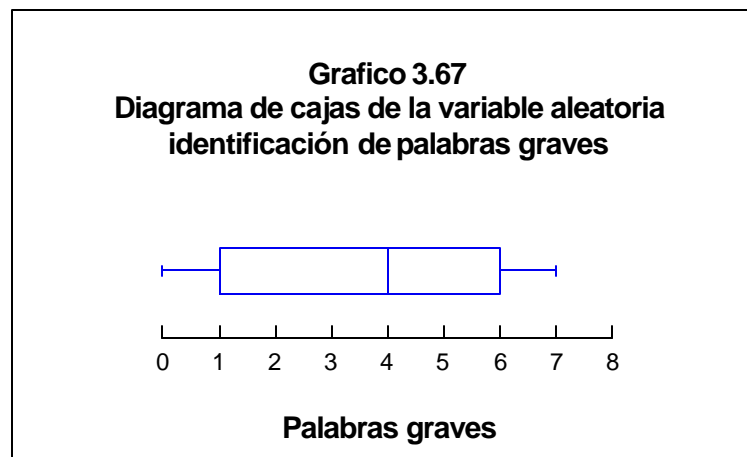
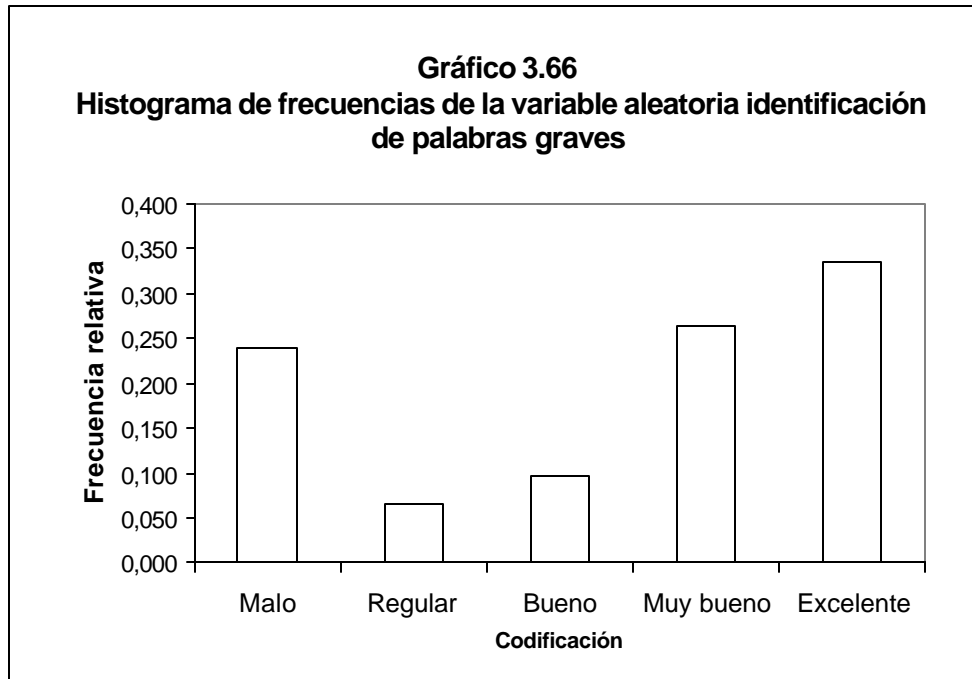
Los valores de los parámetros que definen esta población son mostrados en la tabla LXXXVIII. La moda de la variable aleatoria analizada en esta sección es 7, es decir que el valor que más se repite, es el de número máximo de respuestas correctas de esta variable, el valor de la moda está incluido en la calificación excelente, para la que 34 de cada 100 estudiantes obtuvieron esta calificación. La media y la mediana de esta

población son 3.754 y 4 respectivamente, el valor de la media corresponde a la calificación bueno y el de la mediana está incluido en la calificación muy bueno; por cada 100 estudiantes 9 obtienen la calificación de bueno, mientras que esta proporción aumenta para la calificación muy bueno, pues por cada 100 estudiantes existen 26, que la obtienen.

Tabla LXXXVIII
Parámetros poblacionales de la variable aleatoria
identificación de palabras graves

Media	3,754	Mínimo	0
Mediana	4	Máximo	7
Desviación estándar	2,708	Sesgo	-0,238
Varianza	7,333	Kurtosis	-1,47
Error estándar	0,21	Primer cuartil	1
Rango	7	Tercer cuartil	6
Moda	7	Rango intercuartil	5
Suma	627	Coefficiente de variación	0.721

En el gráfico 3.66 se puede apreciar que la distribución de esta variable aleatoria está sesgada hacia la izquierda, debido a que el coeficiente de asimetría es negativo (-0.238), esto significa que la pregunta es fácil. En lo que respecta al coeficiente de kurtosis cuyo valor es -1.47, este indica que la distribución es platicúrtica, es decir que la forma es más achatada que la de la distribución normal.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.24 + 0.07e^t + 0.04e^{2t} + 0.05e^{3t} + 0.13e^{4t} + 0.13e^{5t} + 0.09e^{6t} + 0.25e^{7t}$$

Existe una proporción alta de estudiantes que no saben identificar palabras que tienen acento grave pues alrededor de 24 de cada 100 estudiantes obtuvieron la calificación malo que corresponde al valor cero. En la tabla LXXXIX se puede apreciar las frecuencias de las observaciones de la variable aleatoria analizada en esta sección.

Tabla LXXXIX
Tabla de frecuencias de la variable aleatoria
identificación de palabras graves

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	40	0,24	40	0,24
1	11	0,07	51	0,31
2	7	0,04	58	0,35
3	9	0,05	67	0,40
4	22	0,13	89	0,53
5	22	0,13	111	0,66
6	15	0,09	126	0,75
7	41	0,25	141	0,84

3.4.6 Variable aleatoria identificación de palabras esdrújulas

La codificación utilizada en esta variable es malo si la respuesta es cero, bueno si es uno y excelente si es 2. Como se puede observar en la tabla de parámetros XC, los valores de las medidas de tendencia central, la

media y la mediana son respectivamente 1.03 y 1, lo que significa que en promedio los estudiantes identificaron un acento esdrújulo correctamente, la moda es de 2, es decir que el valor que más se repite corresponde al mayor número de aciertos posibles. La desviación estándar de esta variable aleatoria 0.888 es alta ya que representa el 86.2% de variación con respecto a la media.

Tabla XC
Parámetros poblacionales de la variable aleatoria
identificación de palabras esdrújulas

Media	1,03	Mínimo	0
Mediana	1	Máximo	2
Desviación estándar	0,888	Sesgo	-0,059
Varianza	0,788	Kurtosis	-1,737
Error estándar	0,069	Primer cuartil	0
Rango	2	Tercer cuartil	2
Moda	2	Rango intercuartil	2
Suma	172	Coefficiente de variación	0.862

El coeficiente de asimetría es negativo(-0.059) por lo tanto la distribución está sesgada hacia la izquierda, lo que significa que la pregunta es fácil, este hecho se puede apreciar en el gráfico 3.68. En lo que respecta al coeficiente de kurtosis (-1.737) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

En la tabla XCI se puede apreciar las frecuencias de la variable identificación de palabras esdrújulas, con estos valores se puede afirmar

que tanto para la codificación malo y para la codificación excelente existe una proporción similar de estudiantes, la cual es 38 de cada 100 estudiantes y 41 de cada 100 estudiantes respectivamente.

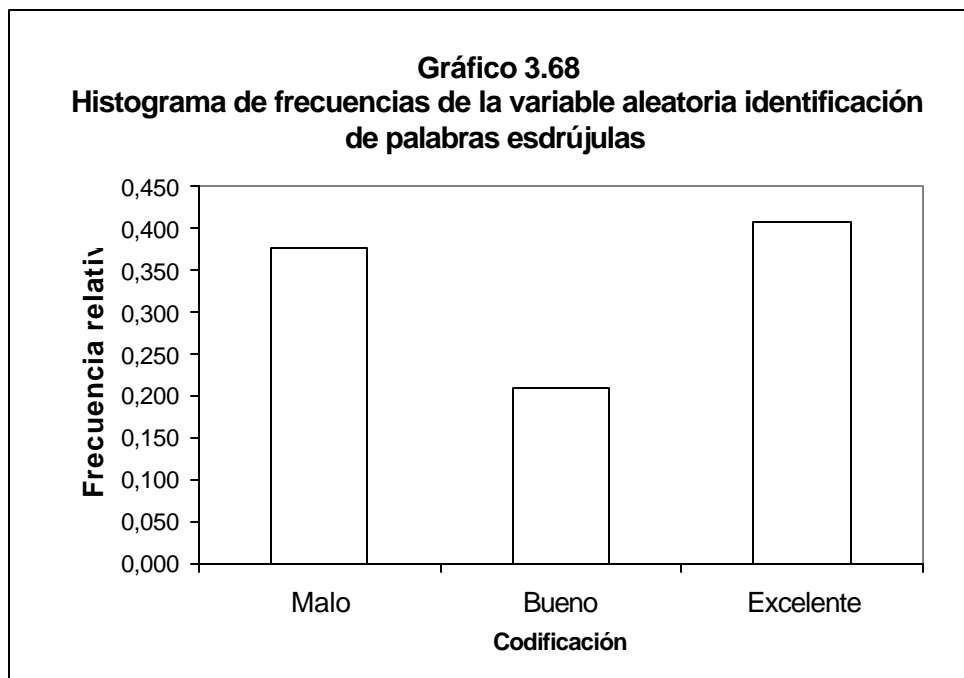
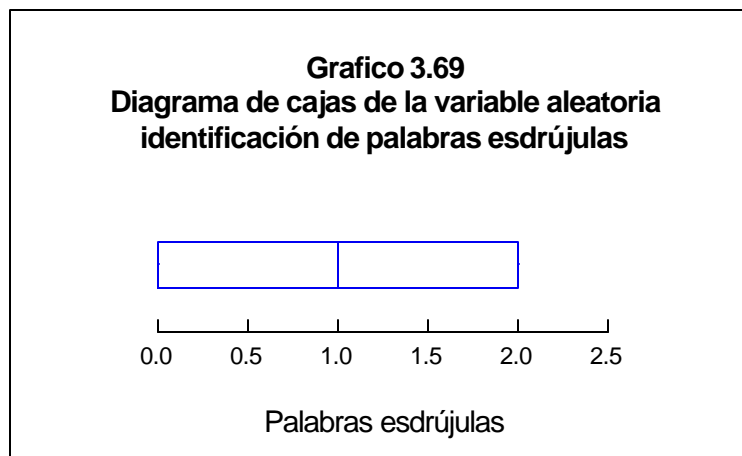


Tabla XCI
Frecuencias de la variable aleatoria identificación de palabras esdrújulas

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	63	0,38	63	0,38
1	35	0,21	98	0,59
2	68	0,41	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.38 + 0.21e^t + 0.41e^{2t}$$



3.4.7 Variable aleatoria identificación de palabras sobresdrújulas

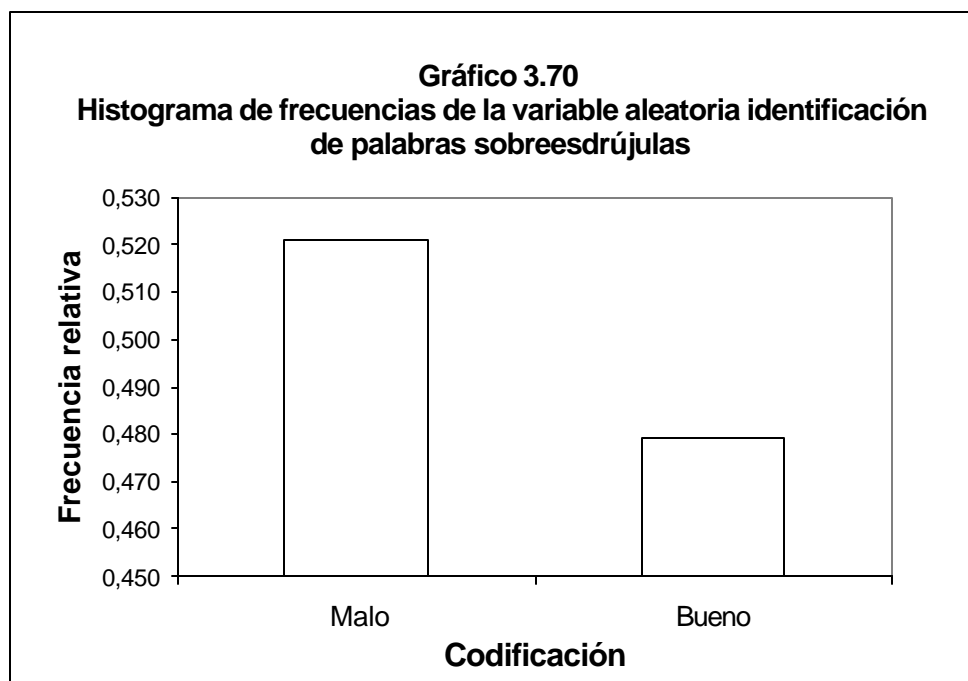
Los valores de tendencia central moda y mediana de esta variable son iguales a cero lo que indica que existe una gran proporción de estudiantes que no saben identificar palabras con acento sobresdrújulo, esto se confirma si se observa el valor de la media que es 0.479, pues alrededor de este valor se agrupan las observaciones.

Como se puede observar en la tabla XCII el coeficiente de variación es mayor a uno lo cual indica que la variación de las observaciones es de más del 100% con respecto a la media.

Tabla XCII
Parámetros poblacionales de la variable aleatoria
identificación de palabras sobresdrújulas

Media	0,479	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,501	Sesgo	0,085
Varianza	0,251	Kurtosis	-2,017
Error estándar	0,039	Primer cuartil	0
Rango	1	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	80	Coefficiente de variación	1.045

El coeficiente de sesgo es positivo (0.085) lo que indica que la distribución está sesgada hacia la derecha como se observa en el gráfico 3.70, esto quiere decir que la pregunta es difícil. Por otro lado el coeficiente de kurtosis que es -2.017 este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.



La variable aleatoria identificación de palabras sobresdrújulas tiene dos resultados posibles 0 (si no identificó correctamente) o 1 (si identificó correctamente). La probabilidad de obtener éxito, es decir 0, es $p = 0.52$ la probabilidad de fracaso que corresponde al valor 1, es $q = 1-p = 0.48$, entonces X es una variable aleatoria Bernoulli, tal que

$$P(X=x) = \binom{1}{x} 0.52^x (0.48)^{1-x} \quad x=0, 1$$

La función generadora de momentos de esta variable de estudio es:

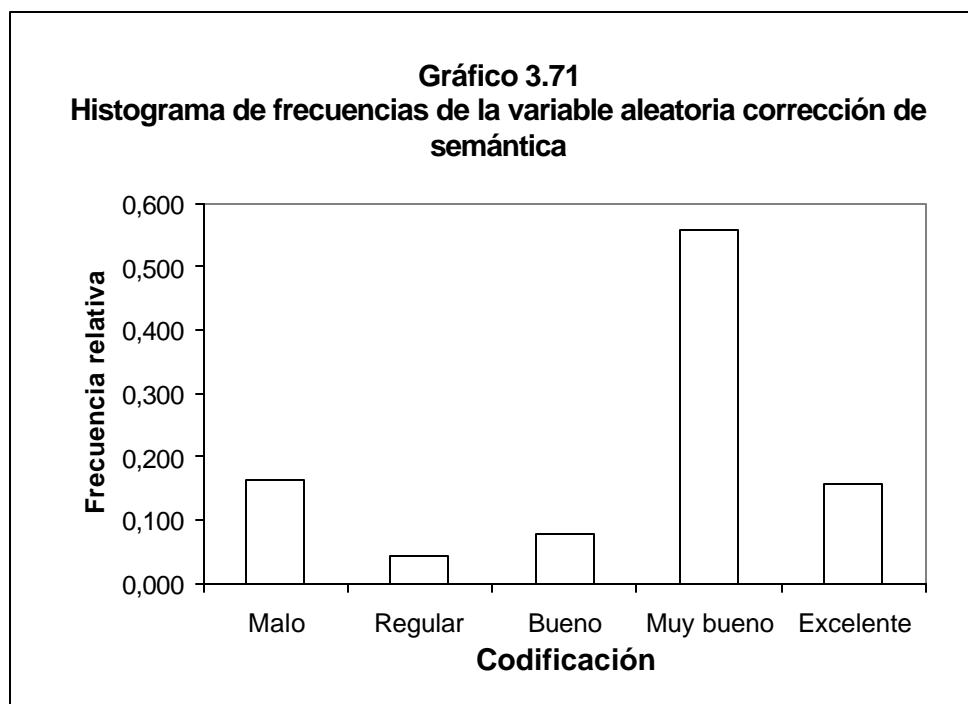
$$M(t) = 0.52 + 0.48e^t$$

3.4.8 Variable aleatoria sentido de la oración

Como se puede observar en la tabla de parámetros XCIII los valores de tendencia central obtenidos son la media igual a 3.86, mediana y moda igual a 5, es decir que los estudiantes en una gran proporción tienen conocimientos para corregir la semántica de las oraciones. En lo que respecta a las medidas de dispersión se puede observar que el valor de la desviación estándar es 2.073, el cual es un valor alto si se considera que la variación relativa de las observaciones con respecto a la media es 53.9%.

Tabla XCIII
Parámetros poblacionales de la variable aleatoria
corrección de semántica

Media	3,868	Mínimo	0
Mediana	5	Máximo	7
Desviación estándar	2,073	Sesgo	-0,787
Varianza	4,297	kurtosis	-0,415
Error estándar	0,16	Primer cuartil	3
Rango	7	Tercer cuartil	5
Moda	5	Rango intercuartil	2
Suma	641	Coefficiente de variación	0,539



En el gráfico 3.71 se puede observar que debido a que el coeficiente de sesgo es negativo (-0.787), la distribución está sesgada hacia la izquierda, por lo tanto la pregunta es fácil. El coeficiente de kurtosis (-0.415) de esta variable aleatoria indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

Los resultados que se muestran en la tabla de frecuencias XCIV indican que 56 de cada 100 estudiantes entrevistados tienen un muy buen conocimiento para corregir la semántica de las oraciones.

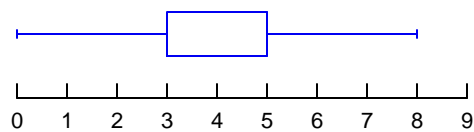
Tabla XCIV
Frecuencias de la variable aleatoria corrección de semántica

Clase	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	27	0,16	27	0,13
1	7	0,04	34	0,20
2	4	0,02	38	0,23
3	9	0,05	47	0,28
4	37	0,22	84	0,50
5	56	0,34	141	0,84
6	19	0,11	160	0,96
7	7	0,04	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.16 + 0.04e^t + 0.02e^{2t} + 0.05e^{3t} + 0.22e^{4t} + 0.34e^{5t} + 0.11e^{6t} + 0.04e^{7t}$$

Grafico 3.72
Diagrama de cajas de la variable aleatoria corrección de semántica



Corrección de semántica

3.4.9 Variable aleatoria identificación de sinónimas

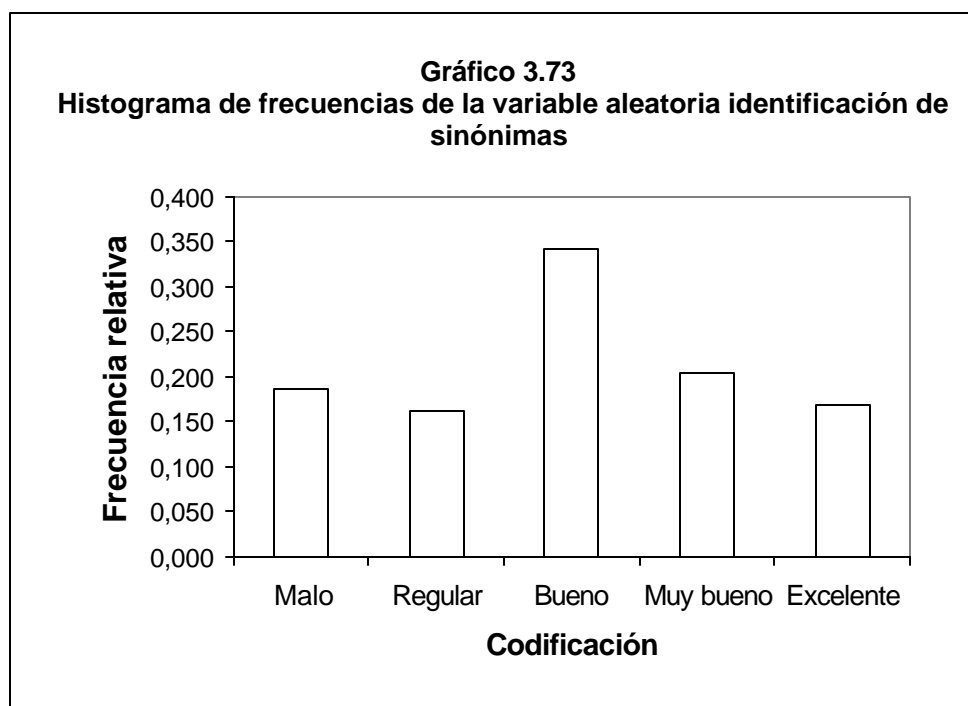
Esta variable cuenta con un máximo de 8 respuestas correctas. De los valores de tendencia central la moda (0), indica que el valor que más se repite es aquel que representa que los estudiantes no saben identificar palabras sinónimas.

La desviación estándar es alta (2.304), si se observa que el porcentaje de variación relativa de las observaciones de la media, es 69.8%.

Media	3,299	Mínimo	0
Mediana	3	Máximo	8
Desviación estándar	2,304	Sesgo	0,071
Varianza	5,308	Kurtosis	-0,951
Error estándar	0,178	Primer cuartil	1
Rango	8	Tercer cuartil	5
Moda	0	Rango intercuartil	4
Suma	551	Coefficiente de variación	69.8

Por otro lado el coeficiente de asimetría (0.071), indica que la distribución está ligeramente sesgada hacia la derecha, por lo tanto la pregunta tiene un grado de dificultad medio. En lo que respecta al coeficiente de kurtosis (-0.951) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. Los valores 3 y 4 que puede tomar la variable aleatoria son aquellos que corresponden a la codificación

bueno, la cual representa que aproximadamente 34 de cada 100 estudiantes tienen este número de respuestas, sin embargo para el número máximo de respuestas que es 8 y entra en la codificación de excelente junto con 7, existen tan solo 10 de cada 100 estudiantes entrevistados que saben identificar sinónimos correctamente



En el gráfico 3.73 se puede apreciar que la mayor concentración de observaciones caen en la codificación bueno y para el resto de categorías las proporciones son parecidas, las frecuencias de esta variable aleatoria se pueden observar en la tabla XCVI.

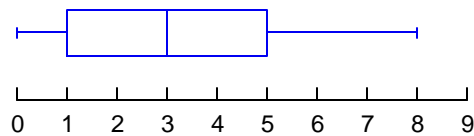
Tabla XCVII
Frecuencias de la variable aleatoria identificación
de sinónimas

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	31	0,19	31	0,19
1	15	0,09	46	0,28
2	12	0,07	58	0,35
3	27	0,16	85	0,51
4	30	0,18	115	0,69
5	23	0,14	138	0,83
6	11	0,07	149	0,89
7	14	0,08	163	0,98
8	4	0,02	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.1 + 0.09e^t + 0.07e^{2t} + 0.16e^{3t} + 0.18e^{4t} + 0.14e^{5t} + 0.07e^{6t} + 0.08e^{7t} + 0.02e^{8t}$$

Grafico 3.74
Diagrama de cajas de la variable aleatoria
identificación de sinónimas



Sinónimas

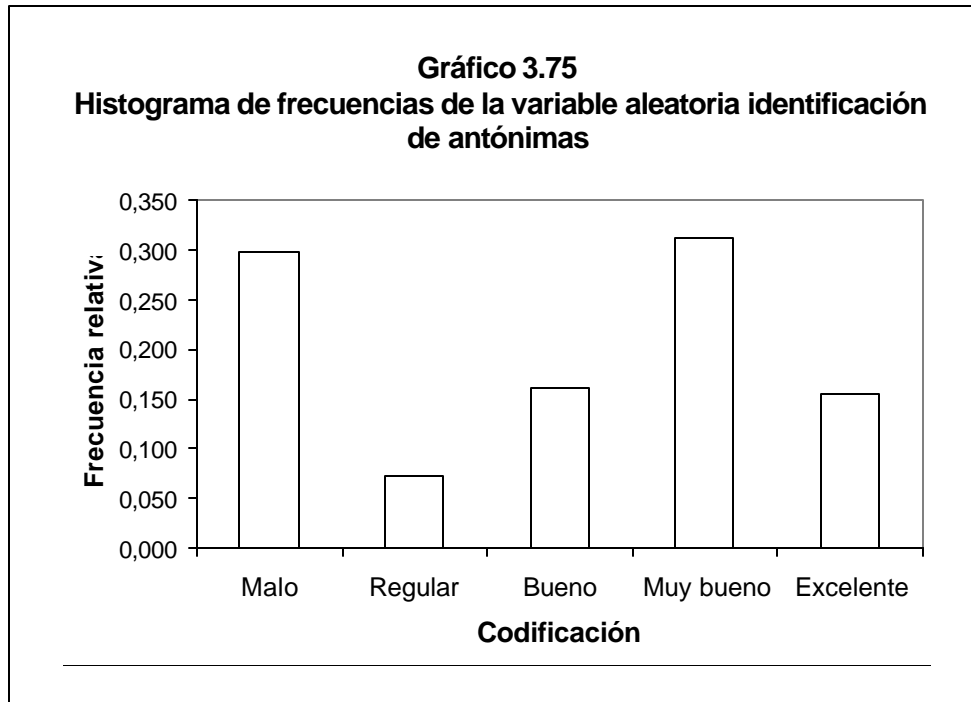
3.4.10 Variable aleatoria identificación de antónimas

Los parámetros obtenidos de esta población se pueden observar en la tabla XCVII, en lo que respecta a las medidas de tendencia central, la media (3.551) y la mediana (4) son valores que están alrededor de 4, que corresponde a la codificación bueno, y la moda (0) está en la categoría malo; esta situación se presenta porque la desviación estándar de las observaciones es alta (2.804), pues la variación relativa de las observaciones con respecto a la media del 78.9%.

Media	3,551	Mínimo	0
Mediana	4	Máximo	8
Desviación estándar	2,804	Sesgo	-0,078
Varianza	7,862	Kurtosis	-1,442
Error estándar	0,217	Primer cuartil	0
Rango	8	Tercer cuartil	6
Moda	0	Rango intercuartil	6
Suma	593	Coefficiente de variación	0.789

Debido a que el coeficiente de asimetría es negativo(-0.078) la distribución está ligeramente sesgada hacia la izquierda, esto quiere decir que la pregunta es medianamente fácil. En el gráfico 3.75 se muestra el histograma de frecuencia relativa de esta variable aleatoria.

La distribución de la variable identificación de sinónimas tiene forma platicúrtica, debido a que el coeficiente de kurtosis (-1.442) es menor a 3 que es el coeficiente de kurtosis de la distribución normal.



Los resultados obtenidos de la distribución de esta variable indican que 30 de cada 100 estudiantes entrevistados no saben identificar palabras antónimas, que es representada por la codificación malo, y que una proporción similar 31 de cada 100 estudiantes identificaron correctamente 3 o 4 sinónimas lo que corresponde a la codificación bueno. La proporción de estudiantes que si sabe identificar sinónimas es de 15 de cada 100.

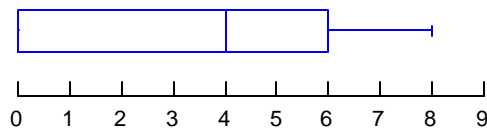
Tabla XCVIII
Frecuencias de la variable aleatoria identificación
de antónimas

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	50	0,30	50	0,30
1	5	0,03	55	0,33
2	7	0,04	62	0,37
3	15	0,09	77	0,46
4	12	0,07	89	0,53
5	25	0,15	114	0,68
6	27	0,16	141	0,84
7	14	0,08	155	0,93
8	12	0,07	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.30 + 0.03e^t + 0.04e^{2t} + 0.09e^{3t} + 0.07e^{4t} + 0.15e^{5t} + 0.016e^{6t} + 0.08e^{7t} + 0.07e^{8t}$$

Grafico 3.76
Diagrama de cajas de la variable aleatoria
identificación de antónimas



3.4.11 Variable aleatoria identificación de palabras aumentativas

Como se puede observar en la tabla IC el valor de la media(1.06) y la mediana (1) están muy cerca, lo quiere decir que existe una agrupación de las observaciones alrededor de este punto. El máximo número de respuestas correctas y la moda es 2, es decir que es el valor que más se repite es el del mayor número de aciertos. La variación existente de las observaciones es muy alta (0.961) si se considera que el coeficiente de variación 0.906 lo cual indica que la variación relativa de las observaciones con respecto a la media es del 90.6%.

Media	1,06	Mínimo	0
Mediana	1	Máximo	2
Desviación estándar	0,961	Sesgo	-0,121
Varianza	0,923	Kurtosis	-1,922
Error estándar	0,074	Primer cuartil	0
Rango	2	Tercer cuartil	2
Moda	2	Rango intercuartil	2
Suma	177	Coefficiente de variación	0.906

El coeficiente de sesgo (-0.121) indica que la distribución está sesgada hacia la izquierda, por lo tanto la pregunta es medianamente fácil;

además la distribución es platicúrtica dado que el coeficiente de kurtosis es -1.922. Esta situación se puede apreciar en el gráfico 3.77.

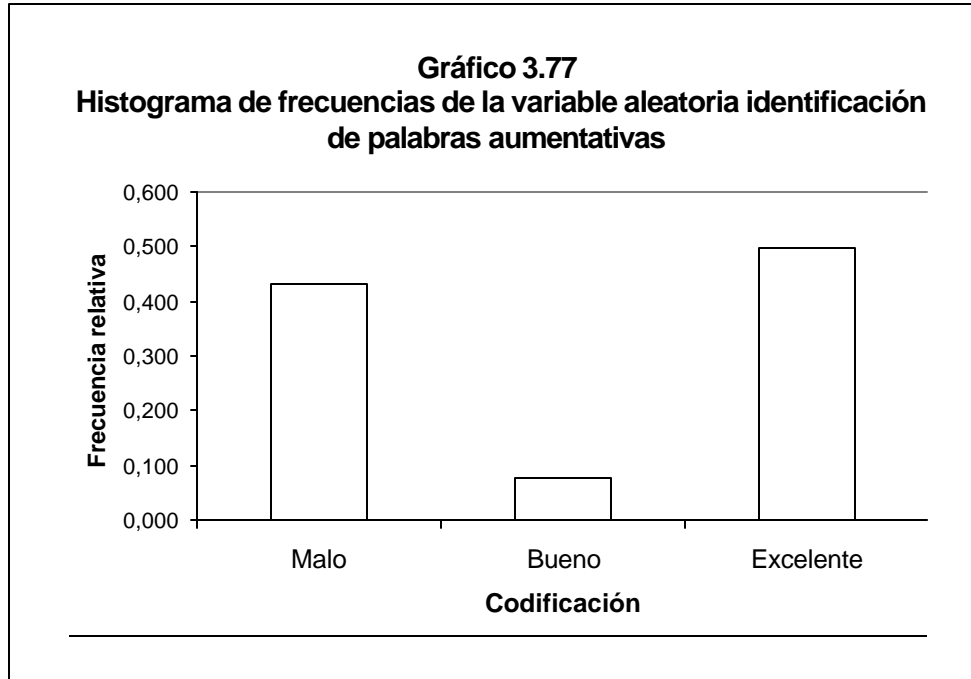
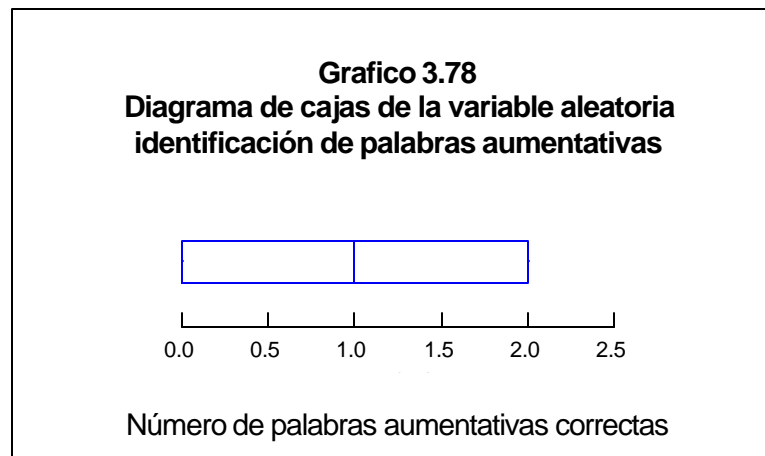


Tabla XCVIII
Frecuencias de la variable aleatoria identificación de palabras aumentativas

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	72	0,43	72	0,43
1	13	0,08	85	0,51
2	82	0,49	167	1,00

En la tabla C se muestran las frecuencias absolutas y relativas; acumuladas y acumuladas relativas de esta variable, a partir de estos resultados se puede afirmar que por cada 100 estudiantes entrevistados 43 no saben identificar palabras aumentativas, mientras que 47 si saben identificar este tipo de palabras.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.43 + 0.08e^t + 0.49^{2t}$$

3.4.12 Variable aleatoria identificación de palabras diminutivas

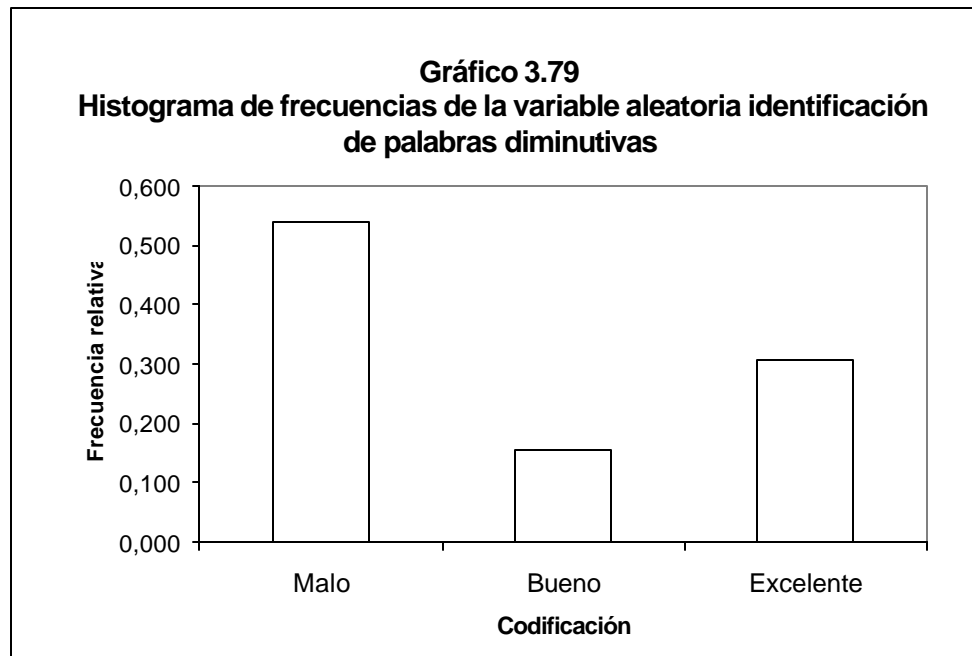
Los resultados obtenidos de los parámetros de esta variable se pueden observar en la tabla Cl. Los valores de la moda y la mediana es 0, que corresponde a cero respuestas correctas y la media es 0.766 que es un valor cercano a una respuesta correcta. La variación existente en las

observaciones es muy alta, la desviación estándar (0.891) representa más del 100% de variación con respecto a la media de la población.

Tabla CI
Parámetros poblacionales de la variable aleatoria
identificación de palabras diminutivas

Media	0,766	Mínimo	0
Mediana	0	Máximo	2
Desviación estándar	0,891	Sesgo	0,478
Varianza	0,793881	Kurtosis	-1,579
Error estándar	0,069	Primer cuartil	0
Rango	2	Tercer cuartil	2
Moda	0	Rango intercuartil	2
Suma	128	Coefficiente de variación	1.163

El coeficiente de asimetría (0.478) indica que la distribución está sesgada hacia la derecha, por lo que la pregunta es medianamente difícil; el coeficiente de kurtosis es -1.579, lo que significa que la distribución es platicúrtica, es decir más achatada que la distribución normal. En la tabla CII se pueden observar los valores de las frecuencias de esta variable aleatoria, de estos valores se puede afirmar que por cada 100 estudiantes evaluados 54 no saben identificar palabras aumentativas y 31 si saben hacerlo.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.54 + 0.16e^t + 0.31e^{2t}$$

Tabla CII
Frecuencias de la variable aleatoria identificación de palabras diminutivas

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	90	0,54	90	0,54
1	26	0,16	116	0,69
2	51	0,31	167	1,00

3.4.13 Variable aleatoria identificación de palabras despectivas

La variable aleatoria identificación de palabras despectivas tiene dos resultados posibles 0 (si no identificó la palabra) o 1 (si identificó la palabra). La probabilidad de obtener éxito, es decir 0, es $p = 0.69$ la probabilidad de fracaso que corresponde al valor 1, es $q = 1-p = 0.31$, entonces X es una variable aleatoria Bernulli, tal que

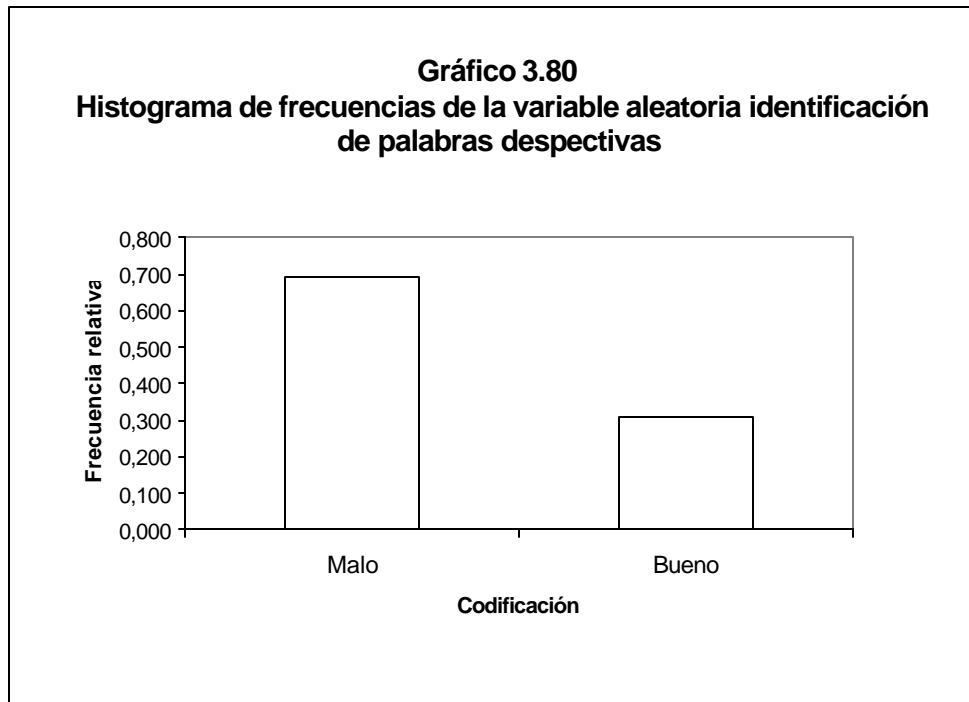
$$P(X=x) = \binom{1}{x} 0.69^x (0.31)^{1-x} \quad x=0, 1$$

Tabla CIII
Parámetros poblacionales de la variable aleatoria
identificación de palabras despectivas

Media	0,305	Mínimo	0
Mediana	0	Máximo	1
Desviación estándar	0,462	Sesgo	0,853
Varianza	0,213444	Kurtosis	-1,288
Error estándar	0,036	Primer cuartil	0
Rango	1	Tercer cuartil	1
Moda	0	Rango intercuartil	1
Suma	51	Coefficiente de variación	1.514

El valor esperado de esta variable es 51.77, como se puede observar en la tabla CIII, además el resultado de la moda y de la mediana es 0, es decir cero respuestas correctas

Debido a que el coeficiente de sesgo es positivo(0.853) la distribución está sesgada hacia la derecha, esto indica que las observaciones tienden a agruparse hacia los valores más pequeños de la variable. En lo que respecta al coeficiente de kurtosis (-1.288) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal. Estos resultados pueden ser apreciados en el gráfico 3.80.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.69 + 0.31e^t$$

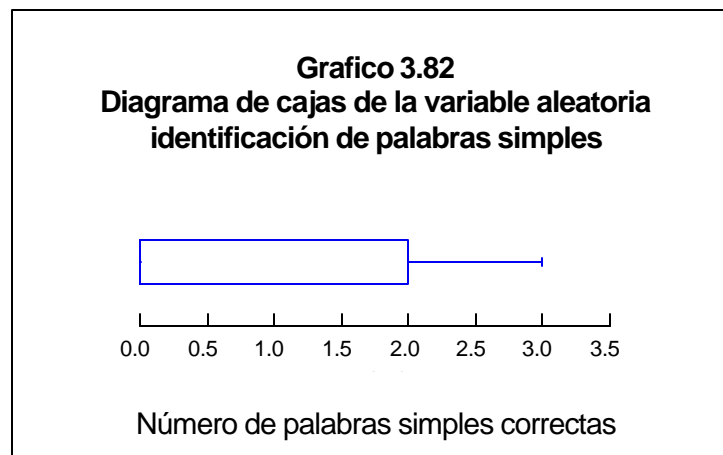
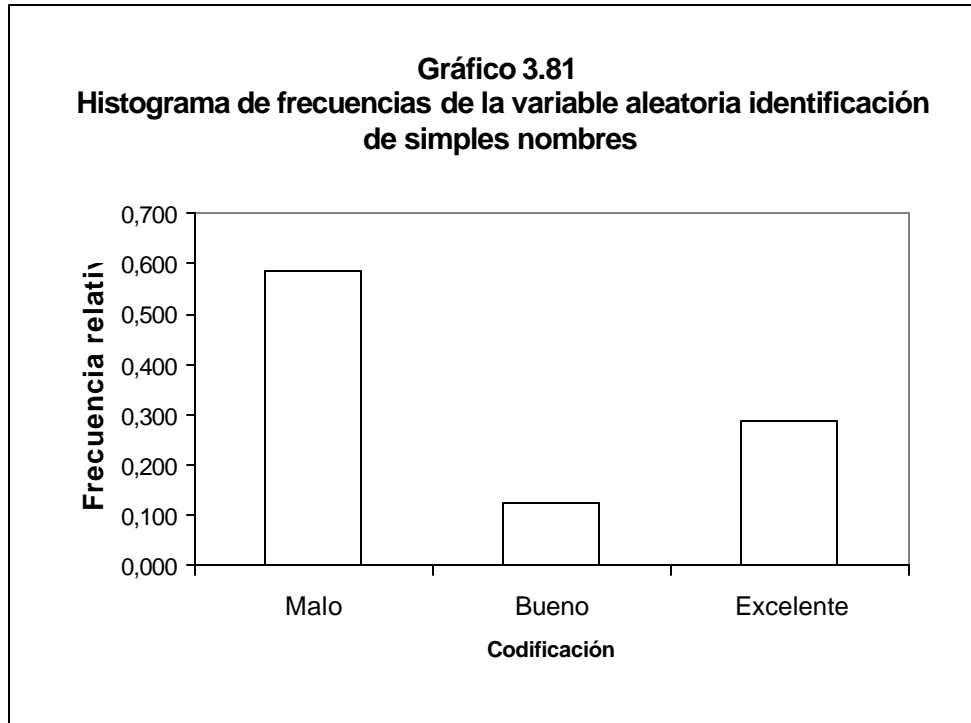
3.4.14 Variable aleatoria identificación de simples nombres

Los resultados posibles de esta variable son cuatro 0, 1, 2 y 3 los cuales indican el número de respuestas correctas. En la tabla CIV muestran los valores de los parámetros obtenidos, la media (0.73), la moda (0) y la mediana (0) indican que las observaciones están agrupadas alrededor de los valores cero y uno. La variación de las observaciones es alta, pues la desviación estándar (1.042), representa una variación relativa de las observaciones con respecto a la media del 141%.

Media	0,734	Mínimo	0
Mediana	0	Máximo	3
Desviación estándar	1,042	Sesgo	0,896
Varianza	1,085	Kurtosis	-0,678
Error estándar	0,081	Primer cuartil	0
Rango	3	Tercer cuartil	2
Moda	0	Rango intercuartil	2
Suma	131	Coefficiente de variación	1.41

En el gráfico 3.81 se puede observar que la distribución está sesgada hacia la derecha, pues el coeficiente de sesgo es positivo(0.896), por lo tanto la pregunta es difícil; además la distribución es platicúrtica dado

que el coeficiente de kurtosis es $-0,678 < 3$, que es el coeficiente de kurtosis de la normal.



Las frecuencias relativa de esta variable se pueden observar en la tabla CV, de esta se obtiene que 59 de cada 100 estudiantes entrevistados no saben clasificar a la palabras como simples nombres y tan solo 8 de cada 100 saben hacerlo.

Tabla CV
Frecuencias de la variable aleatoria identificación de simples nombres

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	98	0,59	98	0,59
1	21	0,13	119	0,71
2	34	0,20	153	0,92
3	14	0,08	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.59 + 0.13e^t + 0.20e^{2t} + 0.08e^{3t}$$

3.4.15 Variable aleatoria vocabulario

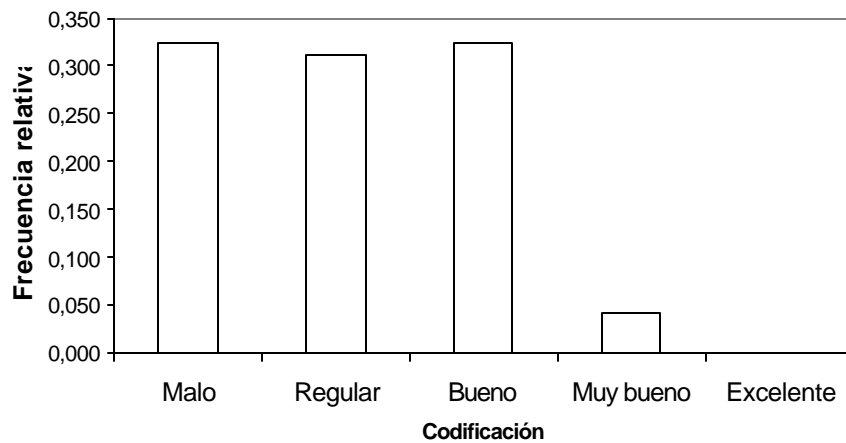
El número de respuestas correctas posibles es 0, 1, 2, 3 y 4. En la tabla CVI se muestran los resultados de los parámetros poblacionales obtenidos, el valor de la media (1.08) y el de la mediana (1) están alrededor de una respuesta correcta, esto indica que las observaciones se agrupan hacia este valor. Por otro lado el valor del rango (3) indica

que ningún estudiante pudo identificar correctamente todos los significados de las palabras.

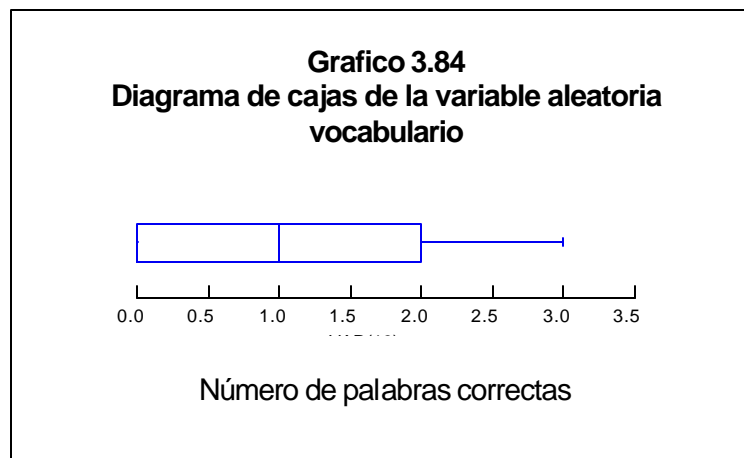
Tabla CVI
Parámetros poblacionales de la variable aleatoria
vocabulario

Media	1,084	Mínimo	0
Mediana	1	Máximo	3
Desviación estándar	0,901	Sesgo	0,183
Varianza	0,811	Kurtosis	-1,096
Error estándar	0,07	Primer cuartil	0
Rango	3	Tercer cuartil	2
Moda	2	Rango intercuartil	2
Suma	181	Coefficiente de variación	0.831

Gráfico 3.83
Histograma de frecuencias de la variable aleatoria vocabulario



La distribución de esta variable está sesgada hacia la derecha, como indica el coeficiente de sesgo (0.183), por lo tanto la pregunta es difícil. Esta situación puede apreciarse en el gráfico 3.83. En lo que respecta al coeficiente de kurtosis (-1.096) este indica que la distribución es platicúrtica.



De los valores de la frecuencia relativa mostrados en la tabla CVII se puede observar que existe una proporción similar de los alumnos que identifican correctamente el significado de 0, 1 y 2 palabras, esto es 32, 31 y 32 de cada 100 estudiantes respectivamente, esta situación se puede apreciar en el gráfico de la distribución.

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.32 + 0.31e^t + 0.32e^{2t} + 0.04e^{3t}$$

Tabla CVII
Frecuencias de la variable aleatoria vocabulario

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	54	0,32	54	0,32
1	52	0,31	106	0,63
2	54	0,32	160	0,96
3	7	0,04	167	1,00

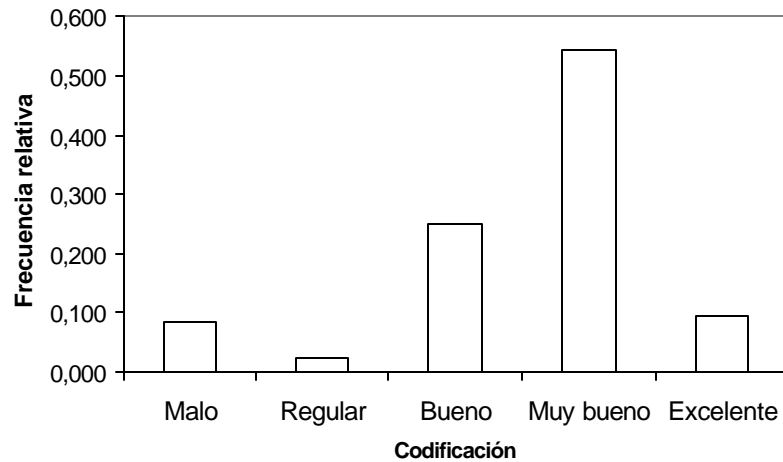
3.4.16 Variable aleatoria conjugación de verbos

El máximo de respuestas correctas posibles en esta variable es 6. De los parámetros poblacionales que se observan en la tabla CVIII, las medidas de tendencia central la media (3.784) y la mediana (4), tienen valores aproximados a cuatro lo cual indica que existe una mayor proporción de observaciones que se agrupan en la codificación muy bueno que también incluye el valor 5 y que es el que más se repite. La variación de las observaciones se puede analizar mediante el coeficiente de variación, el cual indica que la desviación estándar de las observaciones con respecto a la media es del 43%. El rango que es otra medida de dispersión es 6, lo que significa que las observaciones toman valores entre 0 y 6.

Tabla CVIII
Parámetros poblacionales de la variable aleatoria
conjugación de verbos

Media	3,784	Mínimo	0
Mediana	4	Máximo	6
Desviación estándar	1,629	Sesgo	-0,917
Varianza	2,653	Kurtosis	0,253
Error estándar	0,126	Primer cuartil	3
Rango	6	Tercer cuartil	5
Moda	5	Rango intercuartil	2
Suma	632	Coefficiente de variación	0.430

Gráfico 3.85
Histograma de frecuencias de la variable aleatoria
conjugación de verbos



En el gráfico 3.85 se puede apreciar que la distribución de esta variable aleatoria está sesgada hacia la izquierda, por lo tanto la pregunta es fácil. La distribución de esta variable es platicúrtica, es decir que es más

achatada que la distribución normal, debido a que el coeficiente de kurtosis es 0.253.

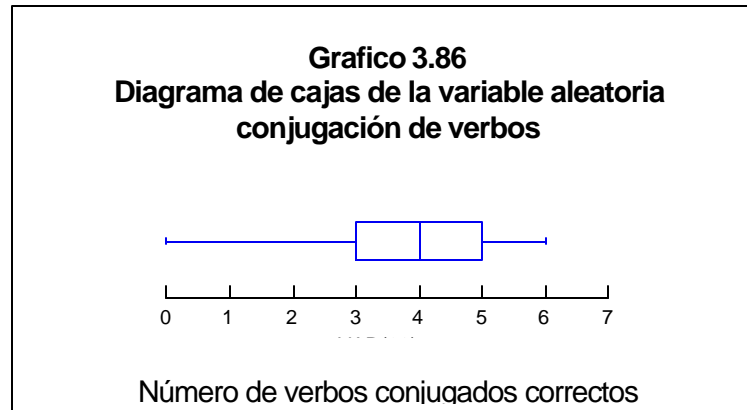


Tabla CIX
Frecuencias de la variable aleatoria conjugación de verbos

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	14	0,08	14	0,08
1	4	0,02	18	0,11
2	10	0,06	28	0,17
3	32	0,19	60	0,36
4	39	0,23	99	0,59
5	52	0,31	151	0,90
6	16	0,10	167	1,00

En la tabla CIX se muestran las frecuencias de la variable aleatoria conjugación de verbos, de la frecuencia relativa se obtuvo que 54 de cada 100 estudiantes saben conjugar verbos con una calificación de muy

bueno, y que la proporción de estudiantes que saben conjugar correctamente los verbos (10 de cada 100) es mayor a la de los que no saben conjugar verbos (8 de cada 100).

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.08 + 0.02e^t + 0.06e^{2t} + 0.19e^{3t} + 0.23e^{4t} + 0.31e^{5t} + 0.10e^{6t}$$

3.4.17 Variable aleatoria identificación del sujeto

Los valores que puede tomar esta variable son 0, 1, 2, 3 y 4. Los parámetros obtenidos son mostrados en la tabla CX. La media de esta población es 2.371, es decir que alrededor de este valor se agrupan las observaciones, la mediana que es también una medida de tendencia central es 2 e indica que sobre este valor se encuentran el 50% de las observaciones y la moda es 3, es decir que es el valor que más se repite. La dispersión de las observaciones es alta ya que existe una variación relativa del 52.2% de la desviación estándar con respecto a la media.

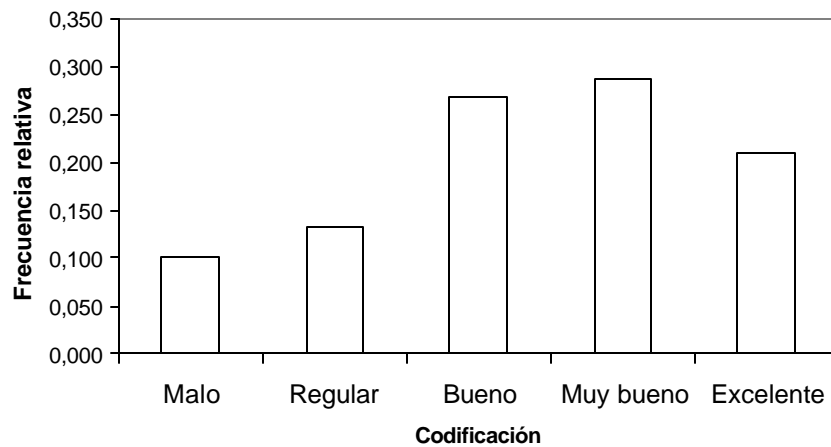
En el gráfico 3.87 se puede observar que la distribución está sesgada hacia la izquierda dado que el coeficiente de sesgo es negativo(-0.392), por lo tanto la pregunta es medianamente fácil; en lo que respecta al

coeficiente de kurtosis (-0.752) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

Tabla CX
Parámetros poblacionales de la variable aleatoria
identificación del sujeto

Media	2,371	Mínimo	0
Mediana	2	Máximo	4
Desviación estándar	1,239	Sesgo	-0,392
Varianza	1,535121	Kurtosis	-0,752
Error estándar	0,096	Primer cuartil	2
Rango	4	Tercer cuartil	3
Moda	3	Rango intercuartil	1
Suma	396	Coeficiente de variación	0.522

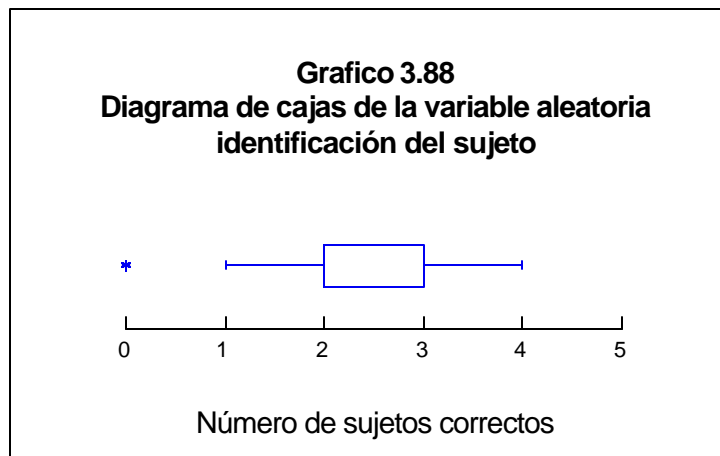
Gráfico 3.87
Histograma de frecuencias de la variable aleatoria identificación
del sujeto



En esta variable se obtuvo como resultado las siguientes proporciones que 10 de cada 100 estudiantes entrevistados no saben identificar el sujeto de la oración mientras que más del doble es decir que 21 de cada 100 si lo saben hacer. Los valores que corresponden a las categorías de bueno y muy bueno son similares esto es 27 y 29 de cada 100 estudiantes respectivamente. Estos resultados se pueden observar en la tabla CXI.

Tabla CXI
Frecuencias de la variable aleatoria identificación del sujeto

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	17	0,10	17	0,10
1	22	0,13	39	0,23
2	45	0,27	84	0,50
3	48	0,29	132	0,79
4	35	0,21	167	1,00



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.10 + 0.13e^t + 0.27e^{2t} + 0.29e^{3t} + 0.21e^{4t}$$

3.4.18 Variable aleatoria identificación del predicado

Los valores que puede tomar esta variable son 0, 1, 2, 3 y 4. Como se puede observar en la tabla CXII el resultado de la media es 2.341 lo que significa que las observaciones se localizan alrededor de este valor, es decir que los estudiantes en promedio aproximadamente tienen dos respuestas correctas, los resultados de la moda y la mediana que es 2, confirman este hecho. La dispersión de las observaciones se la puede analizar por medio del coeficiente de dispersión, el cual indica que existe una variación relativa de las observaciones del 51% con respecto a la media.

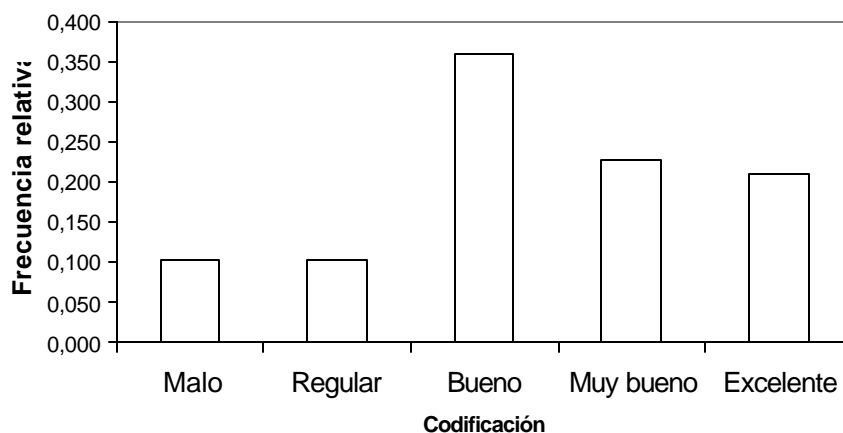
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.10 + 0.10e^t + 0.36e^{2t} + 0.23e^{3t} + 0.21e^{4t}$$

Tabla CXII
Parámetros poblacionales de la variable aleatoria
identificación del predicado

Media	2,341	Mínimo	0
Mediana	2	Máximo	4
Desviación estándar	1,211	Sesgo	-0,312
Varianza	1,466	Kurtosis	-0,621
Error estándar	0,094	Primer cuartil	2
Rango	4	Tercer cuartil	3
Moda	2	Rango intercuartil	1
Suma	391	Coefficiente de variación	0.517

Gráfico 3.89
Histograma de frecuencias de la variable aleatoria identificación
del predicado

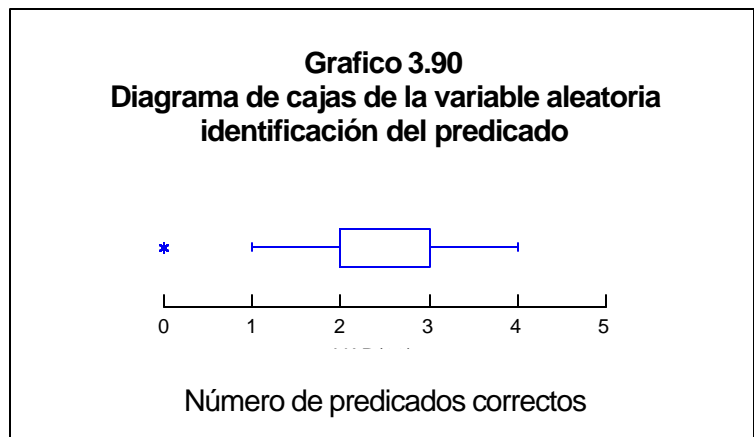


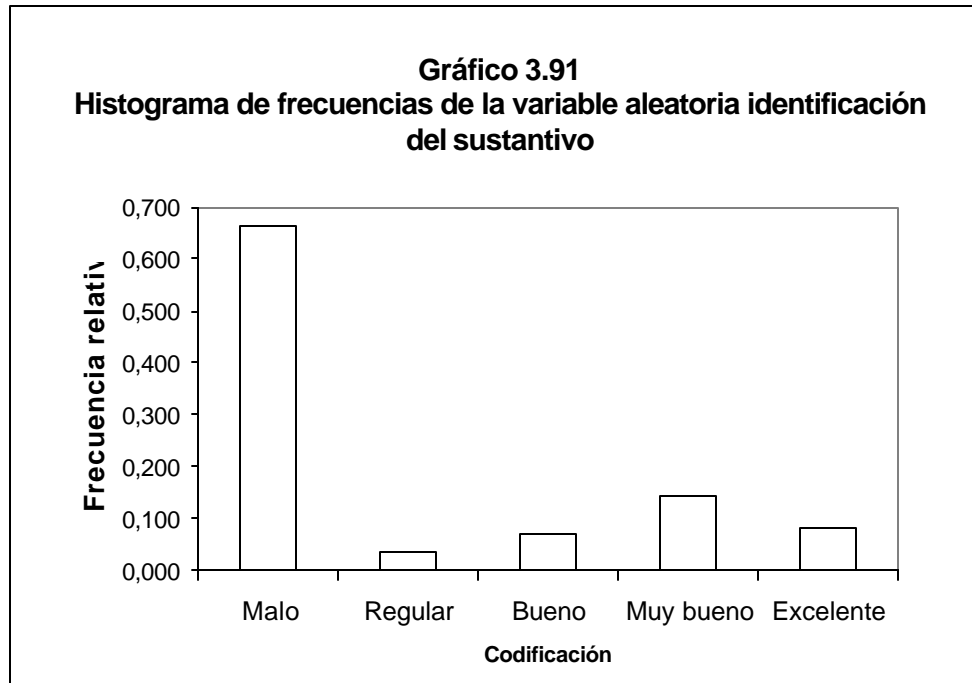
En el gráfico 3.89 se puede observar que la distribución está sesgada hacia la izquierda, pues el coeficiente de sesgo es negativo (-0.312), por lo tanto la pregunta es fácil; además la distribución es platicúrtica, es

decir más achatada que la distribución normal, dado el coeficiente de kurtosis (-0.621) de esta variable aleatoria. En la tabla CXIII se muestran las frecuencias de la variable analizada en esta sección, de los resultados de la frecuencia relativa se puede afirmar que 10 de cada 100 estudiantes entrevistados no saben identificar el predicado de la oración, y 21 de cada 100 si saben hacerlo, además los estudiantes que respondieron correctamente el valor de la mediana son 36 de cada 100.

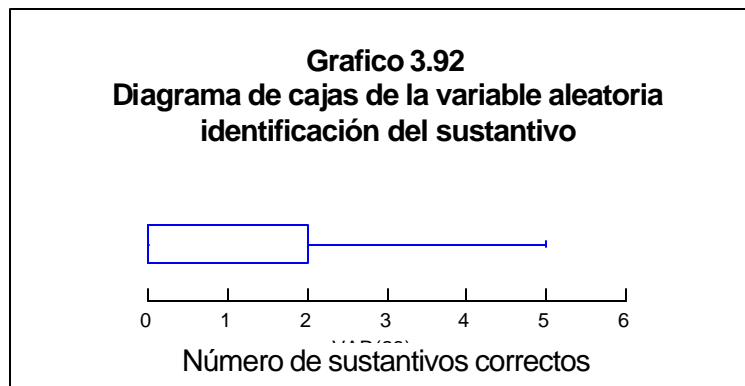
Tabla CXIII
Frecuencias de la variable aleatoria identificación del predicado

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	17	0,10	17	0,10
1	17	0,10	34	0,20
2	60	0,36	94	0,56
3	38	0,23	132	0,79
4	35	0,21	167	1,00





Gráficamente se puede apreciar que la distribución esta sesgada hacia la derecha, el coeficiente de sesgo es positivo (1.204), por lo tanto la pregunta es difícil. En lo que respecta al coeficiente de kurtosis (-0.176) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.



La proporción de estudiantes que no saben identificar sustantivos en la oración que pertenece lo cual está codificado como malo, es de 66 de cada 100 y solamente 8 de cada 100 estudiantes sabe hacerlo. Estos resultados se pueden obtener en base a la frecuencia relativa de la variable que se muestra en la tabla CXV.

Tabla CXV
Frecuencias de la variable aleatoria identificación del sustantivo

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	111	0,66	111	0,66
1	6	0,04	117	0,70
2	12	0,07	129	0,77
3	8	0,05	137	0,82
4	16	0,10	153	0,92
5	14	0,08	167	1,00

La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.66 + 0.04e^t + 0.07e^{2t} + 0.05e^{3t} + 0.10e^{4t} + 0.08e^{5t}$$

3.4.20 Variable aleatoria identificación del verbo

Los resultados posibles de esta variable son 0, 1, 2, 3 y 4. Los parámetros de esta población se pueden observar en la tabla CXVI. Las

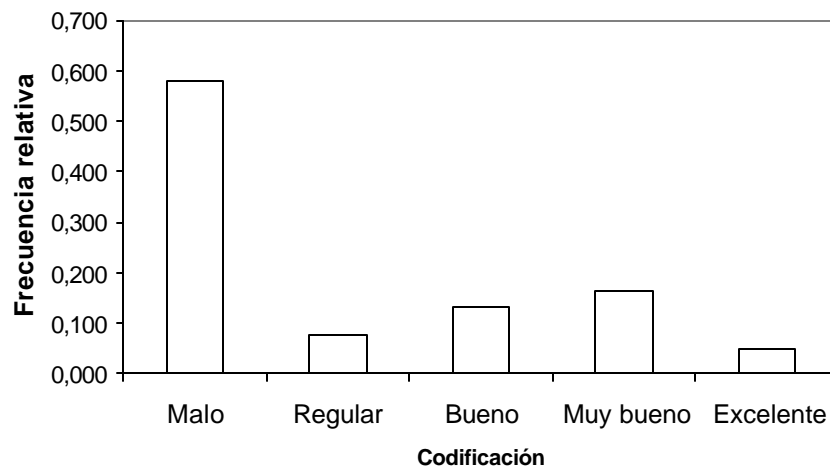
medidas de tendencia central de esta variable indican que el conocimiento de los estudiantes para identificar el verbo de la oración es muy bajo, por ejemplo la media que es 1.018 significa que en promedio los estudiantes solamente tienen una respuesta correcta.

Tabla CXVI
Parámetros poblacionales de la variable aleatoria
identificación del verbo

Media	1,018	Mínimo	0
Mediana	0	Máximo	4
Desviación estándar	1,342	Sesgo	0,86
Varianza	1,800	Kurtosis	0,794
Error estándar	0,104	Primer cuartil	0
Rango	4	Tercer cuartil	2
Moda	0	Rango intercuartil	2
Suma	170	Coefficiente de variación	1.318

La mediana y la moda son iguales a 0, este resultado tiene una proporción de 58 estudiantes por cada 100 entrevistados. La dispersión de las observaciones es muy alta, si se observa el coeficiente de variación este indica que la desviación estándar representa una variación relativa del 131.8% con respecto a la media. En el gráfico 3.93 se puede observar que la distribución esta sesgada hacia la derecha dado que el coeficiente de sesgo es positivo(0.086), en lo que respecta al coeficiente de kurtosis (0.794) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

Gráfico 3.93
Histograma de frecuencias de la variable aleatoria identificación del verbo



Para la codificación excelente que corresponde al máximo de respuestas correctas (4), se tiene que solamente 5 de cada 100 estudiantes saben identificar el verbo en la oración, el número de respuestas correctas que antecede a la codificación excelente es (3) y corresponde al código muy bueno, la situación con este mejora un poco pues por cada 100 estudiantes 16 tienen este número de respuestas.

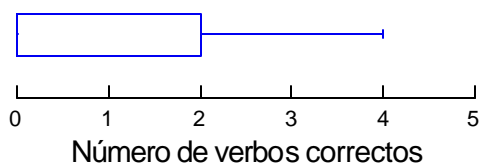
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.58 + 0.08e^t + 0.13e^{2t} + 0.16e^{3t} + 0.05e^{4t}$$

Tabla CXVII
Frecuencias de la variable aleatoria identificación del verbo

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	97	0,58	97	0,58
1	13	0,08	110	0,66
2	22	0,13	132	0,79
3	27	0,16	159	0,95
4	8	0,05	167	1,00

Grafico 3.94
Diagrama de cajas de la variable aleatoria identificación del verbo



3.4.21 Variable aleatoria identificación de frases

Esta variable representa el número de frases identificadas correctamente por los estudiantes que cursan el décimo año de educación básica. Para la evaluación se planteó un máximo de 2 frases para identificar.

Los parámetros que definen esta población se muestran en la tabla CXVIII, entre estos resultados se encuentran la media (1.401), la

mediana (2) y la moda (2) que son medidas de tendencia central que indican hacia que valores tienden a localizarse las observaciones, en este caso se puede decir que existe una gran proporción de estudiantes que saben identificar frases, exactamente por cada 100 estudiantes entrevistados 62 están en esta situación.

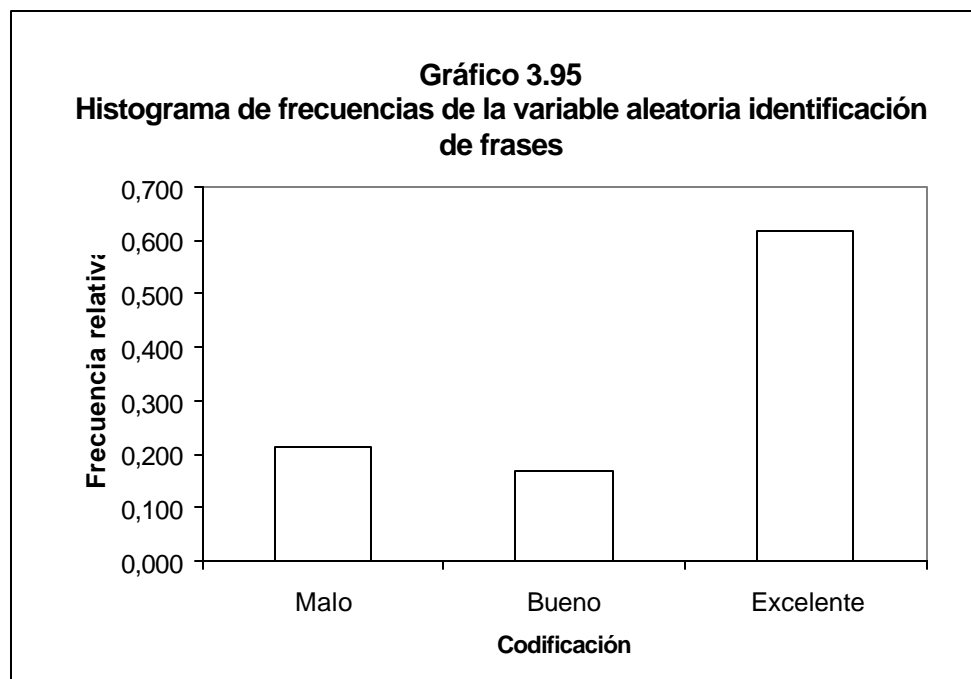
Tabla CXVIII
Parámetros poblacionales de la variable aleatoria
identificación de frases

Media	1,401	Mínimo	0
Mediana	2	Máximo	2
Desviación estándar	0,822	Sesgo	-0,865
Varianza	0,675	Kurtosis	-0,974
Error estándar	0,064	Primer cuartil	1
Rango	2	Tercer cuartil	2
Moda	2	Rango intercuartil	1
Suma	234	Coefficiente de variación	0.586

De las medidas de dispersión la desviación estándar que es 0.822 representa una variación del 58.6% del valor de la media de la población (1.401), lo cual es un porcentaje alto si se considera que el rango de la variable es 2.

El coeficiente de asimetría es negativo(-0.865) por lo tanto la distribución está sesgada hacia la izquierda, por lo tanto la pregunta es fácil, además es platicúrtica es decir más achatada que la distribución normal dado que

el coeficiente de kurtosis es -0.974 menor a tres que es el coeficiente de kurtosis de una variable aleatoria normal. Esta situación se puede apreciar en el gráfico 3.95 que muestra el histograma de frecuencia absoluta de la variable analizada.



De los resultados que se muestran en la tabla CXIX se obtuvo que por cada 100 estudiantes 22 no saben identificar frases lo que es considerado como malo y 17 de cada 100 identifican tan solo una frase correctamente, lo cual es un resultado considerado bueno.

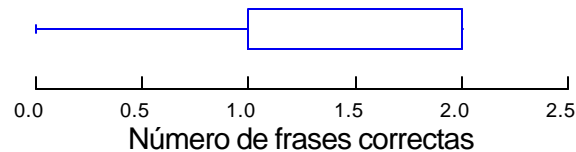
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.22 + 0.17e^t + 0.62e^{2t}$$

Tabla CXIX
Frecuencias de la variable aleatoria identificación de frases

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	36	0,22	36	0,22
1	28	0,17	64	0,38
2	103	0,62	167	1,00

Grafico 3.96
Diagrama de cajas de la variable aleatoria identificación de frases



3.4.22 Variable aleatoria identificación de oraciones

Los valores que puede tomar esta variable son 0, 1, 2 o 3 respuestas correctas. En la tabla CXX se muestra los valores de los parámetros poblacionales. Para describir hacia que punto tienden a agruparse las

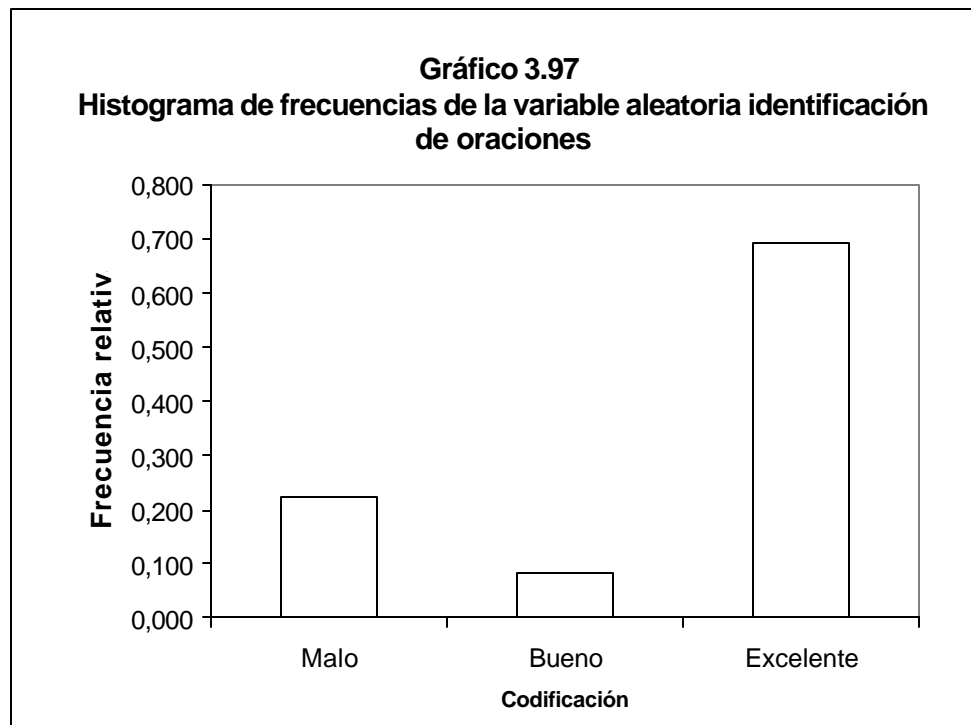
observaciones se utiliza la media cuyo valor 1.898, además el valor de la mediana que es 2 y el de la moda 3, indican que la localización de las observaciones es hacia un intervalo lo que conlleva a que existe mucha variación en las observaciones. La desviación estándar es 1.18 y el coeficiente de variación indica que existe una variación relativa del 62.1%.

En el gráfico 3.97 se puede apreciar que la distribución está sesgada hacia la izquierda dado que el coeficiente de sesgo es negativo (-0,624), por lo tanto la pregunta es fácil. En lo que respecta al coeficiente de kurtosis (-1.143) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

Tabla CXX
Parámetros poblacionales de la variable aleatoria
identificación de oraciones

Media	1,898	Mínimo	0
Mediana	2	Máximo	3
Desviación estándar	1,18	Sesgo	-0,624
Varianza	1,392	Kurtosis	-1,143
Error estándar	0,091	Primer cuartil	1
Rango	3	Tercer cuartil	3
Moda	3	Rango intercuartil	2
Suma	317	Coefficiente de variación	0,621

La codificación excelente contiene los números de respuesta 2 y 3 y es la que mayor proporción alcanza, pues por cada 100 estudiantes entrevistados 70 obtuvieron esta codificación, esta afirmación se da en base a las frecuencias de la variable aleatoria mostradas tabla CXXI.



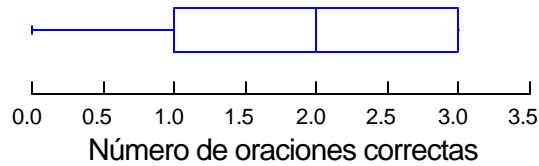
La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.22 + 0.08e^t + 0.27e^{2t} + 0.43e^{3t}$$

Tabla CXXI
Frecuencias de la variable aleatoria identificación de oraciones

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	37	0,22	37	0,22
1	14	0,08	51	0,31
2	45	0,27	96	0,57
3	71	0,43	167	1,00

Grafico 3.98
Diagrama de cajas de la variable aleatoria identificación de oraciones



3.4.23 Variable aleatoria clasificación de oraciones

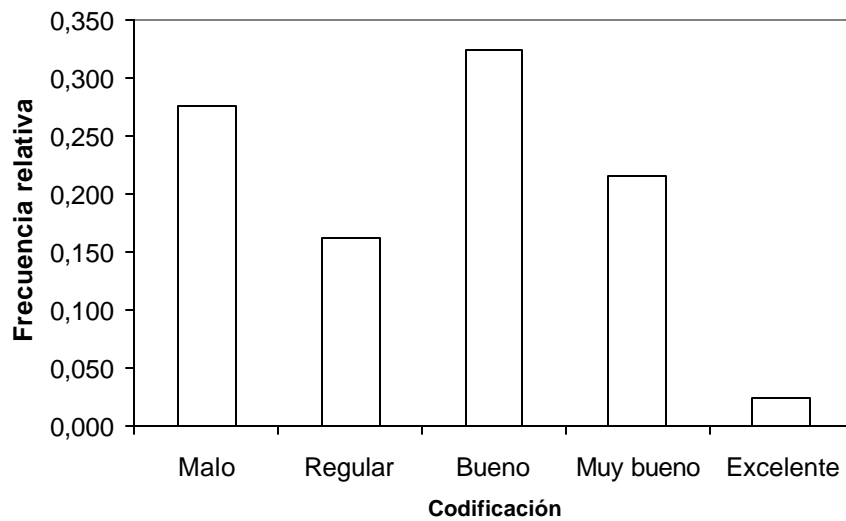
El número máximo de respuestas correctas es 7, sin embargo como se puede observar en la tabla CXXII ningún estudiante alcanzó este puntaje. La media de esta variable es 1.989, la mediana es 2 y la moda 0, esto significa que alrededor de los valores más bajo se localizan las observaciones. Para tener una idea de la dispersión de los datos se

utiliza el valor de la desviación estándar 1.742 es un valor alto ya que representa 87.5% del valor de la media.

Tabla CXXII
Parámetros poblacionales de la variable aleatoria
clasificación de oraciones

Media	1,989	Mínimo	0
Mediana	2	Máximo	6
Desviación estándar	1,742	Sesgo	0,516
Varianza	3,034	kurtosis	-0,813
Error estándar	0,135	Primer cuartil	0
Rango	6	Tercer cuartil	3
Moda	0	Rango intercuartil	3
Suma	332	Coefficiente de variación	0.875

Gráfico 3.99
Histograma de frecuencias de la variable aleatoria clasificación
de oraciones



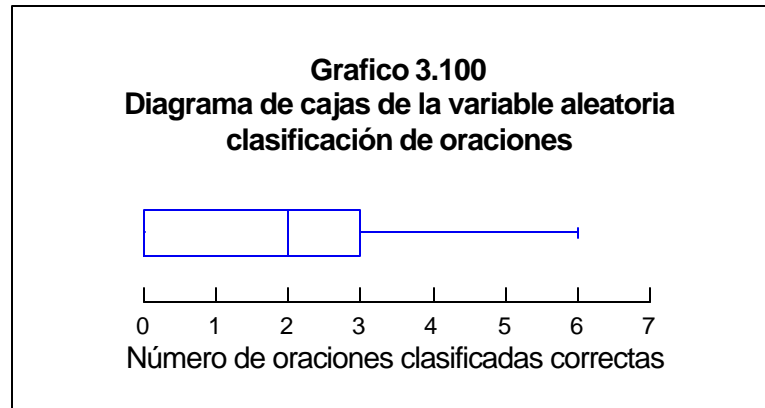
La distribución de esta variable aleatoria tiene forma platicúrtica debido a que el coeficiente de kurtosis es -0.813 menor a 3, es decir más achatada que la distribución normal. Por otro lado el coeficiente de sesgo es positivo (0.516) lo que indica que la distribución está sesgada hacia la derecha, por lo tanto la pregunta es difícil; esta situación se puede apreciar mejor si se observa el gráfico 3.99.

Tabla CXXIII
Frecuencias de la variable aleatoria clasificación de oraciones

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	46	0,28	46	0,28
1	27	0,16	73	0,44
2	39	0,23	112	0,67
3	15	0,09	127	0,76
4	22	0,13	149	0,89
5	14	0,08	163	0,98
6	4	0,02	167	1,00

En la tabla CXXIII se muestran las frecuencias absolutas y relativas; acumuladas y acumuladas relativas. De estos resultados se puede decir que tan solo 2 de cada 100 estudiantes entrevistados saben clasificar todo tipo de oraciones, que 21 de cada 100 tienen muy buen

conocimiento para clasificar oraciones, que 28 de cada 100 no saben clasificar las oraciones.



La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.28 + 0.16e^t + 0.23e^{2t} + 0.09e^{3t} + 0.13e^{4t} + 0.08e^{5t} + 0.02e^{6t}$$

3.4.24 Variable aleatoria corrección de sintaxis

El número máximo posible de oraciones corregidas es 6, sin embargo el valor máximo calculado en esta variable es 4, lo que significa que ningún estudiante pudo corregir correctamente todas las oraciones.

Como se puede observar en la tabla CXXIV, los valores de las medidas de tendencia central son para la media 0.341, la mediana 0 y la moda 0 esto significa el conocimiento de los estudiantes para corregir la sintaxis de las frases u oraciones es muy bajo, pues tan solo 1 de cada 100 de los entrevistados obtuvo codificación muy bueno.

Tabla CXXIV
Parámetros poblacionales de la variable aleatoria
corrección de sintaxis

Media	0,341	Mínimo	0
Mediana	0	Máximo	4
Desviación estándar	0,734	Sesgo	2,319
Varianza	0,538	Kurtosis	5,326
Error estándar	0,057	Primer cuartil	0
Rango	4	Tercer cuartil	0
Moda	0	Rango intercuartil	0
Suma	57	Coefficiente de variación	2.152

Los parámetros del coeficiente de sesgo y de kurtosis indican que la distribución está sesgada hacia la derecha, por lo tanto la pregunta es difícil, y que es leptocúrtica, es decir que tiene un pico más alto que la distribución normal, respectivamente, lo cual se puede apreciar en el gráfico 3.101. En base a las frecuencias relativas calculadas para esta variables se puede decir que 78 de cada 100 de los estudiantes entrevistados no tiene conocimiento para corregir la sintaxis de las

oraciones, en la tabla CXXV se muestran las frecuencias absolutas y relativas; acumuladas y acumuladas relativas.

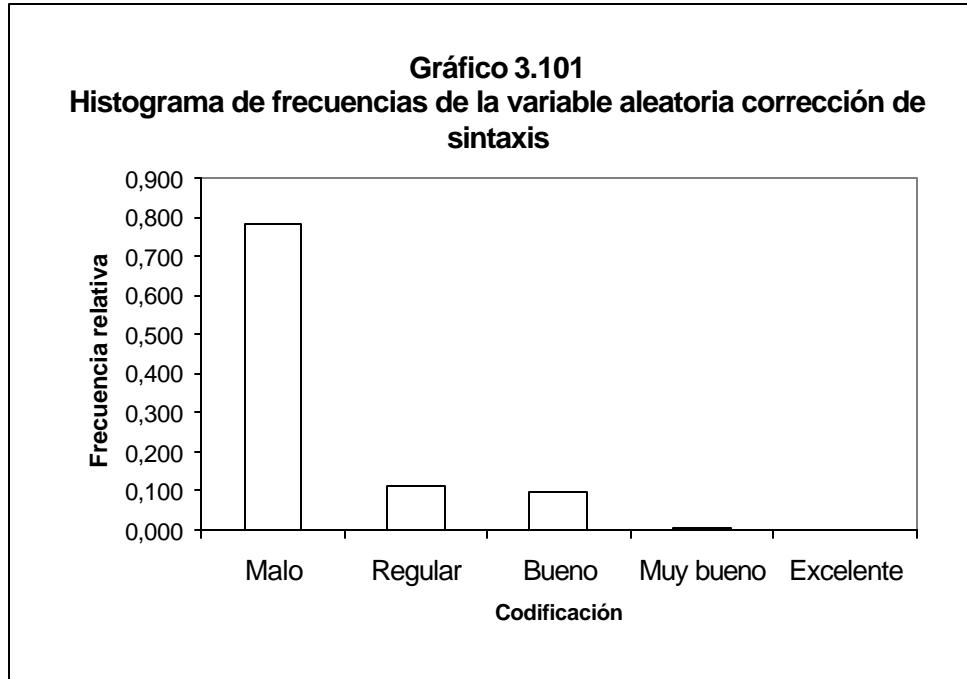
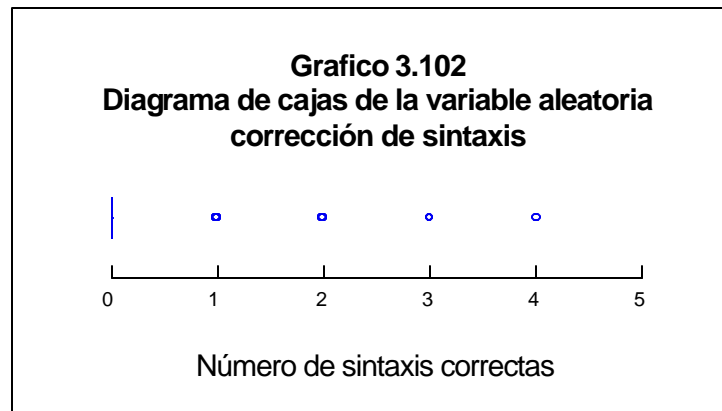


Tabla CXXV
Frecuencias de la variable aleatoria corrección de sintaxis

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	131	0,78	131	0,78
1	19	0,11	150	0,90
2	14	0,08	164	0,98
3	2	0,01	166	0,99
4	1	0,01	167	1,00

La función generadora de momentos de esta variable de estudio es :

$$M(t) = 0.78 + 0.11e^t + 0.08e^{2t} + 0.01e^{3t} + 0.01e^{4t}$$



3.4.25 Variable aleatoria palabras tildadas correctamente

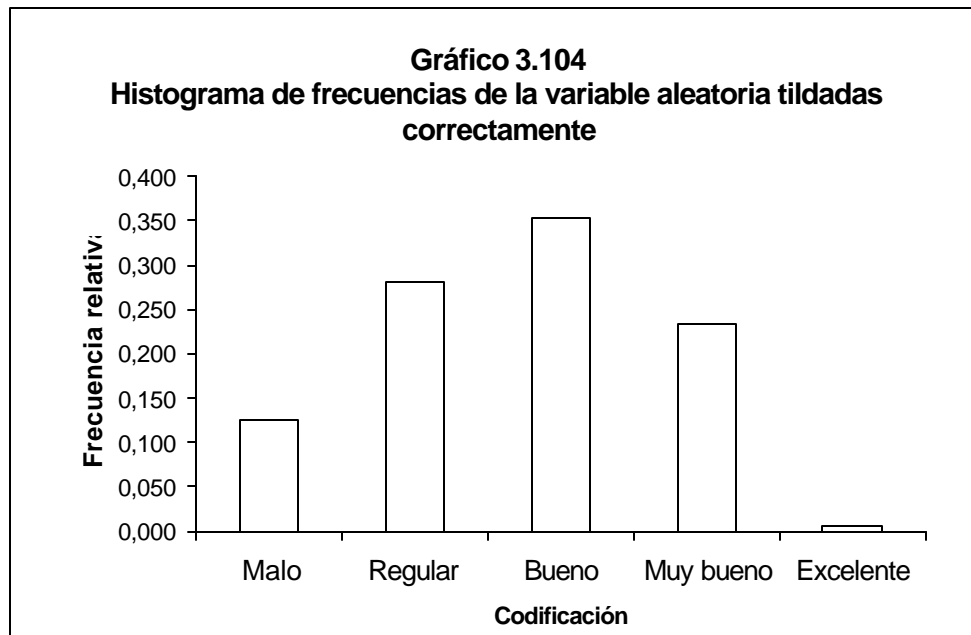
En la tabla CXXVI los valores de la media (7.82) y de la mediana (8) están muy cerca y corresponden al código bueno, para el cual 28 de cada 100 estudiantes entre 7 y 11 palabras. El máximo de respuestas posibles es 21, sin embargo el máximo obtenido en esta variables es 18, donde uno de cada 100 estudiantes tienen este número de respuestas. Los valores del primer cuartil (4) y tercer cuartil (11) indican que el 25% de la observaciones están por debajo 4 palabras correctamente tildadas y que el 75% están por debajo de 11 palabras correctamente tildadas. Debido a que el coeficiente de asimetría es negativo(-0.059) la

distribución está sesgada hacia la derecha, por lo tanto la pregunta es difícil. En lo que respecta al coeficiente de kurtosis (-0.972) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.

Tabla CXXVI
Parámetros poblacionales de la variable aleatoria
palabras tildadas correctamente

Media	7,82	Mínimo	0
Mediana	8	Máximo	18
Desviación estándar	4,932	Sesgo	-0,059
Varianza	24,324	Kurtosis	-0,972
Error estándar	0,392	Primer cuartil	4
Rango	18	Tercer cuartil	11
Moda	0	Rango intercuartil	7,5
Suma	1266	Coefficiente de variación	0.63

En la tabla CXXVII se muestran las frecuencias de esta variable aleatoria, en base a la frecuencia absoluta relativa se puede decir que de cada 100 estudiantes 13 no saben tildar los acentos en las palabras, 28 tienen conocimientos regulares y 22 tienen un conocimiento muy bueno.



La función generadora de momentos de esta variable de estudio es:

$$\begin{aligned}
 M(t) = & 0.13 + 0.03e^t + 0.04e^{2t} + 0.04e^{3t} + 0.05e^{4t} + 0.05e^{5t} + 0.07e^{6t} + 0.08e^{7t} \\
 & + 0.08e^{8t} + 0.06e^{9t} + 0.07e^{10t} + 0.06e^{11t} + 0.04e^{12t} + 0.04e^{13t} + 0.05e^{14t} + 0.04e^{15t} \\
 & + 0.04e^{16t} + 0.01e^{17t} + 0.01e^{18t}
 \end{aligned}$$

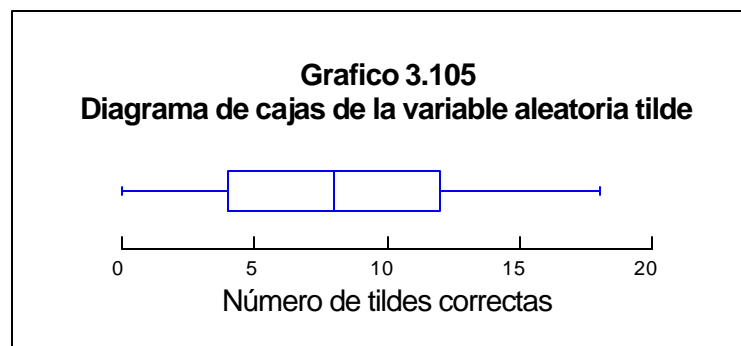


Tabla CXXVII
Frecuencias de la variable aleatoria tildadas
correctamente

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	21	0,13	21	0,13
1	5	0,03	26	0,16
2	7	0,04	33	0,20
3	6	0,04	39	0,23
4	8	0,05	47	0,28
5	9	0,05	56	0,34
6	12	0,07	68	0,41
7	13	0,08	81	0,49
8	14	0,08	95	0,57
9	10	0,06	105	0,63
10	12	0,07	117	0,70
11	10	0,06	127	0,76
12	7	0,04	134	0,80
13	7	0,04	141	0,84
14	9	0,05	150	0,90
15	7	0,04	157	0,94
16	7	0,04	164	0,98
17	2	0,01	166	0,99
18	1	0,01	167	1,00

3.4.26 Variable aleatoria lectura comprensiva

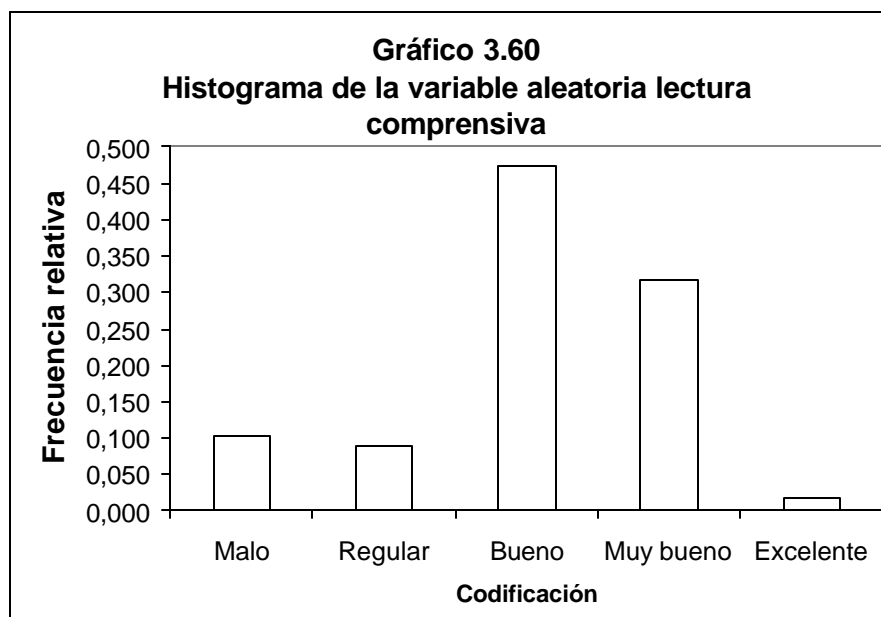
Como se puede observar en la tabla CXXVIII se muestran los resultados de los parámetros de esta población, entre estos resultados están las medidas de tendencia central que indican alrededor de que valor se agrupan las observaciones.

La media (2.868), la mediana (3) y la moda (3) indican que los estudiantes en gran proporción respondieron correctamente 3 de las 6 preguntas, el valor 3 está incluido en la clase bueno. De los estudiantes entrevistados 47 de cada 100 tienen buenos conocimientos para leer comprensivamente.

Tabla CXXVIII
Parámetros poblacionales de la variable aleatoria
lectura comprensiva

Media	2,868	Mínimo	0
Mediana	3	Máximo	6
Desviación estándar	1,543	Sesgo	-0,245
Varianza	2,380	Kurtosis	-0,644
Error estándar	0,119	Primer cuartil	2
Rango	6	Tercer cuartil	4
Moda	3	Rango intercuartil	2
Suma	475	Coefficiente de variación	0.538

El coeficiente de asimetría es negativo(-0.245) lo que significa que la distribución está sesgada hacia la izquierda, por lo tanto la pregunta es medianamente difícil. En lo que respecta al coeficiente de kurtosis (-0.644) este indica que la distribución es platicúrtica, es decir más achatada que la distribución normal.



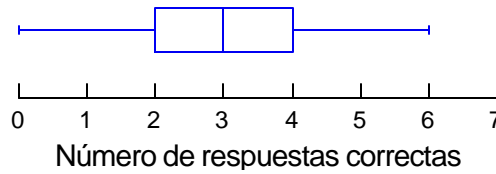
Los estudiantes cuyas respuestas caen en la clasificación mala, es decir 0 respuestas correctas son 10 de cada 100, los que tienen un conocimiento regular 9 de cada 100, los que responden correctamente 4 o 5 preguntas son 32 de cada 100 y los que tienen un conocimiento excelente para leer comprensivamente son 2 de cada 100. Estos resultados se obtienen a partir de la frecuencia absoluta relativa de la variable aleatoria mostrada en la tabla CXXIX. La función generadora de momentos de esta variable de estudio es:

$$M(t) = 0.10 + 0.09e^t + 0.18e^{2t} + 0.29e^{3t} + 0.18e^{4t} + 0.14e^{5t} + 0.02e^{6t}$$

Tabla CXXIX
Frecuencias de la variable aleatoria lectura comprensiva

Valor	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia acumulada relativa
0	17	0,10	17	0,10
1	15	0,09	32	0,19
2	30	0,18	62	0,37
3	49	0,29	111	0,66
4	30	0,18	141	0,84
5	23	0,14	164	0,98
6	3	0,02	167	1,00

Gráfico 3.107
Diagrama de cajas de la variable aleatoria lectura comprensiva



3.4.27 Variable aleatoria calificación de lenguaje

Al realizar el análisis univariado de esta variable se obtuvo los resultados que se muestran en la tabla CXXX. La media de esta variable es 43.144, la mediana 44 y la moda 40, esto indica que las observaciones están localizadas alrededor de este valor.

Tabla CXXX
Parámetros poblacionales de la variable aleatoria
calificación de lenguaje

Media	43,144	Mínimo	0
Mediana	44	Máximo	83
Desviación estándar	14,664	Sesgo	-0,361
Varianza	215,032	Kurtosis	-1,073
Error estándar	1,135	Primer cuartil	36
Rango	83	Tercer cuartil	52
Moda	40	Rango intercuartil	16
Suma	7193	Coefficiente de variación	0.339

Para analizar la dispersión de los datos se utiliza el valor de la desviación estándar de esta variable que es 14.664, este valor es relativamente bajo dado que la variación relativa de las observaciones con respecto a la media es del 33.9%. Como se puede apreciar en el gráfico 3.108 la distribución está ligeramente sesgada hacia la izquierda pues el coeficiente de asimetría es negativo(-0.361).

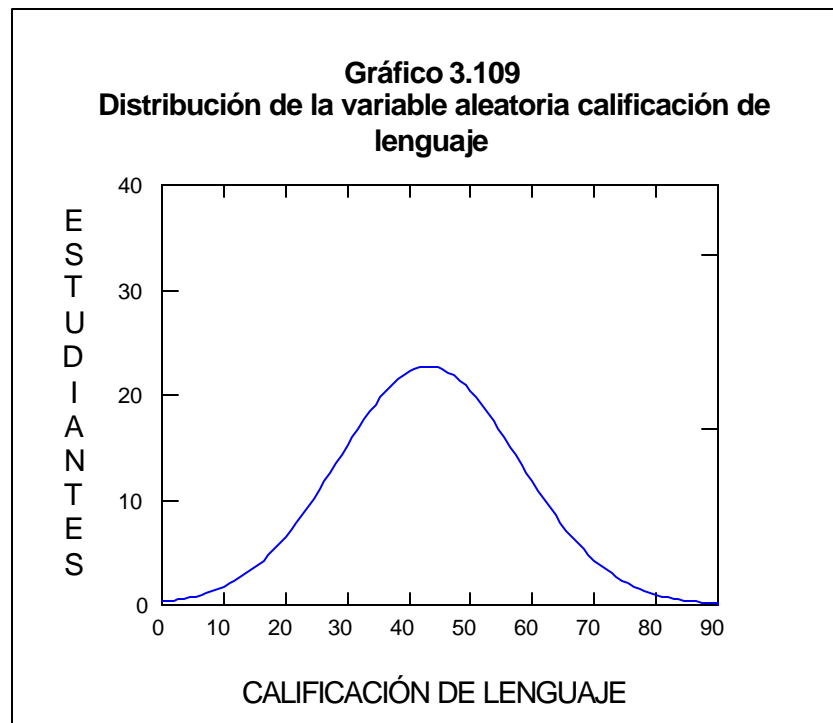
Para determinar si la variable calificación de matemáticas está distribuida normalmente con una media 43.14 y varianza 215.032, se realizó el siguiente contraste de hipótesis:

$$H_0: X \sim N(43.14, 215.032)$$

vs.

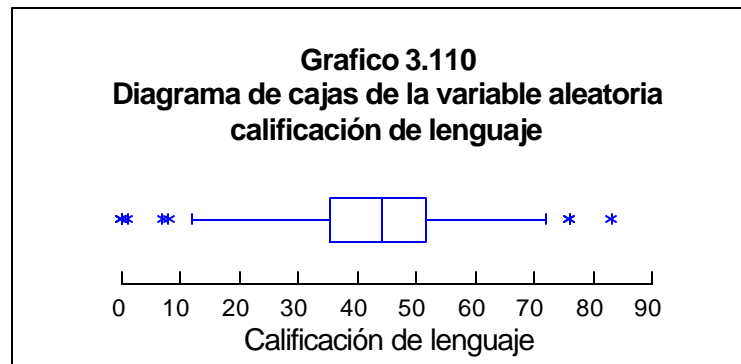
$$H_1: \neg H_0$$

El valor p obtenido al realizar la prueba Kolmogorov-Smirnov fue 0.361 y la máxima diferencia 0.071, por lo tanto existe evidencia estadística para no rechazar la hipótesis nula H_0 .



La función generadora de momentos de la variable aleatoria calificación de lenguaje es:

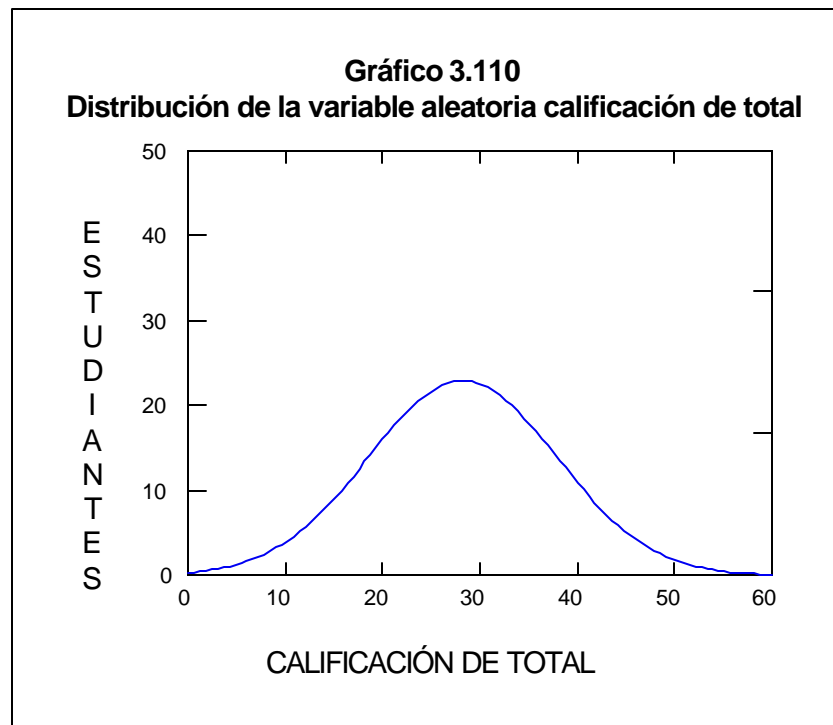
$$M_X(t) = e^{43.144 t + \frac{1}{2}(215.032 t^2)}$$



3.4.28 Variable aleatoria calificación total

Los parámetros de esta variable aleatoria se muestran en la tabla CXXXI. La media, la mediana y la moda indican alrededor de que valor se agrupan las observaciones como se puede observar estas calificaciones son muy bajas, pues están entre 26 y 28 puntos sobre cien.

El coeficiente de variación de esta variable indica la variación relativa de las observaciones con respecto a la media que es del 33.9%. Como se puede apreciar en el gráfico 3.110 la distribución está ligeramente sesgada hacia la izquierda.



Para determinar si la variable calificación de matemáticas está distribuida normalmente con una media 28.22 y varianza 93.83, se realizó el siguiente contraste de hipótesis:

$$H_0: X \sim N(28.22, 93.83)$$

Vs.

$$H_1: \neg H_0$$

El valor p obtenido al realizar la prueba Kolmogorov-Smirnov fue 0.432 y la máxima diferencia 0.068, por lo tanto existe evidencia estadística para no rechazar la hipótesis nula H_0 .

Tabla CXXVIII
Parámetros poblacionales de la variable aleatoria
calificación total

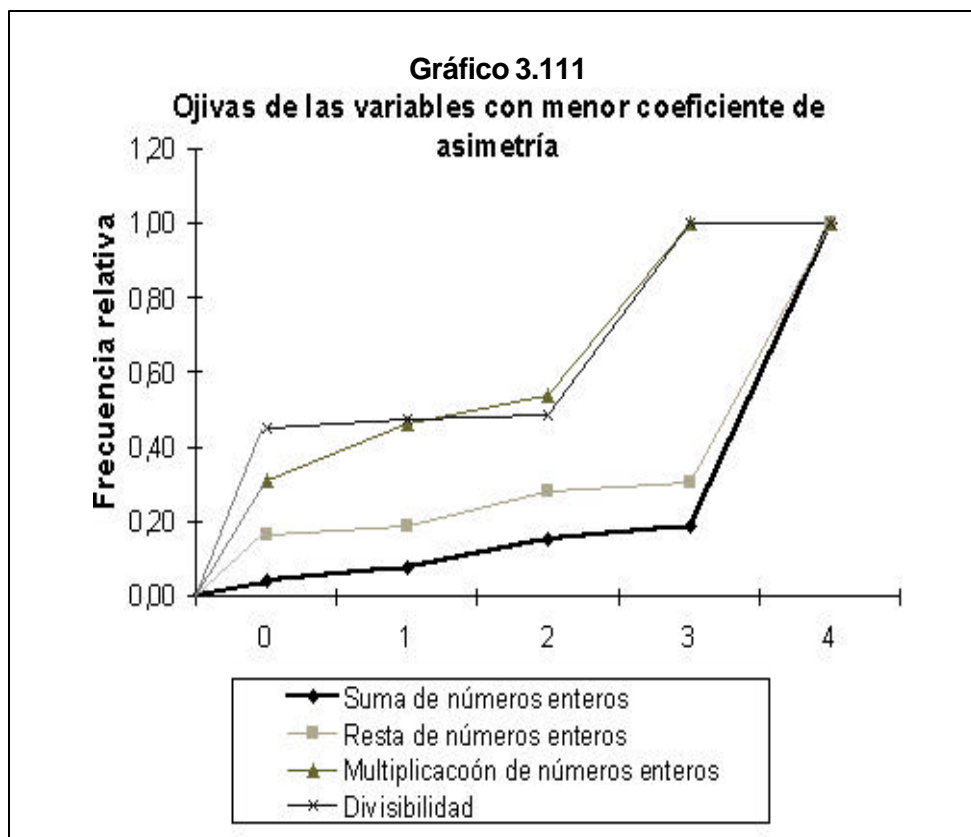
Media	28.22	Mínimo	0.5
Mediana	28.5	Máximo	54
Desviación estándar	9.687	Sesgo	-0,244
Varianza	93.83	Kurtosis	0.966
Error estándar	0.75	Primer cuartil	21.5
Rango	53.5	Tercer cuartil	34.75
Moda	26	Rango intercuartil	13.25
Suma	4714	Coefficiente de variación	0.374

La función generadora de momentos de la variable aleatoria calificación de total es:

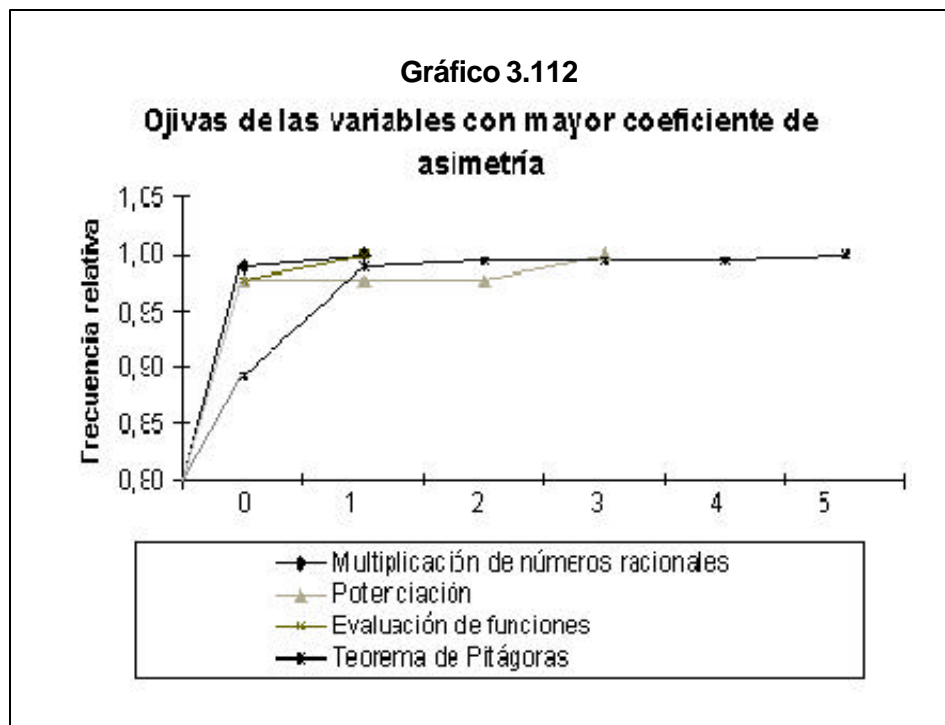
$$M_x(t) = e^{28.22 t + \frac{1}{2}(93.83 t^2)}$$

3.5 Determinación del grado de dificultad de las preguntas de las pruebas de matemáticas y lenguaje.

La dificultad de las preguntas de las pruebas de matemáticas y lenguaje, se la determinó utilizando el coeficiente de asimetría de cada variable aleatoria y observando la rapidez con que crecía la función. En el gráfico 3.111 se puede observar las ojivas de las cuatro variables de la prueba de matemáticas con menor coeficiente de asimetría.



El menor valor que puede tomar una variable aleatoria es cero el cual indica que el estudiante no respondió la pregunta, las escalas en que se codificó las variables aleatorias van de un menor a un mayor puntaje, por lo tanto la ojiva de la variable que crece más lentamente corresponde a la pregunta con mayor número de respuestas correctas.



En el gráfico 3.112 se puede observar el rápido crecimiento de las variables aleatorias que corresponden a las preguntas más difíciles de la prueba de matemáticas.

En la prueba de matemáticas las preguntas con menor grado de dificultad determinadas, son suma de números enteros, resta de números enteros, multiplicación de números enteros y divisibilidad, los porcentajes de estudiantes que respondieron correctamente estas preguntas son 82.03%, 69.46%, 51.49% y 46.10% respectivamente.

Otros resultados obtenidos en estas preguntas son: en la suma de números enteros el 13.78% de los estudiantes sumó correctamente de 1 a 3 cifras y el 4.19% no realizó la suma; en la resta de números enteros el 14.38% de los estudiantes restó correctamente de 1 a 3 cifras y el 16.165 no realizó la resta; en la multiplicación de números enteros el 22.77% de los estudiantes multiplicó correctamente 1 ó 2 cifras y el 31.13% no respondió y en el ejercicio de divisibilidad el 44.91% de los estudiantes no respondió y el 3.6% realizó correctamente 1 ó 2 ejercicios de divisibilidad.

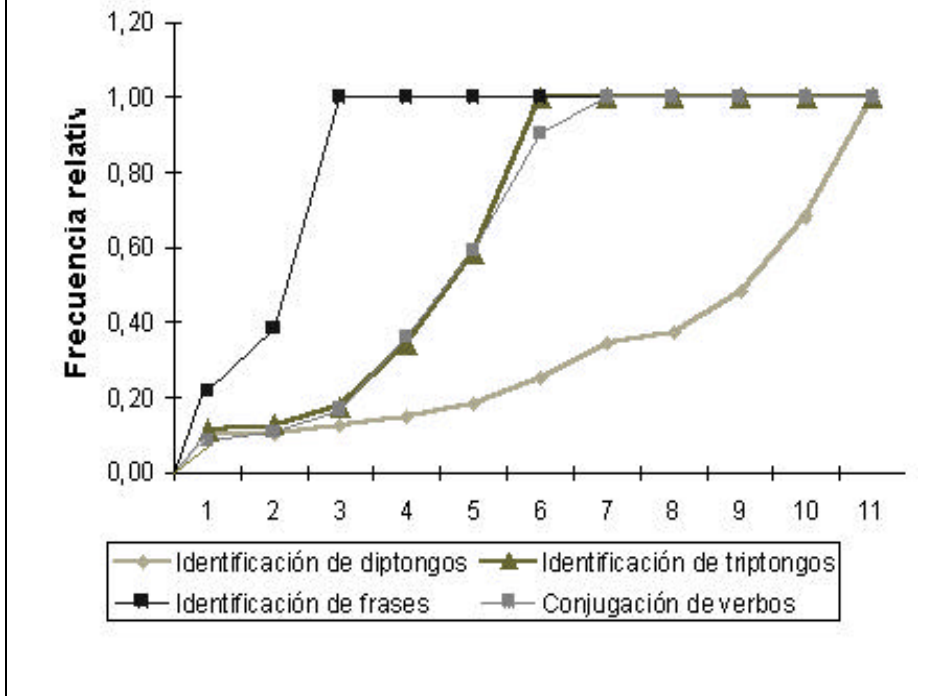
Las preguntas de las variables de la prueba de matemáticas que obtuvieron los mayores coeficientes de asimetría, es decir las de mayor grado de dificultad en la prueba de matemáticas son: multiplicación de números racionales, potenciación y radicación, evaluación de funciones y teorema de Pitágoras, los porcentajes de estudiantes que no

respondieron estas preguntas son respectivamente 98.79%, 98,19%, 98,79% y 89,21%.

Otros resultados obtenidos en preguntas con mayor grado de dificultad en el cuestionario de matemáticas son: en la multiplicación de números racionales el 1.21% de los estudiantes respondió correctamente; en la pregunta potenciación y radicación el 1.21% de los estudiantes respondió correctamente; en la evaluación de funciones el 2.405 de los estudiantes contestó correctamente y en la pregunta sobre el teorema de Pitágoras el 10.78% de los estudiantes contestó correctamente.

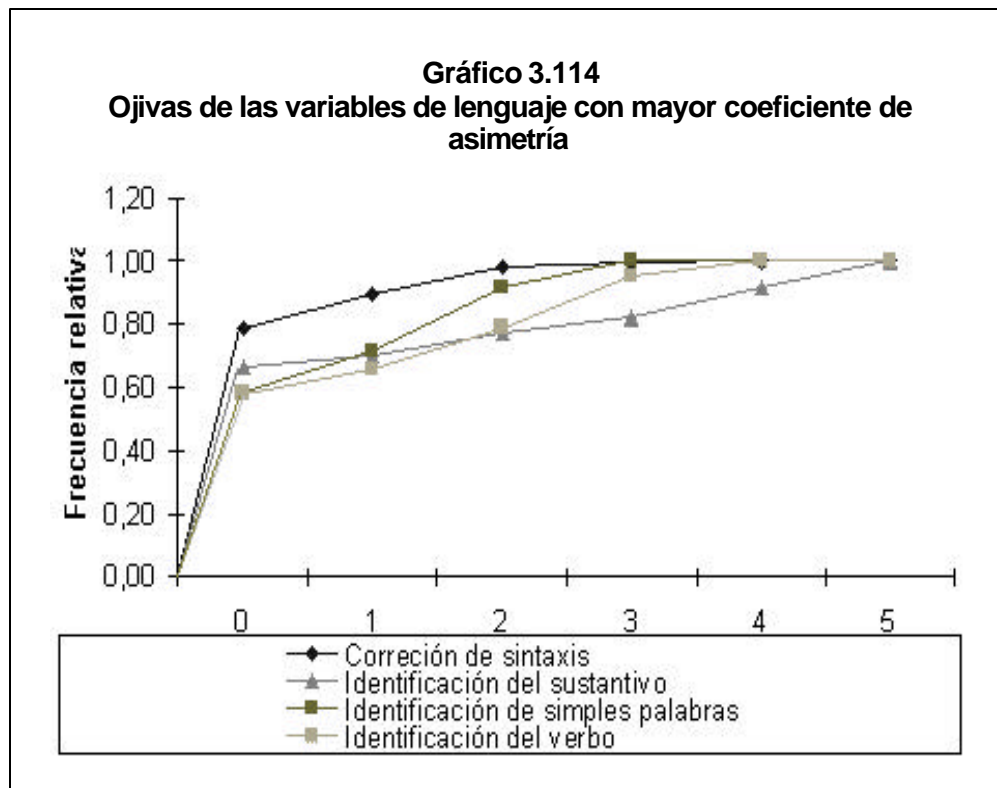
Las preguntas más fáciles correspondientes al cuestionario de lenguaje son conjugación de verbos, identificación de frases, identificación de diptongos e identificación de triptongos. En el gráfico 3.113 se puede observar las ojivas de las variables correspondientes a las preguntas más fáciles de esta prueba. Para cada una de estas preguntas se calculó el porcentaje de respuestas correctas de los estudiantes entrevistados, los cuales son 9.57%, 61.67%, 31.73% y 41.31%.

Gráfico 3.113
Ojivas de las variables de lenguaje con menor coeficiente de
asimetría



Otros resultados obtenidos en las preguntas con menor grado de dificultad del cuestionario de lenguaje son: en la identificación de triptongos el 11.37% de los estudiantes no respondió la pregunta y el 47.32% identificó de 1 a 4 triptongos correctamente; en la pregunta identificación de diptongos el 10.17% de los estudiantes no respondió y el 58.10% identificó de 1 a 9 diptongos correctamente; en la identificación de frases el 21.55% de los estudiantes no respondió y el 16.78% identificó 1

frase correctamente y en la pregunta sobre la conjugación de verbos el 8.38% de los estudiantes no contestó y el 82.05% conjugó de 1 a 5 verbos correctamente.



En el gráfico 3.114 se puede observar el rápido crecimiento de las variables aleatorias que corresponden a las preguntas con mayor grado de dificultad de la prueba de lenguaje.

Las preguntas más difíciles de la prueba de lenguaje son corrección de la sintaxis de la oración, identificación del verbo, identificación del sustantivo de la oración e identificación de palabras simples. Los porcentajes de estudiantes que no respondieron estas preguntas son 78.44%, 58.08%, 66.46%, y 58.68%, respectivamente.

Capítulo IV

4.- Análisis Multivariado

4.1 Introducción

El análisis multivariado está constituido por un conjunto de técnicas estadísticas diseñadas para extraer simultáneamente información de un grupo de variables aleatorias. En particular las técnicas multivariadas que se aplicarán en este capítulo son cinco: correlación lineal, análisis de varianza, análisis de componentes principales, correlación canónica y análisis de contingencia.

4.2 Definiciones

Para definir matemáticamente los métodos multivariados que se utilizarán en este capítulo, es necesario partir de definición de la matriz de datos. La información de las variables debe ser representada en forma de un arreglo rectangular de p columnas por n filas denotado por \mathbf{X} , donde cada columna es un vector en \mathbb{R}^n y cada fila es un vector en \mathbb{R}^p ; el valor p corresponde al número de variables aleatorias (X_1, X_2, \dots, X_p) y el valor n al número de unidades de investigación, entonces a la matriz \mathbf{X} se la denomina matriz de datos.

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1k} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2k} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{j1} & X_{j2} & \cdots & X_{jk} & \cdots & X_{jp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nk} & \cdots & X_{np} \end{bmatrix}$$

La matriz \mathbf{X} contiene todas las observaciones de todas las variables, el valor de X_{jk} representa la j -ésima observación de la k -ésima variable aleatoria, para $j=1,2,\dots,n$ y $k=1,2,\dots,p$.

Sea $\mathbf{X}^T = (X_1, X_2, \dots, X_p)$, un vector aleatorio, se define el vector de medias $\hat{\mathbf{i}}$, correspondientes al vector \mathbf{X} como sigue,

$$\hat{\mathbf{i}} = E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_k) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \hat{i}_1 \\ \hat{i}_2 \\ \vdots \\ \hat{i}_k \\ \vdots \\ \hat{i}_p \end{bmatrix} \in M_{p \times 1}$$

Donde el valor \hat{i}_k representa el valor esperado de la variable aleatoria X_k , es decir $\hat{i}_k = E[X_k]$, para $k = 1, 2, \dots, p$.

Luego de definir el vector de medias $\hat{\mathbf{i}}$, se define la matriz de varianzas y covarianzas $\hat{\mathbf{O}}$, como se muestra a continuación,

Sea $\sigma_{ij} = E(X_i - \hat{i}_i)(X_j - \hat{i}_j)$, para $i, j = 1, 2, \dots, p$, σ_{ij} representa la covarianza entre X_i y X_j , entonces $\hat{\mathbf{O}} = E(\mathbf{X} - \hat{\mathbf{i}})(\mathbf{X} - \hat{\mathbf{i}})^T$, lo cual se puede expresar como

$$\hat{\mathbf{O}} = E \left\{ \begin{bmatrix} X_1 - \hat{i}_1 \\ X_2 - \hat{i}_2 \\ \vdots \\ X_p - \hat{i}_p \end{bmatrix} \begin{bmatrix} X_1 - \hat{i}_1 & X_2 - \hat{i}_2 & \dots & X_p - \hat{i}_p \end{bmatrix} \right\} \in M_{p \times p}$$

$$\mathbf{O} = \begin{bmatrix} E[(X_{1-1})^2] & E[(X_{1-1})(X_{2-2})] & \cdots & E[(X_{1-1})(X_{p-p})] \\ E[(X_{1-1})(X_{2-2})] & E[(X_{2-2})^2] & \cdots & E[(X_{2-2})(X_{p-p})] \\ \vdots & \vdots & \ddots & \vdots \\ E[(X_{1-1})(X_{p-p})] & E[(X_{2-2})(X_{p-p})] & \cdots & E[(X_{p-p})^2] \end{bmatrix} \in M_{p \times p}$$

$$\mathbf{O} = \begin{bmatrix} \acute{o}_{11} & \acute{o}_{12} & \cdots & \acute{o}_{1j} & \cdots & \acute{o}_{1p} \\ \acute{o}_{21} & \acute{o}_{22} & \cdots & \acute{o}_{2j} & \cdots & \acute{o}_{2p} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \acute{o}_{i1} & \acute{o}_{i2} & \cdots & \acute{o}_{ij} & \cdots & \acute{o}_{ip} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \acute{o}_{p1} & \acute{o}_{p2} & \cdots & \acute{o}_{pj} & \cdots & \acute{o}_{pp} \end{bmatrix} \in M_{p \times p}$$

Cuando $i=j$ se tiene que $\sigma_{ij} = \acute{o}_i^2$, que es la varianza de la variable aleatoria X_i . La matriz $\tilde{\mathbf{n}}$ muestra los valores de \tilde{n}_{ij} que representan los coeficientes de correlaci3n entre las variables X_i y X_j ,

$$\tilde{\mathbf{n}} = \begin{bmatrix} 1 & \tilde{n}_{12} & \cdots & \tilde{n}_{1j} & \cdots & \tilde{n}_{1p} \\ \tilde{n}_{21} & 1 & \cdots & \tilde{n}_{2j} & \cdots & \tilde{n}_{2p} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \tilde{n}_{i1} & \tilde{n}_{i2} & \cdots & 1 & \cdots & \tilde{n}_{ip} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \tilde{n}_{p1} & \tilde{n}_{p2} & \cdots & \tilde{n}_{pj} & \cdots & 1 \end{bmatrix} \in M_{p \times p}$$

El coeficiente indica si existe o no, relaci3n lineal entre estas dos variables, y se obtiene con

$$\tilde{\rho}_{ij} = \frac{\text{cov}(X_i, X_j)}{\sigma_i \sigma_j}$$

donde σ_i y σ_j son las desviaciones estándar de X_i y X_j respectivamente, cuando $i=j$ el coeficiente de correlación es igual a 1, además este coeficiente se encuentra en el intervalo $(-1,1)$.

Prueba de que $-1 \leq \tilde{\rho}_{ij}$

Sabemos que la varianza de cualquier variable aleatoria es mayor a cero, $\text{Var}(X) \geq 0$, entonces

$$0 \leq \text{Var}\left(\frac{X_i}{s_i} + \frac{X_j}{s_j}\right) = \text{Var}\left(\frac{X_i}{s_i}\right) + \text{Var}\left(\frac{X_j}{s_j}\right) + 2 \frac{\text{Cov}(X_i, X_j)}{s_i s_j}$$

$$0 \leq 2 + 2 r_{ij} \Rightarrow -1 \leq r_{ij}$$

Prueba de que $\tilde{\rho}_{ij} \leq 1$

Sabemos que la varianza de cualquier variable aleatoria es mayor a cero, $\text{Var}(X) \geq 0$, entonces

$$0 \leq \text{Var}\left(\frac{X_i}{s_i} - \frac{X_j}{s_j}\right) = \text{Var}\left(\frac{X_i}{s_i}\right) + \text{Var}\left(\frac{X_j}{s_j}\right) - 2 \frac{\text{Cov}(X_i, X_j)}{s_i s_j}$$

$$0 \leq 2 - 2 r_{ij} \Rightarrow r_{ij} \leq 1$$

Por lo tanto se cumple que $-1 \leq \tilde{\rho}_{ij} \leq 1$.

4.3 Análisis de la matriz de correlación

De los resultados que se obtuvieron de la matriz de correlación, de las sesenta variables de estudio utilizadas para medir el conocimiento en matemáticas y lenguaje, de los estudiantes de décimo año de educación básica, de los colegios fiscales rurales del cantón Guayaquil, en esta sección se analizarán los mayores y menores coeficientes de correlación entre las variables aleatorias estudiadas.

4.3.1 Análisis de los mayores coeficientes de correlación

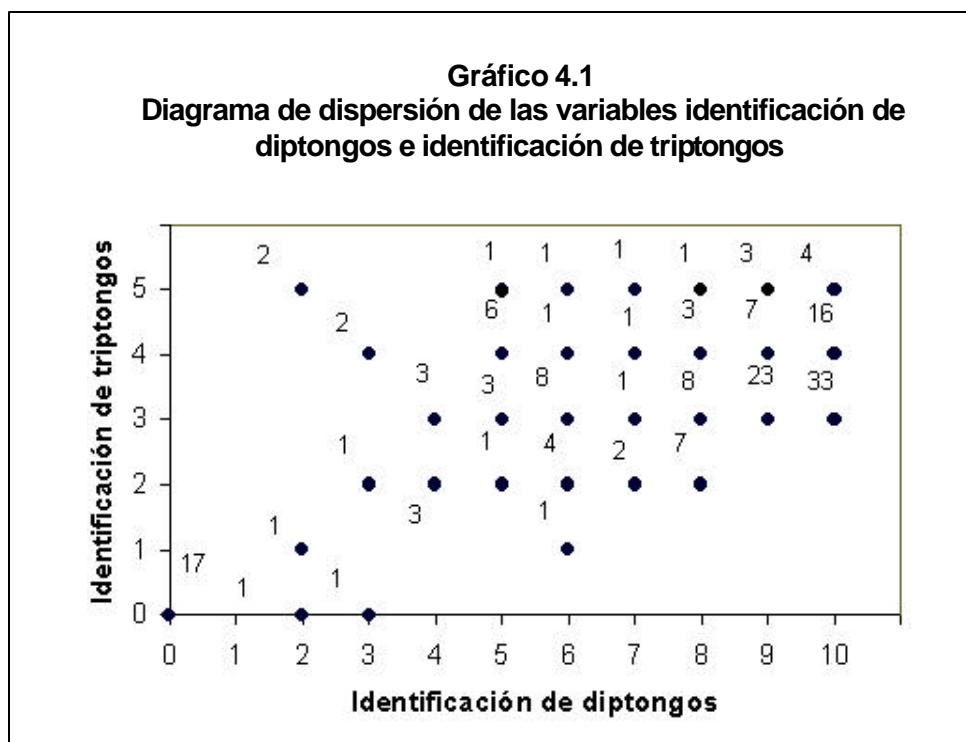
Para realizar este análisis se consideró como significativos los coeficientes de correlación que en valor absoluto fuesen mayores a 0.6. La matriz de correlación obtenida, de todas las variables de estudio se muestra en el anexo número dos. En éste análisis se determinó que existen tres pares de variables aleatorias correlacionadas entres sí.

- **Variable identificación de diptongos con la variable identificación de triptongos**

Se determinó que las variables observables X_{34} identificación de diptongos y X_{35} identificación de triptongos, son linealmente dependientes, pues el coeficiente de correlación lineal entre estas variables es 0.831.

La matriz de correlación de las variables X_{34} y X_{35} es:

	X_{34}	X_{35}
X_{34}	1	0.833
X_{35}	0.833	1



La dispersión entre las variables X_{34} y X_{35} se puede observar en el gráfico 4.1, en el cual se ilustra el número de observaciones en cada uno de los puntos graficados, la relación lineal entre estas dos variables es directa, es decir que a medida que los valores de la variable X_{34} aumentan, los valores de X_{35} también aumentan. La variable aleatoria identificación de diptongos puede tomar valores de 0 a 10 los cuales

representan el número de diptongos que el estudiante identificó correctamente, mientras que la variable aleatoria identificación de triptongos puede tomar valores de 0 a 5. Los puntos graficados indican que existen estudiantes que identificaron correctamente el número de diptongos representados por el valor de la abscisa y el número de triptongos representados por el valor de la ordenada. En el caso contrario, por ejemplo la coordenada (4,0) que no está graficada indica que ningún estudiante identificó correctamente 4 diptongos y ningún triptongo.

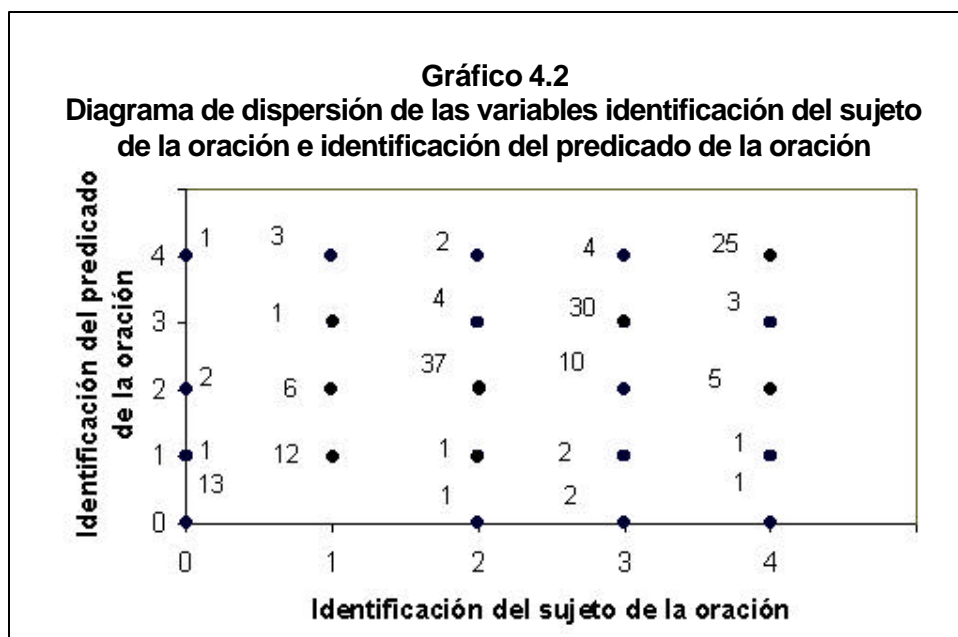
▪ **Variable identificación del sujeto de la oración con la variable identificación del predicado de la oración**

Las variables X_{50} identificación del sujeto de la oración y X_{61} identificación del predicado de la oración, son linealmente dependientes, el coeficiente de correlación lineal entre estas variables es de 0.665. A pesar que éste valor es mayor a 0.6, la dependencia lineal entre las dos variables, no es fuerte.

La matriz de correlación de las variables X_{50} y X_{61} es:

	X_{50}	X_{61}
X_{50}	1	0.665
X_{61}	0.665	1

La dependencia lineal existente entre las variables X_{60} y X_{61} determinadas por el coeficiente de correlación, el cual es mayor a 0.6, no se puede apreciar claramente en el gráfico 4.2, donde se muestra la dispersión entre estas dos variables, debido a que el valor de este coeficiente no es muy cercano a uno.



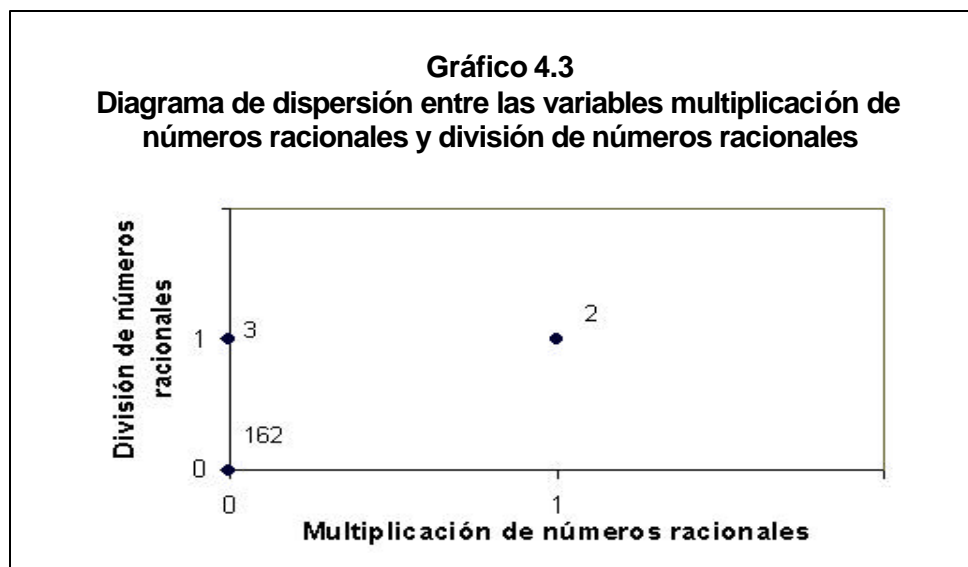
Tanto la variable aleatoria identificación del sujeto como la variable aleatoria identificación del predicado pueden tomar valores entre 0 y 4, los cuales representan el número de sujetos o predicados identificados correctamente. Como se puede observar en el gráfico 4.2 a excepción de los puntos con coordenadas (0,3) y (1,0), todos los demás resultados posibles de respuestas se presentan entre esta dos variables aleatorias, con el respectivo número de observaciones.

- **Variable multiplicación de números racionales con la variable división de números racionales**

Existe dependencia lineal entre las variables X_8 multiplicación de fracciones y X_{10} división de números racionales, el coeficiente de correlación entre ambas variables es 0.627.

La matriz de correlación de las variables X_8 y X_{10} es:

	X_8	X_{10}
X_8	1	0.627
X_{10}	0.627	1



Los valores que pueden tomar las observaciones son 0 (respuesta incorrecta) o 1 (respuesta correcta) para ambas variables aleatorias, en el gráfico 4.3 se muestra el diagrama de dispersión entre las variables X_8 y X_{10} . En este gráfico se puede observar que no existe el punto con coordenadas (1,0), lo que significa que ningún estudiante entrevistado realizó correctamente la multiplicación de números racionales e incorrectamente la división de números racionales.

4.3.2 Análisis de los menores coeficientes de correlación

Cuando dos variables aleatorias X y Y son independientes, entonces la covarianza y el coeficiente de correlación entre estas variables, es igual a cero; sin embargo, si el valor de la covarianza y el coeficiente de correlación entre estas dos variables es igual a cero, no implica, que las variables aleatorias son independientes, sino que, no existe dependencia lineal entre ambas variables.

Otros resultados obtenidos a partir del análisis de la matriz de correlación, fueron los de los pares de variables aleatorias menos correlacionadas, para lo cual se consideraron los coeficientes de correlación lineal menores a 0.000, a continuación se detallan los tres pares de variables aleatorias que cumplen con esta condición.

- **Variable identificación de palabras agudas con la variable identificación de palabras diminutivas**

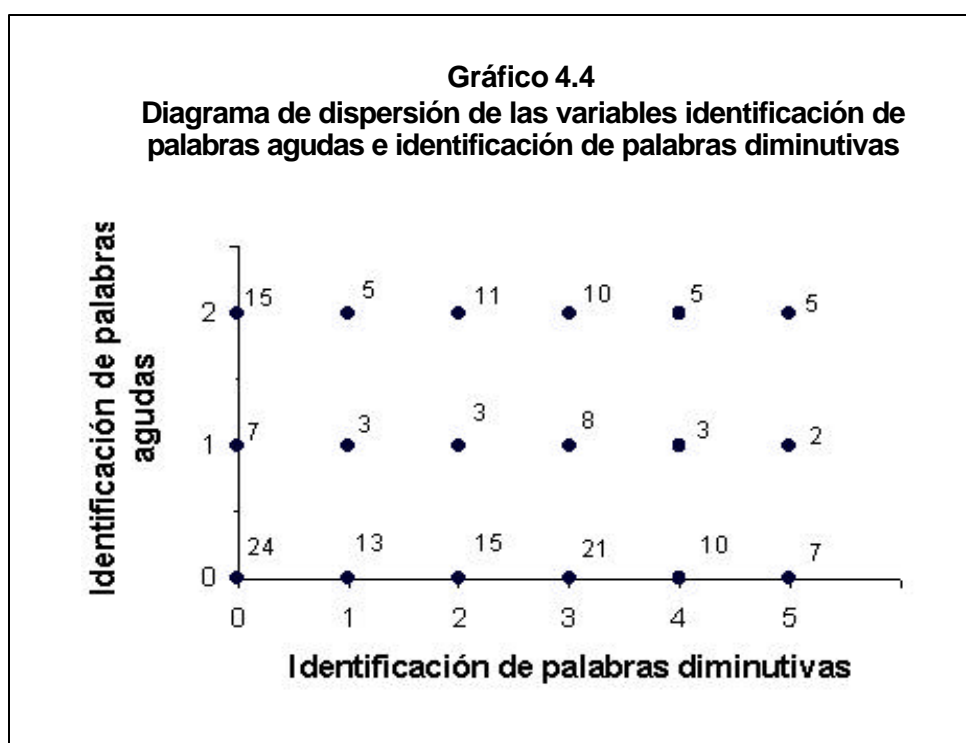
El coeficiente de correlación entre las variables observables X_{37} identificación de palabras agudas y X_{45} identificación de palabras diminutivas es 0.000, el cual indica que estas variables aleatorias no son linealmente dependientes. La matriz de correlación de las variables X_{37} y X_{45} es:

La matriz de correlación de las variables X_{37} y X_{45} es:

	X_{37}	X_{45}
X_{37}	1	0
X_{45}	0	1

La dispersión entre las variables X_{37} y X_{45} se puede observar en el gráfico 4.4. La variable aleatoria identificación de palabras agudas puede tomar valores de 0 a 5 los cuales representan el número de palabras agudas que el estudiante identificó correctamente, mientras que la variable aleatoria identificación de palabras diminutivas puede tomar valores de 0 a 2. Los puntos graficados indican que existen estudiantes que identificaron correctamente el número de palabras agudas representados por el valor de la accisa y el número de palabras

diminutivas representados por el valor de la ordenada. Se puede observar en el gráfico que el mayor número de observaciones corresponden a las coordenadas del punto (0,0), que corresponde a los estudiantes que no identificaron correctamente las palabras agudas y las palabras diminutivas.



- **Variable número de relaciones de orden correctas con la variable número de palabras definidas correctamente.**

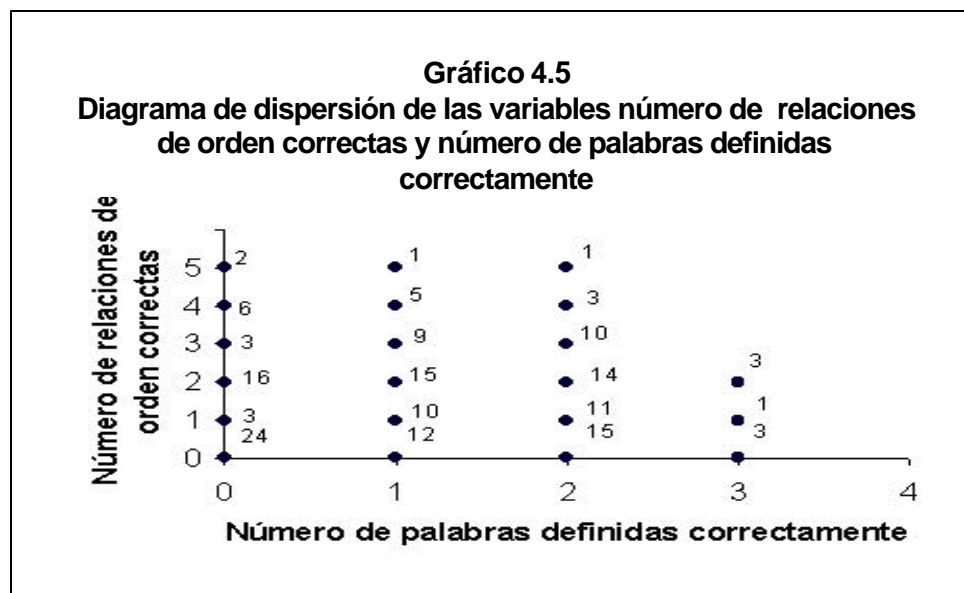
Entre las variables X_{13} número de relaciones de orden correctas y X_{48} número de palabras definidas correctamente, el coeficiente de correlación

lineal es 0, éste valor indica que las variables aleatorias X_{13} y X_{48} no son linealmente dependientes.

La matriz de correlación de las variables X_{13} y X_{48} es:

	X_{13}	X_{48}
X_{13}	1	0
X_{48}	0	1

La dispersión existente entre las variables X_3 y X_{48} , se ilustra claramente en el gráfico 4.5.



Las variables aleatorias número de relaciones de orden correctas y número de palabras definidas correctamente pueden tomar valores entre

0 y 5, y entre 0 y 4, respectivamente. Como se puede observar en el gráfico 4.5 entre los puntos que no están graficados, se encuentra (4,5) que corresponden al máximo número de definiciones de palabras y relaciones de orden correctas, es decir que ningún estudiante entrevistado pudo contestar correctamente las preguntas correspondientes a las variables X_{13} y X_{48} .

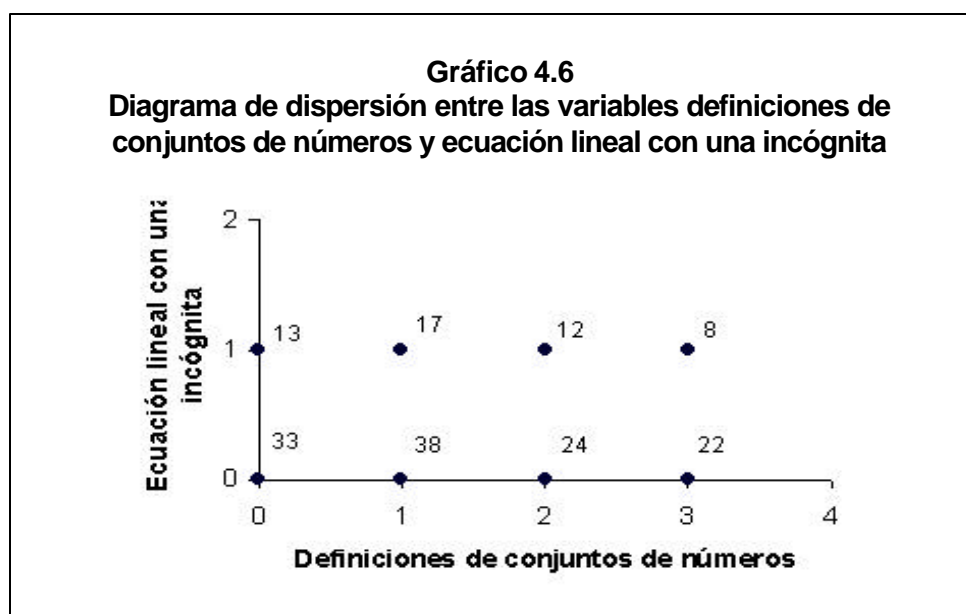
- **Variable definiciones de conjuntos de números con la variable ecuación lineal con una incógnita**

De acuerdo al valor del coeficiente de correlación obtenido entre las variables X_{11} definiciones de conjuntos de números y X_{29} ecuación lineal con una incógnita, se determinó que no existe dependencia lineal entre ambas variables aleatorias.

La matriz de correlación de las variables X_{11} y X_{29} es:

	X_{11}	X_{29}
X_{11}	1	0
X_{29}	0	1

Las variables aleatorias definiciones de conjuntos de números y ecuación lineal con una incógnita, pueden tomar los siguientes valores, para la primera variable de 0 a 4, dependiendo del número de respuestas correctas que obtuvo el estudiante y para la segunda variable 0 si la respuesta es incorrecta o 1 si es correcta, en el gráfico 4.6 se muestra el diagrama de dispersión entre las variables X_{11} y X_{29} . En este gráfico se puede observar claramente que ningún estudiante entrevistado respondió correctamente la pregunta correspondiente a la variable aleatoria X_{11} o a la pregunta correspondiente a la variable X_{29} .



4.4 Análisis de contingencia

En el análisis de la matriz de correlación de las variables aleatorias que se utilizaron para, medir el conocimiento en matemáticas y lenguaje de los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil, se detectaron en base a los coeficientes de correlación lineal, aquellos pares de variables aleatorias que son independientes ó linealmente dependientes; mientras que el análisis de contingencia se realiza con el objetivo de determinar la dependencia ó independencia lineal o no, de dos métodos o criterios de clasificación de las variables aleatorias observadas.

El análisis de contingencia se lo define de la siguiente forma, sean X y Y dos variables aleatorias observables y sea n el número de observaciones de cada variable, para cada una de las variables se define una clasificación con las cuales se construye un arreglo matricial de r filas y c columnas, donde c es el número de criterios de clasificación de la variable X y r el número de criterios de clasificación de la variable Y , este arreglo matricial se denomina tabla de contingencia.

En las tablas de contingencia se contabilizan las observaciones que cumplen simultáneamente con cada criterio de clasificación de cada

variable. Deben de haber al menos dos criterios de clasificación por variable los cuales deben ser exhaustivos y mutuamente excluyentes.

Para realizar el análisis de contingencia se construye un contraste de hipótesis para probar la hipótesis nula de que los criterios de clasificación de la variable aleatoria X son independientes de los criterios de clasificación de la variable aleatoria Y .

A continuación se muestra un modelo de tabla de contingencia en el que se puede observar como se contabilizan las observaciones de acuerdo a los criterios de clasificación de cada una de las variables aleatorias.

		Criterios de clasificación de la variable aleatoria Y				
Criterios de clasificación de la variable aleatoria X	1	2	...	c	Total	
1	n_{11} $E [n_{11}]$	n_{12} $E [n_{12}]$...	n_{1c} $E [n_{1c}]$	r_1	
2	n_{21} $E [n_{21}]$	n_{22} $E [n_{22}]$...	n_{2c} $E [n_{2c}]$	r_2	
\vdots	\vdots	\vdots		\vdots	\vdots	
r	n_{r1} $E [n_{r1}]$	n_{r2} $E [n_{r2}]$...	n_{rc} $E [n_{rc}]$	r_r	
Total	c_1	c_2	...	c_c	n	

Los valores de n_{ij} representan el número de observaciones que cumplen con el criterio de clasificación i , para $i=1,2,\dots,r$ de la variable aleatoria X y con el criterio de clasificación j , para $j=1,2,\dots,c$ de la variable aleatoria Y .

Los valores de r_i representan el total de observaciones que están sujetas al criterio de observación i de la variable aleatoria X , análogamente los valores de c_j representan el total de observaciones que caen en el criterio de observación j de la variable aleatoria Y . El valor de n representa el total de observaciones de la muestra o de la población.

El valor esperado de n_{ij} denotado por $E[n_{ij}]$ es igual al producto del total de observaciones del renglón i (r_i) por el total de observaciones de la columna j (c_j) dividido entre el total de observaciones (n).

$$E[n_{ij}] = (r_i * c_j) / n.$$

El contraste de hipótesis que se construye para probar la independencia de los criterios de clasificación de las variables aleatorias X y Y es el siguiente:

H_0 : Los criterios de clasificación son independientes

vs.

H_a : Los criterios de clasificación son dependientes

El estadístico de prueba que utiliza la información contenida en la tabla de contingencias, es una variable aleatoria ji-cuadrado denotada por χ^2 con $(r-1)(c-1)$ grados de libertad y se lo calcula como se muestra a continuación:

Estadístico de prueba: $c^2 \sim c^2_{(r-1)(c-1)}$

donde,

$$c^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - E[n_{ij}])^2}{E[n_{ij}]}$$

Entonces se rechaza H_0 en favor de H_a , con $(1-\alpha)100\%$, si:

$$c^2 > c^2_{\alpha} \quad (r-1)(c-1)$$

Se determina valor p de la prueba o nivel de significancia alcanzado, que es el mínimo valor de α para el cual los datos observados indican que se tendría que rechazar la hipótesis nula.

A continuación se realizará análisis de contingencia entre las variables generales (edad, sexo, ubicación y sección) y las variables de calificación de lenguaje y calificación de matemáticas, utilizando el procedimiento descrito en esta sección.

4.4.1 Análisis de contingencia entre las variables edad y calificación de lenguaje

El número de criterios de clasificación que se utilizará para cada variable es dos, debido a que al realizar la prueba con más criterios existían celdas de la tabla de contingencia que contienen menos de cinco observaciones lo cual no es recomendable.

Definimos primeramente los criterios de clasificación de la variable aleatoria edad, el criterio **A** corresponde a las edades entre 10 y 15 años y el criterio **B** contiene a las edades de 16 a 22 años.

Criterios de clasificación para la edad	Criterios de clasificación para la calificación de lenguaje		Total
	C	D	
A	60 65.209	61 55.79	121
B	30 24.79	16 21.209	46
Total	90	77	167

Para la variable aleatoria calificación de lenguaje el criterio de **C** es el de las calificaciones desde 0 a 44 puntos sobre cien, y el criterio **D** corresponde a las calificaciones entre 45 y 83 puntos. En la tabla CXXXII

se muestran las observaciones agrupadas de acuerdo a los criterios de clasificación antes mencionados.

Los valores esperados n_{ij} obtenidos son mostrados en la tabla CXXXII, estos valores son utilizados para calcular el estadístico de prueba. A continuación se construye el contraste de hipótesis para determinar la independencia de las variables edad y calificación de lenguaje.

H₀: Los criterios de clasificación de las variables aleatorias
edad y calificación de lenguaje son independientes

vs.

H_a: Los criterios de clasificación de las variables aleatorias
edad y calificación de lenguaje son dependientes

Valor del estadístico de prueba $\chi^2 = 3.277$

Grados de libertad 1

Valor p = 0.070256

En vista de que el valor p, es $0.10 \leq p \leq 0.05$, simplemente mostramos el mismo.

4.4.2 Pruebas entre las variables sexo y calificación de lenguaje

Para realizar el análisis de contingencia entre estas dos variables se definieron los siguientes criterios de clasificación. Para la variable sexo el criterio **A** corresponde al sexo femenino y el criterio **B** al sexo masculino; y los criterios para la variable aleatoria calificación de lenguaje son **C** que corresponde a las calificaciones entre 0 y 44 puntos, y el criterio **D** que contiene las calificaciones de 45 a 83 puntos. Las observaciones que cumplen con estos criterios de clasificación son mostradas en la tabla CXXXIII.

Tabla CXXXIII
Criterios de clasificación para las variables sexo y calificación de lenguaje

Criterios de clasificación para la variable sexo	Criterios de clasificación para la variable calificación de lenguaje		Total
	C	D	
A	45 46.317	40 38.682	85
B	46 44.682	36 37.317	82
Total	91	76	167

Para determinar la independencia de los criterios de clasificación de las variables sexo y calificación de lenguaje se realizó el siguiente contraste de hipótesis:

H₀: Los criterios de clasificación de las variables sexo y calificación de lenguaje son independientes

vs.

H_a: Los criterios de clasificación de las variables sexo y calificación de lenguaje son dependientes

Valor del estadístico de prueba $\chi^2 = 3.84146$

Grados de libertad 1

Valor p = 0.16767

Existe evidencia estadística para no rechazar H₀, por lo tanto se concluye que los criterios de clasificación son independientes.

4.4.3 Pruebas entre las variables ubicación del establecimiento educativo y calificación de lenguaje

Los criterios de clasificación utilizados para el análisis de contingencia de estas dos variables son para la variable ubicación del establecimiento educativo el criterio **A** que corresponde a las zonas de Progreso, Sabana

grande y Puná , y el criterio **B** contiene a las zonas de Posorja y Tenguel.

En lo que respecta a la variable calificación de lenguaje los criterios de clasificación son: **C** que contiene a las calificaciones de 0 a 44 puntos, y el criterio **D** que corresponde a las calificaciones entre 45 y 83 puntos. Los totales de las observaciones que cumplen con cada uno de estos criterios son mostrados en la tabla CXXXIV.

Criterios de clasificación para la ubicación	Criterios de clasificación para la calificación de lenguaje		Total
	C	D	
A	17 17.784	37 36.215	54
B	38 37.215	75 75.784	113
Total	55	112	167

Para determinar la independencia de las variables ubicación del establecimiento educativo y calificación de lenguaje se realizó el siguiente contraste de hipótesis:

H₀: Los criterios de clasificación de las variables ubicación
y calificación de lenguaje son independientes

vs.

H_a: Los criterios de clasificación de las variables ubicación
y calificación de lenguaje son dependientes

Valor del estadístico de prueba $\chi^2 = 3.84146$

Grados de libertad 1

Valor p = 0.78245

Existe evidencia estadística para no rechazar H₀, por lo tanto se concluye que los criterios de clasificación son independientes.

4.4.4 Pruebas entre las variables jornada de trabajo del establecimiento educativo y calificación de lenguaje

Los criterios de clasificación definidos para las variable analizadas en esta sección son para la variable jornada el criterio **A** corresponde a la jornada matutina y el criterio **B** contiene a las jornadas vespertina y nocturna.

Para la variable calificación de lenguaje los criterios son **C** que corresponde a las calificaciones entre 0 y 44 puntos; y el criterio **D** que contiene a las calificaciones desde 45 a 83. En la tabla CXXXV se

pueden observar los totales de las observaciones que cumplen con cada criterio de clasificación.

Tabla CXXXV			
Criterios de clasificación para las variables jornada y calificación de lenguaje			
Criterios de clasificación para la variable jornada	Criterios de clasificación para la variable calificación de lenguaje		Total
	C	D	
A	79 78.467	12 12.539	101
B	65 65.532	11 10.467	66
Total	144	23	167

Para determinar la independencia de los criterios de clasificación definidos, de las variables jornada de trabajo del establecimiento educativo y calificación de lenguaje se realizó el siguiente contraste de hipótesis:

H₀: Los criterios de clasificación de las variables jornada y calificación de lenguaje son independientes

vs.

H_a: Los criterios de clasificación de las variables jornada y calificación de lenguaje son dependientes

Valor del estadístico de prueba $\chi^2 = 0.057749$

Grados de libertad 1

Valor $p = 0.81008$

Existe evidencia estadística para no rechazar H_0 , por lo tanto se concluye que los criterios de clasificación son independientes.

4.4.5 Pruebas entre las variables edad y calificación de matemáticas

Para realizar el análisis de contingencia con estas dos variables se definieron los criterios de clasificación de la siguiente manera, para la variable edad se utilizó el mismo criterio del análisis de la sección 4.5.1, es decir el criterio **A** corresponde a las edades de 10 a 15 años y el criterio **B** contiene a las edades de 16 a 22 años.

Tabla CXXXVI
Criterios de clasificación para las variables edad y calificación de matemáticas

Criterios de clasificación para la variable edad	Criterios de clasificación para la variable calificación de matemáticas		Total
	C	D	
A	76 73.175	45 47.82	121
B	25 27.82	21 18.178	46
Total	101	66	167

En lo que respecta a la variable calificación de matemáticas el criterio de **C** contiene a las calificaciones de 0 a 14 puntos, y el criterio **D** corresponde a las calificaciones entre 15 y 32 puntos.

Para determinar la independencia de los criterios de clasificación entre las variables aleatorias edad y calificación de matemáticas se realizó el siguiente contraste de hipótesis:

H₀: Los criterios de clasificación de las variables edad y calificación de matemáticas son independientes

vs.

H_a: Los criterios de clasificación de las variables edad y calificación de matemáticas son dependientes

Valor del estadístico de prueba $\chi^2 = 0.9985$

Grados de libertad 1

Valor p = 0.31767

Existe evidencia estadística para no rechazar H_0 , por lo tanto se concluye que los criterios de clasificación son independientes.

4.4.6 Pruebas entre las variables sexo y calificación de matemáticas

La definición de los criterios de clasificación para las variables analizadas en esta sección son, para la variable sexo el criterio **A** corresponde al

sexo femenino y el criterio **B** al sexo masculino, y para la variable calificación de matemáticas el criterio de **C** es el de las calificaciones entre 0 y 14 puntos y el criterio **D** corresponde a la calificaciones entre 15 y 32 puntos

Tabla CXXXVII
Criterios de clasificación para las variables sexo y calificación de matemáticas

Criterios de clasificación para la variable sexo	Criterios de clasificación para la variable calificación de matemáticas		Total
	C	D	
A	51 52.425	34 32574	85
B	52 50.574	30 31.425	82
Total	103	64	167

Para determinar la independencia de los criterios de clasificación de las variables sexo y calificación de matemáticas se realizó el siguiente contraste de hipótesis:

H₀: Los criterios de clasificación de las variables sexo y calificación de matemáticas son independientes

vs.

H_a: Los criterios de clasificación de las variables sexo y calificación de matemáticas son dependientes

Valor del estadístico de prueba $\chi^2 = 0.20588$

Grados de libertad 1

Valor p = 0.65001

Existe evidencia estadística para no rechazar H_0 , por lo tanto se concluye que los criterios de clasificación son independientes.

4.4.7 Pruebas entre las variables ubicación del establecimiento educativo y calificación de matemáticas

Los criterios de clasificación de las variables que se analizarán en esta sección son para la variable ubicación el criterio **A** corresponde a las zonas de Progreso, Sabana grande y Puná , el criterio **B** contiene a las zonas de Posorja y Tenguel. Para la variable calificación de matemáticas se definieron los criterios **C** que corresponde a las calificaciones entre 0 y 14 puntos y **D** que corresponde a las calificaciones de 15 a 32 puntos.

Una vez definidos los criterios de clasificación de las variables para el análisis de contingencia se realizó el siguiente contraste de hipótesis:

H_0 : Los criterios de clasificación de las variables ubicación
y calificación de matemáticas son independientes

vs.

H_a : Los criterios de clasificación de las variables ubicación y calificación de matemáticas son dependientes

Tabla CXXXVIII
Criterios de clasificación para las variables ubicación y calificación de matemáticas

Criterios de clasificación para la variable ubicación	Criterios de clasificación para la variable calificación de matemáticas		Total
	C	D	
A	24 33.592	31 21.407	55
B	78 68.407	34 43.592	112
Total	102	65	167

Valor del estadístico de prueba $\chi^2 = 0.10494$

Grados de libertad 1

Valor p = 0.00119775

Existe evidencia estadística para rechazar H_0 , en favor de H_a por lo tanto se concluye que los criterios de clasificación son dependientes.

4.4.8 Pruebas entre las variables jornadas del establecimiento educativo y calificación de matemáticas

Para realizar el análisis de contingencia entre las variables jornada y calificación de matemáticas se definieron los siguientes criterios de clasificación, para la variable jornada el criterio **A** corresponde a la jornada matutina y el criterio **B** contiene a la jornada vespertina y nocturna; en lo que respecta a la variable calificación de matemáticas el criterio **C** es el de las calificaciones entre 0 y 14 puntos y el criterio **D** corresponde a las calificaciones desde 15 a 32 puntos. En la tabla CXXXIX se muestran las observaciones que cumplen estos criterios.

Criterios de clasificación para la variable jornada	Criterios de clasificación para la variable calificación de matemáticas		Total
	C	D	
A	83	61	144
B	19	4	23
Total	102	65	167

Para determinar la independencia de los criterios de clasificación de las variables jornada y calificación de matemáticas se realizó el siguiente contraste de hipótesis:

H₀: Los criterios de clasificación de las variables jornada y calificación de matemáticas son independientes

vs.

H_a: Los criterios de clasificación de las variables jornada y calificación de matemáticas son dependientes

Valor del estadístico de prueba $\chi^2 = 5.2014$

Grados de libertad 1

Valor p = 0.022568

Existe evidencia estadística para rechazar H₀, en favor de H_a por lo tanto se concluye que los criterios de clasificación son dependientes.

4.4.9 Pruebas entre las variables calificación de lenguaje y calificación de matemáticas

Para el análisis de contingencia realizado con estas dos variables se definieron los siguiente criterios de clasificación, para la variable calificación de lenguaje los criterios son **C** que corresponde a las

calificaciones desde 0 a 44 puntos y el criterio **D** que contiene a las calificaciones entre 45 y 83 puntos.

Tabla CXL
Criterios de clasificación para las variables calificación de lenguaje y calificación de matemáticas

Criterios de clasificación para la variable calificación de matemáticas	Criterios de clasificación para la variable calificación de lenguaje		Total
	C	D	
A	66 54.97	36 47.029	102
B	24 35.029	41 29.97	75
Total	90	77	167

Los criterios de la variable calificación de matemáticas son **A** que corresponde a las calificaciones desde 0 a 14 puntos, el criterio **B** que contiene a las calificaciones entre 15 y 32 puntos. Las observaciones que cumplen con estos criterios se muestran en la tabla CXL.

Para determinar la independencia de los criterios de clasificación de las variables calificación de lenguaje y calificación de matemáticas se realizó el siguiente contraste de hipótesis:

H₀: Los criterios de clasificación de las variables calificación de lenguaje y calificación de matemáticas son independientes

vs.

H_a: Los criterios de clasificación de las variables calificación de lenguaje y calificación de matemáticas son dependientes

Valor del estadístico de prueba $\chi^2 = 12.332$

Grados de libertad 1

Valor p = 0.0004455

Se concluye que existe evidencia estadística para rechazar H₀, en favor de H_a por lo tanto los criterios de clasificación son dependientes.

4.4.10 Resultados del análisis de contingencia entre otras variables aleatorias

Los resultados obtenidos al realizar el análisis de contingencia entre otras variables aleatorias, se muestran en la tabla CXLI, en la cual se indica entre que variables aleatorias se efectuó el análisis de contingencia, el valor p de la prueba y la conclusión.

Tabla CXLI
Resultados del análisis de contingencia realizado a otras variables
aleatorias

Variable X	Variable Y	Valor p	Conclusión
Suma de enteros	Resta de enteros	0.004	Dependientes
División de fracciones	Multiplicación de fracciones	0.000	Dependientes
Área del círculo	Área del triángulo	0.000	Dependientes
Ecuación con una incógnita	Ecuación con dos incógnitas	0.033	Dependientes
Estadística	Probabilidad	0.104	Independientes
Resta de enteros	Lectura Comprensiva	0.713	Independientes
Suma de enteros	Lectura Comprensiva	0.000	Dependientes
Palabras agudas	Palabras graves	0.000	Dependientes
Palabras sinónimas	Palabras antónimas	0.000	Dependientes
Palabras diminutivas	Palabras despectivas	0.000	Dependientes
Identificación del sujeto	Identificación del predicado	0.000	Dependientes
Identificación del sustantivo	Identificación del verbo	0.000	Dependientes
Identificación de frases	Identificación de oraciones	0.000	Dependientes
Palabras Agudas	Palabras tildadas	0.037	Dependientes

Como se puede observar en la tabla CXLI, de los pares de variables aleatorias analizados, se obtuvo que las variables correspondientes a los ejercicios de estadística y de probabilidad son independientes; y entre las variables correspondientes a la resta de números enteros y lectura comprensiva también se obtuvo que son independencia no lineal.

4.5 Análisis de varianza

El diseño de experimentos es una técnica estadística en la cual una variable aleatoria cuantitativa es explicada en términos de una o más variables aleatorias cualitativas. A las variables aleatorias cualitativas se las denomina **factores**, y a los distintos valores que pueden tomar dichos factores se los denomina **niveles o tratamientos**, el número de tratamientos de un factor se denota por a . La respuesta que se observa en cada uno de los tratamientos es una variable aleatoria.

Los datos típicos para un modelo unifactorial se tabulan como se muestra en la siguiente tabla:

Tratamiento (nivel)	Observaciones				Total	Promedio
1	Y_{11}	Y_{12}	...	Y_{1n}	$Y_{1.}$	\bar{Y}_1
2	Y_{21}	Y_{22}	...	Y_{2n}	$Y_{2.}$	\bar{Y}_2
\vdots	\vdots	\vdots		\vdots	\vdots	\vdots
a	Y_{a1}	Y_{a2}	...	Y_{an}	$Y_{a.}$	\bar{Y}_a
					$Y_{..}$	$\bar{Y}_{..}$

Estas observaciones se pueden escribir por medio un modelo estadístico lineal como el siguiente:

$$y_{ij} = \mathbf{m} + \mathbf{t}_i + \epsilon_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases} \quad \epsilon_{ij} \sim N(0, \mathbf{s}^2) \text{ y } \text{Cov}(\epsilon_{ij}, \epsilon_{lm}), i \neq l \text{ y } j \neq m$$

donde y_{ij} es la j -ésima observación del i -ésimo tratamiento, μ es un parámetro común de todos los tratamientos denominado media global, τ_i es un parámetro único para el i -ésimo tratamiento llamado efecto del tratamiento i -ésimo, y ϵ_{ij} es la componente aleatoria del error, donde $\epsilon_{ij} \sim N(0, \sigma^2)$ y se supone es constante para todos los tratamientos.

Una vez planteado el modelo matemático el objetivo es construir contrastes de hipótesis con respecto a los efectos de los tratamientos y hacer una estimación de estos. A este modelo se lo denomina *análisis de varianza* de clasificación en un sentido, dado que es un modelo unifactorial. Este diseño es completamente aleatorizado, por lo que se requiere que el experimento se realice en orden aleatorio, de tal forma que el medio ambiente en el que se usan los tratamientos tengan las mismas condiciones.

En el modelo unifactorial $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$, cuando σ^2 , la varianza del error ϵ_{ij} permanece constante y se incluyen todos los i tratamientos del factor, el modelo se denomina *de efectos fijos*. En este modelo los efectos de los tratamiento τ_i se definen como desviaciones con respecto a la media general μ , por lo que:

$$\sum_{i=1}^a t_i = 0$$

Sea y_i el total de las observaciones bajo el i -ésimo tratamiento, \bar{y}_i el promedio de las observaciones bajo el i -ésimo tratamiento. De igual forma, sea $y_{..}$ la suma de todas las observaciones y $\bar{y}_{..}$ la media general de las observaciones, expresado matemáticamente:

$$y_i = \sum_{j=1}^n y_{ij} \quad \bar{y}_i = y_i / n$$

$$y_{..} = \sum_{i=1}^a \sum_{j=1}^n y_{ij} \quad \bar{y}_{..} = y_{..} / N$$

En donde $i = 1, 2, \dots, a$ y $N = an$ es el número total de observaciones. La notación del punto en el subíndice, indica la suma del subíndice que reemplaza.

La media del i -ésimo tratamiento es $E(y_{ij}) = \mu_i = \mu + \tau_i$, $i = 1, 2, \dots, a$. Es decir que el valor medio del i -ésimo tratamiento es la suma de la media general y el efecto del i -ésimo tratamiento. El contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_a$$

vs.

$$H_1 : \mu_i = \mu_j, \text{ para al menos un par de } (i, j)$$

El contraste se lo puede escribir en términos de los efectos de los tratamientos de la siguiente forma

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_a$$

vs.

$$H_1 : \tau_i \neq 0, \text{ para al menos un } i$$

Por lo tanto se desea probar la igualdad de las medias de los tratamientos, o que los efectos de los tratamientos τ_i son iguales a cero. El procedimiento que se utiliza para esta prueba es el análisis de varianza, es decir la prueba se basa en un análisis de la variabilidad total de los datos, la cual está dada por la *suma total de cuadrados SCT*, si se divide SCT para $an-1$ se obtiene la varianza de la población, la cual es una medida de la variabilidad de las observaciones.

$$SCT = \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2, \text{ donde } \bar{Y}_{..} = \frac{1}{na} \sum_{i=1}^a \sum_{j=1}^n Y_{ij}$$

La suma total de cuadrados se puede expresar como la *suma de cuadrados de los tratamientos SCTr* y la *suma de cuadrados del error SCE*, como se muestra a continuación,

$$\sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 = n \sum_{i=1}^a (\bar{Y}_i - \bar{Y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2$$

$$SCT = SCTr + SCE$$

La suma cuadrática de los tratamientos SCT_r , mide la variación existente entre las medias de los a tratamientos, mientras que la suma cuadrática del error SCE , mide la variación existente dentro de cada tratamiento.

Dado que los $Y_{i1}, Y_{i2}, \dots, Y_{in}$, para $i = 1, 2, \dots, a$ representan una muestra aleatoria de tamaño n , tomada de una población normal con media $\mu + \tau_i$ y varianza σ^2 , se tiene que para cada i

$$\frac{1}{s^2} \cdot \sum_{j=1}^n (Y_{ij} - \bar{Y}_{i.})^2 = \frac{[(n-1)S^2]}{s^2}$$

es una variable aleatoria ji-cuadrado χ^2 con $n-1$ grados de libertad.

Además puesto que los a tratamientos son variables aleatorias independientes, se tiene que

$$\frac{1}{s^2} \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{i.})^2$$

es una variable aleatoria ji-cuadrado con $a(n-1)$ grados de libertad.

Luego supongamos que X es una variable aleatoria ji-cuadrado con v grados de libertad, entonces $E(X) = v$, en base a esto el valor esperado de la variable aleatoria $(1/\sigma^2) SCE$ es $a(n-1)$, donde $\frac{SSE}{a(n-1)}$ es un

estimador de σ^2 , esta cantidad se denomina *cuadrado medio del error* y se denota por *CME*

Por otro lado se tiene que las \bar{Y}_i son variables aleatorias independientes, distribuidas normalmente con media μ y varianza $\frac{s^2}{n}$, se tiene que

$$\frac{n}{s^2} \sum_{i=1}^a (\bar{Y}_i - \bar{Y}_{..})^2$$

es una variable aleatoria ji-cuadrado con a-1 grados de libertad. El valor esperado de esta variable aleatoria es igual a (a-1), por lo tanto $\frac{SSTr}{a-1}$ es otro estimador de σ^2 , este estimador se denomina *cuadrado medio de los tratamientos* y se denota por *CMTr*.

En este análisis se supone que los estimadores de σ^2 , CMTr y CME son variables aleatorias independientes con a-1 y a(n-1) grados de libertad respectivamente, entonces $f = \frac{SCTr / (a-1)s^2}{SCE / a(n-1)s^2} = \frac{CMTr}{CME}$ es una variable aleatoria que tiene distribución F con (a-1) y a(n-1), grados de libertad.

La hipótesis nula H_0 (de que los efectos de los a tratamientos son iguales a cero) se rechaza si el valor f calculado en base a las observaciones, es mayor que $F_{\alpha, k-1, k(n-1)}$, donde α es el nivel de significancia de la prueba. Si la hipótesis nula es verdadera CMTr es un estimador insesgado de σ^2 .

Tabla CXLII
Análisis de varianza de clasificación en un sentido

Fuente	Grados de libertad	Sumas cuadráticas	Cuadrado medio	F
Tratamientos	$a-1$	SCTr	CMTr	$\frac{CMTr}{CME}$
Error	$a(n-1)$	SCE	CME	
Total	$an-1$	STC		

El procedimiento descrito anteriormente se denomina análisis de varianza de clasificación en un sentido y se lo representa mediante una *tabla de análisis de varianza ANOVA*, como se muestra en la tabla CXLII. Los diferentes métodos de hacer comparaciones múltiples se emplean sólo cuando el resultado del ANOVA resulta significativo. En tal caso, se sabe que existen diferencias entre las muestras, pero sin poder especificar entre cuales de ellas. Se necesita, entonces, alguna forma de

poder compararlas entre sí, y alcanzar así el objetivo final del ANOVA. Los métodos de comparaciones múltiples son la herramienta adecuada para este fin, pues permiten un análisis muy sutil de la significación encontrada. Para los casos en que se determine que los efectos de los tratamientos no son todos iguales a cero, se utilizará el método de la *mínima diferencia significativa (LSD)*, para determinar en que parejas de tratamientos existe diferencias de medias.

Con el método *LSD* se desea probar el siguiente contraste de hipótesis

$$H_0 : \mu_i - \mu_j = 0, \text{ para } i \neq j, \text{ e } i, j = 1, 2, \dots, a$$

vs.

$$H_1 : \mu_i - \mu_j \neq 0, \text{ para } i \neq j, \text{ e } i, j = 1, 2, \dots, a$$

Para lo cual se utiliza el estadístico de prueba:

$$t_0 = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{CME \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}$$

Para probar esta hipótesis se calcula la mínima diferencia significativa *LSD*, tal como sigue,

$$LSD = t_{\alpha/2, (N-a)} \sqrt{MCE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}, \text{ donde } N = a \times n, \text{ entonces se}$$

rechaza la hipótesis nula H_0 en favor de la hipótesis alterna H_a sí,

$$|\bar{Y}_i - \bar{Y}_j| > LSD .$$

4.5.1 Análisis de varianza de la variable aleatoria calificación de lenguaje con respecto a la ubicación de los planteles educativos fiscales rurales del cantón Guayaquil.

El número de tratamientos del factor ubicación es cinco, los cuales corresponden a las parroquias rurales Progreso, Sabana Grande, Puná, Posorja y Tenguel. Para realizar el análisis de varianza entre la variable aleatoria calificación de lenguaje con respecto a la ubicación de los planteles educativos fiscales rurales del cantón Guayaquil, se planteó el modelo estadístico lineal siguiente:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \begin{cases} i = 1, 2, 3, 4, 5 \\ j = 1, 2, \dots, n \end{cases}$$

$$\epsilon_{ij} \sim N(0, \sigma^2) \text{ y } \text{Cov}(\epsilon_{ij}, \epsilon_{lm}), i \neq l \text{ y } j \neq m$$

donde y_{ij} es la j -ésima observación del i -ésimo tratamiento, μ es 43.14 puntos, τ_i es el efecto del i -ésimo tratamiento y ϵ_{ij} es la componente aleatoria del error, donde $\epsilon_{ij} \sim N(0, \sigma^2)$ y se supone, es constante para todos los tratamientos.

El contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

vs.

$$H_1 : \mu_i \neq \mu_j, \text{ para al menos un par de } (i, j)$$

El contraste en términos de los efectos de los tratamientos es

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5$$

vs.

$$H_1 : \tau_i \neq 0, \text{ para al menos un } i$$

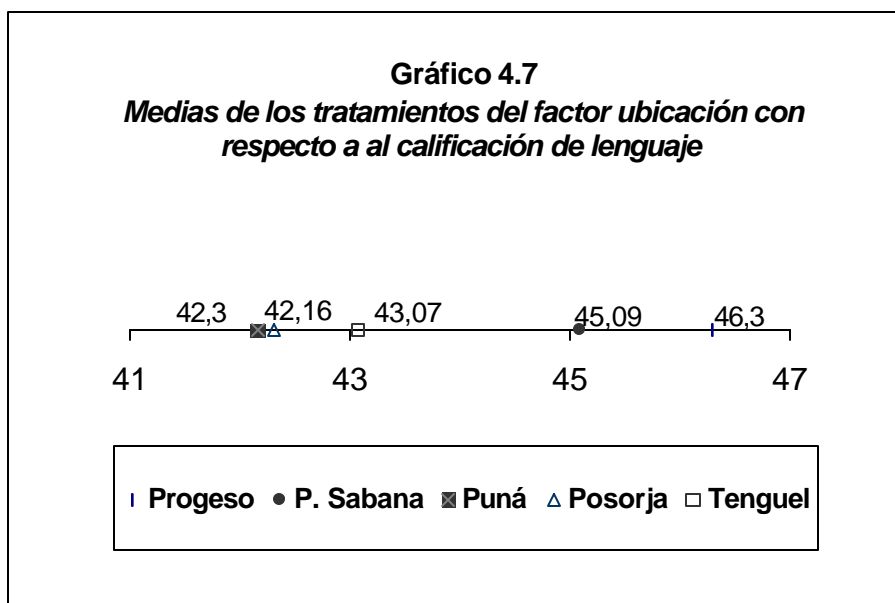
Por lo tanto se desea probar la igualdad de las medias de los tratamientos, o que los efectos de los tratamientos τ_i son iguales a cero.

En el análisis de varianza de la variable aleatoria calificación de lenguaje con respecto al factor ubicación se obtuvo los siguientes resultados:

Tabla de análisis de varianza

<i>Fuente</i>	<i>Grados de libertad</i>	<i>Sumas cuadráticas</i>	<i>Cuadrados medios</i>	<i>F</i>	<i>P</i>
<i>Tratamientos</i>	4	355.86	88.97	0.408	0.803
<i>Error</i>	162	35340.655	218.152		
<i>Total</i>	166	35696.515			

El valor $p = 0.803$ de esta prueba indica que existe evidencia estadística para no rechazar la hipótesis nula de que las medias de los a tratamientos son iguales. En el gráfico 4.7 se puede apreciar las medias de cada uno de los tratamientos del factor ubicación, con respecto a la calificación de la prueba de lenguaje.



4.5.2 Análisis de varianza de la variable aleatoria calificación de lenguaje con respecto a la jornada de los planteles educativos fiscales rurales del cantón Guayaquil.

El número de tratamientos del factor jornada es tres, los cuales corresponden a las jornadas matutina, vespertina y nocturna. Para realizar el análisis de varianza entre la variable aleatoria calificación de lenguaje con respecto a la jornada de trabajo de los planteles educativos fiscales rurales del cantón Guayaquil, se planteó el modelo estadístico lineal siguiente:

$$y_{ij} = \mathbf{m} + \mathbf{t}_i + \epsilon_{ij} \begin{cases} i = 1, 2, 3 \\ j = 1, 2, \dots, n \end{cases}$$

$$\epsilon_{ij} \sim N(0, \mathbf{s}^2) \text{ y } Cov(\epsilon_{ij}, \epsilon_{lm}), i \neq l \text{ y } j \neq m$$

donde y_{ij} es la j -ésima observación del i -ésimo tratamiento, μ es 43.14 puntos, τ_i es el efecto del i -ésimo tratamiento y ϵ_{ij} es la componente aleatoria del error, donde $\epsilon_{ij} \sim N(0, \sigma^2)$ y se supone, es constante para todos los tratamientos.

El contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

vs.

$$H_1 : \mu_i = \mu_j, \text{ para al menos un par de } (i, j)$$

El contraste en términos de los efectos de los tratamientos es

$$H_0 : \tau_1 = \tau_2 = \tau_3$$

vs.

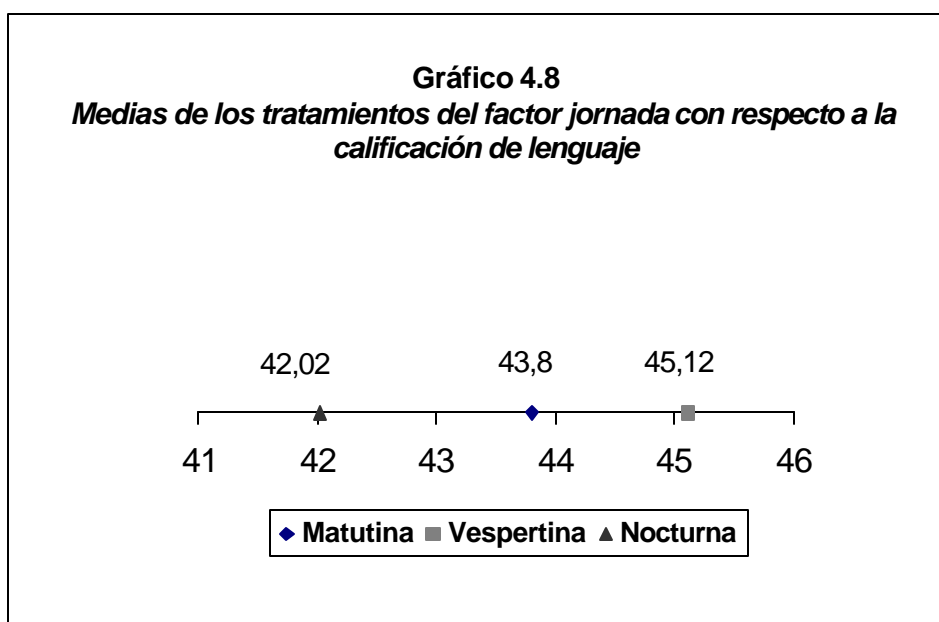
$$H_1 : \tau_i \neq 0, \text{ para al menos un } i$$

Por lo tanto se desea probar la igualdad de las medias de las jornadas matutina, vespertina y nocturna, o que los efectos de los tratamientos τ_i son iguales a cero. En el análisis de varianza de la variable aleatoria calificación de lenguaje con respecto al factor jornada se obtuvo los siguientes resultados:

Tabla de análisis de varianza

<i>Fuente</i>	<i>Grados de libertad</i>	<i>Sumas cuadráticas</i>	<i>Cuadrados medios</i>	<i>F</i>	<i>P</i>
<i>Tratamientos</i>	2	27.405	13.703	0.063	0.939
<i>Error</i>	164	35669.146	217.495		
<i>Total</i>	166	35796.551			

El valor $p = 0.803$ de esta prueba indica que existe evidencia estadística para no rechazar la hipótesis nula de que las medias de los a tratamientos son iguales.



En el gráfico 4.8 se puede apreciar las medias de cada uno de los tratamientos del factor jornada, con respecto a la calificación de la prueba de lenguaje. En el gráfico se puede observar que los planteles educativos que funcionan en jornada vespertina, obtuvieron la más alta media en la calificación de lenguaje.

4.5.3 Análisis de varianza de la variable aleatoria calificación de matemáticas con respecto al factor ubicación.

Para realizar el análisis de varianza entre la variable aleatoria calificación de matemáticas, con respecto a la ubicación de los planteles educativos fiscales rurales del cantón Guayaquil, se planteó el modelo estadístico lineal siguiente:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \begin{cases} i = 1, 2, 3, 4, 5 \\ j = 1, 2, \dots, n \end{cases}$$

$$\epsilon_{ij} \sim N(0, \sigma^2) \text{ y } Cov(\epsilon_{ij}, \epsilon_{lm}), i \neq l \text{ y } j \neq m$$

donde y_{ij} es la j -ésima observación del i -ésimo tratamiento, μ es 13.31 puntos, τ_i es el efecto del i -ésimo tratamiento y ϵ_{ij} es la componente aleatoria del error, donde $\epsilon_{ij} \sim N(0, \sigma^2)$ y se supone, es constante para todos los tratamientos. El número de tratamientos del factor ubicación es cinco, los cuales corresponden a las parroquias Progreso, Sabana Grande, Puná, Posorja y Tenguel.

El contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

vs.

$$H_1 : \mu_i = \mu_j, \text{ para al menos un par de } (i, j)$$

El contraste en términos de los efectos de los tratamientos es

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5$$

vs.

$$H_1 : \tau_i \neq 0, \text{ para al menos un } i$$

Por lo tanto se desea probar la igualdad de las medias de los tratamientos, o que los efectos de los tratamientos τ_i son iguales a cero.

En el análisis de varianza de la variable aleatoria calificación de matemáticas con respecto al factor ubicación se obtuvo los siguientes resultados:

Tabla de análisis de varianza

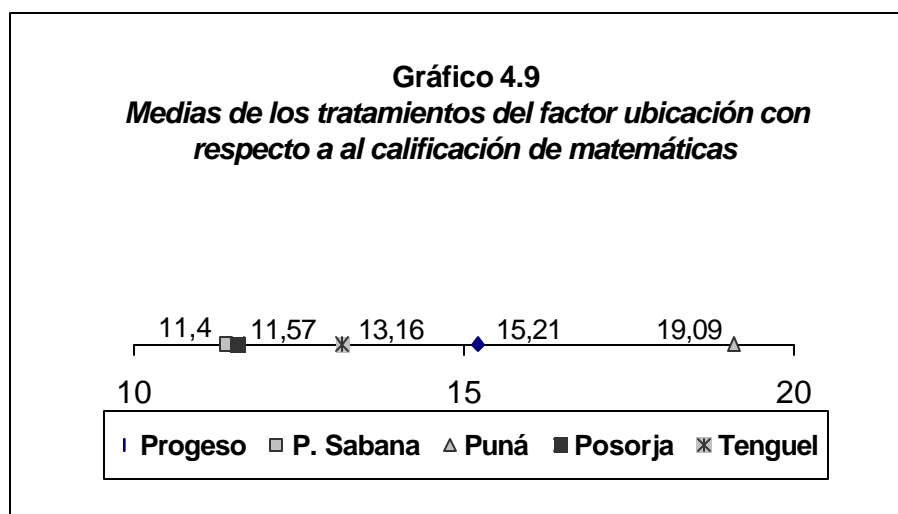
<i>Fuente</i>	<i>Grados de libertad</i>	<i>Sumas cuadráticas</i>	<i>Cuadrados medios</i>	<i>F</i>	<i>P</i>
<i>Tratamientos</i>	4	1028.075	257.019	5.576	0.000
<i>Error</i>	162	7467.7333	46.057		
<i>Total</i>	166	8495.808			

El nivel de significancia alcanzado o valor $p = 0.000$ de esta prueba indica que existe evidencia estadística para rechazar la hipótesis nula de que las medias de los 5 tratamientos son iguales. Por lo tanto se realiza la prueba de mínima diferencia significativa, descrita en la sección 4.5, para determinar cuales pares de medias, de los cinco tratamientos del factor ubicación, difieren entre sí, para lo cual se determinó que el valor de LSD es igual a 6.94.

Las diferencias de las medias de los tratamientos Sabana Grande y Puná es 7.69, que es mayor al LSD, por lo tanto existe evidencia estadística para concluir que las medias de estos dos tratamientos son diferentes. De igual forma se calculó la diferencia de medias de los tratamientos de Puná y Posorja, cuyo valor es $7.52 > \text{LSD}$, por lo tanto se concluye que existe evidencia estadística de que las medias de estos dos tratamientos difieren significativamente

H₀	H₁	Diferencias
$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 \neq 0$	3.81
$\mu_1 - \mu_3 = 0$	$\mu_1 - \mu_3 \neq 0$	3.88
$\mu_1 - \mu_4 = 0$	$\mu_1 - \mu_4 \neq 0$	3.64
$\mu_1 - \mu_5 = 0$	$\mu_1 - \mu_5 \neq 0$	2.05
$\mu_2 - \mu_3 = 0$	$\mu_2 - \mu_3 \neq 0$	7.69
$\mu_2 - \mu_4 = 0$	$\mu_2 - \mu_4 \neq 0$	0.17
$\mu_2 - \mu_5 = 0$	$\mu_2 - \mu_5 \neq 0$	1.76
$\mu_3 - \mu_4 = 0$	$\mu_3 - \mu_4 \neq 0$	7.52
$\mu_3 - \mu_5 = 0$	$\mu_3 - \mu_5 \neq 0$	5.93
$\mu_4 - \mu_5 = 0$	$\mu_4 - \mu_5 \neq 0$	1.59

En el gráfico 4.9 se puede apreciar las medias de cada uno de los tratamientos del factor ubicación, con respecto a la calificación de la prueba de matemáticas.



4.5.4 Análisis de varianza de la variable aleatoria calificación de matemáticas con respecto al factor jornada de trabajo de los colegios fiscales rurales del cantón Guayaquil.

El número de tratamientos del factor jornada es tres, los cuales corresponden a las jornadas matutina, vespertina y nocturna. Para realizar el análisis de varianza entre la variable aleatoria calificación de matemáticas con respecto a la jornada de los planteles educativos fiscales rurales del cantón Guayaquil, se planteó el modelo estadístico lineal siguiente:

$$y_{ij} = \mathbf{m} + \mathbf{t}_i + \epsilon_{ij} \begin{cases} i = 1, 2, 3 \\ j = 1, 2, \dots, n \end{cases}$$

$$\epsilon_{ij} \sim N(0, \mathbf{s}^2) \text{ y } \text{Cov}(\epsilon_{ij}, \epsilon_{lm}), i \neq l \text{ y } j \neq m$$

donde y_{ij} es la j -ésima observación del i -ésimo tratamiento, μ es 13.31 puntos, τ_i es el efecto del i -ésimo tratamiento y ϵ_{ij} es la componente aleatoria del error, donde $\epsilon_{ij} \sim N(0, \sigma^2)$ y se supone, es constante para todos los tratamientos.

El contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

vs.

$$H_1 : \mu_i \neq \mu_j, \text{ para al menos un par de } (i, j)$$

El contraste en términos de los efectos de los tratamientos es

$$H_0 : \tau_1 = \tau_2 = \tau_3$$

vs.

$$H_1 : \tau_i \neq 0, \text{ para al menos un } i$$

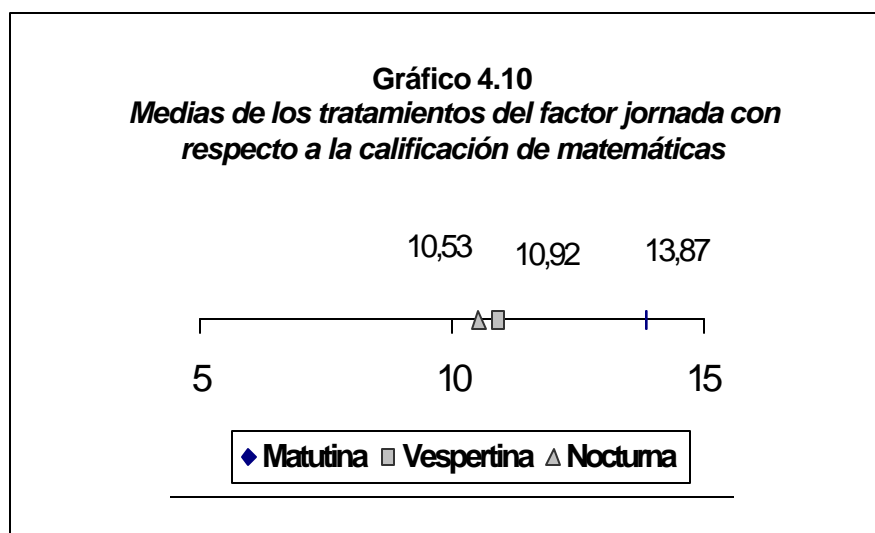
Por lo tanto se desea probar la igualdad de las medias de los tratamientos, o que los efectos de los tratamientos τ_i son iguales a cero.

En el análisis de varianza de la variable aleatoria calificación de matemáticas con respecto al factor jornada se obtuvo los siguientes resultados:

Tabla de análisis de varianza

<i>Fuente</i>	<i>Grados de libertad</i>	<i>Sumas cuadráticas</i>	<i>Cuadrados medios</i>	<i>F</i>	<i>P</i>
<i>Tratamientos</i>	2	208.215	104.108	2.06	0.131
<i>Error</i>	164	8287.593	50.534		
<i>Total</i>	166	8495.808			

El nivel de significancia alcanzado o valor $p = 0.131$ de esta prueba indica que existe evidencia estadística para no rechazar la hipótesis nula de que las medias de los 3 tratamientos son iguales. En el gráfico 4.10 se puede apreciar las medias de cada uno de los tratamientos del factor jornada, con respecto a la calificación de la prueba de matemáticas.



4.5.5 Análisis de varianza de la variable aleatoria calificación total con respecto al factor ubicación de los colegios fiscales rurales del cantón guayaquil

El número de tratamientos del factor ubicación es cinco, los cuales corresponden a las parroquias Progreso, Sabana Grande, Puná, Posorja y Tenguel. Para realizar el análisis de varianza entre la variable aleatoria calificación total con respecto a la ubicación de los planteles educativos fiscales rurales del cantón Guayaquil, se planteó el modelo estadístico lineal siguiente:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \begin{cases} i = 1, 2, 3, 4, 5 \\ j = 1, 2, \dots, n \end{cases}$$

$$\epsilon_{ij} \sim N(0, \sigma^2) \text{ y } Cov(\epsilon_{ij}, \epsilon_{lm}), i \neq l \text{ y } j \neq m$$

donde y_{ij} es la j -ésima observación del i -ésimo tratamiento, μ es 28.22 puntos, τ_i es el efecto del i -ésimo tratamiento y ϵ_{ij} es la componente aleatoria del error, donde $\epsilon_{ij} \sim N(0, \sigma^2)$ y se supone, es constante para todos los tratamientos.

El contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

vs.

$$H_1 : \mu_i = \mu_j, \text{ para al menos un par de } (i, j)$$

El contraste en términos de los efectos de los tratamientos es

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5$$

vs.

$$H_1 : \tau_i \neq 0, \text{ para al menos un } i$$

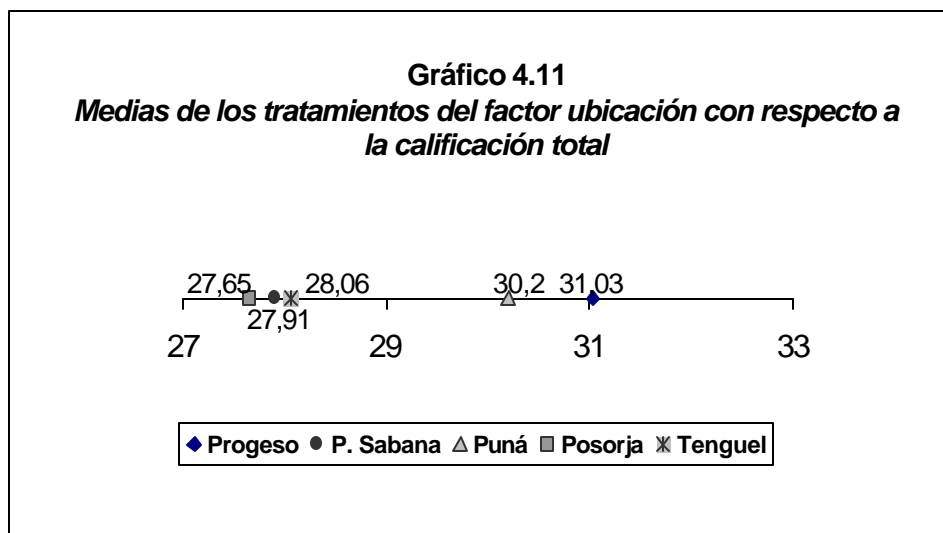
Por lo tanto se desea probar la igualdad de las medias de los tratamientos, o que los efectos de los tratamientos τ_i son iguales a cero.

En el análisis de varianza de la variable aleatoria calificación total con respecto al factor ubicación se obtuvo los siguientes resultados:

Tabla de análisis de varianza

<i>Fuente</i>	<i>Grados de libertad</i>	<i>Sumas cuadráticas</i>	<i>Cuadrados medios</i>	<i>F</i>	<i>P</i>
<i>Tratamientos</i>	4	349.614	87.403	0.935	0.445
<i>Error</i>	162	15140.015	93.457		
<i>Total</i>	166	15489.629			

El nivel de significancia alcanzado o valor $p = 0.445$ de esta prueba indica que existe evidencia estadística para no rechazar la hipótesis nula de que las medias de los 5 tratamientos son iguales. En el gráfico 4.11 se puede apreciar las medias de cada uno de los tratamientos del factor ubicación, con respecto a la calificación total. En este gráfico se observa que los estudiantes cuyo colegio está ubicado en Progreso obtuvieron la más alta media de la calificación total.



4.5.6 Análisis de varianza de la variable aleatoria calificación total con respecto al factor jornada de los colegios fiscales rurales del cantón Guayaquil.

El número de tratamientos del factor jornada es tres, los cuales corresponden a las jornadas matutina, vespertina y nocturna. Para realizar el análisis de varianza entre la variable aleatoria calificación total, con respecto a la jornada de los planteles educativos fiscales rurales del cantón Guayaquil, se planteó el modelo estadístico lineal siguiente:

$$y_{ij} = \mathbf{m} + \mathbf{t}_i + \epsilon_{ij} \begin{cases} i = 1, 2, 3 \\ j = 1, 2, \dots, n \end{cases}$$

$$\epsilon_{ij} \sim N(0, \mathbf{s}^2) \text{ y } Cov(\epsilon_{ij}, \epsilon_{lm}), i \neq l \text{ y } j \neq m$$

donde y_{ij} es la j -ésima observación del i -ésimo tratamiento, μ es 13.31 puntos, τ_i es el efecto del i -ésimo tratamiento y ϵ_{ij} es la componente aleatoria del error, donde $\epsilon_{ij} \sim N(0, \sigma^2)$ y se supone, es constante para todos los tratamientos.

El contraste de hipótesis es:

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

vs.

$$H_1 : \mu_i = \mu_j, \text{ para al menos un par de } (i, j)$$

El contraste en términos de los efectos de los tratamientos es

$$H_0 : \tau_1 = \tau_2 = \tau_3$$

vs.

$$H_1 : \tau_i \neq 0, \text{ para al menos un } i$$

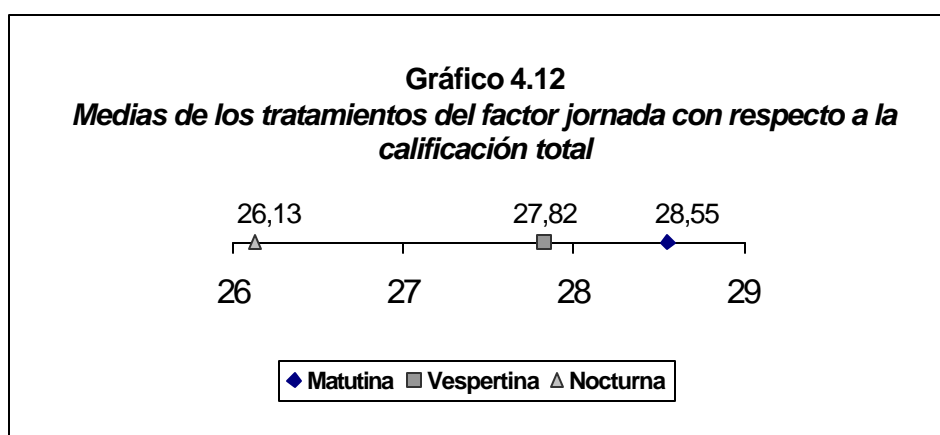
Por lo tanto se desea probar la igualdad de las medias de los tratamientos, o que los efectos de los tratamientos τ_i son iguales a cero.

En el análisis de varianza de la variable aleatoria calificación total con respecto al factor jornada se obtuvo los siguientes resultados:

Tabla de análisis de varianza

<i>Fuente</i>	<i>Grados de libertad</i>	<i>Sumas cuadráticas</i>	<i>Cuadrados medios</i>	<i>F</i>	<i>P</i>
<i>Tratamientos</i>	2	57.358	28.679	0.305	0.738
<i>Error</i>	164	15432.271	94.099		
<i>Total</i>	166	15489.629			

El nivel de significancia alcanzado o valor $p = 0.738$ de esta prueba indica que existe evidencia estadística para no rechazar la hipótesis nula de que las medias de los 3 tratamientos son iguales.



En el gráfico 4.12 se puede apreciar las medias de cada uno de los tratamientos del factor jornada, con respecto a la calificación total de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil. Como se puede observar los estudiantes de colegios, que funcionan con jornada matutina, obtuvieron la mayor media de la calificación total (28.55).

4.6 Análisis de componentes principales

El análisis de componentes principales es una técnica multivariada, que se utiliza con el objetivo explicar la estructura de la matriz de varianzas y covarianzas de un grupo de p variables aleatorias observables, para lo cual se construyen k variables aleatorias no observables, como combinaciones lineales de las p variables.

Los propósitos de la aplicación de esta técnica son dos. El primero es la reducción de datos, para la cual supongamos que se tienen n unidades de investigación y p variables aleatorias observables X_1, X_2, \dots, X_p con las cuales se construyen k componentes principales, la reducción de datos se produce cuando $k < p$, el valor de k se lo selecciona de tal forma que las k componentes principales contengan la mayor cantidad de información de las p variables originales en el menor número de estas, entonces la matriz original de datos de tamaño $n \times p$ se reduce a una de tamaño $n \times k$.

El segundo propósito es la interpretación de las variables no observables, con el fin de identificar alguna representación real de estas.

4.6.1 Determinación de las componentes principales

Sea $\mathbf{X}^T = (X_1, X_2, \dots, X_p)$ un vector aleatorio p-variado, es decir que $\mathbf{X} \in \mathbb{R}^p$, sean la matriz de varianzas y covarianzas y el vector de medias correspondientes a este vector, \mathbf{S} y $\mathbf{\bar{\mu}}$ respectivamente y siendo $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, los valores propios correspondientes a la matriz Σ , se construyen las siguientes combinaciones lineales.

$$Y_1 = \mathbf{a}_1^T \mathbf{X} = a_{11} X_1 + a_{21} X_2 + \dots + a_{p1} X_p$$

$$Y_2 = \mathbf{a}_2^T \mathbf{X} = a_{12} X_1 + a_{22} X_2 + \dots + a_{p2} X_p$$

$$\vdots$$

$$Y_p = \mathbf{a}_p^T \mathbf{X} = a_{1p} X_1 + a_{2p} X_2 + \dots + a_{pp} X_p$$

Donde \mathbf{a}_i^T representa el vector propio correspondiente al valor propio λ_i de la matriz Σ . Se debe cumplir que:

$$\text{Var}(Y_i) = \mathbf{a}_i^T \mathbf{S} \mathbf{a}_i \quad \text{para } i = 1, 2, \dots, p$$

$$\text{Cov}(Y_i, Y_j) = \mathbf{a}_i^T \mathbf{S} \mathbf{a}_j \quad \text{para } i, j = 1, 2, \dots, p$$

Entonces las componentes principales son las combinaciones lineales Y_1, Y_2, \dots, Y_k donde $k \leq p$, que no están correlacionadas entre sí y que maximizan la varianza, así tenemos que:

Y_1 es la primera componente principal y es la combinación lineal $\mathbf{a}_1^T \mathbf{X}$ que máxima la varianza $\text{Var}(\mathbf{a}_1^T \mathbf{X})$, sujeta a la condición de que:

$$\|\mathbf{a}_1\| = 1$$

Y_2 es la segunda componente principal y es la combinación lineal $\mathbf{a}_2^T \mathbf{X}$ que máxima la varianza $\text{Var}(\mathbf{a}_2^T \mathbf{X})$, sujeta a las condiciones de que:

$$\|\mathbf{a}_2\| = 1$$

$$\text{Cov}(\mathbf{a}_1^T \mathbf{X}, \mathbf{a}_2^T \mathbf{X}) = 0$$

$$\langle \mathbf{a}_1, \mathbf{a}_2 \rangle = 0$$

Generalizando tenemos que:

Y_i es la i -ésima componente principal y es la combinación lineal $\mathbf{a}_i^T \mathbf{X}$ que máxima la varianza $\text{Var}(\mathbf{a}_i^T \mathbf{X})$, sujeta a las condiciones de que:

$$\|\mathbf{a}_i\| = 1$$

$$\text{Cov}(\mathbf{a}_i^T \mathbf{X}, \mathbf{a}_j^T \mathbf{X}) = 0, \text{ para } i < j, \text{ donde } i, j = 1, 2, \dots, k \leq p$$

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0, \text{ para } i \neq j$$

Las componentes principales Y_1, Y_2, \dots, Y_k , donde $k \leq p$, en general deben satisfacer lo siguiente:

- $\text{Var}(y_i) > \text{Var}(y_k)$, para todo $i < k$.
- La norma del vector propio \mathbf{a}_i de la matriz \mathbf{S} , para $i = 1, 2, \dots, p$ debe ser igual a 1, es decir $\|\mathbf{a}_i\| = 1$.

- La varianza de la i -ésima componente principal Y_i , para $i = 1, \dots, k \leq p$ es λ_i ,

$$\text{Var}(Y_i) = \mathbf{a}_i^T \hat{\mathbf{O}} \mathbf{a}_i = \lambda_i, \text{ siendo } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq 0.$$

- La covarianza de la i -ésima componente principal Y_i con la j -ésima componente principal Y_j , para $i, j = 1, \dots, k \leq p$, con $i \neq j$, es igual a cero:

$$\text{Cov}(Y_i, Y_j) = \mathbf{a}_i^T \hat{\mathbf{O}} \mathbf{a}_j = 0 \quad i \neq j.$$

- La suma de las varianzas de las variables aleatorias observables X_i es igual a la suma de las varianzas de las componentes principales (variables artificiales) Y_i , para $i = 1, 2, \dots, k \leq p$

$$\hat{\sigma}_{11} + \hat{\sigma}_{22} + \dots + \hat{\sigma}_{pp} = \sum_{i=1}^p \text{Var}(X_i) = \hat{\epsilon}_1 + \hat{\epsilon}_2 + \dots + \hat{\epsilon}_p = \sum_{i=1}^p \text{Var}(Y_i)$$

- La correlación entre la componente principal Y_i y la variable aleatoria observable X_k es igual a

$$\tilde{r}_{Y_i X_k} = \frac{a_{ik} \sqrt{\hat{\epsilon}_i}}{\sqrt{\hat{\sigma}_{kk}}} \quad i, k = 1, 2, \dots, p$$

Una vez que se han construido las componentes principales, el objetivo es obtener la mayor proporción del total de la varianza de la población, explicada en el menor número de componentes principales, así el aporte de cada componente principal está dado por:

$$\frac{I_k}{I_1 + I_2 + \dots + I_p} \quad k = 1, 2, \dots, p$$

que es la proporción del total de la varianza de la población explicada por la k-ésima componente, y λ_i es el i-ésimo valor propio de la matriz \mathbf{S} .

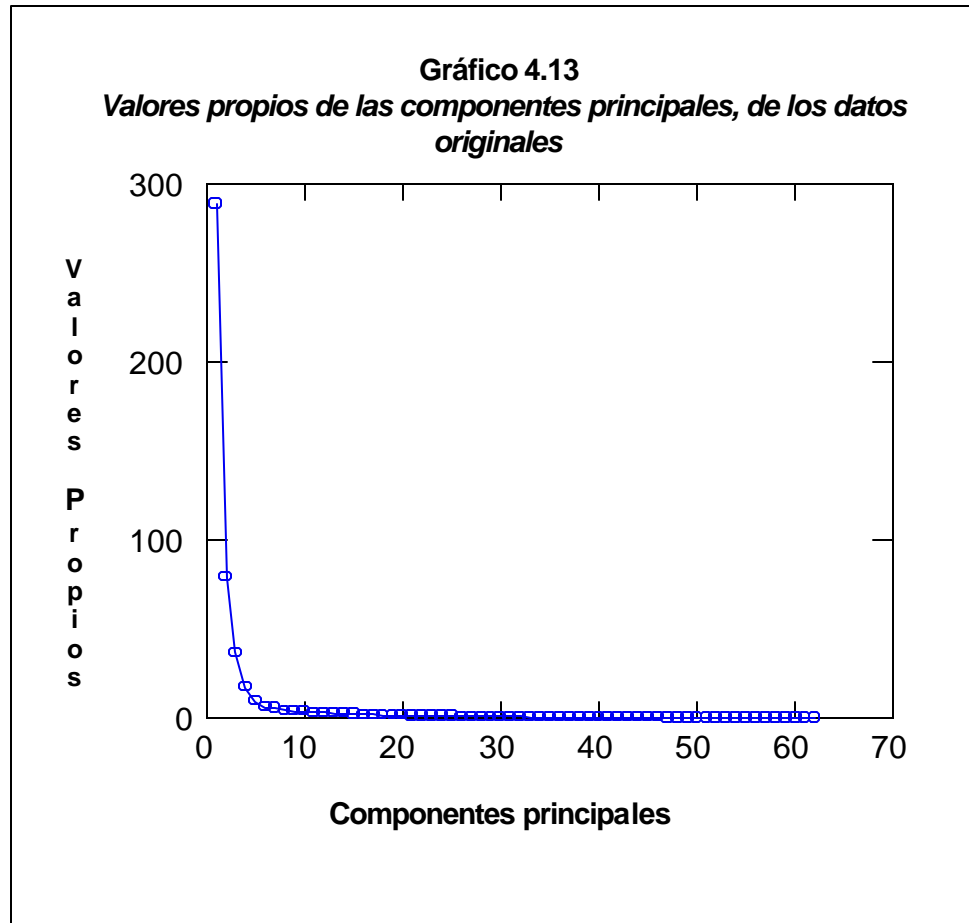
El número de componentes principales escogidas, dependerá del porcentaje de la varianza que se desea que estas expliquen, este porcentaje se determina en base a las necesidades del estudio que se realiza.

4.6.2 Construcción de componentes principales aplicadas a los datos originales

En esta sección se construirán los componentes principales utilizando las variables observables no estandarizadas, efectuando este análisis se obtuvo como resultado que el porcentaje de explicación de la varianza de la población es muy alto, en relación con el de los componentes determinados sobre las variables estandarizadas. Otro resultado es que se produce una mayor reducción de datos.

En el gráfico 4.13 se pueden observar los valores propios obtenidos de la matriz Σ , con respecto a los 62 componentes principales, en este gráfico se puede apreciar que alrededor del quinto componente principal se forma un codo, el cual indica que los valores propios de ahí en adelante son bastante pequeños (varianzas de los componentes), por lo que no

influirán en la proporción de explicación del total de la varianza de la población.



En la tabla CXLIII se muestran las varianzas individuales de los cinco primeros componentes principales, con los que al calcular el porcentaje de explicación del total de la varianzas de la población, de estos componentes se obtiene como resultado 86.121%.

Tabla CXLIII
Varianzas y porcentaje de explicación de las varianzas de las 5 primeras componentes principales de los datos originales

Componente	1	2	3	4	5
Varianza	258.201	37.204	17.263	9.515	6.210
Porcentaje	67.713	9.757	4.527	2.495	1.629

Varianza Total = 381.85

En la Tabla CXLIII se muestran los porcentajes de explicación individuales de los cinco primeros componentes principales, la elección de los componentes principales construidos a partir de las variables aleatorias no estandarizadas con su respectiva matriz de varianzas y covarianzas se basa en el mayor porcentaje de explicación de estas y la considerable reducción de datos al escoger las cinco primeras componentes.

El porcentaje de explicación de los cinco primeros componentes principales se lo obtiene utilizando la siguiente expresión:

$$\frac{\sum_{i=1}^5 \ddot{e}_i}{\sum_{i=1}^{62} \ddot{e}_i} = \frac{328.393}{381.85} = 0.8612$$

Este resultado indica que el 86.125 % del total de la varianza de la población es explicada por las cinco primeras componentes principales. Los porcentajes de explicación individuales de cada componente se muestran en la tabla CXLIII.

Los coeficientes de los cinco primeros componentes obtenidos en base a las variables aleatorias observables no estandarizadas, son mostrados en el anexo 3.

4.6.3 Construcción de componentes principales aplicadas a las variables estandarizadas

Utilizando el procedimiento descrito en la sección 4.4.1 para la determinación de los componentes principales de un vector p -variado observable, donde $p=62$, se realizó el cálculo sobre la base de las variables estandarizadas, es decir que se utilizó la matriz de correlación, en vez de la de varianzas y covarianzas.

La estandarización de las variables aleatorias se la realiza cuando, las escalas de los valores que toman las variables de estudio difieren considerablemente de variable a variable. La estandarización de las variables aleatorias observables se la realiza de la siguiente manera:

$$Z_1 = \frac{(X_1 - \bar{x}_1)}{\sqrt{\hat{\sigma}_{11}}}$$

$$Z_2 = \frac{(X_2 - \bar{x}_2)}{\sqrt{\hat{\sigma}_{22}}}$$

$$\vdots$$

$$\vdots$$

$$Z_p = \frac{(X_p - \bar{x}_p)}{\sqrt{\hat{\sigma}_{pp}}}$$

Donde X_i es la i -ésima variable aleatoria observable, \bar{x}_p es el valor esperado de X_i y $\sqrt{\hat{\sigma}_{pp}}$ representa la desviación estándar de X_i , entonces Z_i es la i -ésima variable aleatoria estandarizada, para $i = 1, 2, \dots, p$.

Una vez que se estandarizaron las variables de estudio, las cuales son representadas por el vector $\mathbf{Z}^T = (Z_1, Z_2, \dots, Z_p)$, donde $p = 62$, se calculó la matriz de varianzas y covarianzas de \mathbf{Z} que es igual a la matriz de correlación $\tilde{\mathbf{\Pi}}$, lo cual se denota como $\text{Cov}(\mathbf{Z}) = \tilde{\mathbf{\Pi}}$.

En la tabla CXLIV se muestran los 20 valores propios de la matriz de correlación $\tilde{\mathbf{\Pi}}$, que corresponden a las varianzas de las primeras veinte componentes principales. El porcentaje de explicación del total de la varianza de la población, de las veinte primeras componentes se lo obtiene mediante el cálculo de la siguiente expresión:

$$\frac{\sum_{i=1}^{20} \ddot{e}_i}{\sum_{i=1}^{62} \ddot{e}_i} = \frac{42.125}{61.903} = 0.6805$$

Este resultado indica que el 68.05% del total de la varianza de la población es explicada con los veinte primeros componentes principales.

Tabla CXLIV

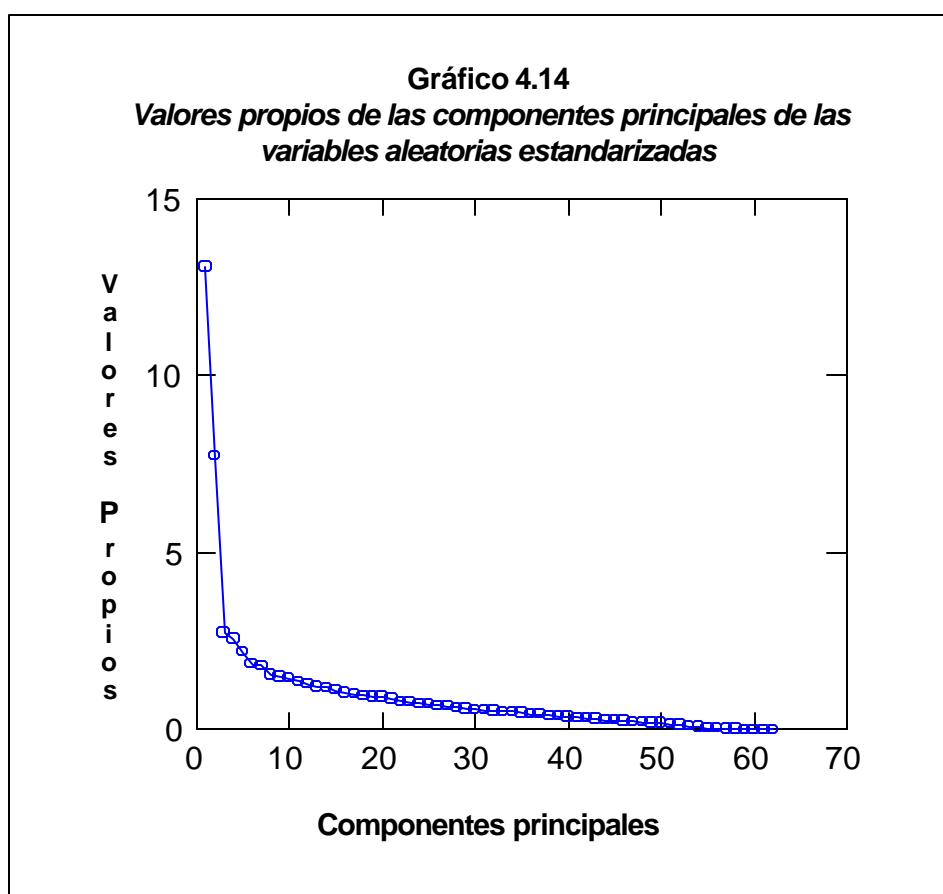
Varianzas y porcentaje de explicación de las varianzas de las 20 primeras componentes principales de las variables estandarizadas

Componente	1	2	3	4	5
Varianza	9.185	3.054	2.955	2.494	2.304
Porcentaje	14.81	4.92	4.76	4.02	3.71
Componente	6	7	8	9	10
Varianza	2.126	1.947	1.819	1.702	1.642
Porcentaje	3.42	3.14	2.93	2.74	2.64
Componente	11	12	13	14	15
Varianza	1.563	1.461	1.410	1.334	1.323
Porcentaje	2.52	2.35	2.27	2.15	2.13
Componente	16	17	18	19	20
Varianza	1.248	1.201	1.156	1.131	1.070
Porcentaje	2.01	1.96	1.90	1.84	1.75

Varianza Total = 61.903

Por otro lado, la determinación del número apropiado de componentes, por medio del gráfico de la magnitud de los valores propios versus el

número de componentes, indica que, se deben elegir solamente los diez primeros componentes, debido a que desde λ_{11} hasta λ_{62} todos los valores propios son muy pequeños y están alrededor de un mismo valor. En el gráfico 4.14 se puede observar esta situación, pues alrededor del décimo componente se forma un codo, luego del cual los valores propios son pequeños.



La elección del número de componentes principales construidos en base a las variables estandarizadas, se la realizó utilizando el criterio del

porcentaje de la varianza de la población que estos explican, de acuerdo a esto el número de componentes elegidos es veinte. Los porcentajes de explicación individuales de cada uno de los veinte componentes elegidos se muestran en la tabla CXLIV. Los coeficientes de los veinte componentes principales se muestran en el anexo 4.

Una vez que se calcularon los coeficientes de los vectores propios e_i correspondientes a los valores propios λ_i , para $i = 1, 2, \dots, k$, donde $k=20$, a continuación se tabulan las componentes principales:

$$Y_1 = a_1^T X = -0.038 X_1 - 0.147 X_2 + \dots + 0.394 X_{62}$$

$$Y_2 = a_2^T X = 0.108 X_1 + 0.028 X_2 + \dots + 0.681 X_{62}$$

$$\vdots \qquad \qquad \qquad \vdots$$

$$\vdots \qquad \qquad \qquad \vdots$$

$$Y_{20} = a_{20}^T X = -0.021 X_1 - 0.103 X_2 + \dots - 0.020 X_{62}$$

Si se estandarizan las variables aleatorias observables, se logra reducir la matriz de datos original de tamaño 167×62 a otra de 167×20 .

Aunque las dos primeras componentes principales obtenidas a partir de las variables aleatorias originales, explican un mayor porcentaje del total de la varianza de la población, que las veinte componentes obtenidas en base a las variables aleatorias estandarizadas, las componentes que se

utilizarán para remplazar a las variables originales son las veinte primeras componentes principales obtenidos a partir de las variables estandarizadas. La selección se basa en el hecho de que al estandarizar las variables, se reduce el efecto de las variables que son medidas en escalas con rangos muy diferentes o si las unidades de medida no son proporcionales.

4.6.4 Rotación de las componentes principales obtenidas en base de las variables aleatorias estandarizadas

Como se mencionó en la sección anterior las componentes principales seleccionadas fueron las que se obtuvieron a partir de las variables aleatorias observables estandarizadas, con esta selección se obtuvo una reducción de datos de una matriz original de tamaño 167 x 62 a una de tamaño 167 x 20.

El siguiente paso que se realizó en este análisis es el de rotar los componentes principales. Las combinaciones lineales geoméricamente representan la selección de un nuevo sistema de coordenadas, luego de rotar el sistema original con axisas X_1, X_2, \dots, X_{62} , las nuevas coordenadas representan las direcciones con máxima variación de las observaciones, que proveen una simple y detallada descripción de la estructura de la covarianza del vector \mathbf{X} . Para rotar los componentes se

utilizó el método varimax, mediante el cual el sistema de coordenadas obtenido es ortogonal. La rotación se la realiza con el objetivo de interpretar las combinaciones lineales (componentes principales) para obtener una representación real de lo que significan.

Tabla CXLV
Varianzas y porcentaje de explicación de las veinte primeras componentes principales obtenidas a partir de las variables estandarizadas y rotadas

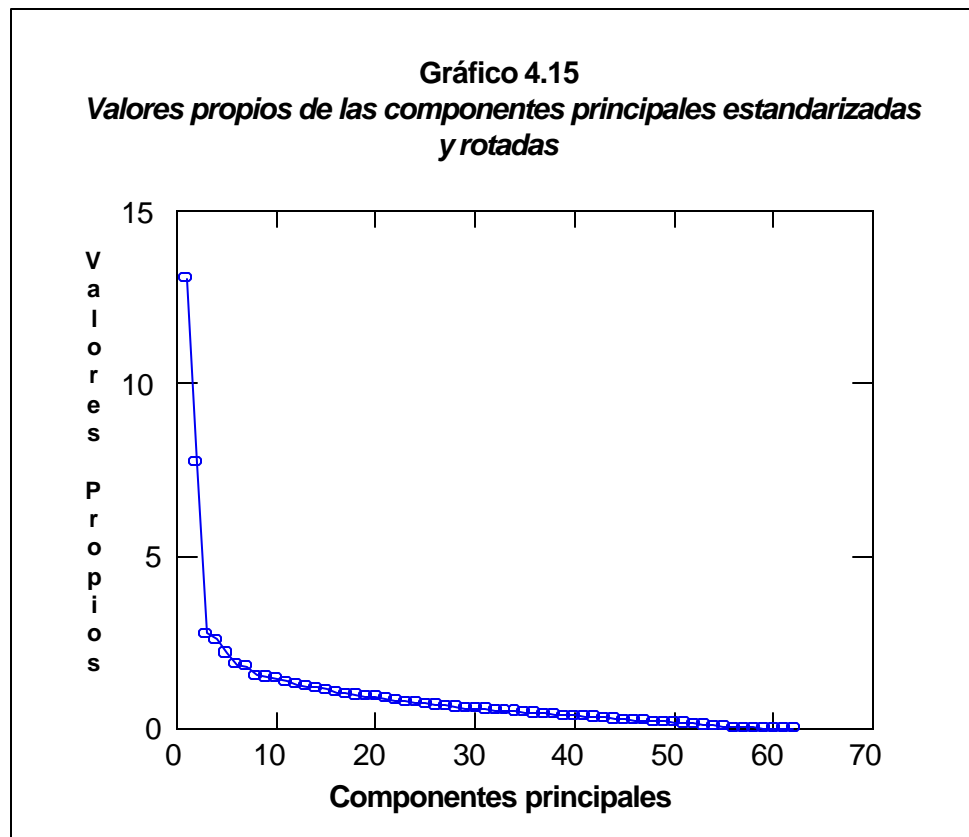
Componente	1	2	3	4	5
Varianza	11.226	2.982	1.671	3.076	2.531
Porcentaje	18.107	4.809	2.695	4.962	4.082
Componente	6	7	8	9	10
Varianza	2.570	1.349	2.488	1.878	1.523
Porcentaje	4.145	2.177	4.013	3.030	2.457
Componente	11	12	13	14	15
Varianza	1.541	1.464	1.683	2.295	1.408
Porcentaje	2.485	2.361	2.715	3.702	2.271
Componente	16	17	18	19	20
Varianza	1.934	1.415	1.653	1.293	1.382
Porcentaje	3.120	2.282	2.665	2.086	2.229

Varianza Total = 61.903

Las varianzas de los componentes obtenidos luego de la rotación de las se muestran en la tabla CXLV, al observar estos valores y al compararlos con las varianzas obtenidas sin rotar se puede afirmar que el porcentaje de explicación del total de la varianza de la población, aumentó al rotar los componentes.

El porcentaje del total de la varianza de la población explicada por los componentes principales rotados es del 76.38%. Calculado de la siguiente forma:

$$\frac{\sum_{i=1}^{20} \ddot{e}_i}{\sum_{i=1}^{62} \ddot{e}_i} = 0.7638$$



En el gráfico 4.15 se pueden observar las magnitudes de los valores propios λ_i , correspondientes a las 62 combinaciones lineales rotadas.

Para los casos de las componentes principales obtenidas a partir de las variables observables estandarizadas sin rotar y de las obtenidas después de la rotación, se puede apreciar que alrededor del valor de λ_{10} se forma un codo, el cual indica que los valores propios ubicados desde ahí en adelante son pequeños. La selección de las componentes se la realizó en base al porcentaje de explicación de la varianza, el cual es mayor al 75% si se utilizan los veinte primeros componentes principales. En la tabla CXLV se puede observar los porcentajes individuales de la explicación del total de la varianza de las veinte primeras componentes principales obtenidas a partir de las variables aleatorias estandarizadas y rotadas. Los coeficientes de las veinte combinaciones lineales obtenidas luego de estandarizar y rotar la variables aleatorias originales, se muestran en el anexo 5.

El programa que se utilizó para calcular las componentes principales es el software estadístico SYSTAT, versión 7.0.

4.6.5 Rotulación de las componentes principales aplicadas a las variables aleatorias estandarizadas y rotadas

La rotulación es el proceso que se realiza para nominar las componentes principales, de tal forma que el nombre corresponda a alguna representación real del componente. Para este efecto se considera la

magnitud de los coeficientes que tienen las variables aleatorias observables (X_1, X_2, \dots, X_{62}) en cada una de las combinaciones lineales.

Para que los coeficientes de las combinaciones lineales sean considerados como significativos estos deben ser mayores a 0.5, en este caso se utiliza la variable para rotular el componente y si son menores a 0.3 se la omite.

Rotulación de la primera componente principal

$$\begin{aligned}
 Y_1 = & 0.495Z_1 + 0.385Z_2 + 0.252Z_3 - 0.517 Z_4 + 0.101Z_5 + 0.094Z_6 + 0.051Z_7 + \\
 & 0.069Z_8 + 0.092Z_9 + 0.159Z_{10} + 0.29Z_{11} + 0.13Z_{12} + 0.08Z_{13} + 0.058Z_{14} + \\
 & 0.154Z_{15} + 0.13Z_{16} + 0.177Z_{17} + 0.052Z_{18} + 0.15Z_{19} + 0.111Z_{20} + 0.128Z_{21} \\
 & + 0.126Z_{22} - 0.008Z_{23} + 0.025Z_{24} + 0.172Z_{25} + 0.134Z_{26} + 0.07Z_{27} - \\
 & 0.538Z_{28} + 0.077Z_{29} + 0.129Z_{30} + 0.098 Z_{31} - 0.047Z_{32} - 0.411 Z_{33} + \\
 & 0.157Z_{34} - 0.96Z_{35} + 0.122Z_{36} - 0.966Z_{37} + 0.192Z_{38} - 0.97Z_{39} + 0.143Z_{40} \\
 & - 0.951Z_{41} + 0.087Z_{42} - 0.962Z_{43} + 0.167Z_{44} - 0.455Z_{45} - 0.152Z_{46} - \\
 & 0.368Z_{47} - 0.228Z_{48} - 0.183Z_{49} + 0.088Z_{50} - 0.967Z_{51} - 0.767 Z_{52} - \\
 & 0.923Z_{53} - 0.503Z_{54} - 0.969Z_{55} - 0.056Z_{56} - 0.164Z_{57} + 0.204Z_{58} - \\
 & 0.46Z_{59} - 0.752Z_{60} + 0.052Z_{61} + 0.051Z_{62}
 \end{aligned}$$

Como se puede observar los coeficientes de la primera componente principal que son mayores a 0.5 son los correspondientes a las variables aleatorias Z_4 (sección), Z_{28} (sintaxis correcta), Z_{35} (resta de fracciones), Z_{37} (multiplicación de fracciones), Z_{39} (división de fracciones), Z_{41} (valor absoluto), Z_{43} (potenciación y radicación), Z_{51} (evaluación de funciones), Z_{52} (perímetro del cuadrado), Z_{53} (área del triángulo), Z_{54} (área del círculo), Z_{55} (teorema de Pitágoras) y Z_{60} (probabilidad), por lo tanto el nombre con el que se rotularía esta componente sería el de **factor**

sección del colegio- sintaxis correcta-conocimientos básicos de matemáticas.

Rotulación de la segunda componente principal

$$\begin{aligned}
 Y_2 = & -0,103Z_1 + 0,018Z_2 + 0,122Z_3 + 0,123Z_4 + 0,106Z_5 + 0,174Z_6 + 0,051Z_7 + \\
 & 0,748Z_8 + 0,764Z_9 + 0,699Z_{10} + 0,652Z_{11} + 0,213Z_{12} + 0,212Z_{13} + \\
 & 0,283Z_{14} + 0,109Z_{15} + 0,036Z_{16} + 0,162Z_{17} + 0,052Z_{18} - 0,067Z_{19} + \\
 & 0,15Z_{20} - 0,002Z_{21} + 0,093Z_{22} + 0,144Z_{23} + 0,204Z_{24} + 0,117Z_{25} + \\
 & 0,168Z_{26} + 0,087Z_{27} - 0,084Z_{28} + 0,151Z_{29} + 0,03Z_{30} + 0,448Z_{31} + \\
 & 0,028Z_{32} + 0,075Z_{33} + 0,126Z_{34} - 0,061Z_{35} + 0,142Z_{36} - 0,1Z_{37} + 0,194Z_{38} \\
 & - 0,091Z_{39} - 0,079Z_{40} - 0,088Z_{41} + 0,08Z_{42} - 0,093Z_{43} + 0,165Z_{44} + \\
 & 0,08Z_{45} + 0,03Z_{46} + 0,093Z_{47} - 0,004Z_{48} + 0,107Z_{49} - 0,008Z_{50} - 0,095Z_{51} \\
 & - 0,023Z_{52} - 0,087Z_{53} + 0,074Z_{54} - 0,093Z_{55} + 0,03Z_{56} - 0,038Z_{57} + \\
 & 0,048Z_{58} - 0,009Z_{59} - 0,055Z_{60} + 0,008Z_{61} + 0,148Z_{62}
 \end{aligned}$$

En la segunda componente principal los coeficientes que son mayores a 0.5 son los correspondientes a las variables aleatorias Z_8 (identificación de palabras agudas), Z_9 (identificación de palabras graves), Z_{10} (identificación de palabras esdrújulas) y Z_{11} (identificación de palabras sobreesdrújulas) por lo tanto el nombre con el que se rotularía esta componente sería el de **factor acento de palabras**.

Rotulación de la tercera componente principal

$$\begin{aligned}
 Y_3 = & -0,538Z_1 + 0,031Z_2 - 0,716Z_3 - 0,23Z_4 - 0,012Z_5 - 0,001Z_6 + 0,049Z_7 - 0,05Z_8 \\
 & - 0,025Z_9 - 0,055Z_{10} + 0,123Z_{11} + 0,213Z_{12} + 0,052Z_{13} - 0,232Z_{14} + \\
 & 0,004Z_{15} + 0,247Z_{16} + 0,037Z_{17} - 0,112Z_{18} + 0,015Z_{19} - 0,136Z_{20} - 0,09Z_{21} \\
 & + 0,032Z_{22} + 0,057Z_{23} - 0,047Z_{24} - 0,026Z_{25} + 0,033Z_{26} + 0,096Z_{27} + \\
 & 0,193Z_{28} - 0,307Z_{29} - 0,003Z_{30} - 0,068Z_{31} - 0,156Z_{32} - 0,061Z_{33} + \\
 & 0,004Z_{34} + 0,052Z_{35} + 0,091Z_{36} + 0,001Z_{37} + 0,342Z_{38} + 0,028Z_{39} + \\
 & 0,023Z_{40} - 0,002Z_{41} + 0,303Z_{42} + 0,003Z_{43} + 0,171Z_{44} + 0,156Z_{45} +
 \end{aligned}$$

$$0.016Z_{46} + 0.137Z_{47} + 0.173Z_{48} + 0.214Z_{49} + 0.001Z_{50} + 0.011Z_{51} + \\ 0.064Z_{52} + 0.024Z_{53} + 0.079Z_{54} + 0.014Z_{55} - 0.122Z_{56} + 0.024Z_{57} - \\ 0.075Z_{58} + 0.057Z_{59} + 0.08Z_{60} - 0.089Z_{61} + 0.109Z_{62}$$

En la tercera componente principal los coeficientes que son mayores a 0.5 son los correspondientes a las variables aleatorias Z_1 (edad) y Z_3 (ubicación), por lo tanto el nombre con el que se rotularía esta componente sería el de **factor edad-ubicación del colegio**.

Rotulación de la cuarta componente principal

$$Y_4 = 0.06Z_1 - 0.273Z_2 + 0.157Z_3 - 0.124Z_4 + 0.073Z_5 - 0.022Z_6 + 0.012Z_7 + \\ 0.026Z_8 - 0.018Z_9 + 0.064Z_{10} + 0.043Z_{11} - 0.06Z_{12} + 0.16Z_{13} + 0.009Z_{14} - \\ 0.048Z_{15} + 0.025Z_{16} + 0.197Z_{17} + 0.086Z_{18} + 0.155Z_{19} - 0.028Z_{20} + \\ 0.148Z_{21} + 0.113Z_{22} + 0.001Z_{23} + 0.211Z_{24} + 0.007Z_{25} + 0.021Z_{26} + \\ 0.121Z_{27} - 0.04Z_{28} + 0.24Z_{29} + 0.041Z_{30} + 0.121Z_{31} + 0.061Z_{32} + 0.124 \\ Z_{33} + 0.036Z_{34} + 0.03Z_{35} + 0.045Z_{36} - 0.049Z_{37} + 0.215Z_{38} - 0.023Z_{39} + \\ 0.002Z_{40} - 0.01Z_{41} - 0.077Z_{42} - 0.024Z_{43} + 0.161Z_{44} + 0.272Z_{45} + \\ 0.095Z_{46} + 0.166Z_{47} + 0.058Z_{48} + 0.256Z_{49} + 0.158Z_{50} - 0.037Z_{51} + \\ 0.082Z_{52} - 0.009Z_{53} + 0.338Z_{54} - 0.04Z_{55} + 0.783Z_{56} + 0.786Z_{57} + \\ 0.773Z_{58} + 0.211Z_{59} + 0.028Z_{60} + 0.059Z_{61} + 0.601Z_{62}$$

Los coeficientes que son mayores a 0.5 en la cuarta componente principal son los correspondientes a las variables aleatorias Z_{56} (aplicación de dos casos de factorización), Z_{57} (factorización de un trinomio), Z_{58} (solución de ecuación de una incógnita) y Z_{62} (calificación de matemáticas) por lo tanto el nombre con el que se rotularía esta componente sería el de **factor factorización-ecuaciones-calificación de matemáticas**.

Rotulación de la quinta componente principal

$$\begin{aligned}
 Y_5 = & 0.008Z_1 - 0,147Z_2 + 0.075Z_3 + 0,194Z_4 + 0,136Z_5 + 0,152Z_6 + 0,195Z_7 + \\
 & 0,03Z_8 - 0,013Z_9 + 0,074Z_{10} + 0,093Z_{11} + 0,19Z_{12} + 0,17Z_{13} + 0,104Z_{14} + \\
 & 0.032Z_{15} + 0,197Z_{16} - 0,054Z_{17} + 0,029Z_{18} - 0,03Z_{19} + 0,503Z_{20} + \\
 & 0.754Z_{21} + 0.788Z_{22} + 0.25Z_{23} + 0.146Z_{24} + 0.187Z_{25} + 0.031Z_{26} + \\
 & 0.149Z_{27} - 0.044Z_{28} + 0.19Z_{29} + 0.325Z_{30} + 0.379 Z_{31} + 0.101Z_{32} + 0.011 \\
 & Z_{33} + 0.144Z_{34} - 0.07Z_{35} + 0.102Z_{36} - 0.064Z_{37} - 0.019Z_{38} - 0.062Z_{39} + \\
 & 0.066Z_{40} - 0.074Z_{41} - 0.044Z_{42} - 0.042Z_{43} + 0.311Z_{44} - 0.125Z_{45} - \\
 & 0.109Z_{46} - 0.089Z_{47} + 0.117Z_{48} + 0.089Z_{49} + 0.152Z_{50} - 0.069Z_{51} + 0.07 \\
 & Z_{52} - 0.054Z_{53} - 0.187Z_{54} - 0.07Z_{55} - 0.015Z_{56} + 0.132Z_{57} + 0.093Z_{58} - \\
 & 0.015Z_{59} + 0,023Z_{60} + 0,101Z_{61} + 0,212Z_{62}
 \end{aligned}$$

El nombre con el que se rotularía esta componente sería el de **factor conjugación de verbos-identificación de las partes de la oración**, debido a que los coeficientes que son mayores a 0.5 en esta componente principal son los correspondientes a las variables aleatorias Z_{20} (conjugación de verbos), Z_{21} (identificación del sujeto) y Z_{22} (identificación del predicado).

Rotulación de la sexta componente principal

$$\begin{aligned}
 Y_6 = & 0.052Z_1 + 0,102Z_2 - 0.209Z_3 + 0,013Z_4 + 0,143Z_5 + 0,064Z_6 - 0,057Z_7 - \\
 & 0,045Z_8 + 0,063Z_9 + 0,082Z_{10} + 0,1Z_{11} + 0,078Z_{12} + 0,166Z_{13} + 0,275Z_{14} \\
 & + 0.766Z_{15} + 0,733Z_{16} + 0,295Z_{17} + 0,539Z_{18} + 0,476Z_{19} + 0,196Z_{20} + \\
 & 0.111Z_{21} + 0.017Z_{22} - 0.038Z_{23} + 0.092Z_{24} + 0.096Z_{25} + 0.056Z_{26} + \\
 & 0.191Z_{27} + 0.02Z_{28} + 0.155Z_{29} + 0.105Z_{30} + 0.297 Z_{31} + 0.069Z_{32} + 0.432 \\
 & Z_{33} + 0.065Z_{34} - 0.019Z_{35} + 0.274Z_{36} - 0.083Z_{37} + 0.068Z_{38} - 0.069Z_{39} + \\
 & 0.102Z_{40} - 0.052Z_{41} - 0.098Z_{42} - 0.076Z_{43} + 0.024Z_{44} + 0.033Z_{45} - \\
 & 0.001Z_{46} + 0.067Z_{47} + 0.038Z_{48} + 0.007Z_{49} + 0.057Z_{50} - 0.079Z_{51} + 0.03
 \end{aligned}$$

$$Z_{52} - 0.038Z_{53} - 0.077Z_{54} - 0.088Z_{55} - 0.074Z_{56} + 0.088Z_{57} + 0.055Z_{58} - 0.023Z_{59} + 0,072Z_{60} + 0,097Z_{61} + 0,103Z_{62}$$

El nombre con el que se rotularía esta componente sería el de **factor clasificación de palabras**, debido a que los coeficientes que son mayores a 0.5 en esta componente principal son los correspondientes a las variables aleatorias Z_{15} (identificación de palabras aumentativas), Z_{16} (identificación de palabras diminutivas) y Z_{18} (identificación de simples nombres).

Rotulación de la séptima componente principal

$$Y_7 = -0.232Z_1 + 0,125Z_2 - 0.026Z_3 - 0,15Z_4 + 0,053Z_5 + 0,097Z_6 + 0,713Z_7 + 0,134Z_8 + 0,013Z_9 - 0,094Z_{10} - 0,02Z_{11} + 0,015Z_{12} - 0,073Z_{13} + 0,179Z_{14} + 0,071Z_{15} - 0,207Z_{16} - 0,096Z_{17} - 0,002Z_{18} + 0,005Z_{19} + 0,006Z_{20} + 0,079Z_{21} + 0,085Z_{22} - 0,037Z_{23} + 0,064Z_{24} - 0,007Z_{25} + 0,034Z_{26} + 0,135Z_{27} + 0,157Z_{28} + 0,173Z_{29} - 0,109Z_{30} + 0,114 Z_{31} - 0,253Z_{32} + 0,122 Z_{33} + 0,106Z_{34} + 0,017Z_{35} - 0,168Z_{36} - 0,05Z_{37} - 0,125Z_{38} - 0,037Z_{39} + 0,095Z_{40} + 0,025Z_{41} - 0,031Z_{42} - 0,05Z_{43} - 0,072Z_{44} + 0,052Z_{45} - 0,087Z_{46} - 0,042Z_{47} - 0,022Z_{48} - 0,093Z_{49} - 0,113Z_{50} - 0,039Z_{51} + 0,06 Z_{52} - 0,038Z_{53} + 0,042Z_{54} - 0,029Z_{55} + 0,071Z_{56} - 0,039Z_{57} + 0,04Z_{58} + 0,504Z_{59} + 0,173Z_{60} + 0,051Z_{61} - 0,04Z_{62}$$

Las variables aleatorias que en esta componente principal tienen coeficientes mayores a 0.5 y por lo tanto son escogidas para rotular la componentes son Z_7 (identificación de hiatos) y Z_{69} (solución de ecuación con dos incógnitas), por lo tanto la componente se rotularía como **factor identificación de hiatos- solución de ecuación**.

Rotulación de la octava componente principal

$$\begin{aligned}
 Y_8 = & 0.037Z_1 - 0.149Z_2 + 0.02Z_3 - 0.174Z_4 + 0.838Z_5 + 0.858Z_6 + 0.288Z_7 + \\
 & 0.079Z_8 + 0.076Z_9 + 0.032Z_{10} + 0.134Z_{11} + 0.175Z_{12} + 0.014Z_{13} + \\
 & 0.097Z_{14} + 0.074Z_{15} + 0.111Z_{16} + 0.245Z_{17} + 0.098Z_{18} + 0.27Z_{19} + \\
 & 0.2333Z_{20} + 0.165Z_{21} + 0.098Z_{22} - 0.004Z_{23} + 0.123Z_{24} + 0.114Z_{25} + \\
 & 0.07Z_{26} + 0.152Z_{27} - 0.033Z_{28} + 0.097Z_{29} + 0.078Z_{30} + 0.348 Z_{31} + \\
 & 0.125Z_{32} - 0.224 Z_{33} + 0.086Z_{34} - 0.107Z_{35} + 0.055Z_{36} - 0.032Z_{37} + \\
 & 0.147Z_{38} - 0.041Z_{39} + 0.104Z_{40} - 0.077Z_{41} + 0.001Z_{42} - 0.046Z_{43} + \\
 & 0.012Z_{44} - 0.149Z_{45} + 0.082Z_{46} + 0.007Z_{47} + 0.165Z_{48} + 0.072Z_{49} + \\
 & 0.301Z_{50} - 0.042Z_{51} + 0.071 Z_{52} + 0.035Z_{53} + 0.118Z_{54} - 0.028Z_{55} + \\
 & 0.057Z_{56} - 0.048Z_{57} + 0.008Z_{58} - 0.165Z_{59} - 0.082Z_{60} + 0.163Z_{61} - \\
 & 0.155Z_{62}
 \end{aligned}$$

Las variables aleatorias Z_5 (identificación diptongos) y Z_6 (identificación triptongos) tienen coeficientes mayores a 0.5 y por lo tanto son escogidas para rotular la componente, en base a esto la componente se rotularía como **factor identificación de diptongos y triptongos**.

Rotulación de la novena componente principal

$$\begin{aligned}
 Y_9 = & 0.135Z_1 - 0.162Z_2 - 0.129Z_3 - 0.268Z_4 + 0.156Z_5 + 0.056Z_6 - 0.073Z_7 + \\
 & 0.023Z_8 + 0.096Z_9 + 0.038Z_{10} - 0.05Z_{11} + 0.084Z_{12} + 0.037Z_{13} + 0.074Z_{14} \\
 & + 0.115Z_{15} - 0.051Z_{16} + 0.09Z_{17} + 0.106Z_{18} + 0.01Z_{19} - 0.001Z_{20} + \\
 & 0.082Z_{21} + 0.048Z_{22} + 0.007Z_{23} + 0.008Z_{24} + 0.029Z_{25} + 0.184Z_{26} - \\
 & 0.069Z_{27} - 0.012Z_{28} - 0.054Z_{29} + 0.156Z_{30} + 0.084 Z_{31} + 0.085Z_{32} + 0.061 \\
 & Z_{33} + 0.137Z_{34} - 0.046Z_{35} + 0.513Z_{36} - 0.064Z_{37} + 0.187Z_{38} - 0.076Z_{39} + \\
 & 0.061Z_{40} + 0.009Z_{41} + 0.591Z_{42} - 0.084Z_{43} + 0.111Z_{44} - 0.062Z_{45} + \\
 & 0.03Z_{46} + 0.1Z_{47} + 0.003Z_{48} + 0.024Z_{49} + 0.122Z_{50} - 0.075Z_{51} + 0.097 \\
 & Z_{52} - 0.028Z_{53} + 0.075Z_{54} - 0.038Z_{55} + 0.157Z_{56} - 0.107Z_{57} - 0.077Z_{58} + \\
 & 0.175Z_{59} + 0.068Z_{60} + 0.756Z_{61} + 0.39Z_{62}
 \end{aligned}$$

En la novena componente principal las variables aleatorias Z_{36} (multiplicación de enteros), Z_{42} (relación de orden) y Z_{61} (ejercicio de estadística) tienen coeficientes mayores a 0.5 y por lo tanto son utilizadas para rotular la componente, el nombre es **factor conocimientos matemáticos**.

Rotulación de la décima componente principal

$$Y_{10} = 0.27Z_1 + 0.15Z_2 - 0.031Z_3 - 0.112Z_4 + 0.034Z_5 + 0.095Z_6 + 0.116Z_7 + 0.136Z_8 - 0.077Z_9 + 0.046Z_{10} + 0.062Z_{11} + 0.005Z_{12} + 0.019Z_{13} - 0.017Z_{14} + 0.053Z_{15} + 0.09Z_{16} - 0.14Z_{17} - 0.155Z_{18} + 0.221Z_{19} - 0.002Z_{20} + 0.052Z_{21} + 0.146Z_{22} + 0.038Z_{23} - 0.086Z_{24} - 0.061Z_{25} + 0.104Z_{26} + 0.21Z_{27} + 0.013Z_{28} - 0.051Z_{29} - 0.155Z_{30} + 0.055Z_{31} + 0.526Z_{32} - 0.049Z_{33} + 0.731Z_{34} - 0.025Z_{35} + 0.142Z_{36} + 0.033Z_{37} + 0.032Z_{38} + 0.012Z_{39} + 0.084Z_{40} - 0.053Z_{41} - 0.092Z_{42} + 0.017Z_{43} + 0.169Z_{44} - 0.038Z_{45} + 0.056Z_{46} + 0.17Z_{47} + 0.132Z_{48} + 0.036Z_{49} - 0.017Z_{50} + 0.015Z_{51} - 0.089Z_{52} - 0.013Z_{53} - 0.16Z_{54} + 0.032Z_{55} + 0.052Z_{56} - 0.167Z_{57} + 0.143Z_{58} - 0.173Z_{59} - 0.101Z_{60} + 0.14Z_{61} + 0.231Z_{62}$$

Como se puede observar en la tabulación de la décima componente principal las variables aleatorias Z_{32} (suma de enteros) y Z_{34} (resta de enteros) tienen coeficientes mayores a 0.5 y por lo tanto son consideradas para rotular la componente, el nombre es **factor suma y resta de números enteros**.

Rotulación de la décima primera componente principal

$$Y_{11} = 0.032Z_1 + 0.566Z_2 - 0.04Z_3 + 0.051Z_4 + 0.072Z_5 + 0.092Z_6 - 0.149Z_7 - 0.094Z_8 - 0.05Z_9 + 0.094Z_{10} + 0.128Z_{11} - 0.107Z_{12} - 0.058Z_{13} + 0.116Z_{14}$$

$$\begin{aligned}
& + 0.031Z_{15} + 0.08Z_{16} + 0.091Z_{17} + 0.118Z_{18} - 0.054Z_{19} - 0.009Z_{20} - \\
& 0.066Z_{21} + 0.107Z_{22} + 0.066Z_{23} - 0.037Z_{24} + 0.062Z_{25} + 0.06Z_{26} - \\
& 0.019Z_{27} + 0.005Z_{28} - 0.016Z_{29} + 0.122Z_{30} + 0.028Z_{31} + 0.046Z_{32} - 0.097 \\
& Z_{33} + 0.047Z_{34} - 0.049Z_{35} - 0.087Z_{36} - 0.061Z_{37} + 0.216Z_{38} - 0.058Z_{39} + \\
& 0.013Z_{40} - 0.06Z_{41} + 0.246Z_{42} - 0.051Z_{43} + 0.04Z_{44} + 0.028Z_{45} - \\
& 0.032Z_{46} + 0.384Z_{47} + 0.004Z_{48} + 0.193Z_{49} + 0.697Z_{50} - 0.061Z_{51} + \\
& 0.085Z_{52} - 0.02Z_{53} + 0.131Z_{54} - 0.043Z_{55} - 0.175Z_{56} + 0.029Z_{57} + \\
& 0.166Z_{58} + 0.262Z_{59} + 0.063Z_{60} - 0.028Z_{61} + 0.235Z_{62}
\end{aligned}$$

Como se puede observar en la tabulación de la décima primera componente principal las variables aleatorias Z_2 (sexo) y Z_{50} (ejercicios de lógica) tienen coeficientes mayores a 0.5 y por lo tanto son consideradas para rotular la componente, el nombre es **factor sexo-lógica matemática**.

Rotulación de la décima segunda componente principal

$$\begin{aligned}
Y_{12} = & 0.011Z_1 - 0.034Z_2 - 0.141Z_3 - 0.187Z_4 + 0.084Z_5 + 0.015Z_6 - 0.08Z_7 + \\
& 0.016Z_8 + 0.083Z_9 - 0.05Z_{10} + 0.036Z_{11} + 0.239Z_{12} - 0.056Z_{13} + 0.005Z_{14} \\
& + 0.023Z_{15} - 0.273Z_{16} + 0.122Z_{17} - 0.028Z_{18} + 0.103Z_{19} + 0.031Z_{20} + \\
& 0.124Z_{21} - 0.07Z_{22} + 0.032Z_{23} + 0.008Z_{24} - 0.116Z_{25} + 0.195Z_{26} + \\
& 0.415Z_{27} + 0.185Z_{28} + 0.138Z_{29} - 0.034Z_{30} - 0.086Z_{31} + 0.032Z_{32} - 0.005 \\
& Z_{33} + 0.215Z_{34} + 0.025Z_{35} - 0.252Z_{36} + 0.017Z_{37} + 0.05Z_{38} + 0.027Z_{39} - \\
& 0.094Z_{40} - 0.06Z_{41} + 0.246Z_{42} + 0.194Z_{44} - 0.454Z_{45} - 0.089Z_{46} + \\
& 0.106Z_{47} + 0.663Z_{49} + 0.134Z_{50} + 0.021Z_{51} - 0.107Z_{52} + 0.02Z_{53} - \\
& 0.029Z_{54} + 0.011Z_{55} + 0.024Z_{56} + 0.089Z_{57} - 0.008Z_{58} + 0.034Z_{59} - \\
& 0.001Z_{60} + 0.004Z_{61} + 0.186Z_{62}
\end{aligned}$$

En la décima segunda componente principal solamente la variable aleatoria Z_{49} (propiedades de conjuntos) es mayor a 0.5, por lo tanto la rotulación para esta componente es **factor propiedades de conjuntos**.

Rotulación de la décima tercera componente principal

$$\begin{aligned}
 Y_{13} = & -0.128Z_1 - 0,102Z_2 + 0.012Z_3 + 0,016Z_4 + 0,062Z_5 + 0,045Z_6 - 0,173Z_7 + \\
 & 0,065Z_8 + 0,097Z_9 + 0,021Z_{10} - 0,122Z_{11} + 0,144Z_{12} - 0,005Z_{13} + 0,173Z_{14} \\
 & - 0,071Z_{15} - 0,026Z_{16} + 0,068Z_{17} + 0,099Z_{18} + 0,27Z_{19} + 0,327Z_{20} - \\
 & 0,126Z_{21} - 0,098Z_{22} + 0,096Z_{23} - 0,01Z_{24} - 0,023Z_{25} + 0,113Z_{26} + 0,057Z_{27} \\
 & + 0,294Z_{28} + 0,158Z_{29} + 0,083Z_{30} + 0,135Z_{31} - 0,095Z_{32} + 0,222Z_{33} + \\
 & 0,137Z_{34} + 0,055Z_{35} - 0,087Z_{36} - 0,022Z_{37} + 0,258Z_{38} - 0,012Z_{39} + \\
 & 0,106Z_{40} - 0,005Z_{41} + 0,144Z_{42} + 0,01Z_{43} + 0,368Z_{44} + 0,19Z_{45} + \\
 & 0,784Z_{46} + 0,343Z_{47} + 0,093Z_{48} - 0,073Z_{49} - 0,014Z_{50} - 0,012Z_{51} + 0,012 \\
 & Z_{52} - 0,002Z_{53} - 0,005Z_{54} - 0,007Z_{55} + 0,04Z_{56} + 0,05Z_{57} + 0,01Z_{58} + \\
 & 0,062Z_{59} + 0,023Z_{60} + 0,006Z_{61} + 0,197Z_{62}
 \end{aligned}$$

En la décima tercera componente principal solamente la variable aleatoria Z_{46} (ejercicio de proporcionalidad) es mayor a 0.5, por lo tanto la rotulación para esta componente es **factor ejercicio de proporcionalidad**.

Rotulación de la décima cuarta componente principal

$$\begin{aligned}
 Y_{14} = & 0.118Z_1 + 0,105Z_2 + 0.016Z_3 - 0,005Z_4 + 0,064Z_5 + 0,131Z_6 + 0,11Z_7 + \\
 & 0,161Z_8 + 0,166Z_9 - 0,075Z_{10} - 0,091Z_{11} - 0,033Z_{12} + 0,237Z_{13} + 0,045Z_{14} \\
 & + 0,105Z_{15} + 0,085Z_{16} - 0,1Z_{17} - 0,023Z_{18} + 0,17Z_{19} + 0,267Z_{20} + 0,078Z_{21} \\
 & + 0,097Z_{22} + 0,127Z_{23} + 0,019Z_{24} + 0,824Z_{25} + 0,772Z_{26} + 0,016Z_{27} + \\
 & 0,052Z_{28} + 0,414Z_{29} + 0,124Z_{30} + 0,353Z_{31} + 0,247Z_{32} - 0,024Z_{33} - \\
 & 0,053Z_{34} - 0,045Z_{35} + 0,095Z_{36} - 0,032Z_{37} + 0,128Z_{38} - 0,042Z_{39} + \\
 & 0,049Z_{40} - 0,02Z_{41} + 0,159Z_{42} - 0,019Z_{43} + 0,172Z_{44} + 0,198Z_{45} + \\
 & 0,059Z_{46} - 0,001Z_{47} + 0,072Z_{48} + 0,114Z_{49} + 0,076Z_{50} - 0,039Z_{51} - \\
 & 0,072Z_{52} - 0,056Z_{53} + 0,003Z_{54} - 0,021Z_{55} - 0,071Z_{56} + 0,018Z_{57} + \\
 & 0,073Z_{58} - 0,098Z_{59} - 0,162Z_{60} + 0,053Z_{61} + 0,215Z_{62}
 \end{aligned}$$

En esta componente principal las variables Z_{25} (identificación de frases) y Z_{26} (identificación de oraciones) tienen coeficientes mayores a 0.5, por lo que su aportación en esta componente es significativa, por lo tanto la rotulación sería **factor identificación de frases y oraciones**.

Rotulación de la décima quinta componente principal

$$Y_{15} = 0.027Z_1 - 0.057Z_2 + 0.043Z_3 + 0.057Z_4 - 0.05Z_5 + 0.053Z_6 + 0.088Z_7 - 0.169Z_8 + 0.134Z_9 - 0.107Z_{10} + 0.012Z_{11} - 0.06Z_{12} + 0.008Z_{13} - 0.109Z_{14} + 0.038Z_{15} + 0.075Z_{16} - 0.264Z_{17} - 0.049Z_{18} - 0.242Z_{19} + 0.016Z_{20} - 0.099Z_{21} - 0.123Z_{22} - 0.775Z_{23} - 0.166Z_{24} + 0.067Z_{25} - 0.188Z_{26} - 0.161Z_{27} + 0.069Z_{28} + 0.065Z_{29} + 0.035Z_{30} - 0.137Z_{31} + 0.188Z_{32} + 0.079Z_{33} - 0.063Z_{34} - 0.007Z_{35} - 0.202Z_{36} + 0.043Z_{37} + 0.156Z_{38} + 0.04Z_{39} + 0.016Z_{40} + 0.015Z_{41} + 0.011Z_{42} + 0.0291Z_{43} + 0.408Z_{44} - 0.015Z_{45} - 0.097Z_{46} + 0.141Z_{47} + 0.024Z_{48} - 0.138Z_{49} - 0.062Z_{50} + 0.035Z_{51} - 0.124Z_{52} - 0.033Z_{53} - 0.087Z_{54} + 0.035Z_{55} + 0.016Z_{56} + 0.134Z_{57} - 0.147Z_{58} - 0.138Z_{59} - 0.037Z_{60} + 0.055Z_{61} + 0.042Z_{62}$$

En la décima quinta componente principal solamente la variable aleatoria Z_{23} (identificación del sustantivo) es mayor a 0.5, por lo tanto la rotulación para esta componente es **factor identificación de sustantivos**.

Rotulación de la décima sexta componente principal

$$Y_{16} = 0.121Z_1 + 0.017Z_2 - 0.057Z_3 - 0.245Z_4 + 0.063Z_5 + 0.085Z_6 + 0.054Z_7 - 0.164Z_8 + 0.113Z_9 + 0.051Z_{10} + 0.087Z_{11} + 0.487Z_{12} + 0.713Z_{13} + 0.618Z_{14} + 0.132Z_{15} + 0.14Z_{16} + 0.026Z_{17} + 0.261Z_{18} + 0.071Z_{19} + 0.162Z_{20} + 0.04Z_{21} + 0.163Z_{22} + 0.12Z_{23} + 0.014Z_{24} + 0.045Z_{25} - 0.143Z_{26} + 0.055Z_{27} - 0.138Z_{28} - 0.002Z_{29} + 0.201Z_{30} + 0.312Z_{31} + 0.107Z_{32} - 0.321Z_{33} - 0.033Z_{34} - 0.04Z_{35} - 0.137Z_{36} + 0.012Z_{37} - 0.023Z_{38} - 0.003Z_{39}$$

$$\begin{aligned}
& - 0.018Z_{40} - 0.025Z_{41} - 0.017Z_{42} + 0.009Z_{43} + 0.163Z_{44} + 0.078Z_{45} + \\
& 0.087Z_{46} + 0.186Z_{47} - 0.007Z_{48} + 0.092Z_{49} - 0.057Z_{50} - 0.003Z_{51} - \\
& 0.099Z_{52} + 0.009Z_{53} - 0.027Z_{54} + 0.017Z_{55} + 0.138Z_{56} - 0.023Z_{57} - \\
& 0.032Z_{58} - 0.008Z_{59} - 0,126Z_{60} + 0,145Z_{61} + 0,097Z_{62}
\end{aligned}$$

Las variables aleatorias cuyos coeficientes son mayores a 0.5 en esta componente principal son Z_{13} (identificación de sinónimos) y Z_{14} (identificación de antónimos), por lo tanto la representación real de este componente se entendería si se rotula como **factor identificación de sinónimos y antónimos**.

Rotulación de la décima séptima componente principal

$$\begin{aligned}
Y_{17} = & 0.01Z_1 + 0,234Z_2 - 0.212Z_3 + 0,12Z_4 + 0,108Z_5 + 0,063Z_6 - 0,022Z_7 + \\
& 0,084Z_8 + 0,037Z_9 + 0,002Z_{10} - 0,15Z_{11} - 0,184Z_{12} + 0,007Z_{13} + 0,095Z_{14} \\
& - 0,011Z_{15} + 0,029Z_{16} + 0,295Z_{17} + 0,322Z_{18} - 0,267Z_{19} - 0,283Z_{20} + \\
& 0,135Z_{21} + 0,05Z_{22} - 0,006Z_{23} - 0,096Z_{24} + 0,008Z_{25} + 0,07Z_{26} + 0,007Z_{27} \\
& - 0,073Z_{28} - 0,122Z_{29} + 0,024Z_{30} - 0,004Z_{31} - 0,237Z_{32} - 0,021Z_{33} + 0,178 \\
& Z_{34} + 0,049Z_{35} - 0,035Z_{36} + 0,045Z_{37} + 0,247Z_{38} + 0,067Z_{39} - 0,035Z_{40} + \\
& 0,032Z_{41} + 0,082Z_{42} + 0,043Z_{43} + 0,057Z_{44} - 0,011Z_{45} + 0,103Z_{46} - \\
& 0,031Z_{47} + 0,747Z_{48} + 0,007Z_{49} - 0,056Z_{50} + 0,043Z_{51} - 0,025Z_{52} + \\
& 0,059Z_{53} - 0,022Z_{54} + 0,041Z_{55} + 0,039Z_{56} + 0,037Z_{57} - 0,033Z_{58} - \\
& 0,04Z_{59} - 0,125Z_{60} - 0,018Z_{61} + 0,156Z_{62}
\end{aligned}$$

En esta componente principal solamente la variable Z_{48} (sistema métrico) tiene un coeficiente mayor a 0.5, por lo tanto la representación real de este componente se entendería si se rotula como **factor sistema métrico**.

Rotulación de la décima octava componente principal

$$\begin{aligned}
 Y_{18} = & 0.014Z_1 - 0.06Z_2 - 0.12Z_3 + 0.112Z_4 + 0.066Z_5 + 0.095Z_6 + 0.051Z_7 + \\
 & 0.003Z_8 + 0.145Z_9 + 0.069Z_{10} - 0.123Z_{11} - 0.021Z_{12} + 0.117Z_{13} + \\
 & 0.064Z_{14} + 0.081Z_{15} + 0.079Z_{16} + 0.172Z_{17} + 0.194Z_{18} + 0.014Z_{19} - \\
 & 0.114Z_{20} + 0.189Z_{21} + 0.125Z_{22} + 0.089Z_{23} + 0.108Z_{24} + 0.074Z_{25} - \\
 & 0.004Z_{26} + 0.7Z_{27} + 0.431Z_{28} + 0.125Z_{29} + 0.059Z_{30} + 0.267Z_{31} + 0.293Z_{32} \\
 & - 0.025Z_{33} + 0.011Z_{34} - 0.026Z_{35} + 0.125Z_{36} - 0.036Z_{37} + 0.023Z_{38} - \\
 & 0.035Z_{39} + 0.008Z_{40} - 0.033Z_{41} + 0.147Z_{42} - 0.064Z_{43} + 0.022Z_{44} + \\
 & 0.097Z_{45} + 0.118Z_{46} + 0.018Z_{47} - 0.029Z_{48} + 0.041Z_{49} + 0.075Z_{50} - \\
 & 0.044Z_{51} + 0.013Z_{52} + 0.039Z_{53} + 0.188Z_{54} - 0.006Z_{55} - 0.003Z_{56} + \\
 & 0.018Z_{57} + 0.056Z_{58} + 0.143Z_{59} - 0.107Z_{60} - 0.085Z_{61} + 0.139Z_{62}
 \end{aligned}$$

Las variables cuyos coeficientes son mayores a 0.5 en esta componente son Z_{27} (clasificación de oraciones) y Z_{30} (lectura comprensiva), en base a esto la rotulación de esta componente es **factor clasificación de oraciones-lectura comprensiva**.

Rotulación de la décima novena componente principal

$$\begin{aligned}
 Y_{19} = & 0.055Z_1 - 0.013Z_2 + 0.035Z_3 - 0.2942Z_4 + 0.077Z_5 + 0.017Z_6 + 0.084Z_7 - \\
 & 0.111Z_8 + 0.125Z_9 + 0.359Z_{10} - 0.123Z_{11} + 0.15Z_{12} - 0.066Z_{13} + 0.034Z_{14} \\
 & + 0.171Z_{15} + 0.005Z_{16} + 0.174Z_{17} - 0.059Z_{18} - 0.135Z_{19} + 0.002Z_{20} + \\
 & 0.141Z_{21} - 0.011Z_{22} + 0.165Z_{23} + 0.679Z_{24} + 0.15Z_{25} - 0.036Z_{26} + \\
 & 0.005Z_{27} + 0.025Z_{28} + 0.142Z_{29} + 0.129Z_{30} + 0.191Z_{31} - 0.046Z_{32} - \\
 & 0.112Z_{33} - 0.034Z_{34} - 0.024Z_{35} - 0.211Z_{36} + 0.021Z_{37} - 0.044Z_{38} + \\
 & 0.011Z_{39} - 0.021Z_{40} - 0.005Z_{41} + 0.12Z_{42} - 0.01Z_{43} + 0.111Z_{44} + \\
 & 0.027Z_{45} - 0.02Z_{46} + 0.162Z_{47} - 0.088Z_{48} - 0.101Z_{49} - 0.054Z_{50} + 0.01Z_{51} - \\
 & 0.14Z_{52} - 0.051Z_{53} - 0.327Z_{54} + 0.02Z_{55} + 0.077Z_{56} + 0.128Z_{57} - \\
 & 0.033Z_{58} - 0.081Z_{59} + 0.057Z_{60} + 0.008Z_{61} + 0.02Z_{62}
 \end{aligned}$$

En esta componente principal solamente la variable Z_{24} (identificación del verbo en la oración) tiene un coeficiente mayor a 0.5, lo que implica que la aportación de esta variable a la componente indica la representación real de la misma, entonces se la rotula como **factor identificación del verbo de la oración**.

Rotulación de la vigésima componente principal

$$\begin{aligned}
 Y_{20} = & 0.041Z_1 - 0.098Z_2 - 0.038Z_3 + 0.067Z_4 + 0.079Z_5 + 0.065Z_6 + 0.091Z_7 + \\
 & 0.031Z_8 - 0.135Z_9 + 0.002Z_{10} + 0.043Z_{11} + 0.028Z_{12} - 0.081Z_{13} + \\
 & 0.061Z_{14} + 0.102Z_{15} + 0.017Z_{16} + 0.293Z_{17} + 0.168Z_{18} - 0.17Z_{19} + \\
 & 0.013Z_{20} + 0.061Z_{21} + 0.017Z_{22} - 0.021Z_{23} - 0.017Z_{24} + 0.017Z_{25} + \\
 & 0.024Z_{26} + 0.02Z_{27} - 0.088Z_{28} + 0.159Z_{29} + 0.048Z_{30} + 0.045Z_{31} + \\
 & 0.278Z_{32} - 0.017Z_{33} + 0.027Z_{34} - 0.022Z_{35} + 0.204Z_{36} - 0.033Z_{37} + \\
 & 0.248Z_{38} - 0.033Z_{39} + 0.865Z_{40} - 0.019Z_{41} + 0.17Z_{42} - 0.037Z_{43} + \\
 & 0.017Z_{44} + 0.166Z_{45} + 0.108Z_{46} + 0.196Z_{47} - 0.06Z_{48} + 0.091Z_{49} + \\
 & 0.039Z_{50} - 0.036Z_{51} - 0.035Z_{52} + 0.009Z_{53} + 0.003Z_{54} - 0.025Z_{55} - \\
 & 0.07Z_{56} + 0.109Z_{57} - 0.052Z_{58} + 0.083Z_{59} - 0.045Z_{60} - 0.062Z_{61} + \\
 & 0.166Z_{62}
 \end{aligned}$$

En esta última componente la única variable cuyo componente es mayor a 0.5 y que por lo tanto es considerada para rotular el componente es Z_{40} (conjunto de números), el nombre sería **factor conjuntos de números**.

4.6.6 Aplicación de la prueba de Bartlett

Los resultados obtenidos al realizar el análisis de componentes principales utilizando la matriz de datos estandarizados proporcionan un total de veinte componentes principales que explican el 68.05% del total

de la varianza de las 62 variables observables, debido a estos resultados se procedió a aplicar una prueba, para determinar si la matriz de correlación de las 62 variables aleatorias es factorizable, Bartlett con este fin, desarrolló una prueba basada en la distribución ji-cuadrado, la cual es sensible al tamaño de la muestra, pues mientras mayor sea el tamaño de la muestra es más seguro que la matriz de correlación sea factorizable.

La prueba para la matriz de correlación de las variables de la prueba de matemáticas, de las variables de la prueba de lenguaje y de las variables edad y sexo de los estudiantes de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil, es la siguiente:

H_0 : La matriz de correlación es factorizable

vs.

H_a : La matriz de correlación no es factorizable

Valor del estadístico de prueba = 3084.797

Grados de libertad = 1540

Valor p = 0.000

Se concluye que existe evidencia estadística para rechazar H_0 en favor de H_a , por lo tanto la matriz de correlación no es factorizable.

Además de realizar al prueba de Bartlett a la matriz de correlación de las 62 variables aleatorias, se la aplicó a las matrices de correlación de las variables de la prueba de lenguaje y de las variables de la prueba de matemáticas por separado, sin embargo los resultados de estas pruebas indican que estas matrices tampoco son factorizables. Los resultados de la aplicación de la prueba de Bartlett para ambas matrices se muestran en la tabla CXLVI.

	Matriz de correlación de las variables de matemáticas	Matriz de correlación de las variables de lenguaje
Estadístico de prueba	994.345	1331.971
Grados de libertad	435	325
Valor p	0.000	0.000

4.7 Análisis de correlación canónica

Las correlaciones canónicas constituyen una generalización de las correlaciones simples y múltiples. Las correlaciones simples estiman la relación lineal existente entre dos variables aleatorias X y Y . Las correlaciones múltiples estiman la relación lineal existente entre dos conjuntos de variables aleatorias artificiales U_k y V_k las cuales se construyen a partir de dos grupos de variables aleatorias observables.

Sean $\mathbf{X}^{(1)} = (X_1, X_2, \dots, X_p)$ y $\mathbf{X}^{(2)} = (X_1, X_2, \dots, X_q)$ dos vectores observables aleatorios p -variado y q -variado respectivamente con $p < q$, y siendo Σ_{12} la matriz de varianzas y covarianzas de estos vectores, se construyen las variables artificiales U y V de la siguiente manera:

$$U = \mathbf{a}^T \mathbf{X}^{(1)}$$

$$V = \mathbf{b}^T \mathbf{X}^{(2)}$$

Las variables artificiales U y V deben satisfacer las siguientes condiciones:

$$\text{Var} (U) = \mathbf{a}^T \text{Cov} (\mathbf{X}^{(1)}) \mathbf{a} = \mathbf{a}^T \mathbf{S}_{11} \mathbf{a}$$

$$\text{Var} (V) = \mathbf{b}^T \text{Cov} (\mathbf{X}^{(2)}) \mathbf{b} = \mathbf{b}^T \mathbf{S}_{22} \mathbf{b}$$

$$\text{Cov} (U , V) = \mathbf{a}^T \text{Cov} (\mathbf{X}^{(1)} , \mathbf{X}^{(2)}) \mathbf{b} = \mathbf{a}^T \mathbf{S}_{12} \mathbf{b}$$

Los vectores \mathbf{a} y \mathbf{b} se deben determinar de tal forma que maximicen la correlación entre las variables canónicas U y V .

$$-1 \leq \mathbf{r}_{U,V} = \frac{\mathbf{a}^T \Sigma_{12} \mathbf{b}}{\sqrt{\mathbf{a}^T \Sigma_{11} \mathbf{a}} \sqrt{\mathbf{b}^T \Sigma_{22} \mathbf{b}}} \leq 1$$

Entonces en base a lo descrito anteriormente se puede definir a la variables canónicas como sigue:

El primer par de variables canónicas, es el par de combinaciones lineales U_1 y V_1 que tienen varianza igual a uno y que maximizan la correlación $\text{Corr}(U_1, V_1)$.

El segundo par de variables canónicas es el par de combinaciones lineales U_2 y V_2 que tienen varianza igual a uno y que maximizan la correlación $\text{Corr}(U_2, V_2)$ y que no están correlacionadas con el primer par de variables aleatorias canónicas

El k-ésimo par de variables canónicas es el par de combinaciones lineales U_k y V_k que tienen varianza igual a uno y que maximizan la correlación $\text{Corr}(U_k, V_k)$ y que no están correlacionadas con ninguno de los pares anteriores.

4.7.1 Determinación de las variables canónicas

Dados los vectores aleatorios $\mathbf{X}^{(1)} \in \mathbb{R}^p$ y $\mathbf{X}^{(2)} \in \mathbb{R}^q$, donde $p < q$, entonces se define lo siguiente:

- Las matrices de varianza y covarianza

$$\text{Cov}(\mathbf{X}^{(1)}) = \mathbf{S}_{11}$$

$$\text{Cov}(\mathbf{X}^{(2)}) = \mathbf{S}_{22}$$

- Los vectores de media

$$E(\mathbf{X}^{(1)}) = \mathbf{m}^{(1)}$$

$$E(\mathbf{X}^{(2)}) = \mathbf{m}^{(2)}$$

- La covarianza entre $\mathbf{X}^{(1)}$ y $\mathbf{X}^{(2)}$

$$\text{Cov}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}) = \mathbf{S}_{12} = \mathbf{S}_{21}^T$$

Siendo los vectores de coeficientes $\mathbf{a} \in \mathbb{R}^p$ y $\mathbf{b} \in \mathbb{R}^q$ de las combinaciones lineales $U = \mathbf{a}^T \mathbf{X}^{(1)}$ y $V = \mathbf{b}^T \mathbf{X}^{(2)}$ entonces,

$$\max_{\mathbf{a}, \mathbf{b}} \text{Corr}(U, V) = \rho_1^*$$

a partir de lo cual se construye el primer par de variables canónicas

$$U_1 = e_1^T \mathbf{\Sigma}_{11}^{-1/2} \mathbf{X}^{(1)} \quad \text{y} \quad V_1 = f_1^T \mathbf{\Sigma}_{22}^{-1/2} \mathbf{X}^{(2)},$$

El k-ésimo par de variables canónicas para $k = 2, 3, \dots, p$ es

$$U_k = e_k^T \mathbf{\Sigma}_{11}^{-1/2} \mathbf{X}^{(1)} \quad \text{y} \quad V_k = f_k^T \mathbf{\Sigma}_{22}^{-1/2} \mathbf{X}^{(2)}$$

que maximiza $\text{Corr}(U_k, V_k) = \rho_k^*$, donde las k-ésimas variables canónicas no están correlacionadas con las anteriores k-1 variables canónicas.

El propósito es encontrar un par de vectores U_k y V_k , tal que el coeficiente de correlación entre estos vectores sea el mayor posible para $k, l = 1, 2, \dots, p$. Donde $\mathbf{r}_1^{*2} \geq \mathbf{r}_2^{*2} \geq \dots \geq \mathbf{r}_p^{*2}$, representan los valores propios de la matriz $\mathbf{S}_{11}^{-1/2} \mathbf{S}_{12} \mathbf{S}_{22}^{-1} \mathbf{S}_{21} \mathbf{S}_{11}^{-1/2}$ con vectores propios e_1, e_2, \dots, e_p de tamaño $(p \times 1)$, los cuales son iguales a los valores propios de la matriz $\mathbf{S}_{22}^{-1/2} \mathbf{S}_{21} \mathbf{S}_{11}^{-1} \mathbf{S}_{12} \mathbf{S}_{22}^{-1/2}$ con vectores propios asociados f_1, f_2, \dots, f_p de tamaño $(q \times p)$.

Donde $\mathbf{a}_k^T = e_k^T \mathbf{\Sigma}_{11}^{-1/2}$ y $\mathbf{b}_k^T = f_k^T \mathbf{\Sigma}_{22}^{-1/2}$; los \mathbf{e}_k^T y \mathbf{f}_k^T son los vectores propios de la matriz $\Sigma \in M_{(p+q) \times (p+q)}$ que como se explica posteriormente se forma por la unión de las matrices \mathbf{S}_{11} , \mathbf{S}_{22} , \mathbf{S}_{12} y \mathbf{S}_{21}^T , y $\mathbf{S}_{11}^{-1/2}$ y $\mathbf{S}_{22}^{-1/2}$ son

las matrices de varianza y covarianza de cada uno de los grupos de variables observables elevados a la -1/2.

El objetivo de la correlación canónica es medir la fuerza de la relación entre dos conjuntos de variables. Para el análisis es conveniente definir el vector \mathbf{X} que esta formado por la unión de dos vectores de dos grupos de variables observables consideradas en el análisis como se muestra a continuación:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}^{(1)} \\ \dots \\ \mathbf{x}^{(2)} \end{bmatrix} = \begin{bmatrix} X_1^{(1)} \\ X_2^{(1)} \\ \vdots \\ X_p^{(1)} \\ \dots \\ X_{p+1}^{(2)} \\ X_{p+2}^{(2)} \\ \vdots \\ X_q^{(2)} \end{bmatrix} \in \mathfrak{R}^{p+q}$$

El vector de medias del vector $\mathbf{X} \in \mathfrak{R}^{p+q}$ está formada por el valor esperado de cada uno de los vectores componentes como se muestra a continuación:

$$\hat{\mathbf{i}} = E[\mathbf{X}] = \begin{bmatrix} E[\mathbf{X}^{(1)}] \\ \dots \\ E[\mathbf{X}^{(2)}] \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{i}}^{(1)} \\ \dots \\ \hat{\mathbf{i}}^{(2)} \end{bmatrix} \in \mathfrak{R}^{p+q}$$

La matriz de varianzas y covarianzas del vector \mathbf{X} se calcula de la siguiente manera:

$$\hat{\mathbf{O}} = E [(\mathbf{X} - \hat{\mathbf{i}})(\mathbf{X} - \hat{\mathbf{i}})^T] \in M_{(p+q) \times (p+q)}$$

La expresión anterior se puede expresar como la unión de las matrices de varianzas y covarianzas de los vectores $\mathbf{X}^{(1)}$ y $\mathbf{X}^{(2)}$ respectivamente y la matriz de covarianzas entre estos mismos vectores como se muestra a continuación:

$$\Sigma = \begin{bmatrix} E[(X^{(1)} - \mathbf{m}^{(1)})(X^{(1)} - \mathbf{m}^{(1)})^T] & \vdots & E[(X^{(1)} - \mathbf{m}^{(1)})(X^{(2)} - \mathbf{m}^{(2)})^T] \\ \dots & \vdots & \dots \\ E[(X^{(2)} - \mathbf{m}^{(2)})(X^{(1)} - \mathbf{m}^{(1)})^T] & \vdots & E[(X^{(2)} - \mathbf{m}^{(2)})(X^{(2)} - \mathbf{m}^{(2)})^T] \end{bmatrix}$$

La matriz $\mathbf{S} \in M_{(p+q) \times (p+q)}$ es simétrica y definida positiva, se puede abreviar de la siguiente manera:

$$\hat{\mathbf{O}} = \left[\begin{array}{c|c} \sum_{11} & \sum_{12} \\ \hline \sum_{21} & \sum_{22} \end{array} \right] \in M_{(p+q) \times (p+q)}$$

$(p \times p)$ $(p \times q)$
 $(q \times p)$ $(q \times q)$

4.7.2 Correlación canónica para los conjuntos de variables de la prueba de lenguaje y de la prueba de matemáticas

El procedimiento descrito en la sección 4.6.1 para la determinación de las variables canónicas, se aplicó a los conjuntos de variables de la prueba de lenguaje y de la prueba de matemáticas. Siendo el vector $\mathbf{X}^{(1)} \in \mathbb{R}^p$ el que corresponde a las variables de lenguaje, donde $p=26$ y el vector $\mathbf{Y}^{(2)} \in \mathbb{R}^q$ el que corresponde a las variables de matemáticas, donde $q=30$. Los resultados de las correlaciones canónicas y los coeficientes de las variables canónicas se los obtuvo utilizando el software de aplicaciones estadísticas SPSS (Social purpose statistical system) versión 8.0, el procedimiento es el siguiente se abre la matriz de datos, se va al menú principal, se abre la ventana de New y se entra a la ventana Syntax en la que se realiza la siguiente función:

```
INCLUDE 'Canonical Correlation.sps'
```

CANCORR SET 1 = Nombres de las variables de la prueba de lenguaje (ó códigos de las variables) /Set 2 = Nombres de las variables de la prueba de matemáticas (ó códigos de las variables). Luego desde el menú inicial se corre esta función con la que se obtiene los coeficientes de las variables canónicas y los coeficientes de correlación canónica.

Las correlaciones canónicas ρ_k^* obtenidas entre cada par de las variables canónicas U_k, V_k para $k = 1, 2, \dots, 26$ son mostradas en la tabla CXLVII. A continuación se analizan los cuatro primeros pares de variables canónicas, que son los que tienen los mayores coeficientes de correlación canónica.

Tabla CXLVII
Correlaciones canónicas

Variable	Correlación	Variable	Correlación
1	0,745	14	0,362
2	0,706	15	0,35
3	0,657	16	0,288
4	0,636	17	0,268
5	0,617	18	0,253
6	0,615	19	0,245
7	0,553	20	0,231
8	0,546	21	0,182
9	0,516	22	0,168
10	0,5	23	0,143
11	0,441	24	0,088
12	0,428	25	0,058
13	0,383	26	0,03

El primer par de variables canónicas es U_1 y V_1

$$U_1 = -0.517X_1 + -0.419X_2 + 0.066X_3 -0.237 X_4 -0.361X_5 -0.261X_6 -0.218X_7 -0.314X_8 -0.478X_9 -0.31X_{10} -0.274X_{11} -0.502X_{12} -0.452X_{13} -0.477X_{14} -0.307X_{15} -0.309X_{16} -0.429X_{17} -0.425X_{18} -0.161X_{19} -0.216X_{20} -0.418X_{21} -0.478X_{22} -0.414X_{23} -0.385X_{24} -0.442X_{25} -0.487X_{26}$$

$$V_1 = -0.449Y_1 - 0.18Y_2 + 0.236Y_3 - 0.27 Y_4 - 0.419Y_5 - 0.172Y_6 - 0.469Y_7 - 0.303Y_8 - 0.171Y_9 - 0.033Y_{10} - 0.249Y_{11} - 0.093Y_{12} - 0.504Y_{13} - 0.249Y_{14} - 0.446Y_{15} - 0.402Y_{16} - 0.332Y_{17} - 0.518Y_{18} - 0.409Y_{19} + 0.083Y_{20} - 0.255Y_{21} - 0.335Y_{22} - 0.255Y_{23} - 0.253Y_{24} - 0.319Y_{25} - 0.458Y_{26} - 0.317Y_{27} - 0.074Y_{28} + 0.039Y_{29} - 0.308Y_{30}$$

Si se analizan los coeficientes (de las variables observables del conjunto de la prueba de lenguaje y de la prueba de matemáticas) canónicos asociados a la primera correlación canónica resulta que X_1 (identificación de diptongos) = -0.517 y X_{12} (identificación de palabras diminutivas) = -0.502, Y_{13} (ejercicio de divisibilidad) = -0.504 y Y_{18} (propiedades de conjuntos) = -0.518 son las variables que más aportación ofrecen a la correlación existente $\rho^*_1 = 0.745$. Tal que la $\text{Var}(U_1) = 1$, $\text{Var}(V_1) = 1$ y la $\text{Cov}(U_1, V_1) = 0.745$.

El segundo par de variables canónicas es U_2 y V_2

$$U_2 = 0.036X_1 - 0.035X_2 + 0.183X_3 + 0.011 X_4 - 0.217X_5 - 0.12X_6 - 0.011X_7 - 0.037X_8 - 0.282X_9 + 0.035X_{10} + 0.296X_{11} - 0.013X_{12} + 0.017X_{13} + 0.063X_{14} + 0.221X_{15} + 0.286X_{16} + 0.465X_{17} + 0.228X_{18} + 0.173X_{19} - 0.082X_{20} - 0.108X_{21} - 0.08X_{22} + 0.175X_{23} + 0.117X_{24} + 0.417X_{25} + 0.13X_{26}$$

$$V_2 = 0.12Y_1 + 0.287Y_2 + 0.193Y_3 + 0.07 Y_4 + 0.281Y_5 - 0.169Y_6 - 0.295Y_7 - 0.118Y_8 + 0.343Y_9 + 0.159Y_{10} - 0.182Y_{11} + 0.174Y_{12} + 0.061Y_{13} - 0.253Y_{14} - 0.073Y_{15} - 0.087Y_{16} - 0.057Y_{17} + 0.146Y_{18} + 0.05Y_{19} + 0.149Y_{20} + 0.258Y_{21} + 0.03Y_{22} - 0.061Y_{23} - 0.003Y_{24} + 0.002Y_{25} + 0.157Y_{26} + 0.205 Y_{27} + 0.302Y_{28} + 0.379Y_{29} - 0.031Y_{30}$$

Los coeficientes del segundo par de variables canónicas de la prueba de lenguaje y de la prueba de matemáticas más significativos son X_{17} (identificación del sujeto de la oración) = 0.465 y X_{25} (palabras tildadas correctamente) = 0.417, Y_9 (ejercicio de valor absoluto) = 0.343 y Y_{29} (ejercicio de probabilidad) = 0.379, estas son las variables que más aportación ofrecen a la correlación existente entre U_2 y V_2 , donde $\rho_{22}^* = 0.706$. Tal que la $\text{Var}(U_2) = 1$, $\text{Var}(V_2) = 1$ y la $\text{Cov}(U_2, V_2) = 0.706$.

El tercer par de variables canónicas es U_3 y V_3

$$U_3 = -0.014X_1 + 0.008X_2 + 0.004X_3 + 0.354X_4 - 0.254X_5 - 0.035X_6 + 0.015X_7 - 0.102X_8 + 0.159X_9 + 0.083X_{10} - 0.034X_{11} - 0.13X_{12} - 0.206X_{13} - 0.206X_{14} + 0.129X_{15} + 0.036X_{16} + 0.142X_{17} - 0.018X_{18} + 0.096X_{19} + 0.051X_{20} - 0.016X_{21} + 0.032X_{22} + 0.044X_{23} - 0.342X_{24} + 0.004X_{25} + 0.299X_{26}$$

$$V_3 = 0.08Y_1 - 0.078Y_2 + 0.191Y_3 + 0.086Y_4 - 0.119Y_5 - 0.002Y_6 - 0.007Y_7 - 0.038Y_8 + 0.051Y_9 + 0.016Y_{10} - 0.215Y_{11} + 0.451Y_{12} - 0.106Y_{13} + 0.099Y_{14} + 0.043Y_{15} - 0.135Y_{16} + 0.104Y_{17} - 0.054Y_{18} - 0.054Y_{19} + 0.264Y_{20} + 0.188Y_{21} - 0.024Y_{22} + 0.188Y_{23} - 0.167Y_{24} + 0.339Y_{25} + 0.214Y_{26} + 0.292Y_{27} - 0.24Y_{28} - 0.324Y_{29} - 0.057Y_{30}$$

En el tercer par de variables canónicas, los coeficientes canónicos más significativos son los asociados a las variables X_4 (identificación palabras agudas) = 0.354 y X_{24} (corrección de sintaxis) = -0.342, Y_{12} (ejercicio de divisibilidad) = 0.451 y Y_{25} (factorización de dos polinomios) = 0.339, estas son las variables que más aportación ofrecen a la correlación existente

entre U_3 y V_3 , donde $\rho_3^* = 0.657$. Tal que la $\text{Var}(U_3) = 1$, $\text{Var}(V_3) = 1$ y la $\text{Cov}(U_3, V_3) = 0.657$.

El cuarto par de variables canónicas es U_4 y V_4

$$U_4 = -0.431X_1 - 0.488X_2 - 0.31X_3 - 0.158X_4 + 0.147X_5 + 0.083X_6 + 0.082X_7 + 0.085X_8 + 0.134X_9 - 0.074X_{10} + 0.092X_{11} + 0.115X_{12} - 0.165X_{13} - 0.279X_{14} + 0.149X_{15} + 0.231X_{16} - 0.03X_{17} - 0.247X_{18} - 0.049X_{19} + 0.296X_{20} + 0.179X_{21} + 0.059X_{22} - 0.212X_{23} - 0.012X_{24} + 0.132X_{25} + 0.005X_{26}$$

$$V_4 = 0.064Y_1 + 0.25Y_2 - 0.337Y_3 + 0.204Y_4 + 0.071Y_5 - 0.103Y_6 - 0.166Y_7 - 0.129Y_8 + 0.112Y_9 + 0.112Y_{10} - 0.096Y_{11} + 0.185Y_{12} + 0.096Y_{13} + 0.26Y_{14} - 0.002Y_{15} + 0.026Y_{16} - 0.36Y_{17} - 0.081Y_{18} - 0.166Y_{19} + 0.139Y_{20} - 0.127Y_{21} - 0.309Y_{22} - 0.316Y_{23} - 0.02Y_{24} - 0.015Y_{25} + 0.208Y_{26} + 0.035Y_{27} - 0.342Y_{28} + 0.173Y_{29} - 0.131Y_{30}$$

Los coeficientes canónicos más altos en el cuarto par de variables canónicas, son los asociados a las variables X_1 (identificación diptongos) = -0.431 y X_2 (identificación de triptongos) = -0.488, Y_3 (resta de números enteros) = -0.337 y Y_{23} (área del círculo) = -0.316, estas son las variables que más aportación ofrecen a la correlación existente entre U_4 y V_4 , donde $\rho_4^* = 0.636$. Tal que la $\text{Var}(U_4) = 1$, $\text{Var}(V_4) = 1$ y la $\text{Cov}(U_4, V_4) = 0.657$.

CONCLUSIONES

1. En el décimo año de educación básica, de los colegios fiscales rurales del cantón Guayaquil, analizados en esta tesis se obtuvo que, el número de estudiantes por profesor en el año lectivo 2000-2001, fue de aproximadamente 13 estudiantes, de los cuales seis son mujeres y siete hombres, este resultado se obtiene a partir de los valores mostrados en las tablas XXI y XXII.
2. La edad promedio de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil en el año lectivo 2000-2001, es de 14.93 años (tabla XXV), el promedio de la edad de los estudiantes de sexo masculino y de sexo femenino son 14.87 y 14.99 años, respectivamente.
3. La edad promedio de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil en el año lectivo 2000-2001, por jornada de trabajo matutina, vespertina y nocturna son 14.82, 14.34 y 15.63 años, respectivamente.
4. Por cada 100 estudiantes entrevistados, 51 son de sexo femenino y 49 son de sexo masculino (sección 3.2.2).

5. De la jornada en que funcionan los colegios, se puede decir que por cada 100 estudiantes entrevistados, 86 asisten a colegios matutinos, 5 a colegios vespertinos y 9 a colegios nocturnos.

6. Utilizando el coeficiente de asimetría de las variables aleatorias, se estimó el grado de dificultad de las preguntas de la prueba de lenguaje, (sección 3.5). Las características investigadas que obtuvieron los menores coeficientes de asimetría son identificación de triptongos (-1.228), identificación de diptongos (-1.122), conjugación de verbos (-0.917) e identificación de frases (-0.865).

7. El 41.31% de los estudiantes entrevistados de décimo año de los colegios fiscales rurales del cantón Guayaquil, respondió correctamente a la pregunta, correspondiente a la identificación de triptongos, (sección 3.4.2). De este porcentaje el 50.72% son estudiantes de sexo masculino y el 49.28% son de sexo femenino. El 11.37% de los estudiantes no respondió la pregunta y el 47.32% identificó de 1 a 4 triptongos correctamente.

8. El 31.73% de los estudiantes entrevistados de décimo año de los colegios fiscales rurales del cantón Guayaquil, respondió correctamente a la pregunta, correspondiente a la identificación de

diptongos (sección 3.4.1). De este porcentaje el 37.72% son estudiantes de sexo masculino y el 62.28% son de sexo femenino. El 10.17% de los estudiantes no respondió y el 58.10% identificó de 1 a 9 diptongos correctamente.

9. El 61.67% de los estudiantes entrevistados de décimo año de los colegios fiscales rurales del cantón Guayaquil, respondió correctamente a la pregunta identificación de frases (sección 3.4.21). De este porcentaje el 51.45% son estudiantes de sexo masculino y el 48.55% son de sexo femenino. El 21.55% de los estudiantes no respondió y el 16.78% identificó 1 frase correctamente.

10. El 9.57% de los estudiantes entrevistados de décimo año de los colegios fiscales rurales del cantón Guayaquil, respondió correctamente a la pregunta conjugación de verbos (sección 3.4.16). De este porcentaje el 56.21% son estudiantes de sexo masculino y el 43.79% son de sexo femenino. El 8.38% de los estudiantes no contestó y el 82.05% conjugó de 1 a 5 verbos correctamente.

11. Utilizando el coeficiente de asimetría de las variables aleatorias, se estimó el grado de dificultad de las preguntas de la prueba de lenguaje (sección 3.4). Las preguntas que obtuvieron los mayores

coeficientes de asimetría son corrección de la sintaxis de la oración (2.319), identificación del sustantivo (1.204), identificación de simples palabras (0.896) e identificación del verbo (0.86).

12. El 78.44% de los estudiantes entrevistados, no respondió la pregunta sintaxis de la oración (sección 3.4.25). De este porcentaje el 50.38% son estudiantes de sexo masculino y el 49.62% son de sexo femenino. El 21.55% de los estudiantes realizó de 1 a 3 respuestas correctas y el 0.01% realizó 4 correcciones.

13. El 66.46% de los estudiantes entrevistados, no respondió la pregunta, correspondiente a la identificación del sustantivo (sección 3.4.19). De este porcentaje el 48.64% son estudiantes de sexo masculino y el 51.36% son de sexo femenino. El 25.16% de los estudiantes identificó de 1 a 4 sustantivos correctamente y el 8.38% identificó todos los sustantivos correctamente.

14. El 58.68% de los estudiantes entrevistados, no respondió la pregunta, correspondiente a la identificación de simples palabras (sección 3.4.14). De este porcentaje el 49.94% son estudiantes de sexo masculino y el 53.06% son de sexo femenino. El 32.94% de los

estudiantes identificó correctamente 1 ó 2 simples palabras y el 8.38% identificó correctamente todas las palabras simples.

15. El 58.08% de los estudiantes entrevistados, no respondió la pregunta, correspondiente a la identificación del verbo de la oración (sección 3.4.20). De este porcentaje el 54.63% son estudiantes de sexo masculino y el 45.37% son de sexo femenino. El 37.13% de los estudiantes entrevistados identificó correctamente de 1 a 3 verbos y el 4.79% identificó correctamente todos los verbos.

16. Utilizando el coeficiente de asimetría de las variables aleatorias, se estimó el grado de dificultad de las preguntas de la prueba de matemáticas (sección 3.5). Las preguntas que obtuvieron los menores coeficientes de asimetría son suma de números enteros (-2.289), resta de números enteros (-1.124), multiplicación de números enteros (-0.219) y divisibilidad (-0.121).

17. El 82.03% de los estudiantes entrevistados, respondió correctamente a la pregunta, correspondiente a la suma de números enteros (sección 3.3.1). De este porcentaje el 47.44% son estudiantes de sexo masculino y el 52.56% son de sexo femenino. El 13.78% de los

estudiantes sumó correctamente de 1 a 3 cifras y el 4.19% no realizó la suma.

18. El 69.46% de los estudiantes entrevistados, respondió correctamente a la pregunta resta de números enteros (sección 3.3.3). De este porcentaje el 51.72% son estudiantes de sexo masculino y el 48.28% son de sexo femenino. El 14.38% de los estudiantes restó correctamente de 1 a 3 cifras y el 16.165 no realizó la resta.

19. El 46.10% de los estudiantes entrevistados, respondió correctamente a la pregunta multiplicación de números enteros (sección 3.3.5). De este porcentaje el 46.74% son estudiantes de sexo masculino y el 53.26% son de sexo femenino. El 22.77% de los estudiantes multiplicó correctamente 1 ó 2 cifras y el 31.13% no respondió.

20. El 51.49% de los estudiantes entrevistados, respondió correctamente a la pregunta divisibilidad (sección 3.3.14). De este porcentaje el 41.85% son estudiantes de sexo masculino y el 58.15% son de sexo femenino. El 44.91% de los estudiantes no respondió y el 3.6% realizó correctamente 1 ó 2 ejercicios de divisibilidad.

21. Utilizando el coeficiente de asimetría de las variables aleatorias, se estimó el grado de dificultad de las preguntas de la prueba de matemáticas (sección 3.5). Las preguntas que obtuvieron los mayores coeficientes de asimetría son multiplicación de números racionales (9.054), potenciación (6.284), evaluación de funciones (6.284) y teorema de Pitágoras (6.387).

22. El 98.79% de los estudiantes entrevistados, no respondió a la pregunta multiplicación de números racionales (sección 3.3.6). De este porcentaje el 48.48% son estudiantes de sexo masculino y el 51.52% son de sexo femenino. El 1.21% de los estudiantes respondió correctamente.

23. El 98.19% de los estudiantes entrevistados, no respondió a la pregunta potenciación (sección 3.3.12). De este porcentaje el 48.17% son estudiantes de sexo masculino y el 51.83% son de sexo femenino. El 1.21% de los estudiantes respondió correctamente

24. El 98.79% de los estudiantes entrevistados, no respondió a la pregunta, correspondiente a la evaluación de funciones (sección 3.3.20). De este porcentaje el 48.48% son estudiantes de sexo

masculino y el 51.52% son de sexo femenino. El 2.405 de los estudiantes contestó correctamente.

25. El 89.21% de los estudiantes entrevistados, no respondió a la pregunta teorema de Pitágoras (sección 3.3.24). De este porcentaje el 49.55% son estudiantes de sexo masculino y el 50.54% son de sexo femenino. El 10.78% de los estudiantes contestó correctamente.

26. La calificación de los estudiantes en lenguaje, fue en promedio 43.14 sobre 100 puntos, (sección 3.3.27). La calificación máxima obtenida por un estudiante fue 83, y la mínima fue 0.

27. La calificación de los estudiantes en matemáticas (sección 3.3.31), fue en promedio 13.31 sobre 100 puntos. La calificación máxima obtenida por un estudiante fue 32, y la mínima fue 0.

28. La calificación total de los estudiantes de décimo año de los colegios fiscales rurales del cantón Guayaquil (sección 3.3.28), es en promedio 28.22 sobre 100 puntos.

29. Existe una fuerte relación lineal directa entre las preguntas identificación de diptongos e identificación de triptongos (sección 4.3.1). El coeficiente de correlación entre estas variables es 0.831.
30. Existe una fuerte relación lineal directa entre las características estudiadas identificación del sujeto e identificación del predicado (sección 4.3.1). El coeficiente de correlación entre estas variables es 0.665.
31. Existe una fuerte relación lineal directa entre las preguntas multiplicación de fracciones y división de fracciones (sección 4.3.1). El coeficiente de correlación entre estas variables es 0.627.
32. El coeficiente de correlación de las características investigadas identificación de palabras agudas e identificación de palabras diminutivas es cero (sección 4.3.2), por lo tanto estas variables son linealmente independientes.
33. El coeficiente de correlación de las preguntas número de relaciones de orden correctas y número de palabras definidas correctamente es cero (sección 4.3.2), por lo tanto estas variables son linealmente independientes.

34. El coeficiente de correlación de las características investigadas definiciones de conjuntos de números y ecuación lineal de una incógnita es cero (sección 4.3.2), por lo tanto estas variables son linealmente independientes.
35. Utilizando la prueba de Bartlett, se obtuvo que la matriz de correlación correspondiente a las 62 variables aleatorias, no es factorizable (sección 4.6.6). Además aplicando esta misma prueba a las matrices de correlación, de las variables aleatorias correspondientes a la prueba de lenguaje y a la prueba de matemáticas por separado, se obtuvo que estas matrices no son factorizables.
36. En el análisis de contingencia realizado en esta tesis se obtuvo que existe dependencia de la variable aleatoria calificación de matemáticas con respecto a la jornada de trabajo de los planteles educativos (sección 4.4.8), el valor p de la prueba es 0.022.
37. Otro resultado del análisis de contingencia realizado, es que existe dependencia de la variable aleatoria calificación de matemáticas con respecto a la ubicación del establecimiento educativo (sección 4.4.7) y, el valor p de la prueba es 0.0011.

38. Existe dependencia no lineal entre las variables aleatorias calificación de lenguaje y calificación de matemáticas (sección 4.4.9), con un valor $p = 0.00044$.
39. En la prueba para determinar la igualdad de las medias de la variable aleatoria calificación de lenguaje con respecto a la ubicación de los planteles educativos (sección 4.5.1), no se rechazó la hipótesis nula H_0 de que las medias de la calificación de lenguaje de los planteles educativos ubicados en Progreso, P. de Sabana Grande, Puná, Posorja y Tenguel son iguales, con un valor p de 0.803.
40. En el análisis de varianza realizado a la variable aleatoria calificación de lenguaje con respecto a la jornada de trabajo de los planteles educativos (sección 4.5.2), no se rechazó la hipótesis nula H_0 de que las medias de la calificación de lenguaje de los planteles con jornadas matutina, vespertina y nocturna son iguales, con un valor p de 0.939.
41. En el análisis de varianza realizado se rechazó la hipótesis nula H_0 : la medias de la variable aleatoria calificación de matemáticas con respecto a la ubicación de los planteles educativos son iguales, en

favor de H_1 : al menos un par de medias no son iguales, con un valor p de 0.000, (sección 4.5.3).

42. Al aplicar el método de la mínima diferencia significativa (LSD), para determinar cuales pares de medias de la calificación de matemáticas, con respecto a la ubicación de los colegios son diferentes (sección 4.5.3), se obtuvo que, las medias de los colegios ubicados en Puná y P. de Sabana Grande, y la de los colegios ubicados en Puná y Posorja difieren significativamente.

43. En otro análisis de varianza que se realizó, se determinó que las medias de la variable aleatoria calificación de matemáticas con respecto a la jornada de trabajo de los planteles educativos (sección 4.5.4), son iguales, el valor p de la prueba es 0.131.

44. Las medias de la calificación total de los estudiantes, de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil, con respecto a la ubicación de los colegios son iguales, (sección 4.5.5), esto se determinó, mediante un análisis de varianza en el que el valor p es 0.445.

45. Con respecto a la jornada de trabajo de los colegios, se probó que, las medias de la calificación total de los estudiantes de décimo año de educación básica (sección 4.5.6), de los colegios fiscales rurales del cantón Guayaquil, son iguales, el valor p de esta prueba es 0.738.

46. Los coeficientes canónicos más significativos asociados a la primera correlación canónica, (sección 4.7.2), son identificación de diptongos (0.517) y identificación de palabras diminutivas (-0.502), ejercicio de divisibilidad (-0.504) y propiedades de conjuntos (-0.518). El primer par de variables canónicas tienen un coeficiente de correlación $\rho^*_1 = 0.745$.

47. Los coeficientes del segundo par de variables canónicas de la prueba de lenguaje y de la prueba de matemáticas, (sección 4.7.2), más significativos son identificación del sujeto de la oración (0.465) y palabras tildadas correctamente (0.417), ejercicio de valor absoluto (0.343) y ejercicio de probabilidad (0.379), estas son las variables que más aportación ofrecen a la correlación existente entre U_2 y V_2 , donde $\rho^*_2 = 0.706$.

48. En el tercer par de variables canónicas (sección 4.7.2), los coeficientes canónicos más significativos son los asociados a las

variables identificación palabras agudas (0.354) y corrección de sintaxis (-0.342), ejercicio de divisibilidad (0.451) y factorización de dos polinomios (0.339), estas son las variables que más aportación ofrecen a la correlación existente entre U_3 y V_3 , donde $\rho_3^* = 0.657$.

49. Las variables que más aportación ofrecen a la correlación existente entre el cuarto par de variables canónicas (sección 4.7.2), donde $\rho_4^* = 0.636$, son identificación diptongos (-0.431) y identificación de triptongos (-0.488), resta de números enteros (-0.337) y área del círculo (-0.316).

RECOMENDACIONES

1. En base a los resultados obtenidos en las preguntas de la prueba de matemáticas, se recomienda que se mejore la enseñanza en los temas operaciones con números racionales, potenciación y radicación, evaluación de funciones y en la aplicación del teorema de Pitágoras,
2. En lo que respecta a las preguntas con mayor grado de dificultad de la prueba de lenguaje, se recomienda que se mejore la enseñanza en los temas, sintaxis de la oración, partes de la oración y clasificación de las palabras.
3. Existe un desnivel significativo en el promedio de las calificaciones de los estudiantes con respecto a matemáticas y a lenguaje, los promedios son 13.31 y 43.14 puntos sobre cien respectivamente, por lo tanto se recomienda que se analicen las causas de éste desnivel para mejorar los niveles de conocimiento de los estudiantes en ambas materias.

4. De acuerdo a los resultados obtenidos, se determinó que existe dependencia no lineal entre la calificación de matemáticas con la ubicación y la jornada de los colegios ubicación de los colegios fiscales rurales, por lo que se recomienda analizar las causas positivas y negativas que influyen en esta dependencia.

5. Además del análisis realizado del nivel de conocimiento de los estudiantes, de décimo año de educación básica de los colegios fiscales rurales del cantón Guayaquil, debería realizarse un análisis similar en las demás zonas rurales del cantón y a nivel nacional, con el objetivo de comparar el nivel del conocimiento de los estudiantes en las diferentes zonas, además de investigar posibles causas del bajo nivel de conocimiento de los estudiantes.

Bibliografía

1. Poveda C, Díaz G, Avendaño A, Araujo S, Arboleda R, (1994), Sistemas Educativos Nacionales, MEC, Quito-Ecuador
2. Vargas J, (1985), Historia de la Cultura ecuatoriana, tomos I y II, Publicaciones educativas Ariel, Quito-Ecuador
3. Saltos N, Vázquez L,(1999), Ecuador su realidad, Fundación José Peralta, séptima edición, Quito-Ecuador
4. Santamaría Andrés, (1995), Diccionario de sinónimos antónimos e ideas afines, Sopena, Barcelona
5. Dávila J, Orbe J, Gómez M, (1990), Ciencias Sociales: historia, geografía y cívica, Libresa, Quito-Ecuador
6. La Educación, (1977), Revista Interamericana de Desarrollo Educativo, Secretaria General de la OEA, Washington, Diccionario Enciclopédico Larousse, (1985), tomo 6, Planeta, Barcelona-España,

7. Nassif R, Rama G, Tedesco J, (1984), El sistema educativo en América Latina, Kapelusz, Buenos Aires -Argentina
8. CEPAL, (1992), Anuario Estadístico de América Latina y el Caribe, Santiago de Chile
9. Sistema nacional de estadísticas educativas del Ecuador, (1993,1994,1995,1996,1997 y 1998), Boletines estadísticos: La educación en cifras, Quito-Ecuador.
10. Posso A, (1998), Universidad ecuatoriana: pasado, presente y futuro, Editorial universitaria, Ibarra-Ecuador
11. Mendenhall W, Wackerly D, Sheaffer R, (1994), Estadística Matemática con aplicaciones, Iberoamérica, segunda edición, México
12. Jhonson R, Wichern D, (1990), Applied Multivariate Statistical Analysis, Prentice Hall, cuarta edición, New Jersey-USA
13. Miller I, Miller M, Estadística matemática con aplicaciones, Pearson educación, sexta edición

14. Montgomery D, (1991) Diseño y análisis de experimentos, Iberoamérica, primera edición, México
15. Freund J., Walpole R. (1990), Estadística Matemática con Aplicaciones, Prentice Hall, cuarta edición, México
16. Software Estadístico SPSS (Social Purpose Statistical System), versión 8.0
17. Software estadístico SYSTAT, versión 7.0, SPSS INC.

PRUEBA DE MATEMÁTICAS PARA LOS ESTUDIANTES DE DECIMO AÑO DE EDUCACIÓN BÁSICA

Operaciones básicas

1.- Efectuar las siguientes operaciones:

$$\begin{array}{r} 4728 \\ + 3197 \\ \hline \end{array}$$

$$\begin{array}{r} 2571 \\ - 1450 \\ \hline \end{array}$$

$$\begin{array}{r} 27418 \\ \times 37 \\ \hline \end{array}$$

$$87450 \overline{) 35}$$

$$\frac{3}{5} + \frac{1}{8} =$$

$$\frac{7}{9} - \frac{1}{3} =$$

$$\frac{7}{5} \times \frac{2}{21} =$$

$$\frac{2}{5} \div \frac{6}{7} =$$

Conjuntos de Números

2.- Los números reales se clasifican en: racionales e irracionales. A su vez los racionales se dividen en enteros y fraccionarios y los enteros se clasifican en enteros positivos, enteros negativos y el cero.

Clasifique los siguientes literales con falso (**F**) o verdadero (**V**) según corresponda:

- a) 3 es un número real, racional y entero no negativo ()
- b) $\sqrt{2}$ es un número real, racional y fraccionario ()
- c) $\frac{2}{9}$ es un número real, racional y entero positivo ()
- d) 15 es un número real e irracional ()

Valor absoluto

3.- Si $|a|$ representa el valor absoluto del número a, siendo el conjunto referencial: los reales. Determine el conjunto solución de la desigualdad $|x - 3| < 5$:

Relaciones de orden

4.- Llene la línea con una de los siguientes signos de relación de orden: igual (=), mayor que (>), menor que (<), según corresponda:

$$4/5 \text{ } 0.6$$

$$3/4 \text{ } 0.75$$

$$5/2 \text{ } 1.2$$

$$1/4 \text{ } 0.5$$

$$3/2 \text{ } 2$$

Potenciación y radicación

- 5.- Determine el valor de a para que se cumplan las siguientes igualdades:

$$7^2 \times 7^{-8} = 7^{-a} \text{ _____}; a =$$

$$\left[\left(\frac{2}{5} \right)^a \right]^2 = \left(\frac{2}{5} \right)^6 \text{ _____}; a =$$

$$\sqrt{\sqrt[3]{2}} = \sqrt[6]{2} \text{ _____}; a =$$

$$\sqrt[3]{7^5} = 7 \text{ _____}; a =$$

Divisibilidad

- 6.- María tiene 35 manzanas, 21 naranjas y 7 peras. Llegan a visitarla siete amigos y ella quiere repartirles las frutas en partes iguales. ¿Cuántas manzanas, naranjas y peras le toca a cada uno de sus amigos?

Proporcionalidad

- 7.- Las edades de un hijo y su padre están en relación de 1 a 5. Si la suma de sus edades es 42. ¿Cuál es la edad de cada uno?
- 8.- A Juan le hacen un préstamo de 900 dólares, y le dicen que si paga dentro de un mes, le rebajan el 5% de su deuda. Si logra cancelar la deuda en ese plazo, ¿cuánto tendría que pagar?
- 9.- En la construcción de una casa, 6 albañiles logran hacer 18 paredes diariamente. Si la cantidad de albañiles aumentara a 10, ¿cuántas paredes hicieran por día?

Sistema métrico

- 10.- Complete:

10 kilómetros equivalen a _____ metros

1 metro tiene _____ decímetros

2 hectómetros tienen _____ metros

Conjuntos

11.- Dados los siguientes conjuntos:

$A = \{\text{perro, gato, conejo, león}\}$; $B = \{\text{pera, manzana, conejo, naranja}\}$; $C = \{\text{gato, pera}\}$
Determine lo siguiente:

$$A \cup B =$$

$$A \cap B =$$

$$A \cap C =$$

$$B - C =$$

Lógica

12.- Toda proposición tiene dos posibles valores de verdad: Verdadero (**V**) o Falso (**F**).
¿Qué valor de verdad corresponde a cada una de las siguientes proposiciones?

- Perú es una ciudad ó $\sqrt{9} = 4$ ()
- La capital de Ecuador es Cuenca y Cristóbal Colón descubrió América ()
- Si febrero es el primer mes del año, entonces $3+5=8$ ()

Funciones

13.- Si $A = \{3, 5, 7, 9\}$, $B = \{3, 7, 9, 11, 13, 15\}$ y $f : A \rightarrow B$ donde $f(x) = x + 4$.
Encuentre:

$$f(3) =$$

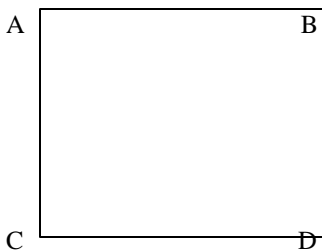
$$f(5) =$$

$$f(7) =$$

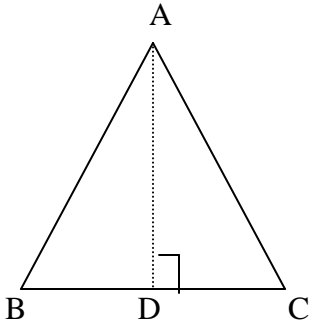
$$f(9) =$$

Polígonos

14.- ABCD es un cuadrado cuyo perímetro es 68 metros, ¿cuál es la longitud del lado de este cuadrado?

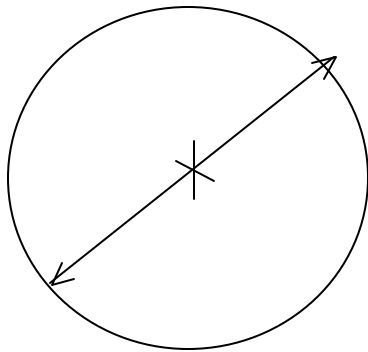


- 15.- ABC es un triángulo isósceles tal que $AB=AC=5$ metros, $AD=4$ metros, ¿cuál es el área del triángulo?



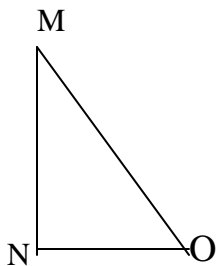
Círculo

- 16.- En la figura se presenta un círculo, cuyo diámetro mide 6 centímetros. Determine su área.



Teorema de Pitágoras

- 17.- Determine la longitud de la hipotenusa del triángulo MNO que se muestra a continuación, donde $MN=4$ y $NO=3$



Factorización

- 18.- Factorice las siguientes expresiones:

$$x^2 + 8x + 16 =$$

$$(x^2 - y^2) =$$

- 19.- Factorice la siguiente expresión: $x^2 + 10x + 21$

Ecuaciones

- 20.- Resuelva la siguiente ecuación: $4x + 12 = x - 7$
- 21.- La edad actual de Jorge es el triple que la de Pedro y dentro de 10 años será el doble, ¿cuál es la edad actual de cada uno de ellos?

Probabilidad

- 22.- ¿Cuál es la probabilidad de que salga el número 3 al lanzar un dado que tiene seis caras?

Estadística

- 23.- A Jorge le gustan mucho las manzanas y durante una semana se le contabilizó la cantidad de manzanas que había comido diariamente, tal contabilidad se muestra a continuación:

Lunes	5 manzanas
Martes	4 manzanas
Miércoles	6 manzanas
Jueves	8 manzanas
Viernes	3 manzanas
Sábado	4 manzanas
Domingo	5 manzanas

- ¿Cuántas manzanas comió Jorge en promedio, durante esa semana?

PRUEBA DE LENGUAJE PARA LOS ESTUDIANTES DE DÉCIMO AÑO DE EDUCACIÓN BÁSICA

Nombre del Plantel _____
Jornada _____
Edad _____
Sexo _____

1. Del siguiente grupo de palabras identifique cuáles tienen **diptongos (D)** y **triptongos (T)**

cuaderno ()	hiena ()	insinuáis ()	vidrio ()	hierro ()
miau ()	Hierba ()	tierra ()	acariciáis ()	ciudad ()
Aprecio ()	Licuéis ()	lluvia ()	Buey ()	pronunciéis ()

2. De las palabras que se muestran a continuación subraye las que contienen vocales en **hiato**

aeropuerto	trofeo	venía	aire	cuidado
canoa	reúne	faena	María	cuaderno
remedio	aprecio	cambio	viento	ciervo

3. Del siguiente grupo de palabras identifique cuáles son **agudas (A)**, **graves (G)**, **esdrújulas (E)** y **sobreesdrújulas (S)**

azúcar ()	césped ()	metálico ()	López ()	farol ()
arriba ()	jueves ()	cántaro ()	cárcel ()	papá ()
metal ()	ágil ()	estudiarás ()	dígamelo ()	feliz ()

4. En cada una de las oraciones que se muestran a continuación **subraye la palabra correcta** dentro del paréntesis que completa el sentido.

Ej: Esta (medicina, medecina) es muy efectiva para mi enfermedad.

Esta mujer está (medio loca, media loca).

En Canadá (neva , nieva) mucho en los meses de invierno.

Cuando (trona, truena) los perros ladran.

Todas las mañanas salgo a (pasear, pasiarme) con mi perro.

(Pásemen, pásenme) las libretas de español.

El (pavimento, pavimento) de la calle está en malas condiciones.

El automóvil levantó una gran (polvareda, polvorín) en el camino.

5. Subraye **dos sinónimas** de cada una de las palabras que se muestran a continuación:

1. grande:	enorme	uniforme	voluminoso	poroso
2. nativo:	altivo	aborigen	indígena	salvaje
3. gélido:	felino	congelado	rígido	frío
4. soberbio:	flamante	altivo	sedante	arrogante

6. Subraye **dos antónimas** de cada una de las palabras que se muestran a continuación:

1. gigante:	andante	enano	diminuto	astuto
2. estéril:	fecundo	hostil	frágil	productivo
3. delgado:	aviso	grueso	gordo	pegajoso
4. pobre:	tacaño	acaudalado	osco	opulento

7. Las palabras pueden clasificarse como: **augmentativas (1), diminutivas (2), despectivas (3), simples nombres (4)**, de acuerdo a esto clasifique cada una de las palabras que se muestran a continuación:

casota ()	banderín ()	librito ()	corazón ()
gentuza ()	Carmencita ()	cuchillo ()	perrote ()

8. Señale el **correcto significado** con una **X**

De súbito	de apuro ()	a diestra y siniestra	a golpes ()
	de corrido ()		sin orden ()
	de repente ()		duramente ()
óbolo	una fruta ()	antiquísimo	muy antiguo ()
	de forma de huevo ()		muy pobre ()
	contribución ()		muy rápido ()

9. Complete con el **verbo** adecuado, el sentido de las siguientes oraciones

Tú _____ bellas canciones. (cantar)
Ellas _____ el puente. (cruzar)
Nosotros _____ al desfile. (ir)
Vicente y tú _____ juntos. (salir)
Mi madre tú y yo _____ ayer. (pasear)
El trabajo y el estudio _____ al hombre. (dignificar)

10. Identifique de las siguientes oraciones el **sujeto y el predicado**, con sus respectivos **componentes**

Claudia, la enfermera, te ama
Las luces como estrellas comenzaron a brillar

Su hacienda se llamaba Sinesia
La pera y la manzana son frutas de la sierra

11. A continuación identifique las **frases (F)** y **oraciones (O)**:

- Los cañaverales del trópico ().
- Los niños juegan y ríen ().
- La responsabilidad de los padres ().
- Mañana iremos al cine ().
- Pedrito juega a todas horas ()

12. Las oraciones pueden ser clasificadas como: admirativas (1), interrogativas (2), enunciativas (3), exhortativas (4), exclamativas (5), negativas (6), desiderativas (7), afirmativas (8), sobre esta base clasifique las oraciones que se muestran a continuación

- 1. Ven acá _____
- 2. ¿Llegará el hombre a Marte? _____
- 3. De hecho los exámenes fueron muy fáciles _____
- 4. Caballero, déme una limosna _____
- 5. No tengo hambre _____
- 6. ¡A dormir! _____
- 7. Ojalá pudiera volver al mar _____

13. Corrija de ser necesario

- Hubieron muchos invitados _____
- Habrán fiestas _____
- Han habido disgustos _____
- Se premiaron a los triunfadores _____
- Se vende estos lotes _____
- Se compone zapatos _____
- Temo de que hagan mal _____

14. Coloque la **tilde** donde corresponda

Te llame cuando aun no salias de la casa
Aun no se si te habran servido el te
Ni tu ni el se interesaron por mi
Jose me pide que le de una amplia explicacion
Aquel dia si estuve alli, se eso deseas saber

Recibi el mensaje, mas lo he olvidado
Aun tu deberas entonar el si en el piano
Se responsable y caritativo y seras apreciado por Dios

15. Lectura comprensiva

Einstein y el átomo

Albert Einstein fue uno de los mas grandes científicos de la historia. Nació en Ulm (Alemania) en el año de 1879 y murió en Princenton (Estados Unidos) en 1955. Con su teoría de la relatividad introdujo una idea nueva de espacio-tiempo y abrió el camino al desarrollo de la Física Nuclear, que ha creado no sólo las terribles bombas atómicas capaces de destruir el mundo entero, sino también esos útiles reactores nucleares de los que se obtiene energía eléctrica.

La base de toda esta revolución científica radica en un ser microscópico que se conoce con el nombre de átomo. Cada átomo es como un sistema solar infinitamente pequeño. Tan pequeño que casi lo podríamos comparar con la nada. Su centro o núcleo es semejante al sol, que es el centro del sistema solar. En torno al núcleo giran los electrones a velocidades fantásticas, dando millones de vueltas por segundo. Cuando se rompen los núcleos, los átomos sueltan sus "planetas", despidiendo su energía. Esta energía liberada conjuntamente por trillones de átomos es terrible y descomunal, si no es controlada.

Conteste las siguientes preguntas:

¿A qué edad murió Einstein y dónde?

¿Qué cambios introdujo Einstein con su teoría de la relatividad?

¿Cuál es la base de la revolución científica introducida por Einstein?

¿Cuáles son los componentes de átomo?

¿Cuándo sueltan su energía los átomos?

¿Por qué se dice que es infinito el poder del átomo?

ANEXO 3

Coeficientes de las cinco primeras componentes principales de los datos originales

	1	2	3	4	5		1	2	3	4	5
Var 1	-0,08	0,207	-0,289	0,034	0,348	Var 31	0,02	0,02	-0,012	-0,024	0,009
Var 2	-0,072	-0,034	0,05	0,032	0,017	Var 32	0,104	0,138	0,07	-0,019	0,01
Var 3	0,281	0,221	-0,035	0,002	0,065	Var 33	4.378	5.643	0,172	0,263	-0,064
Var 4	0,04	0,035	-0,048	0,023	-0,005	Var 34	1.820	0,218	0,701	-2.354	-0,164
Var 5	0,316	0,352	0,187	-0,124	0,09	Var 35	0,893	-0,067	0,326	-1.013	-0,142
Var 6	0,044	0,042	0,007	0,042	0,006	Var 36	0,552	-0,341	0,055	-0,483	-0,269
Var 7	0,356	0,364	0,076	-0,032	-0,015	Var 37	0,78	-0,204	0,311	0,435	-0,448
Var 8	0,012	0,003	0,016	0,005	-0,005	Var 38	1.410	-0,503	0,309	0,998	-1.728
Var 9	0,113	0,172	0,083	-0,014	-0,041	Var 39	0,391	-0,119	0,141	0,214	-0,257
Var 10	0,027	0,025	0,025	0	-0,011	Var 40	0,183	-0,05	0,048	0,059	-0,111
Var 11	0,158	0,234	-0,118	-0,203	0,118	Var 41	0,865	-0,241	0,361	0,246	0,191
Var 12	0,029	0,057	-0,013	0,019	0,013	Var 42	1.204	-0,146	0,34	0,788	0,823
Var 13	0,216	0,513	0,256	0,081	-0,212	Var 43	1.640	-0,441	0,351	0,621	1.243
Var 14	0,035	0,027	-0,015	0,033	0,073	Var 44	0,406	-0,07	0,065	-0,007	0,15
Var 15	0,589	0,481	-0,056	0,232	-0,141	Var 45	0,318	0,023	0,169	-0,005	0,162
Var 16	0,093	0,087	0,041	0,109	-0,009	Var 46	0,163	0,068	0,09	-0,049	-0,004
Var 17	0,181	0,142	0,034	0,033	0,007	Var 47	0,46	0,067	0,055	0,001	0,208
Var 18	0,303	0,33	0,055	0,081	0,019	Var 48	0,323	0,02	0,023	-0,077	0,121
Var 19	0,167	0,257	0,187	-0,081	-0,011	Var 49	0,856	-0,229	-0,065	0,006	0,126
Var 20	0,334	0,334	-0,046	-0,015	0,021	Var 50	0,649	0,032	0,016	-0,17	0,168
Var 21	0,282	0,347	0,025	-0,181	0,035	Var 51	0,623	0,039	0,114	-0,085	0,222
Var 22	-0,004	-0,006	-0,005	0,004	-0,004	Var 52	0,729	-0,207	0,143	0,079	0,245
Var 23	0,107	0,092	0,04	-0,047	0	Var 53	0,634	-0,168	-0,07	0,004	-0,027
Var 24	0,099	0,065	0,061	-0,071	0,034	Var 54	0,409	-0,05	-0,14	0,023	-0,072
Var 25	0,065	0,103	0,004	-0,022	-0,006	Var 55	0,572	0,056	0,011	0,142	-0,059
Var 26	0,101	0,088	0,038	0,011	-0,017	Var 56	0,793	0,027	0,193	-0,063	0,122
Var 27	0,269	0,446	-0,13	0,032	0,03	Var 57	0,215	0,033	-0,032	0,036	-0,086
Var 28	0,116	0,187	-0,094	0,019	0,017	Var 58	2.972	0,005	-3.902	-0,245	-0,022
Var 29	0,131	0,218	-0,047	0,032	0,023	Var 59	0,785	0,071	0,058	0,031	0,19
Var 30	0,068	0,077	-0,009	0,001	-0,002	Var 60	14.559	-1.644	0,459	0,072	-0,017

ANEXO 4

Coeficientes de las veinte primeras componentes principales de las variables aleatorias estandarizadas

	1	2	3	4	5		1	2	3	4	5
Var 1	-0,038	0,108	-0,386	-0,192	0,199	Var 31	0,073	0,126	-0,103	0,186	-0,458
Var 2	-0,147	0,028	0,393	0,005	0,139	Var 32	0,322	0,225	0,016	0,028	-0,078
Var 3	0,352	0,121	-0,086	-0,124	0,17	Var 33	0,787	0,421	-0,22	-0,082	0,181
Var 4	0,148	0,112	-0,171	0,207	0,047	Var 34	0,573	-0,005	0,163	-0,277	-0,193
Var 5	0,283	0,132	0,11	-0,107	0,13	Var 35	0,542	-0,111	0,228	-0,233	-0,148
Var 6	0,266	0,239	-0,184	0,476	-0,088	Var 36	0,204	-0,221	-0,028	-0,011	-0,349
Var 7	0,372	0,308	0,132	-0,041	-0,155	Var 37	0,417	-0,353	0,069	0,373	0,152
Var 8	0,155	0,05	0,384	-0,042	0,159	Var 38	0,425	-0,374	0,145	0,38	0,17
Var 9	0,381	0,323	0,195	0,084	0,389	Var 39	0,378	-0,308	0,06	0,267	0,137
Var 10	0,228	0,101	0,233	-0,15	0,222	Var 40	0,324	-0,349	0,133	0,162	0,204
Var 11	0,24	0,284	0,087	-0,116	0,002	Var 41	0,384	-0,245	0,148	0,062	-0,108
Var 12	0,108	0,207	-0,174	0,343	-0,358	Var 42	0,49	-0,259	0,021	0,043	0,159
Var 13	0,26	0,392	0,27	0,274	-0,014	Var 43	0,526	-0,258	0,061	0,122	-0,002
Var 14	0,081	-0,111	-0,152	0,038	0,278	Var 44	0,425	-0,054	0,21	-0,039	-0,107
Var 15	0,449	-0,015	0,006	-0,048	0,3	Var 45	0,42	-0,007	0,301	-0,229	0,106
Var 16	0,262	0,168	-0,163	0,439	0,259	Var 46	0,433	0,168	0,124	-0,071	-0,013
Var 17	0,336	0,183	0,022	0,213	0,292	Var 47	0,483	0,08	0,165	-0,209	-0,014
Var 18	0,436	0,282	0,029	0,14	0,239	Var 48	0,365	-0,024	-0,048	-0,116	0,011
Var 19	0,265	0,13	0,296	-0,278	0,212	Var 49	0,457	-0,341	-0,13	0,017	-0,066
Var 20	0,418	0,076	-0,044	-0,205	-0,056	Var 50	0,492	-0,19	-0,163	-0,391	-0,319
Var 21	0,364	0,269	0,017	-0,216	-0,073	Var 51	0,503	-0,216	-0,078	-0,311	-0,205
Var 22	-0,067	-0,157	-0,076	0,061	0,154	Var 52	0,351	-0,195	0,025	0,05	-0,226
Var 23	0,257	0,242	-0,018	0,073	-0,295	Var 53	0,391	-0,277	-0,232	0,145	-0,113
Var 24	0,253	0,206	0,11	-0,103	-0,191	Var 54	0,454	-0,249	-0,007	0,06	0,105
Var 25	0,276	0,298	-0,13	0,047	0,038	Var 55	0,487	-0,094	0,078	0,196	0,148
Var 26	0,294	0,321	0,052	0,388	-0,122	Var 56	0,462	0,024	0,052	-0,04	-0,106
Var 27	0,31	0,171	-0,574	-0,093	0,122	Var 57	0,321	0,109	0,047	0,322	-0,188
Var 28	0,365	0,153	-0,571	-0,244	0,216	Var 58	0,5	-0,209	-0,351	-0,036	-0,06
Var 29	0,357	0,127	-0,56	-0,209	0,209	Var 59	0,506	0,012	0,008	-0,086	-0,185
Var 30	0,258	0,305	-0,188	0,291	-0,387	Var 60	0,915	-0,368	0,021	0,035	-0,099

	6	7	8	9	10		6	7	8	9	10
Var 1	-0,275	0,212	-0,065	0,21	0,187	Var 31	-0,096	0,028	-0,026	0,133	-0,093
Var 2	-0,026	0,125	0,13	-0,039	0,382	Var 32	0,04	-0,069	0,002	0,281	-0,157
Var 3	-0,192	0,089	0,234	-0,101	-0,074	Var 33	0,151	-0,045	-0,007	0,002	0,057
Var 4	-0,55	0,159	-0,023	0,119	-0,148	Var 34	0,195	-0,039	0,11	0,274	-0,308
Var 5	0,063	0,092	0,057	0,058	-0,172	Var 35	0,247	0,017	0,22	0,214	-0,358
Var 6	-0,257	0,104	-0,199	-0,098	-0,018	Var 36	0,233	0,166	0,076	-0,12	-0,425
Var 7	-0,164	0,081	-0,03	0,225	0,162	Var 37	0,236	0,154	-0,006	0,216	0,061
Var 8	0,092	0,52	-0,29	-0,109	-0,014	Var 38	0,173	-0,017	-0,221	0,048	0,04
Var 9	0,109	-0,14	0,18	-0,048	-0,157	Var 39	0,174	0,087	-0,329	0,094	0,116
Var 10	0,069	0,558	-0,361	-0,172	-0,094	Var 40	0,164	0,075	0,078	0,295	0,123
Var 11	-0,137	0,049	0,128	-0,074	-0,234	Var 41	0,053	-0,256	-0,176	0,095	-0,069
Var 12	-0,166	0,112	-0,035	0,113	0,063	Var 42	-0,069	-0,049	-0,165	-0,004	0,084
Var 13	0,17	-0,227	0,131	-0,106	0,062	Var 43	-0,109	-0,084	-0,146	0,156	0,033
Var 14	-0,096	0,199	0,423	0,155	0,03	Var 44	-0,527	-0,075	-0,199	0,122	-0,051
Var 15	0,037	-0,194	0,021	-0,211	-0,164	Var 45	-0,501	-0,018	-0,139	-0,035	-0,03
Var 16	-0,083	0,148	0,03	-0,23	-0,165	Var 46	-0,025	-0,04	-0,192	0,139	0,051
Var 17	-0,074	-0,09	0,096	0,039	-0,211	Var 47	-0,284	-0,188	-0,222	0,077	0,173
Var 18	0,002	-0,25	0,029	0,035	0,034	Var 48	-0,308	-0,148	0,109	0,306	-0,074
Var 19	0,137	0,155	-0,012	-0,055	-0,049	Var 49	-0,204	-0,093	0,278	-0,005	-0,179
Var 20	0,151	-0,268	0,025	0,069	0,407	Var 50	-0,012	0,205	0,005	-0,245	-0,011
Var 21	0,126	-0,044	0,225	0,051	0,319	Var 51	0,1	0,175	0,136	-0,253	0,095
Var 22	-0,044	0,159	0,241	0,334	0,013	Var 52	0,125	0,058	0,01	0,016	0,391
Var 23	0,051	0,343	0,128	0,171	0,02	Var 53	0,029	-0,057	-0,199	-0,147	-0,063
Var 24	0,002	0,282	0,004	0,206	0,094	Var 54	-0,08	-0,021	0,328	-0,336	0,106
Var 25	0,233	0,238	0,072	0,241	0,134	Var 55	-0,019	0,009	0,315	-0,058	0,151
Var 26	0,142	0,109	0,166	-0,158	-0,007	Var 56	-0,083	0,299	-0,036	-0,139	-0,026
Var 27	0,255	-0,027	-0,34	0,089	-0,146	Var 57	-0,066	0,001	0,059	-0,269	-0,024
Var 28	0,074	-0,087	-0,188	-0,16	-0,016	Var 58	-0,108	-0,1	0,117	-0,096	0,039
Var 29	0,155	0,045	-0,082	0,023	0,126	Var 59	-0,083	-0,143	-0,024	-0,32	0,195
Var 30	0,156	0,015	0,027	-0,064	0,125	Var 60	-0,021	-0,002	-0,006	-0,011	0,009

	11	12	13	14	15		11	12	13	14	15
Var 1	-0,031	0,128	-0,094	-0,029	-0,259	Var 31	-0,149	-0,041	0,26	-0,292	0,147
Var 2	0,072	-0,067	0,04	-0,334	-0,121	Var 32	-0,405	-0,058	0,018	0,005	-0,11
Var 3	0,03	0,094	0,194	0,079	0,357	Var 33	-0,162	-0,016	0,103	0,005	-0,04
Var 4	-0,078	0,011	0,257	0,05	-0,09	Var 34	0,167	0,176	-0,052	-0,066	-0,148
Var 5	-0,041	-0,344	0,396	0,193	-0,064	Var 35	0,2	0,205	-0,028	-0,071	-0,068
Var 6	-0,062	-0,115	0,113	0,079	-0,254	Var 36	0,001	0	0,006	-0,108	-0,024
Var 7	-0,26	0,123	0,163	0,233	0,071	Var 37	0,013	-0,038	0,168	0,12	-0,045
Var 8	-0,246	0,08	-0,159	-0,082	0,251	Var 38	0,059	0,226	0,188	0,031	0,054
Var 9	0,099	-0,028	0,166	-0,147	-0,151	Var 39	0,191	0,07	0,263	-0,05	0,024
Var 10	-0,181	0,133	-0,107	0	0,049	Var 40	-0,051	0,034	0,252	-0,107	0,198
Var 11	0,013	0,222	0,005	0,045	0,105	Var 41	-0,162	-0,157	-0,121	0,118	0,125
Var 12	-0,288	0,292	-0,077	0,066	-0,248	Var 42	-0,182	-0,119	-0,269	-0,025	0,03
Var 13	-0,109	0,106	0,033	-0,022	-0,127	Var 43	0,09	-0,128	-0,143	-0,09	-0,148
Var 14	-0,116	-0,345	-0,194	-0,102	-0,065	Var 44	0,003	0,038	0,039	-0,308	0,076
Var 15	-0,289	0,05	0,106	0,127	0,05	Var 45	0,064	-0,105	0,036	-0,219	0,134
Var 16	0,144	-0,092	-0,18	-0,201	0,055	Var 46	0,347	0,014	-0,058	-0,025	-0,239
Var 17	0,179	-0,077	-0,249	0,352	-0,093	Var 47	0,233	0,041	-0,066	0,034	-0,146
Var 18	0,092	-0,005	0,091	-0,042	0,049	Var 48	0,037	-0,229	-0,193	0,032	0,25
Var 19	-0,08	-0,069	0,035	0,03	-0,482	Var 49	-0,041	-0,033	-0,007	-0,002	0,005
Var 20	-0,18	0,042	0,114	0,161	0,14	Var 50	-0,146	-0,127	0,12	0,053	-0,172
Var 21	0,156	0,154	0,092	-0,195	0,028	Var 51	-0,131	-0,285	0,172	0,031	-0,086
Var 22	-0,124	-0,135	-0,076	0,094	0,033	Var 52	0,012	-0,305	-0,154	0,133	-0,057
Var 23	0,029	-0,104	0,019	-0,241	0,129	Var 53	0,032	0,082	-0,066	-0,126	-0,168
Var 24	0,125	0,195	-0,265	0,262	0,024	Var 54	-0,192	0,309	-0,189	-0,19	-0,047
Var 25	0,102	0,161	-0,181	0,042	0,228	Var 55	-0,287	0,132	-0,204	-0,046	-0,08
Var 26	0,142	-0,2	-0,149	-0,115	0,022	Var 56	0,29	-0,223	-0,093	0,199	0,15
Var 27	-0,045	-0,036	-0,038	-0,001	0,012	Var 57	0,176	0,035	0,043	0,356	0,197
Var 28	0,045	0,02	-0,001	-0,195	0,106	Var 58	-0,028	0,281	-0,025	0,021	0,118
Var 29	0,11	-0,146	0,034	-0,085	0,082	Var 59	0,128	0,051	-0,062	0,142	-0,009
Var 30	0,087	-0,149	0,047	-0,188	0,015	Var 60	0,051	0,025	-0,037	0,003	-0,018

	16	17	18	19	20		16	17	18	19	20
Var 1	-0,096	0,023	0,12	0,014	0,027	Var 31	0,191	-0,031	0,139	-0,299	-0,025
Var 2	-0,147	0,167	0,365	-0,184	-0,075	Var 32	0,034	-0,116	-0,123	0,157	-0,046
Var 3	0,055	0,142	-0,058	0,06	0,276	Var 33	0,005	-0,015	0,034	0	-0,067
Var 4	-0,045	0,01	0,003	-0,066	-0,228	Var 34	-0,046	0,016	-0,041	-0,113	-0,092
Var 5	-0,202	0,32	0,034	-0,118	0,213	Var 35	-0,055	0,071	-0,008	-0,05	-0,106
Var 6	-0,151	0,033	-0,006	-0,155	0,002	Var 36	-0,268	0,146	0,165	0,195	0,074
Var 7	0,002	-0,005	-0,246	0,02	0,111	Var 37	-0,023	0,145	-0,041	0,14	0,114
Var 8	-0,073	-0,085	0,055	-0,09	0,045	Var 38	0,066	0,04	-0,075	0,066	-0,163
Var 9	0,056	-0,113	-0,133	0,06	-0,002	Var 39	0,079	-0,163	-0,082	-0,157	0,07
Var 10	-0,004	-0,023	0,127	-0,143	-0,073	Var 40	-0,034	-0,126	0,051	0,166	0,006
Var 11	-0,126	-0,345	0,272	0,028	0,506	Var 41	0,203	-0,087	0,19	-0,114	0,052
Var 12	0,05	0,104	0,093	0,115	0,027	Var 42	0,227	0,189	0,093	0,227	0,042
Var 13	0,081	-0,176	-0,189	0,089	0,075	Var 43	0,15	0,151	0,271	0,122	0,179
Var 14	0,234	-0,21	0,092	0,022	-0,092	Var 44	-0,124	-0,016	0,008	0,103	-0,025
Var 15	0,134	-0,067	0,201	0,025	-0,161	Var 45	0,068	0,049	-0,183	0,055	-0,041
Var 16	0,05	0,034	-0,101	0,014	0,187	Var 46	-0,075	-0,255	-0,148	0,033	0,124
Var 17	0,151	0,01	0,117	-0,305	0,024	Var 47	-0,044	0,04	0,081	0,272	0,027
Var 18	-0,015	-0,13	0,346	-0,134	-0,018	Var 48	-0,207	0,206	-0,156	-0,12	-0,081
Var 19	0,016	0,161	0,051	-0,01	-0,206	Var 49	0,227	0,015	0,013	-0,111	0,039
Var 20	-0,178	0,01	0,124	-0,006	-0,043	Var 50	0,095	-0,153	-0,036	0,012	-0,017
Var 21	0,115	0,058	0,075	-0,24	-0,041	Var 51	0,129	-0,055	0,026	0,097	0,083
Var 22	-0,359	-0,493	0,077	0,131	-0,212	Var 52	-0,099	-0,106	-0,164	-0,303	0,237
Var 23	0,303	-0,006	-0,154	-0,099	-0,129	Var 53	-0,183	-0,228	-0,144	-0,274	0,008
Var 24	0,23	0,033	0,056	-0,022	0,113	Var 54	-0,139	0,099	-0,208	-0,029	-0,018
Var 25	0,063	0,117	-0,027	0,116	-0,12	Var 55	-0,214	0,172	-0,159	-0,106	0,112
Var 26	-0,028	0,008	-0,031	0,218	-0,066	Var 56	-0,2	-0,136	-0,011	0,087	-0,166
Var 27	0,02	0,145	-0,082	0,058	0,002	Var 57	-0,092	0,113	0,099	-0,061	-0,34
Var 28	0,072	-0,123	-0,036	0,016	-0,075	Var 58	-0,162	-0,004	0,171	-0,066	-0,059
Var 29	-0,164	0,121	-0,094	-0,029	0,039	Var 59	0,174	-0,127	-0,001	0,11	-0,15
Var 30	-0,133	0,029	0,248	0,245	0,043	Var 60	-0,031	0,002	-0,004	-0,008	-0,014