



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
INGENIERÍA EN ESTADÍSTICA INFORMÁTICA

“Análisis socio-económico de las principales variables categóricas
de la Encuesta Nacional de Ingresos y Gastos de los hogares del
Cantón Guayaquil 2011-2012”

TESIS DE GRADO

Previa a la obtención del Título de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

Presentado por:

Jhon Henry Quimí Loza

Guayaquil - Ecuador

2015

AGRADECIMIENTO

Esta tesis es el resultado muchos meses de trabajo, esfuerzo y sacrificio que gracias a la ayuda de dios lo culminé con gran éxito. Para así poder obtener mi título de tercer nivel y ser un profesional al servicio de la patria.

Agradezco a mi director de tesis el Matemático John Ramírez Figueroa por haber confiado en mis capacidades y por haberme dado la oportunidad de realizar este proyecto de tesis bajo su supervisión.

A mis padres y hermanos que siempre confiaron y esperaron con ansias este gran momento de mi vida.

Estoy muy contento por haber formado parte de esta grandiosa institución como lo es la Escuela Superior Politécnica del Litoral que forma a los mejores profesionales del país.

Jhon Henry Quimí Loza

DEDICATORIA

Dedico este proyecto de tesis a Dios y a mi familia.

A Dios por darme la fortaleza, la sabiduría y el empeño de superarme en la vida para lograr así ser un profesional. Él quién me ayuda en los momentos complicados dando me fe y el aliento para seguir adelante superando cualquier adversidad que se presente.

A mis padres y hermanos quienes siempre me han apoyado en cada meta que establezco en mi vida, ellos han sido los pilares fundamentales para alcanzar todos mis objetivos y lograr el éxito.

Jhon Henry Quimí Loza

TRIBUNAL DE GRADUACIÓN

MÁSTER VANESSA SALAZAR
PRESIDENTE TRIBUNAL

MÁSTER JOHN RAMÍREZ
DIRECTOR DE TESIS

PH.D. OMAR RUÍZ
VOCAL PRINCIPAL

DECLARACIÓN EXPRESA

La responsabilidad del contenido de esta Trabajo final de graduación de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la Escuela Superior Politécnica del Litoral

Jhon Henry Quimí Loza

RESUMEN

La finalidad de esta tesis es presentar un análisis de la situación socio económica de los hogares que habitan en el cantón Guayaquil en base a las variables cualitativas y cuantitativas obtenidas durante la Encuesta Nacional de Ingresos y Gastos de los Hogares Urbanos y Rurales 2011- 2012 realizada en el país por el Instituto Nacional de Estadística y Censo del Ecuador.

Para el proyecto se seleccionaron las principales variables categóricas de la ENIGHUR 2011-2012, las mismas que fueron elegidas por su alto nivel de significancia que presentaron en esta investigación.

Se utilizó la técnica multivariada el Análisis de componentes principales categórico para reducir el número de variables utilizadas en esta encuesta, con el propósito conservar la mayor cantidad de variabilidad de los datos proporcionados. También se desarrolló un modelo de regresión logística cuyo objetivo fue predecir los ingresos menores a los 334,72 dólares, correspondientes al percentil 30 del total de ingresos que poseen los hogares del cantón Guayaquil.

El estudio mostró tres agrupaciones donde los hogares fueron aglomerados por el tipo de características de las viviendas en que habitan y el nivel calidad de los materiales con las que fueron construidas.

Mediante el análisis de las variables de ingresos, gastos y estilo de vida se obtuvo un gráfico en forma de pirámide, donde se aprecia en la cima a los individuos que poseen un mejor de estilo vida dado a los ingresos que perciben y a los hogares con los menores ingresos en la base de la pirámide.

Se elaboró un modelo de regresión logística que clasificó correctamente el 81.4% de los casos que presentaron una característica de interés analizada entre los individuos.

La información utilizada se encuentra disponible en el sitio web del Instituto Nacional de Estadísticas y Censos (INEC) en la sección Estadística Sociales e ingresando a la opción Ingresos y Gastos.

En lo referente a los sistemas de información los medios de consulta utilizados fueron el sitio web del INEC, el sitio web de las Naciones Unidas y algunos sitios relacionados en la red. Las principales herramientas utilizadas fueron el software SPSS Statistics 17.0, Microsoft Excel y Word 2010.

Contenido

RESUMEN	vi
ÍNDICE DE GRÁFICOS	xi
ÍNDICE DE TABLAS	xiii
INTRODUCCIÓN.....	1
CAPÍTULO 1	3
INTRODUCCIÓN Y OBJETIVOS	3
1.1 GENERALIDADES DEL NIVEL SOCIOECONÓMICO	4
1.2 ÍNDICE DE PRECIOS AL CONSUMIDOR - IPC	6
1.2.1 METODOLOGÍA GENERAL DEL CÁLCULO	7
1.3 ESTRUCTURA DE LA CANASTA BÁSICA DE BIENES Y SERVICIOS.....	8
1.4 ENCUESTA DE ESTRATIFICACIÓN DEL NIVEL DE SOCIOECONÓMICO.	9
1.4.1 ESTRATOS DEL NIVEL SOCIOECONÓMICO	11
1.5 CENSO DE POBLACIÓN Y VIVIENDA	12
1.6 ENCUESTA NACIONAL DE INGRESOS Y GASTOS EN HOGARES URBANOS Y RURALES	13
1.7 ESTUDIOS RELACIONADOS A LA ENIGHUR.....	19
1.8 JUSTIFICACIÓN DE ESTUDIO Y OBJETIVOS	21
1.8.1 JUSTIFICACIÓN DEL ESTUDIO	21
1.9 ORIGEN DE LA INFORMACIÓN	22
1.9.1 FUENTE DE DATOS	22
CAPÍTULO 2	23
MARCO TEÓRICO, METODOLOGÍA Y FUENTE DE DATOS	23
2.1 POBLACIÓN	24
2.2 POBLACIÓN OBJETIVO.....	25
2.3 ANÁLISIS EXPLORATORIO.....	25
2.4 ATRIBUTOS.....	27
2.5 VARIABLE CUALITATIVA O ATRIBUTO	28
2.6 VARIABLE CUANTITATIVA.....	28
2.7 ANÁLISIS MULTIVARIADO DE LOS DATOS.....	28

2.7.1	ESTADÍSTICA MULTIVARIADA EN INVESTIGACIÓN NO EXPERIMENTAL	30
2.7.2	ESCALAMIENTO ÓPTIMO	31
2.7.3	NIVEL DEL ESCALAMIENTO ÓPTIMO	32
2.7.4	GRÁFICOS DE TRANSFORMACIÓN	32
2.7.5	TABLAS DE CONTINGENCIA.....	33
2.7.6	MATRIZ DE BURT.....	34
2.7.7	ANÁLISIS DE CORRESPONDENCIAS	35
2.7.8	ANÁLISIS DE CORRESPONDENCIAS MÚLTIPLE	40
2.7.9	ANÁLISIS DE COMPONENTES PRINCIPALES CATEGÓRICOS	40
2.7.10	METODOLOGÍA DEL ANÁLISIS DE COMPONENTES PRINCIPALES	42
2.7.11	GENERALIDADES DEL ANÁLISIS DE FACTORES Y EL ANÁLISIS DE COMPONENTES PRINCIPALES	43
2.7.12	VECTOR CARACTERÍSTICO Y RAÍZ CARACTERÍSTICA O VALOR PROPIO.....	45
2.7.13	ROTACIÓN DE LAS COMPONENTES.....	47
2.7.14	ROTACIÓN DE TIPO ORTOGONAL	48
2.7.15	COMUNALIDADES, VARIANZA Y COVARIANZA	49
2.7.16	APROXIMACIÓN DEL ANÁLISIS DE COMPONENTES PRINCIPALES CON VARIABLES CATEGÓRICAS	50
2.7.17	CARACTERÍSTICAS DEL ANÁLISIS DE COMPONENTES PRINCIPALES CATEGÓRICOS.....	51
2.7.18	REGRESIÓN LOGÍSTICA.....	52
2.8	FLUJO DE LA INFORMACIÓN	60
2.9	¿QUÉ ES UNA VIVIENDA?.....	61
2.10	DESCRIPCIÓN DE HOGAR	61
2.10.1	MIEMBROS Y NO MIEMBROS DEL HOGAR.....	62
2.10.2	INGRESOS DEL HOGAR	63
2.10.3	PERCEPTORES DE INGRESOS	64
2.11	INGRESO CORRIENTE	64
2.12	INGRESOS NO REGULARES DEL HOGARES	65
2.13	GASTOS DEL HOGAR	65
2.13.1	GASTO CORRIENTE.....	66

CAPÍTULO 3	67
ANÁLISIS DESCRIPTIVO DE LA ENIGHUR	67
3.1. INFORMACIÓN ECONÓMICA DE HOGARES.....	68
3.2. DISTRIBUCIÓN DE FRECUENCIAS FACTORES DESCRIPTIVOS DE LAS VIVIENDAS, EQUIPAMIENTO Y PERCEPCIÓN DEL NIVEL DE VIDA DEL HOGAR.....	68
CAPÍTULO 4	88
IDENTIFICACIÓN DE UNA ESTRUCTURA USANDO LAS PRINCIPALES VARIABLES DE LA ENIGHUR 2011-2012 MEDIANTE EL ANÁLISIS DE COMPONENTES PRINCIPALES CATEGÓRICO.....	90
4.1. REDUCCIÓN DE VARIABLES RELACIONADAS A LA VIVIENDA Y EL HOGAR.....	90
4.1.1. REDUCCIÓN DE VARIABLES RELACIONADAS AL EQUIPAMIENTO Y CARACTERÍSTICAS DEL HOGAR.....	95
4.1.2. REDUCCIÓN DE VARIABLES RELACIONADAS A LA INFORMACIÓN DE LOS MIEMBROS DEL HOGAR Y CARACTERÍSTICAS OCUPACIONALES.....	96
4.1.3. REDUCCIÓN DE VARIABLES RELACIONADAS A INGRESOS Y TRANSACCIONES FINANCIERAS DE LOS PERCEPTORES.	100
4.1.4. REDUCCIÓN DE VARIABLES RELACIONADAS AL GASTO DENTRO DEL HOGAR.	101
4.1.5. REDUCCIÓN DE VARIABLES RESULTANTES EN LA DISMINUCIÓN PARCIAL EN CADA SECCIÓN	103
4.2. MODELO DE REGRESIÓN LOGÍSTICA PARA EVALUAR EL DESARROLLO DEL NIVEL SOCIOECONÓMICO DE LOS HOGARES DEL CANTÓN GUAYAQUIL ...	108
4.2.1. IMPLEMENTACIÓN DE UN MODELO DE REGRESIÓN LOGÍSTICA PARA PREDECIR EL INGRESO DE LOS HOGARES.....	109
CAPÍTULO 5	120
CONCLUSIONES Y RECOMENDACIONES.....	120
5.1. CONCLUSIONES.....	120
5.2. RECOMENDACIONES.....	123
REFERENCIAS BIBLIOGRÁFICAS	126

ÍNDICE DE GRÁFICOS

Gráfico 1.1	Distribución por estratos del nivel socioeconómico	11
Gráfico 1.2	Curva de Lorenz para la brecha del ingreso per cápita	17
Gráfico 2.1	Diagrama del flujo de la información en la ENIGHUR 3	60
Gráfico 2.2	Estructura del ingreso corriente 4 Fuente: INEC (2011).....	64
Gráfico 2.3	Estructura del gasto corriente 5 Fuente: INEC (2011).....	66
Gráfico 3.2.	Vía de acceso a la vivienda	68
Gráfico 3.3.	Vía de acceso a la vivienda	69
Gráfico 3.4.	Material predominante del techo	70
Gráfico 3.5.	Material predominante de las paredes	71
Gráfico 3.6.	Material predominante del piso	72
Gráfico 3.7.	Existencia de cuartos para negocio en la vivienda	73
Gráfico 3.8.	Números de cuartos para negocio.....	74
Gráfico 3.9.	Sitio de la vivienda donde se cocinan los alimentos.....	75
Gráfico 3.10.	Combustible con que principalmente cocinan	76
Gráfico 3.11.	Tipo de servicio higiénico del hogar	77
Gráfico 3.12.	Vía de acceso principal al agua de uso diario	78
Gráfico 3.13.	Modo de abastecimiento del agua.....	79
Gráfico 3.14.	Tiene este hogar servicio de telefonía convencional	80
Gráfico 3.15.	Algún miembro del hogar tiene teléfono celular	81
Gráfico 3.16.	Número de líneas de celular.....	83
Gráfico 3.17.	Tipo de alumbrado con que cuenta el hogar	84
Gráfico 3.18.	Percepción del nivel de vida mediante los ingresos	85
Gráfico 3.20.	Situación económica de su hogar.....	86
Gráfico 3.21.	Nivel de vida del hogar	87
Gráfico 4.1	Individuos por características de la vivienda	92
Gráfico 4.2	Individuos por características de la vivienda reducido.....	94

Gráfico 4.3 Individuos por percepción de nivel de vida y características ocupacionales.	97
Gráfico 4.4 Individuos por percepción de nivel de vida y características ocupacionales reducido.	99
Gráfico 4.5 Individuos por monto de gasto destinado para el hogar	102
Gráfico 4.6 Individuos por nivel socio económico	105
Gráfico 4.7 Individuos por nivel socio económico reducido	108

ÍNDICE DE TABLAS

Tabla 1.1:	Divisiones según CCIF pertenecientes a la canasta básica.....	9
Tabla 1.2:	Puntaje por Estratos del Nivel Socioeconómico.....	12
Tabla 2.1:	Tabla de Contingencia con a Filas y b Columnas	33
Tabla 2.2:	Matriz de Contingencia con de Frecuencias Relativas.....	34
Tabla 2.3:	Matriz de Burt $G'G$ para una Matriz G	35
Tabla 3.3:	Vía de acceso principal a la vivienda	68
Tabla 3.4:	Tipo de vivienda.....	69
Tabla 3.5:	Material predominante del techo	70
Tabla 3.6:	Material predominante de las paredes	71
Tabla 3.7:	Material predominante del piso	72
Tabla 3.8:	Existencia de cuartos para negocio en la vivienda	73
Tabla 3.9:	Número de cuartos para negocio	74
Tabla 3.10:	Sitio de la vivienda donde se cocinan los alimentos	75
Tabla 3.11:	Combustible con que principalmente cocinan los alimentos	76
Tabla 3.12:	Tipo de servicio higiénico del hogar	77
Tabla 3.13:	Vía de acceso principal al agua de uso diario.....	78
Tabla 3.14:	Modo de abastecimiento de agua	79
Tabla 3.15:	Tiene este hogar servicio de telefonía convencional.....	80
Tabla 3.16:	Algún miembro del hogar tiene teléfono celular	81
Tabla 3.17:	Distribución de hogares por posesión de telefonía	82
Tabla 3.18:	Número de líneas de celular	82
Tabla 3.19:	Tipo de alumbrado con que cuenta el hogar.....	83
Tabla 3.20:	Percepción del nivel de vida mediante los ingresos.....	84
Tabla 3.22:	Situación económica de su hogar	85
Tabla 3.23:	Nivel de vida del hogar.....	86
Tabla 4.1:	Matriz de Componentes Sección I. Datos de la Vivienda y Hogar (12 Variables).....	91

Tabla 4.2: Matriz de Componentes Sección I. Datos de la Vivienda y Hogar (9 Variables).....	93
Tabla 4.3: Matriz de Componentes Sección III, IV y V. Miembros del Hogar, Percepción del Nivel de Vida y Características Ocupacionales (14 Variables).....	96
Tabla 4.4: Matriz de Componentes Sección III, IV y V. Miembros del Hogar, Percepción del Nivel de Vida y Características Ocupacionales (10 V.)	98
Tabla 4.5: Matriz de Componentes de Gastos Dentro del Hogar (7 Variables).....	101
Tabla 4.6: Matriz de Componentes de las Secciones I, II, III, IV, V	104
Tabla 4.7: Matriz de Componentes de las Secciones I, II, III, IV, V	106
Tabla 5.1: 33 Resumen del procesamiento de los casos.....	110
Tabla 5.2: 34 Codificación de la variable dependiente.....	111
Tabla 5.3: 35 Codificaciones de las variables categóricas.....	112
Tabla 5.4:36 Tabla de Clasificación	113
Tabla 5.5:37 Resumen del modelo	114
Tabla 5.6:38 Prueba Omnibus del Modelo de Coeficientes	114
Tabla 5.6:39 Tabla de Clasificación	116
Tabla 5.7:40 Tabla de Variables en la ecuación	118

INTRODUCCIÓN

En la actualidad, la ciencia y la tecnología son áreas que permanecen en movimiento continuo, cambiando e innovando.

En las últimas décadas, la estadística ha sido aplicada en casi todas las fases de la actividad humana. Existe un incremento del uso de esta ciencia logrando ser aplicada en muchos campos de estudio, tales como: los campos científicos, campos de la medicina, campos en finanzas, tecnología, etc.

Pocas áreas en nuestro entorno han tenido el impacto del desarrollo de la estadística como ha ocurrido con la ingeniería, la medicina y la administración. Sería difícil sobreestimar las contribuciones que realiza la estadística a la producción, la industria, al uso correcto de materiales y fuerza de trabajo, al desarrollo de nuevos productos y a la investigación básica. Hoy en día la estadística se ha convertido en la herramienta vital para los ingenieros, economistas, científicos, médicos, entre muchos otros profesionales., esta le permite al investigador encontrar patrones, relaciones, fenómenos sujetos a variación de las variables analizadas para así poder predecirlos y controlarlos eficazmente.

En enero del 2014, luego de haber culminados los estudios de pregrado de la Escuela Superior Politécnica y de haber cumplido con todos los créditos necesarios de la malla curricular, se pretende proceder con el proyecto final de graduación (tesis), con la base teórica suficiente brindada durante cinco años de estudios en ESPOL, se toma como objetivo dar comienzo a la tesis, que para efectos de supervisión se encontró bajo la coordinación y guía del Matemático John Ramírez Figueroa tutor de esta tesis.

Se procedió a planificar y efectuar una investigación en base a la población objetivo de los hogares urbanos y rurales cuyos ingresos y gastos fueron investigados en la encuesta nacional durante abril 2011 a marzo 2012 por el INEC.

El tema de la investigación: Análisis Socio-Económico de las principales variables categóricas de la Encuesta Nacional de ingresos y gastos de los hogares del Cantón Guayaquil 2011-2012, mediante la técnica estadística análisis de componente principales categóricos.

CAPÍTULO 1

INTRODUCCIÓN Y OBJETIVOS

Este capítulo trata temas generales relacionados con el entorno socioeconómico en la población, los factores que influyen en su crecimiento o disminución, el comportamiento económico y demográfico de las poblaciones, y las relaciones en función al índice de precios al consumidor. Las secciones iniciales muestran temas generales del entorno económico, canasta básica de bienes y servicios, índice de precios al consumidor – IPC. Además presentan temas importantes de la economía que afectan a las poblaciones, tales como: índice de desarrollo humano, la economía del desarrollo y el impacto de la inflación en la canasta familiar para los ecuatorianos.

En las dos últimas secciones, se aclara la justificación de este estudio, los objetivos generales y específicos que tiene este trabajo, la metodología que

se utiliza, la información analizada y las herramientas que ayudaron a su elaboración y desarrollo.

1.1 GENERALIDADES DEL NIVEL SOCIOECONÓMICO

Cuando la economía de un país pobre mejora, puede suceder que ciertos grupos sociales lo aprovechen dejando a gran parte de la población en la miseria.

Los bajos recursos e ingresos de grandes sectores de la sociedad son grandes dificultades que las personas tienen para poder acceder a alimentos, atención médica, ropa, vivienda y educación, para así mejorar su estilo de vida.

La desigualdad en grandes áreas del planeta genera alarmantes brechas entre los ingresos y recursos de la población, factores como la pobreza y la sobrepoblación son causados por una mala distribución de la riqueza entre la población.

La riqueza acumulada en sólo unos pocos grupos económicos privilegiados, refuerza la pobreza y la crisis social para la mayoría de los ecuatorianos.

La pobreza es multidimensional, es decir, no sólo se relaciona con aspectos materiales como vestuario, alimentación y albergues adecuados, también con

derechos de la población como: oportunidad de una vida saludable y larga, libertad, dignidad y la participación política.

De acuerdo al reporte del programa de desarrollo humano de las Naciones Unidas (PNUD), en el 2013 se destaca el desarrollo de América del Sur que va en constante crecimiento y define al Ecuador con un valor de 0,724 ubicándolo en el grupo de países con “desarrollo humano alto” debido a las mejoras en salud, expectativas de vida, educación y el incremento de la renta bruta per cápita la misma que es de 7,471 dólares por ciudadano ecuatoriano.

Otro factor importante que se presenta es la esperanza de vida al nacer. En la actualidad la esperanza de vida al nacer de los ecuatorianos se encuentra alrededor de los 75.8 años, ubicando al país en el puesto 89 de una lista de 187 clasificados.

El Ecuador se muestra en el PNUD como uno de los países que más aumentó en escalas, ascendiendo diez puestos desde el 2007 al 2012 en el índice de desarrollo humano (IDH).

Ecuador es el tercer país de Sudamérica en avanzar a mayores escalas en la tabla del IDH, pero aún con su avance se sigue ubicando por debajo del 0.741 que corresponde al índice promedio actual de la región (Programa de las Naciones Unidas para el Desarrollo, 2013).

1.2 ÍNDICE DE PRECIOS AL CONSUMIDOR - IPC

Se define como un indicador económico que mide la evolución temporal de los precios en el mercado; índice que va en íntima relación con los productos de la Canasta Familiar Básica.

Los efectos causados por la inflación en el país han impactado tremendamente en el conjunto de bienes y servicios.

Durante el primer trimestre del año 2014 según datos del INEC una familia ecuatoriana está compuesta de 4 miembros la misma que tiene ingresos mensuales de 1,6 salarios básicos unificados lo equivalente a 634,67 dólares, existiendo así en el último mes de este periodo un superávit de 2,48 dólares para la obtención de la canasta básica, representando un costo de 632,19 dólares. En marzo del 2013, la canasta básica se ubicó en 604,25 dólares a diferencia del ingreso mensual del hogar que alcanzó los 593,60 dólares, generando un déficit de 10,65 dólares.

La medición de los precios al consumir final en ocho ciudades del país llegó a 0,70 en marzo de este año, lo que muestra una inflación de 1,13% en los tres primeros meses de ese año.

La ciudad que posee la mayor tasa de inflación es Guayaquil, con el 1,22%, seguido por la ciudad de Esmeraldas con una tasa de 0,89%, mientras que en el límite inferior encontramos a dos ciudades que poseen los menores

porcentajes las cuales son Quito y Cuenca que tienen un índice de inflación de 0,48 y 0,23 respectivamente (Programa de las Naciones Unidas para el Desarrollo, 2013).

1.2.1 METODOLOGÍA GENERAL DEL CÁLCULO

El denominado índice de Laspeyres con base fija es la fórmula que utiliza para el cálculo del IPC, conocido por comparar el valor de una canasta a precios del periodo de referencia frente al valor de la misma canasta a precios del periodo base. El IPC trata de obtener cuánto aumenta o disminuye el valor de compra de una misma canasta de bienes y servicios en dos periodos de tiempo manteniendo inmóvil las cantidades del periodo base. Ecuador es uno de los muchos países que utiliza la Fórmula de Laspeyres para el cálculo de la IPC (Instituto Nacional de Estadísticas y Censos, 2012).

Las cantidades de artículos inmóviles en la canasta de artículos corresponden a una estructura definida por consumo en periodo base. Al mantenerse fija esta estructura de consumo, se asume que la variación del valor de la canasta de artículos da como resultado un efecto puro en la variación de precios.

La fórmula para calcular el índice de Laspeyres con base fija, se define a continuación:

$$I_{t,0} = \frac{\sum_{i=1}^n p_{i,t} * q_{i,0}}{\sum_{i=1}^n p_{i,0} * q_{i,0}}$$

Dónde:

$I_{t,0}$: Índice de precios del periodo t con respecto al periodo base 0.

$p_{i,t}$: Precio del artículo i en el periodo t .

$p_{i,0}$: Precio del artículo i en el periodo base 0.

$q_{i,0}$: Precio del artículo i en el periodo base 0.

1.3 ESTRUCTURA DE LA CANASTA BÁSICA DE BIENES Y SERVICIOS

La estructura de la Canasta de artículos (Bienes y Servicios) del IPC, contiene 12 divisiones, 35 grupos, 68 clases, 98 subclases, 151 productos y 299 artículos (INEC, 2013). En la tabla a continuación se detallan las divisiones que fueron analizadas en la ENIGHUR para los sectores Quito, Guayaquil, Cuenca, Machala, Resto Costa y Resto Sierra:

Tabla 1.1:

Divisiones según CCIF pertenecientes a la canasta básica.

#	DIVISIÓN	# ART
1	ALIMENTOS Y BEBIDAS NO ALCOHÓLICAS	90
2	BEBIDAS ALCOHÓLICAS, TABACO Y ESTUPEFACIENTES	4
3	PRENDAS DE VESTIR Y CALZADO	47
4	ALOJAMIENTO, AGUA, ELECTRICIDAD, GAS Y OTROS COMBUSTIBLES	7
5	MUEBLES, ARTÍCULOS PARA EL HOGAR Y PARA LA CONSERVACIÓN ORDINARIA DEL HOGAR	37
6	SALUD	28
7	TRANSPORTE	17
8	COMUNICACIONES	6
9	RECREACIÓN Y CULTURA	30
10	EDUCACIÓN	10
11	RESTAURANTES Y HOTELES	6
12	BIENES Y SERVICIOS DIVERSOS	17
	TOTAL	299

Nota. Fuente: INEC (2013). *ENIGHUR 2011 - 2012. Quito – Ecuador.*
Elaborado por: Sr. Jhon Quimí Loza

1.4 ENCUESTA DE ESTRATIFICACIÓN DEL NIVEL DE SOCIOECONÓMICO.

Durante el año 2011 el INEC realizó la Encuesta de Estratificación de Nivel Socioeconómico a los hogares urbanos pertenecientes a cinco ciudades del Ecuador: Quito, Guayaquil, Cuenca, Ambato y Machala, con el objeto de encontrar los grupos socioeconómicos más relevantes y las características que poseen.

Una de las aplicaciones de esta encuesta es la homologación de herramientas para la estratificación de hogares. Instrumento utilizado para

definir una adecuada segmentación del mercado de consumo, además identificar variables clasificatorias que permiten caracterizar los niveles socioeconómicos en los hogares ecuatorianos. Derivación como insumo para una mejor dispersión en diseño de muestras.

Durante ese periodo 9,744 viviendas fueron entrevistadas, de ésta manera cubriendo así a los 812 sectores censales en el país. Entre las diferentes ciudades del Ecuador; en Guayaquil se visitaron a 3,372 viviendas seguido de Quito con 2,364, luego, en Cuenca, Ambato y Machala a 1,344, 1,344, y 1320 viviendas entrevistadas respectivamente (INEC, 2011).

Dentro de las viviendas que se consideraron para el estudio se obtuvo que, el 1.9% de población se encuentran dentro del estrato A, estrato mejor calificado con respecto a los bienes que posee, sus hábitos de consumo, su educación y su economía. Seguido del grupo B con el 11.2% de personas, otro 22.8% en el estrato medio C+, y por consiguiente el restante 64,2% de las personas ubicadas en los grupos con las peores puntuaciones distribuidas en: 49,3% en el grupo C- y el faltante 14,9 en el grupo D, tal como se observa en el siguiente gráfico (INEC, 2011).

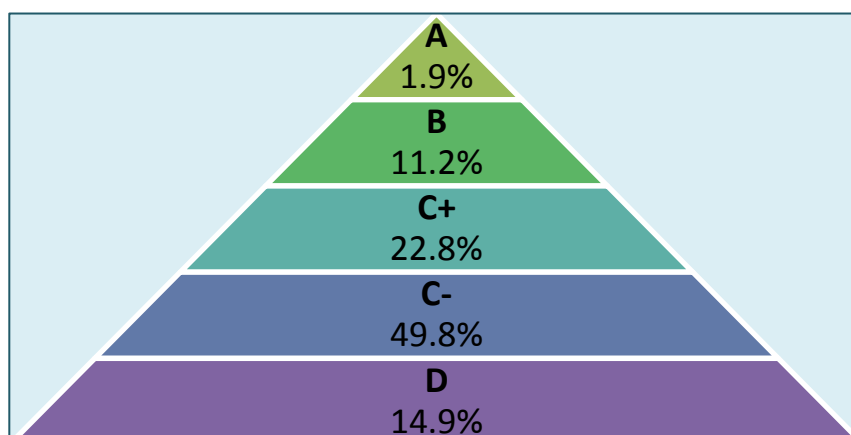


Gráfico 1.1 Distribución por estratos del nivel socioeconómico

Fuente: INEC (2011). *Encuesta de Estratificación del nivel socioeconómico NSE 2011: Presentación agregada. Quito – Ecuador.*

Elaborado: INEC

1.4.1 ESTRATOS DEL NIVEL SOCIOECONÓMICO

Los grupos socioeconómicos obtenidos en la ENIGHUR poseen 5 categorías, de las cuales la de mayor puntaje es la categoría A, mientras la categoría D es la de menor puntaje, también existen categorías medias dependiendo de la puntuación obtenida, esta se ubicará dentro de las categorías B, C+ o C- si el valor en puntos se encuentra dentro del valor situado de alguna de ellas.

Tabla 1.2:

Puntaje por Estratos del Nivel Socioeconómico

Grupos socioeconómicos	Puntaje
A	De 845 a 1,000 puntos
B	De 696 a 845 puntos
C+	De 535 a 696 puntos
C-	De 316 a 535 puntos
D	De 0 a 316 puntos

Nota. Fuente: INEC (2011). *Encuesta de Estratificación del nivel socioeconómico NSE 2011: Presentación agregada. Quito – Ecuador.*

Elaborado por: INEC

1.5 CENSO DE POBLACIÓN Y VIVIENDA

El Censo de Población y Vivienda 2010 es un recuento de la población y las viviendas que habitan en el país, que nos ayuda a generar información estadística confiable, veraz y oportuna acerca de la magnitud, estructura, crecimiento, distribución de la población y de sus características económicas, sociales y demográficas, que sirva de base para la elaboración de planes generales de desarrollo y la formulación de programas y proyectos a cargo de organismos de los sectores público y privado. Determina el volumen y características de las unidades de vivienda en que habita la población ecuatoriana, con la visión de evaluar las condiciones de vida y los requerimientos específicos de vivienda y servicios inherentes.

El empadronamiento se realizó el 28 de noviembre en las áreas urbanas y en las zonas rurales se extendió hasta el 5 de diciembre. Este tipo de censo tiene una cobertura de alcance nacional y es realizada con una periodicidad de 10 años, siendo efectuada la última durante los años 2011 - 2012 (INEC, 2011-2012).

1.6 ENCUESTA NACIONAL DE INGRESOS Y GASTOS EN HOGARES URBANOS Y RURALES

En el 2013 el Instituto Nacional de Estadísticas y Censo (INEC) presentó los resultados obtenidos en la Encuesta Nacional de Ingresos y Gastos en Hogares Urbanos y Rurales (ENIGHUR) durante un año de investigación (abril 2011 a marzo 2012), esta es una de las encuestas más importantes del país; de donde se pudo obtener el comportamiento económico, la situación socio-económica y los factores que influyen en los niveles de riqueza de los hogares en el Ecuador.

Esta encuesta proporciona información para actualizar la base del cálculo de la canasta básica de mercado que propugna el Índice de Precios al Consumidor (IPC). También permite conocer la forma de cómo obtienen los hogares ecuatorianos sus ingresos y qué tipos gastos poseen.

La ENIGHUR representa una de las estadísticas más complejas por la cantidad de temáticas investigadas.

La información recolectada, muestra las características demográficas, ocupacionales, educacionales de los miembros del Hogar, así como la cantidad de ingresos y la forma como adquieren los bienes y servicios cada hogar.

Para realizar la ENIGHUR el INEC consideró las recomendaciones de varias organizaciones, tales como: Organización Internacional del Trabajo, grupo Cambera en el tema ingresos y la del OIT en materia de encuestas de ingresos y gastos.

El Instituto Nacional de Estadísticas y Censos desde el año de 1975 ha realizado investigaciones que le han permitido saber la estructura del ingreso y gasto de los hogares del país, como son las siguientes:

- (1975) Encuesta de Presupuesto Familiares en el Área Urbana.
- (1978-1979) Encuesta de Ingresos y Gastos de los Hogares Rurales.
- (1991) Módulo de Presupuestos Familiares Área Urbana
- (1994-1995) Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos
- (2003-2004) Encuesta Nacional de Ingresos y Gatos de Hogares Urbanos - ENIGHU

- (2011-2012) Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos y Rurales - ENIGHUR

El objetivo general del INEC sobre la ENIGHUR es el proporcionar, la distribución, el monto y la estructura del gasto y el ingreso de los hogares rurales y urbanos, en base a características socioeconómicas y demográficas de los miembros del hogar. Además, conocer la estructura del presupuesto de los hogares, de acuerdo a las diferentes vías origen, montos de ingresos y sus formas de gastos, también detectar la capacidad de ahorro y endeudamiento de los hogares; definir insumos para medir la pobreza, y proporcionar información para la conformación del Sistema de Cuentas Nacionales.

Como unidades de análisis se consideraron a las viviendas del país, los hogares ubicados en su interior y las personas que habitan en ellas (INEC, 2013).

El resultado del análisis de la población objetivo reveló 39.425 viviendas, las mismas que contenían a 39.617 hogares de donde se encontraron a 153.444 personas. Según el área geográfica cada uno de los hogares antes mencionados tiene en promedio 3,9 personas a nivel nacional.

El promedio de personas por hogar es en la actualidad 3,8 durante el 2011 - 2012 en el área urbana frente a las 4,6 personas en los años 1994 - 1995.

Por otro lado el promedio de perceptores pasó de 1,8 a 2,0 entre los años 1994 - 1995 y 2011 - 2012.

En la aplicación de la última ENIGHUR se obtuvo que el ingreso total promedio mensual en el área urbana sea de 1.046,3 dólares, mientras que en el área rural es de 567,1 dólares y a nivel nacional alcanza los 829.9 dólares. Del ingreso total, el 79,3% corresponde al ingreso monetario y el restante 20,7% al no monetario. En el área urbana el ingreso monetario es la principal forma de sustento, cubriendo el 80,3% de los hogares urbanos, también prevalece en los hogares rurales con el 75,4%. El 83% del ingreso monetario a nivel nacional se obtiene del salario obtenido del trabajo, seguido por la renta de la propiedad y del capital, mientras que en el ingreso no monetario el 75,4% se ubica en el ingreso del trabajo asalariado no monetario (salario en especie) y regalos recibidos.

En el Ecuador el ingreso promedio de los hogares a nivel nacional es de 893 dólares, el 10% de hogares posee un ingresos menores o iguales 305 dólares, el 50% de los tienen ingresos por debajo e inclusive a 678 dólares y el máximo ingreso se ubica en 2.459 dólares.

Una manera de medir la inequidad en los ingresos de los ecuatorianos es mediante la curva de Lorenz. La curva de Lorenz muestra la proporción acumulada de los ingresos de países por varias proporciones de la población. Si una persona posee todos los ingresos y el resto no tiene

ninguno, la curva de Lorenz se ajustaría a los dos ejes, mientras que si el ingreso se distribuye por igual la curva de Lorenz sería una línea con un ángulo de 45° grados. Por lo tanto, mientras la curva de Lorenz esté más cercana a la línea de 45° grados mejor será la distribución de la riqueza y más equitativa es la distribución de la renta (Brown & Jackson, 1990).

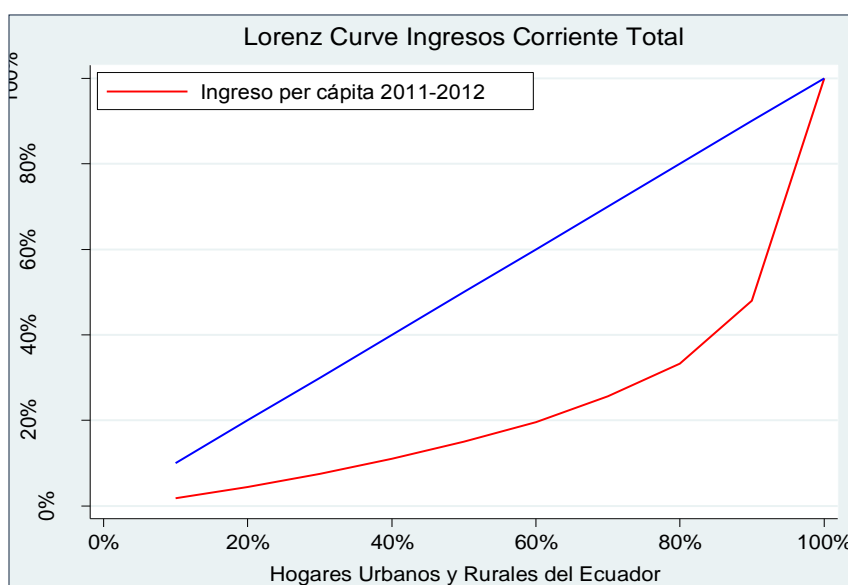


Gráfico 1. 2 Curva de Lorenz para la brecha del ingreso per cápita

Fuente: INEC (2011). *Encuesta de Estratificación del nivel socioeconómico NSE 2011: Presentación agregada. Quito – Ecuador.*

Elaborado: Jhon Quimí Loza

En la curva de Lorenz se observa la existencia de una gran brecha entre los rangos de ingresos de los hogares del Ecuador obtenidos en la ENIGHUR 2011-2012. Se observa que sólo el 20% del ingreso per cápita se distribuye en el 60% de la población y, el restante 80% se reparte apenas en un 40%

de los hogares, verificando de esta manera la gran brecha que existe entre los hogares ecuatorianos respecto a sus ingresos.

El 54,2% de los hogares del país perciben ingresos monetarios inferiores a dos salarios básicos unificados (SBU), el cual está establecido en 528 dólares.

Por otro lado el gasto corriente de consumo mensual se distribuye en monetario con un 77,2% y el no monetario con el 22,8%. Entre los principales gastos corrientes monetarios se encuentran el gasto por alimentación y bebidas no alcohólicas con el 24%, seguido por el gasto en transporte que genera un 14,60%, un 9,90% en bienes y servicios diversos, entre otros.

Los alimentos y bebidas no alcohólicas representan el 32% de los gastos en el área rural frente al 22,6% que representan en los hogares del área urbana. Existe gran similitud en el gasto por transporte en los hogares rurales y urbanos con el 14,5% y 14,6% destinados respectivamente.

1.7 ESTUDIOS RELACIONADOS A LA ENIGHUR

Durante el año 2004 el ingeniero Santamaría Manobanda realizó un estudio a la ENIGHU realizada en el país durante el año 2003 en su tesis titulada “Análisis Estadístico Multivariado De Los Datos Obtenidos En La Encuesta Nacional de Ingresos y Gastos de los Hogares Urbanos 2003”, donde él utilizó la técnica el Análisis de Componentes Principales (ACP) exploratorio para detectar agrupaciones de variables que tiene un comportamiento similar, y elegir una variable representante de cada grupo

Para su estudio utilizó 18 variables de relacionadas a gastos, ingresos y otros tipos de servicios, las cuales son:

1. Distribución de los hogares por provincias
2. Habitantes por hogar
3. Hogares que cuentan con el servicio de alcantarillado
4. Servicio telefónico de red pública
5. Eliminación de la basura por servicio de carro recolectar
6. Obtención de agua por red publica
7. Miembros promedio por hogar
8. Número de perceptores promedio por hogar
9. Grado de dependencia
10. Población que desempeña un trabajo (personas ocupadas)
11. Población económica inactiva (PEI)

12. Población desocupada
13. Distribución de ingresos por provincias de los hogares
14. Ingresos promedios de los hogares
15. Ingresos promedio por perceptores
16. Distribución gasto por provincia de los hogares
17. Gasto promedio de los hogares
18. Gasto promedio de los perceptores.

El 94.5% de las variables que consideraron en el análisis corresponde a variables de tipo cuantitativo y sólo una de las 18 preguntas es cualitativa. Aplicando la técnica multivariada se obtuvo dos nuevas componentes, a la componente 1 se la nombró como “Situación Económica” y está conforma por las variables: gastos promedios en los hogares, personas ocupadas e ingresos promedios de los hogares

Mientras que la segunda componente tuvo el nombre de “Situación Social” esta se conforma por: ingresos promedio de los hogares, personas ocupadas y hogares. (Manobanda, 2005)

1.8 JUSTIFICACIÓN DE ESTUDIO Y OBJETIVOS

1.8.1 JUSTIFICACIÓN DEL ESTUDIO

Existen muchos estudios realizados sobre análisis de variables cuantitativas, gracias a la gran variedad de técnicas que utilizan este tipo de datos y la riqueza que nos brinda al momento de tomar decisiones, como por ejemplo para mostrar la cantidad de personas por provincias, cantones, ciudades, etc... Pero pocos son los análisis realizados sobre las variables cualitativas citando por ejemplo algunas de ellas como: el género de la población, el nivel de educación de la persona, la identificación cultural, el estado civil, si dispone de seguro médico, entre otros; que de igual manera brindan la misma importancia o riqueza que otras variables y son pocos exploradas.

Este tipo de análisis que trabaja con variables cualitativas no se limita en describir un conjunto de datos, sino que trata de descubrir las dimensiones estructurales de las variables. Además pueden ayudarnos a reducir un grupo de variables previamente planteadas y quedarnos con las más relevantes que representen la misma información en tan sólo en unas pocas.

Para este estudio se trabajó en su mayoría con variables cualitativas, pero también se añadió ciertas variables cuantitativas que se fueron utilizadas por ser muy relevantes y explicativas en el análisis.

1.9 ORIGEN DE LA INFORMACIÓN

La información utilizada para este análisis fue seleccionada en base a las características de las variables analizadas, que poseen diferentes tipos de nivel de medida, tales como: nominales, ordinales y escalares. Estos datos serán sometidos a exploración de relaciones entre variables con el fin de encontrar patrones de relación entre los datos.

1.9.1 FUENTE DE DATOS

La fuente de datos que se utilizó para el desarrollo de este trabajo de investigación es la base de la Encuesta Nacional de Ingresos y Gastos de los Hogares Urbanos y Rurales (ENIGHUR) en Ecuador durante el 2011 y a inicios del 2012, realizada por el Instituto Nacional de Estadísticas y Censos (INEC), información de la cual se considera como población objetivo a todos hogares entrevistados del cantón Guayaquil (INEC, 2013).

De la base de datos considerada, se seleccionó una población objetivo la cual fue establecida bajo análisis por la técnica de estadística multivariada el Análisis de Componentes Principales Categórico.

CAPÍTULO 2

MARCO TEÓRICO, METODOLOGÍA Y FUENTE DE DATOS

En este capítulo se detalla el marco teórico y definiciones que son utilizados de una manera u otra para la exploración y estudio analítico que se lleva a cabo en esta tesis.

Conceptos extraídos de fuentes de análisis estadísticos y estudios relacionados con el tema a ser investigado. Las secciones de este capítulo detallan los orígenes de la estadística en el país, y la evolución desde sus inicios. Los proyectos que ha venido realizando el INEC como institución pública al servicio de la población, los mismos que contribuyen a generar información específica del país para tomar decisiones que ayuden al desarrollo económico y al ámbito social del país.

En las próximas secciones encontramos el marco teórico correspondiente al levantamiento y desarrollo multivariado de la información obtenida de la

ENIGHUR 2011 - 2012 realizada por el Instituto nacional de estadísticas y censos en los procesos de producción (consulta, ejecución, veeduría).

2.1 POBLACIÓN

Por población o colectivo se entiende como todo conjunto de personas, animales u objetivos sobre el que se va a llevar a cabo una investigación.

En este sentido, es fundamental que la población esté bien definida, es decir, que no existan ambigüedades.

(Vogt, 2005) Asevera que, la población es un conjunto de personas o individuos (o eventos, instituciones, u otros temas de estudios) que alguien quiere describir o ente al cual se quiere generalizar.

Con el propósito de estimar los parámetros de una población, frecuentemente se estudia una muestra que sea representativa de la población. También llamada población objetivo o universo (Mangin & Mallou, 2003).

2.2 POBLACIÓN OBJETIVO

Es un grupo de personas, instituciones, eventos, u otros sujetos estudiados que el investigador quiere analizar, describir o el cual se desea explorar.

Para poder estimar los parámetros de una población, usualmente se estudia una muestra que represente significativamente a la población. Conocida también como población objetivo y universo (Vogt, 2005).

En este proyecto se estableció como población objetivo o universo a todos los hogares que habitan en el cantón Guayaquil, de donde se analizó a 2,265 hogares de este sector del país.

2.3 ANÁLISIS EXPLORATORIO

El análisis exploratorio se define como el análisis que se realiza antes de aplicar cualquier técnica de análisis multivariante. Es preciso realizar un estudio previo a los datos disponibles antes de profundizar en una análisis multivariado.

Es necesario examinar las variables de manera individual y las relaciones entre ellas, así como evaluar y solucionar problemas en el diseño de la investigación y en la obtención de datos tales como: el tratamiento de información faltante y la presencia de datos anómalos o atípicos.

Como primer paso se suele realizar técnicas gráficas que nos permiten examinar las características de la distribución de las variables implicadas en el análisis, las relaciones bivariantes y multivariantes de estas, y el análisis de la diferencia entre grupos.

En la segunda parte se analizan los datos ausentes, falencias en la información obtenida por problemas al momento de obtención de datos; entes faltantes de estudio, la inaccesibilidad a algunos elementos o los errores en instrumentos de medida.

En el paso siguiente se debe detectar los valores atípicos, investigar cuales son los valores que tiene comportamiento muy diferentes al resto, errores en el ingreso de datos o de codificación. Una vez detectados el investigador tomará la decisión de considerarlos o no en el análisis.

Por otro lado (Pallant, 2011) indica que es primordial revisar los datos para descartar errores, ya que algunos análisis son bastantes sensibles para este tipo de inconsistencias que comúnmente se las conoce como valores aberrantes o datos perdidos los que en muchas ocasiones son producidos por errores de digitación o de escala correspondientes a fallas humanas.

Al inicio del estudio (Pallant, 2011) sugiere dedicar algo de tiempo en verificar la existencia de estos valores para evitar futuros inconvenientes. Dentro de su ideología se define dos pasos principales, los que se detallan a continuación:

1. Chequear errores, es necesario examinar dentro de los valores de cada variables los datos que se encuentren fuera de la escala a la cual se mide.
2. Buscar y corregir los errores en la base de datos. Se debe detectar errores en la base, corregirlos o borrarlos de la base.

Una herramienta básica para exploración de las variables es la obtención de frecuencia mediante estadística descriptiva, además podemos generar gráficos u otros análisis bivariado.

2.4 ATRIBUTOS

Los atributos se definen como una variable cualitativa o característica, que es a menudo usada en contraste con una variable cuantitativa o rasgo. La misma distinción que existe entre este tipo de variables, existe con las variables categóricas y las continuas (Vogt, 2005).

Los caracteres de los elementos de la población pueden ser de dos tipos: cualitativos o atributos y cuantitativos o variables.

2.5 VARIABLE CUALITATIVA O ATRIBUTO

Término generalmente usado para variables categóricas ordinales y nominales, o algún tipo de medición difícil de cuantificar tales como patrones de comportamiento, teorías, conceptos, etc. (Vogt, 2005).

Estas características tratan sobre el estudio sistemático de la experiencia cotidiana. Utilizadas comúnmente en trabajos descriptivos o exploratorios desarrollados en las ciencias sociales.

2.6 VARIABLE CUANTITATIVA

Cuando la variable analizada es expresada numéricamente, se denomina variable cuantitativa.

En estadística estas variables se dividen en dos grupos: las variables discretas y las variables continuas.

2.7 ANÁLISIS MULTIVARIADO DE LOS DATOS

El análisis multivariado constituye una colección de métodos que pueden ser usados cuando varias mediciones son realizadas a los individuos u objetos en una o más muestras. En este análisis los conjuntos de datos multivariados

son comunes, aunque estos no siempre son analizados como tales. El uso exclusivo de los análisis univariados con gran volúmenes datos no brindan mayores resultados, dando así ventaja a las técnicas multivariadas de revelar mejores deducciones, como patrones o relaciones ocultas entre las variables analizadas con la ayuda de software de informáticos (Rencher, 2002).

La mayoría de las aplicaciones de las técnicas multivariadas han sido utilizadas en las ciencias biológicas y del comportamiento.

Los métodos multivariados son de gran ayuda para la construcción de modelos matemáticos aplicados en las ciencias sociales y en numerosos campos de investigación.

Cuando se realiza una investigación se presentan casos donde todas las variables son medidas mediante la misma métrica lo que facilita el uso de las técnicas multivariadas, por otro lado, cuando se presentan casos donde las variables poseen diferentes métricas es necesario utilizar algún método de estandarización de datos.

El análisis multivariado dispone de muchos métodos para examinar varias variables simultáneamente, usualmente dos o más variables independientes y una variable dependiente.

En cualquier etapa que se utilice el análisis multivariado permite al investigador examinar la relación entre más de dos variables.

2.7.1 ESTADÍSTICA MULTIVARIADA EN INVESTIGACIÓN NO EXPERIMENTAL

La investigación no experimental toma muchas formas, por ejemplo la encuesta es una de las herramientas comúnmente utilizadas en este tipo de estudios. Frecuentemente las personas son entrevistadas, y cada una de ellas proporciona información sobre las preguntas que le fueron aplicadas, generando de esta manera un gran número de variables (Tabachnick & Fidell , Using Multivariate Statistics, 1996).

Estas variables están altamente interrelacionadas de manera complejas, por lo que los análisis univariados y bivariados no son sensitivos a este tipo de complejidad. La correlación bivariada entre todos los pares de variables no puede revelar la multicolinealidad entre ellas, esta multicolinealidad puede ser eliminada mediante la reducción de variables que son combinaciones lineales de unas de otras o las consideradas redundantes.

Con el uso de técnicas de estadística multivariada, las interrelaciones complejas entre las variables son reveladas y juzgadas en estadística inferencial.

2.7.2 ESCALAMIENTO ÓPTIMO

El objetivo principal de la técnica del escalamiento óptimo es asignar cuantificaciones numéricas a las categorías de cada variable de estudio, lo cual ayuda a poder utilizar los procedimientos normales de estadística multivariada para lograr resultados con las variables cuantificadas.

Los valores de escalamiento óptimo se asignan a las categorías de cada variable de acuerdo con el criterio de optimización para procedimiento que se vaya a utilizar. A diferencia de las etiquetas originales de las variables nominales u ordinales del análisis, estos valores de escala tienen propiedades métricas.

En la gran mayoría de encuestas e investigación de mercados desarrollados en los estudios sociales y del comportamiento, se puede hacer uso del escalamiento óptimo por las características de las dimensiones que se estén estudiando.

Este concepto evita incurrir en problemas que se presentan con frecuencia en las bases de datos utilizadas, entre este tipo de novedades podemos determinar algunos como: demasiadas variables, demasiados valores por cada variable y observaciones insuficientes.

2.7.3 NIVEL DEL ESCALAMIENTO ÓPTIMO

Existen tres niveles básicos de medida:

- El nivel **nominal** implica que los valores que se asignarán a cada categoría no dispongan de ningún orden específico.
- El nivel **ordinal** establece que las categorías de una variable poseen un orden específico.
- El nivel **numérico** establece a los valores de una variable como categorías ordenadas con una métrica significativa.

Al especificar un nivel óptimo de escalamiento nos referimos al nivel en que se escalan las variables y no al nivel en que se miden.

2.7.4 GRÁFICOS DE TRANSFORMACIÓN

Muestran la relación entre las cuantificaciones y las categorías originales que resultan del nivel del escalamiento óptimo seleccionado que fue aplicado a las variables iniciales.

Cuando se usa a variables numéricas se genera un gráfico de transformación lineal, mientras que con variables escaladas ordinalmente se presentan gráficos no decrecientes. Con variables escaladas nominalmente los gráficos de transformación muestran una relación cuadrática en forma de U o en ocasiones se puede dar el caso de gráficos sin tendencia aparentes mediante el cambio completo del orden de las categorías.

Uno de los objetivos de los gráficos de transformación es evaluar el nivel de escalamiento óptimo seleccionado.

2.7.5 TABLAS DE CONTINGENCIA

Se describe como una tabla que relaciona dos o más variables nominales entre filas y columnas, de modo que cada celda recoge el número de asociaciones que existen entre las correspondientes modalidades que contienen las variables en filas y en columnas. Es la tabla de partida del análisis de correspondencia. Una tabla de contingencia con a filas y b columnas está representada de la siguiente manera:

Tabla 2.1:

Tabla de Contingencia con a Filas y b Columnas

		Columna				Total Fila
		1	2	...	b	
Filas	1	n_{11}	n_{12}	...	n_{1b}	$n_{1.}$
	2	n_{21}	n_{22}		n_{2b}	$n_{2.}$
	⋮	⋮	⋮		⋮	⋮
	a	n_{a1}	n_{a2}	...	n_{ab}	$n_{a.}$
Total Columna		$n_{.1}$	$n_{.2}$...	$n_{.b}$	n

Fuente: Rencher, A. C. (2002).

Elaborado: Jhon Quimí Loza

Los elementos de la matriz de una tabla de contingencia son las frecuencias de la combinación de cada fila con cada columna.

Si se divide a cada una de estas frecuencias para el total marginal respectivo se obtiene la matriz de correspondencia que se muestra a continuación:

Tabla 2.2:

Matriz de Contingencia con de Frecuencias Relativas

		Columna				Total Fila
		1	2	...	<i>b</i>	
Filas	1	p_{11}	p_{12}	...	p_{1b}	$p_{1.}$
	2	p_{21}	p_{22}		p_{2b}	$p_{2.}$
	⋮	⋮	⋮		⋮	⋮
	<i>a</i>	p_{a1}	p_{a2}	...	p_{ab}	$p_{a.}$
Total Columna		$p_{.1}$	$p_{.2}$...	$p_{.b}$	p

Fuente: Rencher, A. C. (2002).

Elaborado: Jhon Quimí Loza

2.7.6 MATRIZ DE BURT

Mangin & Mallou (2003, p.381) la definen como el conjunto de modalidades de las variables estudiadas, que permiten dividir la muestra en diferentes número de clases. Por lo que, con dos variables bajo forma disyunta completa se puede realizar sendas divisiones del conjunto de individuos, obteniendo una matriz G, formada de ceros y unos. El análisis de correspondencia de una tabla de contingencia puede ser generalizado para

una tabla que relaciona más de dos variables mediante el análisis de correspondencias múltiple aplicado sobre una matriz de Burt.

Tabla 2.3:

Matriz de Burt G'G para una Matriz G

Genero		Edad			Estado Civil		Color de Cabello			
7	0	3	1	3	4	3	3	1	1	2
0	5	1	3	1	2	3	1	2	2	0
3	1	4	0	0	1	3	1	1	1	1
1	3	0	4	0	2	2	2	1	1	0
3	1	0	0	4	3	1	1	1	1	1
4	2	1	2	3	6	0	1	3	1	1
3	3	3	2	1	0	6	3	0	2	1
3	1	1	2	1	1	3	4	0	0	0
1	2	1	1	1	3	0	0	3	0	0
1	2	1	1	1	1	2	0	0	3	0
2	0	1	0	1	1	1	0	0	0	2

Fuente: Rencher, A. C. (2002).

Elaborado: Jhon Quimí Loza

2.7.7 ANÁLISIS DE CORRESPONDENCIAS

Cuando hablamos de análisis de correspondencias nos referimos a un método gráfico de representación de relaciones que existen entre tablas bidireccionales, las cuales trabajan sólo con variables categóricas nominales. El análisis de correspondencias es una técnica multivariada que estudia dos o más variables sobre un conjunto de individuos. También puede ser definida como una técnica descriptiva o exploratoria cuya finalidad es resumir una

gran cantidad de datos en un número reducido de dimensiones, con la menor pérdida de información posible (Mangin & Mallou, 2003).

Luque (2000, p.89) menciona que, el análisis de correspondencia también es conocida como análisis factorial de correspondencias. Técnica multivariada de interdependencia la cual no se limita simplemente a describir un conjunto de datos, al contrario trata de descubrir sus dimensiones estructurales, estudiando las similitudes en las formas y no entre las magnitudes absolutas. Además de reducir las dimensiones, conduce a la obtención de mapas perceptuales; por esto puede ser clasificada como una técnica de composición porque proporciona un mapa porcentual basado en la asociación entre objetos y un conjunto de características descriptivas o atributos.

Entre la gran variedad de tablas de datos que pueden ser estudiadas utilizando el análisis de correspondencias, destacan las siguientes:

- Tablas de contingencia de dos variables nominales
- Tablas de datos codificados en forma disyuntiva completa y tablas lógicas de datos binarios de tipo presencia ausencia.
- Tabla de datos ordinales categorizados, como las escalas actitud y tablas de notas de intensidad o puntuaciones desdoblas.
- Tablas de magnitudes homogéneas caracterizando poblaciones y tablas múltiples de números enteros positivos.

- Tablas de proximidad o distancia entre elementos y tablas de preferencias

Para probar la significancia de la asociación de dos variables categóricas en una tabla de contingencia, podemos usar una prueba χ^2 (Rencher, 2002).

Los datos en una tabla de contingencia pueden ser usados para verificar la relación entre dos variables.

El estadístico χ^2 usual para probar independencia entre dos variables está dado por:

$$\chi^2 = n \sum_{i=1}^a \sum_{j=1}^b \frac{(p_{ij} - p_i \cdot p_j)^2}{p_i \cdot p_j}$$

Donde

a = Número de filas en una tabla de contingencia

b = Número de columnas en una tabla de contingencia

p_{ij} = Frecuencia relativa en una tabla de contingencia

$p_{.j}$ = Total marginal de fila de la j – ésima columna

p_i = Total marginal de fila de la i – ésima fila

Este estadístico es aproximadamente distribuido como una variable aleatoria χ^2 con $(a - 1)(b - 1)$ grados de libertad.

Si una tabla de contingencia tiene algunas frecuencias que son muy pequeñas o cero, la aproximación con el estadístico χ^2 no es muy satisfactoria.

Con el análisis de correspondencia se construye un gráfico de puntos que muestra la interacción de dos variables en una tabla de contingencia mediante la relación entre cada fila con cada columna. Para este caso, algunas categorías pueden ser combinadas para incrementar la frecuencia de las celdas.

El análisis de correspondencia puede ser útil identificando las categorías que son similares, las cuales podríamos desear combinar.

En esta técnica graficamos un punto para cada fila y un punto por cada columna de la tabla de contingencia. Estos puntos son, en efecto, proyecciones de las filas y columnas de la tabla de contingencia dentro de un espacio Euclidiano de dos dimensiones. La meta es preservar tan como sea posible la relación de las filas (o columnas) en cada una en un espacio de dos dimensiones.

Si dos puntos que pertenecen a dos filas están muy juntos, las proyecciones de las dos filas (a través de las columnas) son similares. De igual manera, dos puntos pertenecientes a columnas si están muy juntos representan a columnas con similares proyecciones a través de las filas.

Si un punto fila está más cerca de un punto columna, la combinación de categorías entre dos variables se produce de forma más frecuente de lo que ocurre por casualidad si las dos variables son independientes.

Pérez (2004, p.13) menciona que el Análisis de correspondencia al igual que la técnica multivariada Análisis de componentes principales tienen como propósito reducir la dimensión de una tabla de datos formada por variables cualitativas y cuantitativas respectivamente.

Cuando estudiamos el comportamiento de dos variables cualitativas nos encontramos frente al Análisis de correspondencias simple, pero si se generaliza para cuando se dispone de más de dos variables, en este caso se denomina Análisis de correspondencias múltiple.

Para el caso simple las variables cualitativas pueden representarse en una tabla de doble entrada, denominada tabla de contingencia. Pero para el caso del análisis de correspondencias múltiple la tabla de contingencia puede ser una hipertabla en tres o más dimensiones, difícil de representar y se la denomina matriz de Burt.

El objetivo del análisis de correspondencias es establecer relaciones entre las variables no métricas obteniendo la información que ofrecen las tablas de contingencia. Este análisis revela además en qué grado o intensidad contribuyen a esa relación detectada los distintos valores de las variables, información que ser observada y obtenida en forma gráfica.

2.7.8 ANÁLISIS DE CORRESPONDENCIAS MÚLTIPLE

El análisis de correspondencias de una tabla de contingencia de dos dimensiones puede ser extendida a una tabla de tres dimensiones o a una multidimensional de un orden más alto (Rencher, 2002).

Por el método de análisis de correspondencias múltiple, obtenemos un gráfico de dos dimensiones que muestra la información de una tabla de contingencia de multidimensional.

2.7.9 ANÁLISIS DE COMPONENTES PRINCIPALES CATEGÓRICOS

Se define como un procedimiento estadístico que cuantifica las variables categóricas simultáneamente reduciendo la dimensionalidad de los datos. El análisis de componentes principales categórico también es conocido por el acrónimo CATPCA, del inglés Categorical Principal Componentes Analysis.

Se utiliza análisis cuando se quiere conocer los patrones de variación de un conjunto de variables que tienen diferentes niveles de escalamientos óptimos. El CATPCA utiliza gráficos de las puntuaciones de componentes, que muestran patrones de los objetos analizados y pueden revelar resultados pocos habituales en los datos. El análisis de componentes principales categórico maximiza las correlaciones de las puntuaciones de objetos con cada una de las variables cuantificadas para el número de dimensiones.

Si las variables que analizamos son nominales múltiples, el análisis de componentes principales categórico realiza un análisis similar al análisis de correspondencia múltiple.

Para poder hacer uso de esta técnica las variables deben tener como mínimo tres casos válidos, no se utilizará en el análisis de categorías valores fraccionarios solamente números enteros positivos; o también podemos ayudarnos gracias a la opción discretización categórica automática que disponen los software estadísticos.

El principal objetivo del análisis de componentes principales es extraer la varianza máxima del conjunto de datos con cada componente. La primera componente es la combinación lineal de las variables observadas, este análisis separa las variables para maximizar la varianza de los resultados de la componente, mientras que la segunda componente está formada de las correlaciones residuales. Esta componente es la combinación lineal de las variables observadas que extrajeron la variabilidad no correlacionada con la primera componente. Las componentes subsecuentes también extraen la máxima variabilidad de las correlaciones residuales y son ortogonales a todas las componentes previamente extraídas.

La primera componente extrae la mayor varianza mientras que las demás absorben la varianza restante. Si todas las componentes son retenidas

obtendremos la correlación exacta de los datos, es decir se reproduce la matriz de correlación observada.

Las componentes principales son una solución elegida por el investigador, quien es el principal interesado en reducir el número de variables a un número más pequeño de componentes. El Análisis de componentes principales también es usado como un paso inicial en el Análisis de factores donde este revela un gran tratamiento acerca del número máximo y la naturalidad de factores (Tabachnick & Fidell, Using Multivariate Statistics, 2001).

2.7.10 METODOLOGÍA DEL ANÁLISIS DE COMPONENTES PRINCIPALES

El análisis de componentes principales trata con una muestra de n observaciones de un espacio de dimensiones. Este puede ser aplicado a cualquier distribución, pero será más fácil visualizarlo geográficamente si el conjunto de datos tiene forma elipsoidal.

Si las observaciones están correlacionadas, las puntuaciones elipsoidales no se encuentran paralelamente orientadas a ningún eje representado por los datos.

El objetivo de este análisis es encontrar los ejes que representen al elipsoide de datos que tiene como origen la media de todas las observaciones. Si el origen del elipsoide es trasladado y los ejes son rotados, estos serán los ejes naturales del elipsoide, los nuevos ejes o también conocidos como componentes principales deberán no estar correlacionados (Rencher, 2002).

2.7.11 GENERALIDADES DEL ANÁLISIS DE FACTORES Y EL ANÁLISIS DE COMPONENTES PRINCIPALES

En muchas ocasiones se confunde el término componente con factor como si ellos fueron intercambiables, existiendo distinción entre ellos. Las componentes son factores reales que derivan directamente de la matriz de correlación de variables.

En efecto, el método de ejes principales de la técnica análisis de factores es idéntico al de componentes principales excepto que en lugar de la unidad en la diagonal principal de la matriz de correlación se añade las comunalidades.

Es poco probable que unos cuantos factores puedan adsorben toda varianza de una matriz de correlación, por lo que, para obtenerla es necesario seleccionar todos los factores que es equivalente al mismo número de variables originales.

El Análisis de componentes principales maximiza la varianza explicada para un pequeño número de componentes. La varianza explicada es igual a la

suma de los cuadrados de las cargas correspondientes a la componente observada conocida como el valor. Esta varianza fluctúa desde cero a números positivos. El número de valores propios positivos mayores que uno representa el número de componentes necesarios, para explicar la mayor cantidad de varianza de la matriz de correlación.

Existe cierta dificultad para poder diferenciar entre el Análisis de componentes principales y el Análisis de factores, (BERENSON, LEVINE, & GOLDSTEIN, 1983) mencionan que el Análisis de componente principales busca combinaciones lineales de las variables originales. Aunque en pocas componentes pueden acumular gran proporción de la varianza total, se necesita todas las componentes para cubrir la correlación total, diferente al Análisis de factores que define un número más pequeño de factores que las variables originales donde se reproduce la correlación exacta, pero cual no acumula tanta varianza como lo hace el mismo número de componentes principales. Finalmente, es importante notar que mientras en el Análisis de componentes principales los factores son combinaciones lineales de las variables observadas, en el Análisis de factores los factores son combinaciones lineales de las partes comunes de las variables. Por lo tanto, el Análisis de componentes principales es visto de la perspectiva de la varianza orientada, mientras que el Análisis de factor trabaja con la correlación y es visto de la perspectiva de la covarianza orientada.

2.7.12 VECTOR CARACTERÍSTICO Y RAÍZ CARACTERÍSTICA O VALOR PROPIO

A un vector característico se lo define como un vector de pesos aplicable a cada una de las variables originales. Se puede decir que si tenemos seis variables deberíamos tener seis pesos en el vector característico.

Cada factor o componente está asociado a un vector de cargas, las mismas que son obtenidas de la multiplicación de cada elemento del vector con su correspondiente valor propio. La raíz característica conocida como valor propio se deriva de la suma de los cuadrados de las cargas de cada componente y esta refleja la proporción de varianza que es explicada por las componentes, mientras más grande sea el valor propio mayor varianza será explicada por la componente (Kline, 1994).

Denotamos a los ejes de la elipsoide como Z_i y pueden ser rotados multiplicando cada observación por la matriz ortogonal expresada como A , donde la matriz ortogonal es denotada con la letra C .

$$Z_i = Ay_i$$

Encontrar los ejes del elipsoide es equivalente a encontrar la matriz ortogonal A , que rota los ejes sobre la línea con las extensiones naturales del conjunto de datos así que las nuevas variables (componentes principales) $Z = Ay$

están no correlacionadas. De este modo se obtiene la muestra de la matriz de varianza y covarianza de Z (Rencher, 2002) :

$$S_Z = ASA'$$

$$S_Z = \begin{pmatrix} s_{z_1}^2 & \dots & \\ \vdots & s_{z_2}^2 & \vdots \\ & \dots & s_{z_p}^2 \end{pmatrix}$$

Los elementos de la diagonal de S_Z son los valores propios, entonces:

$$C'SC = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$$

Se define a λ_i 's como los valores propios de S (matriz de varianza y covarianza) y C como una matriz ortogonal cuyas columnas son vectores propios ortogonalizados de S . Por lo tanto la matriz ortogonal A es:

$$A = C' = \begin{pmatrix} a'_{11} \\ \vdots \\ a'_{1p} \end{pmatrix}$$

Al ser los valores propios la varianza que explican las componentes principales, podemos expresar a la proporción de varianza, como:

$$\text{Proporción de varianza} = \frac{\lambda_1 + \lambda_1 + \dots + \lambda_k}{\sum_{j=1}^p s_{jj}}$$

2.7.13 ROTACIÓN DE LAS COMPONENTES

La rotación es usualmente usada luego de la extracción de las componentes para maximizar las altas correlaciones y minimizar las bajas. Existen dos tipos métodos de rotación: la rotación ortogonal y la rotación oblicua.

Si a las componentes resultantes no se realiza algún tipo de rotación es probable que exista dificultad al momento de interpretar los resultados, independiente el método de extracción se haya utilizado. Es necesario aplicar rotación luego extraer las componentes para mejorar la interpretabilidad y la utilidad científica de la solución.

Una las características de la rotación es mejorar la calidad del ajuste matemático entre las matrices de correlación observadas. Las soluciones rotadas son matemáticamente equivalentes a unas de otras al igual que las soluciones que resultaron antes de haber aplica rotación.

Obtendremos resultados similares con cualquier método de extracción que utilicemos si disponemos de un buen conjunto de información, lo mismo sucederá con la rotación si los patrones de correlación en los datos son bastante claros. En otras palabras, con información veraz se obtendrá una solución estable independiente de los métodos de extracción o rotación usados.

La decisión está en elegir entre la rotación de tipo ortogonal u oblicua. En la rotación ortogonal los factores no están correlacionados. Las soluciones ortogonales ofrecen aliviar la interpretabilidad, descripción y reporte de los resultados, sin embargo, se esfuerzan a menos que el investigador este convencido que los procesos subyacentes son casi independientes. Mientras que cuando el investigador sospecha que los procesos subyacentes están correlacionados se usa la rotación oblicua, esto genera cierta ventaja conceptual pero una desventaja para la interpretación, descripción y al momento de mostrar los resultados (Tabachnick & Fidell , Using Multivariate Statistics, 1996).

2.7.14 ROTACIÓN DE TIPO ORTOGONAL

Entre las técnicas disponibles para la rotación ortogonal tenemos al Varimax, Quartimax y Equamax.

Varimax es el tipo más usada de todas la rotaciones disponibles, al igual que otros procedimientos de extracción posee diferentes tipos metas estadísticas.

Es un método ampliamente usado de rotación ortogonal de los ejes en el Análisis de factores. Como su nombre indica maximiza la varianza de los factores.

La meta del Varimax es simplificar componentes para maximizar la varianza de las cargas dentro de los factores a través de las variables.

Quartimax realiza para las variables lo que el Varimax realiza para los componentes, este simplifica las variables para incrementar la dispersión de las cargas dentro de las variables a través de la componente.

Equamax es un híbrido entre Varimax y Quartimax que trata simultáneamente de simplificar las variables y componentes.

Por lo tanto, la rotación Varimax simplifica componentes, Quartimax las variables y Equamax a ambas (Tabachnick & Fidell , Using Multivariate Statistics, 1996).

2.7.15 COMUNALIDADES, VARIANZA Y COVARIANZA

La técnica multivariada Análisis de Componentes Principales como muchas otras técnicas analiza patrones de correlación (o covarianza) entre las variables, esto es importante porque las correlaciones entre dos variables continuas o dicotómicas deben ser lo más preciso posible. Bajo ciertas condiciones de investigación, es propicio que las correlaciones sean muy grandes o muy pequeñas de lo que deberían según sea el caso.

La comunalidad para una variable es la varianza representada por todos los factores. Una comunalidad es la suma de las cargas al cuadrado para una variable a través de los factores. Es decir, qué porcentaje de la variabilidad de la información de la variable es explicada dentro del modelo.

La proporción de varianza que explica un factor para el conjunto de variables analizadas es igual a la varianza del factor (valor propio) dividido por la suma de todos los valores propios equivalente al número de variables originales.

La proporción de covarianza es la suma de las cargas al cuadrado para el factor dividido para la suma de las comunalidades (Tabachnick & Fidell , Using Multivariate Statistics, 1996).

2.7.16 APROXIMACIÓN DEL ANÁLISIS DE COMPONENTES PRINCIPALES CON VARIABLES CATEGÓRICAS

El método de Análisis de correspondencias múltiple es la aplicación más antigua del Análisis de componentes principales para variables categóricas. En este caso cada categoría de las variable es representada por una variable dummy, cual toma el valor de uno si una unidad presenta la categoría y cero para otro caso. La técnica de componentes principales es llevada a cabo con la matriz de correlaciones entre estas variables dummy, bajo algunas restricciones de identificabilidad.

La solución del Análisis de correspondencias múltiples es expresada como combinaciones de las categorías de las variables individuales y las cargas correspondientes a las categorías ordenadas que no son necesariamente monótonas, por lo que su interpretación es extremadamente difícil.

Existe cierta analogía en las soluciones del Análisis de correspondencias múltiple con las soluciones del Análisis de componentes principales calculadas para variables escalares de manera no lineal. Por otro lado, el Análisis de componentes principales es equivalente al Análisis de componentes principales para variables categóricas que han sido transformadas a continuas. El escalamiento fue realizado para maximizar el tamaño del valor propio que resulta de la matriz de correlaciones.

2.7.17 CARACTERÍSTICAS DEL ANÁLISIS DE COMPONENTES PRINCIPALES CATEGÓRICOS

Análisis de componentes principales mediante el escalamiento óptimo:

- Especifica el nivel de escalamiento óptimo al cual se quiere analizar cada variable. Se puede elegir entre varios tipo como: LíneaSp ordinal (monótona), LíneaSp nominal (no monótona), ordinal, nominal, o numérica.
- Especifica los pesos de las variables consideradas.

- Discretiza variables continuas o convierte las variables de texto a números enteros por multiplicación, ranking o agrupando valor dentro de una preselección de número de categorías de acuerdo a una distribución óptima (normal o uniforme), o por agrupación de valores por una preselección de intervalos en categorías.
- Especifica la forma para el manejo de datos perdidos. Excluye sólo las celdas de una matriz de datos sin valor válidos, puede imputar para los datos perdidos con la moda de la variable o colocar una categoría extra, o usar una lista de exclusión.
- Especifica el número de dimensiones en la solución

2.7.18 REGRESIÓN LOGÍSTICA

Al trabajar con la técnica multivariada regresión logística, es necesario la inclusión de ciertas definiciones útiles que servirán en la interpretación del análisis tales como: probabilidad o riesgo y oportunidad, los mismos se describen a continuación:

$$\text{Probabilidad o Riesgo} \quad \mathbf{R} = \frac{\text{Número de casos en que el evento ocurre}}{\text{Total de casos}}$$

$$\text{Oportunidad} \quad \mathbf{O} = \frac{\text{Número de casos en que el evento ocurre}}{\text{Número de casos en que no ocurre el evento}}$$

Mediante los términos antes mencionados podemos obtener de dos manera diferentes la explicación de un evento, expresado como:

Riesgo Relativo $RR = \frac{\text{Probabilidad de expuestos a factor de riesgo}}{\text{Probabilidad de no expuestos a factor de riesgo}}$

Oportunidad Relativa $OR = \frac{\text{Oportunidad de expuestos a factor de riesgo}}{\text{Oportunidad de no expuestos a factor de riesgo}}$

Sin embargo, el OR no es tan fácil de interpretar como el RR, estos cocientes toman valores similares en el caso particular de que la probabilidad de que ocurra un evento sea cercana o igual a cero. En el análisis de estas fracciones se debe recordar lo siguiente:

- Un valor de OR igual a uno se interpreta como que no hay tal factor de riesgo, dado que no existe diferencia entre la oportunidad para los que fueron expuestos o no.
- Cuando el OR es mayor que uno, existe un factor de riesgo pues es mayor la oportunidad para los expuestos al factor que a los controles.
- Para el caso ideal donde se desea obtener un OR menor que 1, se interpreta como menor la oportunidad de que ocurra el evento en los individuos expuestos al tratamiento que a los controles.
- Los valores de OR fluctúan entre cero e infinito, esta propiedad lo hace adecuado para ser modelado matemáticamente y puede ser considerado en la regresión logística.
- Se puede determinar intervalos de confianza para el OR utilizando un modelo de regresión logística:

- Si los intervalos contienen un OR igual a uno, no se puede rechazar la existencia de afectación por el tratamiento.
- En otro caso cuando aumenta o disminuye el intervalo de confianza, de igual manera varía la oportunidad de ocurrencia del evento.

Se establece al Odds como una medida de asociación, pero a la vez diferente a otras medidas usuales de asociación, tenemos la escala de “1.0” la que indica que no existe relación entre las variables. El tamaño de la relación es medida por la diferencia de 1.0. Es decir, un valor de odds ratio mayor que 1.0 indica relación directa o positiva; y cuando valor de odds es menor que 1.0 se refiere a una relación inversa o negativa (Vogt, 2005).

Es beneficioso tener las variables estudiadas correctamente codificadas, lo que nos facilitará la interpretación del modelo de regresión logística que obtengamos. La variable de ocurrencia del evento (variable dependiente) debe ser codificada entre 1 y 0, con 1 en el caso de ocurrencia y 0 en el caso de ausencia.

El modelo logístico puede ser representado de la siguiente manera:

$$\log\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + \dots + b_nx_n$$

Donde p es la probabilidad (riesgo) de que ocurra el evento de interés, las variables independientes se encuentran representadas por la letra x , y los coeficientes asociados a cada variable con la letra b .

Con los valores de las variables independientes, se puede calcular directamente la estimación del riesgo de que ocurra el evento de interés, definido como:

$$p = \frac{e^{suma}}{1 + e^{suma}}$$

Donde $suma = b_0 + b_1bx_1 + \dots + b_nx_n$

La oportunidad (*odds*) para los individuos de referencia o control, siguiendo el convenio anterior se representa como $\exp(b_0) = e^{b_0}$

Si tomamos en consideración otro coeficiente del modelo, la cantidad $\exp(b_i) = e^{b_i}$ coincide con la OR del aumento del valor de x_i en una unidad con respecto a aquellos individuos que presentan los valores de todas las demás variables iguales. Si hemos seguido el criterio de codificación recomendado y la variable analizada es dicotómica, esto corresponde a la OR del factor de riesgo de x_i , Si la variable explicada es numérica se estima la OR del factor de riesgo.

El propósito general de la regresión logística es predecir el resultado de una variable de interés categórica dicotómica o que posea más de dos categorías, en base de un grupo de variables independientes que pueden ser continuas, discretas, dicotómicas o mixtas. Esta técnica evalúa el impacto de las variables independientes sobre la variable explicada u observada.

La regresión logística está relacionada y responde las mismas preguntas como, el análisis de la función discriminante, la forma lógica del análisis de frecuencia multivías y el análisis de regresión múltiple con una variable dependiente dicotómica. Esta técnica es mucho más flexible que otras técnicas, a diferencia del análisis de la función discriminante no existen supuestos sobre la distribución de sus variables predictoras, estar linealmente relacionadas o igual varianza dentro de cada grupo (Tabachnick & Fidell , Using Multivariate Statistics, 1996).

Cuando se trabaja con regresión logística ciertos criterios se deben tomar en consideración, por ejemplo, el tamaño de la muestra debe ser lo suficientemente grande para superar significativamente en número a los predictores, lo que evitará tener problemas en el análisis tales como: errores en la estimación de los parámetros y fallas en la convergencia (Pallant, 2011).

Es importante verificar la existencia de altas interrelaciones entre los predictores, al ser la regresión logística muy sensible a correlaciones

extremadamente altas al igual que otros tipos de regresiones, es preferiblemente que los predictores estén fuertemente relacionados con la variable explicada pero no debe existir alta relación entre las variables independientes porque de darse el caso existiría multicolinealidad entre ellas. Una manera de eliminar la multicolinealidad es reemplazar los predictores de tipo discretos por variables dicotómicas, o borrar una o más variables redundantes en el modelo para así eliminar la multicolinealidad.

Otra consideración explorada cuando ocurren inconvenientes en la bondad de ajuste del modelo es la presencia de datos aberrantes, o casos que no contribuyan a la explicación del análisis, estos criterios pueden ser identificados mediante el análisis de los residuales (Pallant, 2011).

Entre los modelos de la regresión logística, tenemos los más simples (peor ajuste) que incluyen solamente la constante y ninguno de los predictores, los modelos más complejos que incluyen la constante, todas las variables predictoras y, tal vez, interacción entre las variables predictoras.

Para alcanzar la meta del análisis de predecir resultados a base de predictores se tiene que establecer si existe alguna relación entre los datos y el conjunto de predictores. Al encontrar alguna relación, usualmente se simplifica el modelo mediante la eliminación de predictores mientras se mantenga fuerte predicción.

Para probar el modelo que hemos obtenido usamos una prueba de bondad de ajuste donde consideramos el modelo reducido y lo comparamos con el modelo completo con todas las variables.

Un estadístico que nos ayuda a medir la significancia de cada coeficiente del modelo es el estadístico de Wald, que sirve para evaluar la significancia de los coeficientes que intervienen en la ecuación.

Existen otras pruebas que nos ayudan a evaluar el modelo como la Omnibus Test, la cual compara el modelo que incluye todas las variables independientes con el modelo donde sólo interviene la constante, ayudándonos a tener una idea de cuales de los predictores obtenidos son capaces de predecir nuestro evento de interés. Además, también está la prueba Hosmer y Lemeshow donde se mide el poco ajuste de los predictores dentro de nuestro modelo. Estas técnicas ayudan como guías en el trayecto del desarrollo de nuestro modelo.

En la práctica los investigadores usan las pruebas de bondad de ajuste para escoger el modelo que realiza el mejor trabajo de predecir con la menor cantidad de predictores.

En el método seleccionado que se encuentra libre de restricciones, la variable resultante pueden ser explicada en base de una mezcla de diferentes tipos de predictoras como son las variables continuas, discretas y

dicotómicas; las misma que se analizan sin ninguna limitación (Tabachnick & Fidell , Using Multivariate Statistics, 1996).

Al igual que en otras técnicas, disponemos de tres principales tipos de regresión logística: el directo, el secuencial y el paso a paso.

Este estudio se enfocará en el desarrollo del modelo con el método directo, que consiste en introducir todos los predictores en la ecuación simultáneamente sin descuidar la tolerancia.

El método permite la evaluación de la contribución por cada variable sobre y bajo la aportación de las demás. Una dificultad ocurre en el momento de la interpretación de los resultados cuando se encuentra alta corrección entre los predictores.

Sin embargo, cuando tratamos de evaluar la bondad de ajuste un modelo, primero debemos seleccionar de un pequeño abanico de modelos el que vamos a utilizar, estos modelos son: el modelo que dispone sólo de la constante, un modelo incompleto que contiene la constante plus y algunos predictores, el modelo completo que incluye la constante y todos predictores, y por último un modelo perfecto hipotético que deberá proveer la fijación exacta de que deseamos medir. Por consecuente tener varias comparaciones: entre el modelo posee sólo la contante y el modelo completo, entre el modelo que posee sólo la constante y el modelo incompleto, entre el modelo completo y el modelo incompleto, entre dos

modelos incompletos, entre modelo escogido y el modelo perfecto, etc. (Tabachnick & Fidell , Using Multivariate Statistics, 1996).

2.8 FLUJO DE LA INFORMACIÓN

Debido a la complejidad de la encuesta para el desarrollo de la ENIGHUR ciertos procesos fueron centralizados, procesos tales como: Revisión, Digitación y Validación en la Administración Central; con el objetivo de evitar sesgos estadísticos cuando se empleen criterios conceptuales en la revisión y validación de la información.

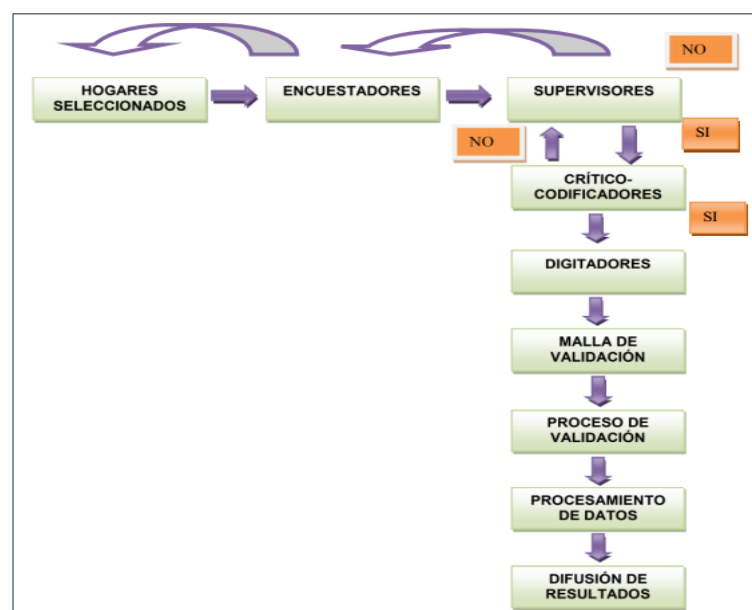


Gráfico 2.1 Diagrama del flujo de la información en la ENIGHUR

Fuente: INEC (2002).

Elaborado: INEC

Los datos recopilados provienen de cada una de las direcciones regionales, con su respectivo archivamiento, procediendo así el proceso de revisión y corrección de inconsistencias, manteniendo paralelamente el levantamiento y manteniendo la alimentación de la información.

2.9 ¿QUÉ ES UNA VIVIENDA?

Vivienda es el espacio delimitado por paredes y techo, de material cualquiera, que posean entrada independiente, destinado a ser habitado por una o más personas; la misma aunque no haya sido construida con esos fines inicialmente, sea destinada para ser vivienda.

2.10 DESCRIPCIÓN DE HOGAR

Es la unidad social que es conformada por una persona o un grupo de personas que deciden compartir alimentación y alojamiento. Es decir, al hogar como un conjunto de personas que habitan en una misma vivienda o en parte de ella, pudiendo tener parentesco alguno o no y que cocinan en común para todos los miembros del hogar.

Los empleados del servicio doméstico y sus familiares forman parte del hogar siempre y cuando duerman bajo el mismo techo y compartan sus alimentos.

Dentro de una misma vivienda pueden encontrarse varios hogares si existe un grupo de personas que viven en una parte de la vivienda y preparan sus

alimentos por separado. Los hogares pueden estar constituidos por personas con ningún grado de consanguinidad entre sí (INEC, 2011-2012).

2.10.1 MIEMBROS Y NO MIEMBROS DEL HOGAR

Se considera miembros del hogar a las siguientes personas;

- a. Los residentes habituales presentes en la vivienda, que viven permanentemente en el hogar, es decir, que duermen la mayor parte del tiempo en él; incluyendo aquellos ausentes en el momento de la entrevista, siempre que su ausencia por sobrepase el período de seis meses.
- b. Las personas sin parentesco con el jefe del hogar o familiares de éste, vivan la mayor parte del tiempo en el hogar, siempre que no tengan otro lugar de residencia.
- c. Los servidores domésticos que residen habitualmente en el hogar y que sus familiares viven con ellos (empleados domésticos puertas adentro).
- d. Persona de las Fuerzas Armadas destacado en el lugar y que vive habitualmente en el hogar, la mayor parte del tiempo.
- e. Los extranjeros que estudian o trabajan dentro del país por lo menos desde hace seis meses y habitan su mayor parte del tiempo en el hogar.

- f. Las personas que habitan como jefes de hogar en dos o más hogares, se considera en cuenta como miembro del hogar donde vive la mayor parte del tiempo.

No son miembros del hogar

- a. Aquellas personas ausentes del hogar por un período mayor a seis meses, durante el último año o tengan otro lugar de residencia.
- b. Las personas que se encuentran temporalmente en el hogar y tienen su residencia habitual en otro lugar.
- c. Los extranjeros que se encuentran en el país en misión de trabajo temporal (menos de seis meses).

2.10.2 INGRESOS DEL HOGAR

El ingreso del hogar es la suma de ingresos en moneda y en especie, consistente, como regla general de tipo periódico destinadas al hogar o a los miembros del hogar por separado, recibidas en periodos anuales o más frecuentes en pequeños intervalos de tiempo. Está constituido por la suma de la Renta Primaria, de la Renta de la Propiedad, y las Transferencias y Prestaciones Recibidas en forma regular por los hogares, antes de deducir los impuestos y el pago de aportaciones a la Seguridad Social y a las Cajas de Pensiones, (se excluye el ingreso de empleados domésticos) (INEC, 2011-2012).

2.10.3 PERCEPTORES DE INGRESOS

Es el miembro del hogar que recibe ingresos de cualquier fuente u origen, sea este por trabajo asalariado, si es patrono o socio activo, si es trabajador por cuenta propia, si es trabajador por cuenta propia, si es trabajador sin remuneración (siempre que reciba transferencias corrientes y otras prestaciones), si es inactivo (siempre que reciba ingresos por transferencias corrientes y otras prestaciones). Se dice por lo tanto, que los perceptores de ingresos es: un jubilado, rentista y puede ser perceptor un estudiante, una ama de casa, etc.

2.11 INGRESO CORRIENTE

Los ingresos y gastos corrientes de los hogares en el Ecuador se dividen en dos tipos: el monetario y el no monetario, de donde se derivan múltiples vías para su desarrollo.

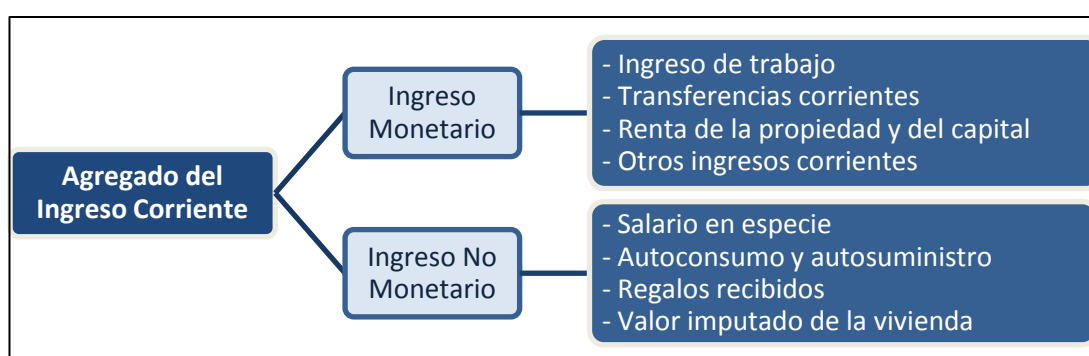


Gráfico 2.2 Estructura del ingreso corriente

Fuente: INEC (2011).

2.12 INGRESOS NO REGULARES DEL HOGARES

Es necesario considerar otra manera de ingresos en el hogar, tales como; la venta de propiedades, retiros de cajas de ahorro, premios de lotería préstamos obtenidos, reembolsos por préstamos (capital), ganancias imprevistas como herencias, vencimiento de pólizas de seguro de vida, indemnizaciones, etc.

2.13 GASTOS DEL HOGAR

Los gastos se dividen en gastos de consumo final y no de consumo final, en consumo final comprenden todos los bienes y servicios que los hogares compran, producen por cuenta propia, incluyen los servicios por viviendas arrendadas por sus arrendatarios. Sin embargo los de no consumo final se refieren a todos los desembolsos de dinero destinados a cubrir impuestos sobre el ingresos y la riqueza, la contribuciones sociales, primas de seguros e intereses pagados.

2.13.1 GASTO CORRIENTE

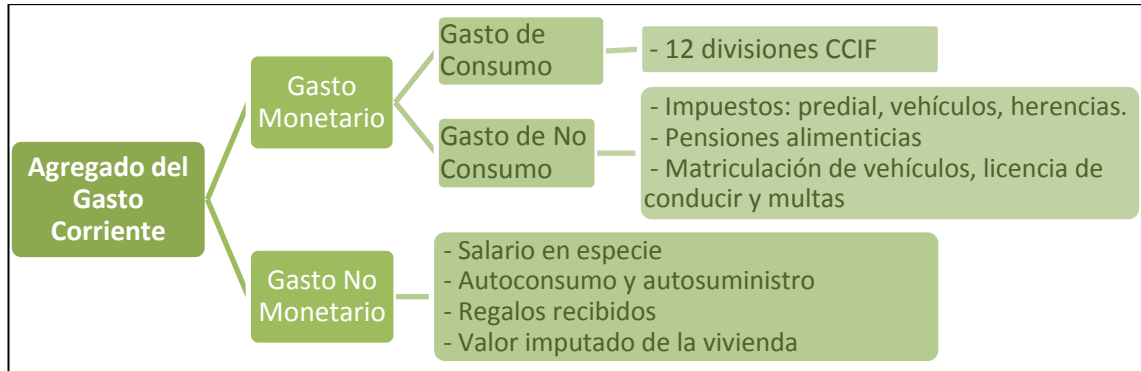


Gráfico 2.3 Estructura del gasto corriente

Fuente: INEC (2011).

CAPÍTULO 3

ANALISIS DESCRIPTIVO DE LA ENIGHUR

En el presente capítulo se realiza un estudio estadístico de las preguntas seleccionadas mediante el análisis univariado, conjunto de técnicas que describen las características de la población objetivo y muestra resultados en diferentes tipos de tablas que resumen la información tabulada de los Hogares Urbanos y Rurales del cantón Guayaquil obtenidos de la aplicación de la ENIGHUR 2011-2012.

3.1. INFORMACIÓN ECONÓMICA DE HOGARES

Dentro de la población objetivo seleccionada encontramos un total de 2,265 hogares ubicados en la ciudad de Guayaquil.

3.2. DISTRIBUCIÓN DE FRECUENCIAS FACTORES DESCRIPTIVOS DE LAS VIVIENDAS, EQUIPAMIENTO Y PERCEPCIÓN DEL NIVEL DE VIDA DEL HOGAR.

Tabla 3.3:

Vía de acceso principal a la vivienda

Vía de acceso principal a la vivienda	Cantidad	Proporción
Carretera, calle pavimentada	1707	0,754
Lastrado, calle de tierra	522	0,230
Empedrado	18	0,008
Sendero, chaquiñan	2	0,001
Otro, cual	16	0,007
Total general	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

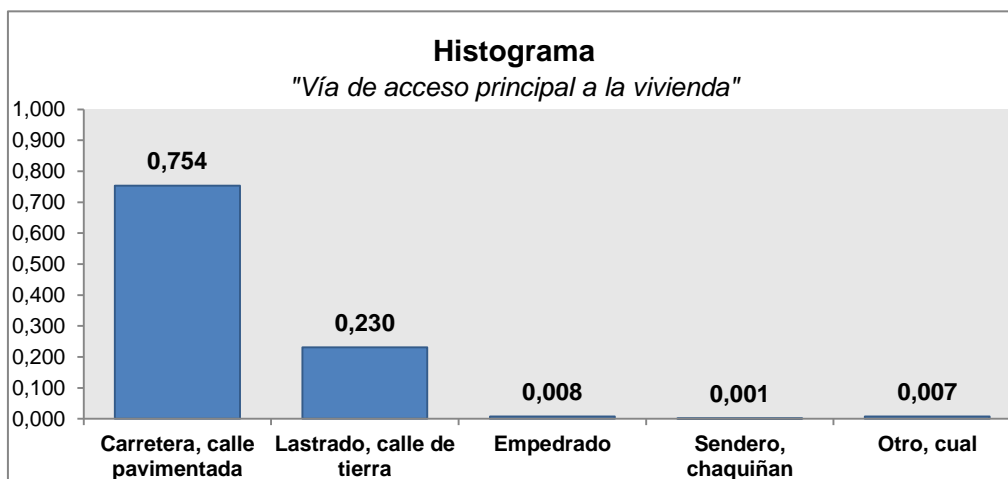


Gráfico 3.2. Vía de acceso a la vivienda

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

En el gráfico anterior se obtienen que las vías de acceso principal a la vivienda son un 75% carreteras o calles pavimentadas, un 23% lastrado o calles de tierra, y el 2% de las vías restantes son senderos, chaquiñán, vías empedradas y otros.

Tabla 3.4:

Tipo de vivienda

Tipo de vivienda	Cantidad	Proporción
Casa villa	1201	0,530
Departamento en casa o edificio	622	0,275
Rancho, choza, covacha	221	0,098
Mediagua	152	0,067
Cuarto (s) en casa de inquilinato	69	0,030
Total general	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

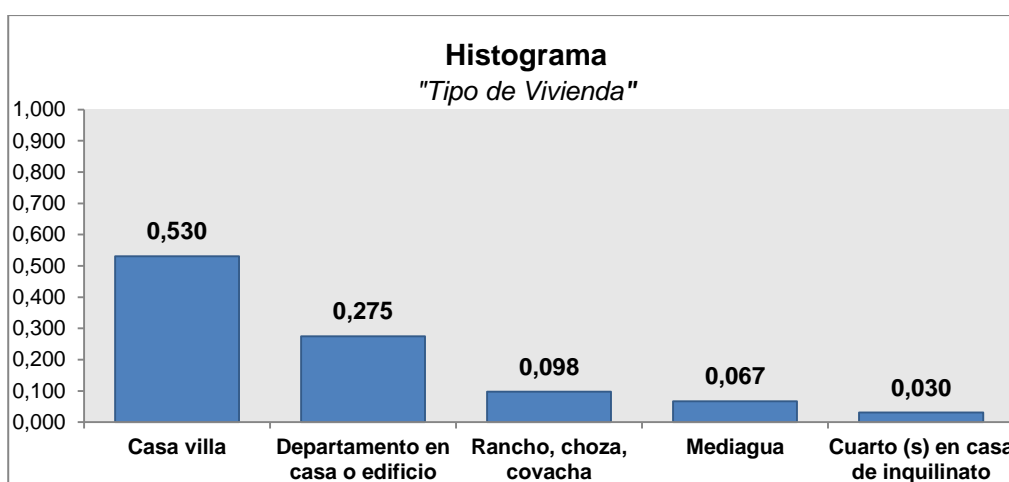


Gráfico 3.3. Vía de acceso a la vivienda

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Un 53% de los hogares estudiados habitan en viviendas Casa Villa

ubicándose en la proporción más alta de la distribución de los tipos de vivienda que son habitables en el cantón Guayaquil, por consiguiente otro 37% se ubican en ranchos, chozas, covacha y departamentos en casas o edificios; el restante 10% en otros tipos no muy frecuente habitados tales como mediagua y cuarto(s) en casa de inquilinato.

Tabla 3.5:

Material predominante del techo

Tipo de Material	Frecuencia	Proporción
Zinc	1325	0,585
Hormigón, losa, cemento	446	0,197
Asbesto (eternit, eurolit)	425	0,188
Madera	63	0,028
Teja	4	0,002
Palma, paja, hoja	1	0,000
Otro	1	0,000
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

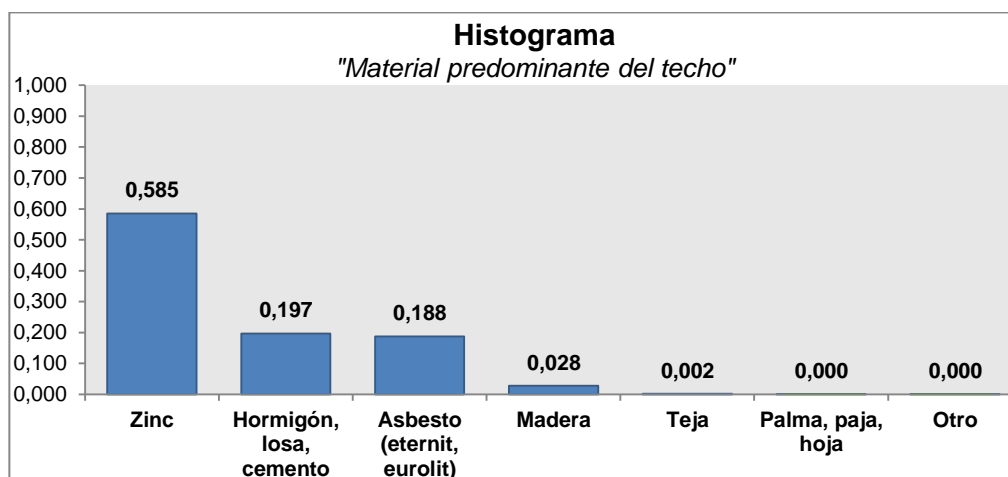


Gráfico 3.4. Material predominante del techo

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

El 59% de los hogares de Guayaquil tiene en el techo de sus viviendas como material predominante las hojas de zinc, seguido por el 38% de las viviendas con techos de asbesto (eternit, eurolit) y hormigón, losa o cemento; el 3% faltante posee techos fabricados de madera, palma, paja, hoja, teja y otros.

Tabla 3.6:

Material predominante de las paredes

Tipo de Material	Frecuencia	Proporción
Hormigón, bloque, ladrillo	2003	0,884
Caña	201	0,089
Asbesto, cemento, fibrolit	31	0,014
Madera	17	0,008
Bahareque	10	0,004
Adobe, tapia	1	0,000
Otro, cual	2	0,001
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

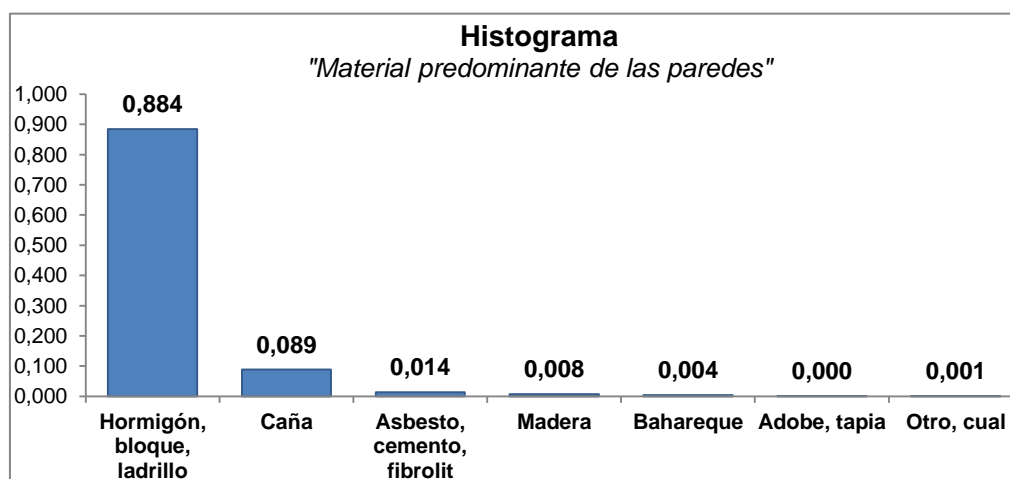


Gráfico 3.5. Material predominante de las paredes

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

En el interior del 88% de viviendas entrevistadas poseen paredes hechas de hormigón, bloque o ladrillo, otro 9% les predomina la caña, mientras que en las restantes viviendas las paredes contienen materiales como madera, asbesto, cemento, fibrolit, bahareque, adobe, tapia entre otros.

Tabla 3.7:

Material predominante del piso

Tipo de Material	Frecuencia	Proporción
Cemento, ladrillo	1045	0,461
Cerámica, baldosa, ninyl	879	0,388
Tabla, tablón no tratado	228	0,101
Tierra	64	0,028
Marmol, marmetón	30	0,013
Otro, cual	10	0,004
Duela, parquet, tabloncillo, piso flotante	6	0,003
Caña	3	0,001
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

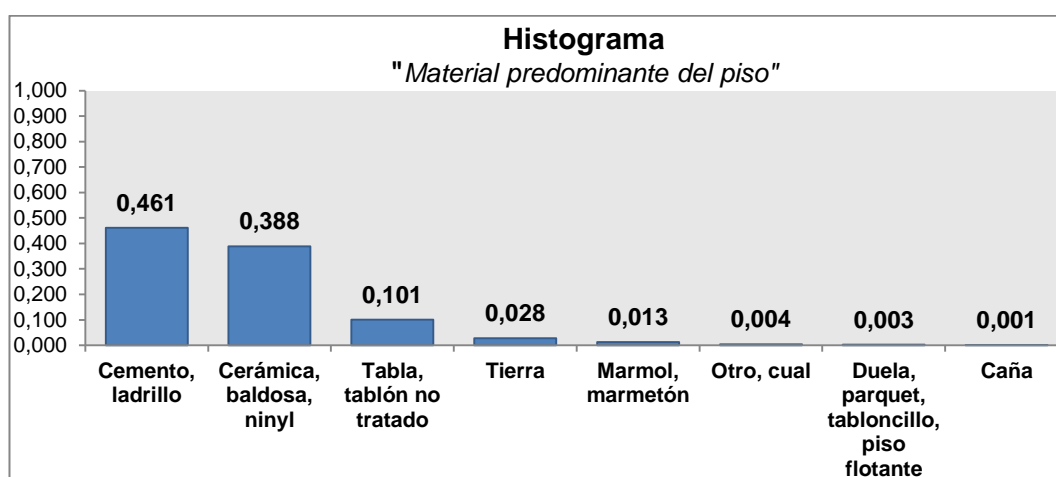


Gráfico 3.6. Material predominante del piso

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Un porcentaje del 46% de los hogares, habitan en casas que poseen como material predominante del piso cemento / ladrillo, seguidos del 39% de los pisos de cerámica, baldosa o ninyl. Los piso hechos de tabla o tablón no tratado son el 10% y el 5% restante de los hogares tienen como material predominante del piso tierra, mármol, marmetón, caña y otros distribuidos en pequeñas proporciones.

Tabla 3.8:

Existencia de cuartos para negocio en la vivienda

Categorías	Frecuencia	Porcentaje
No	2151	0,950
Si	114	0,050
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

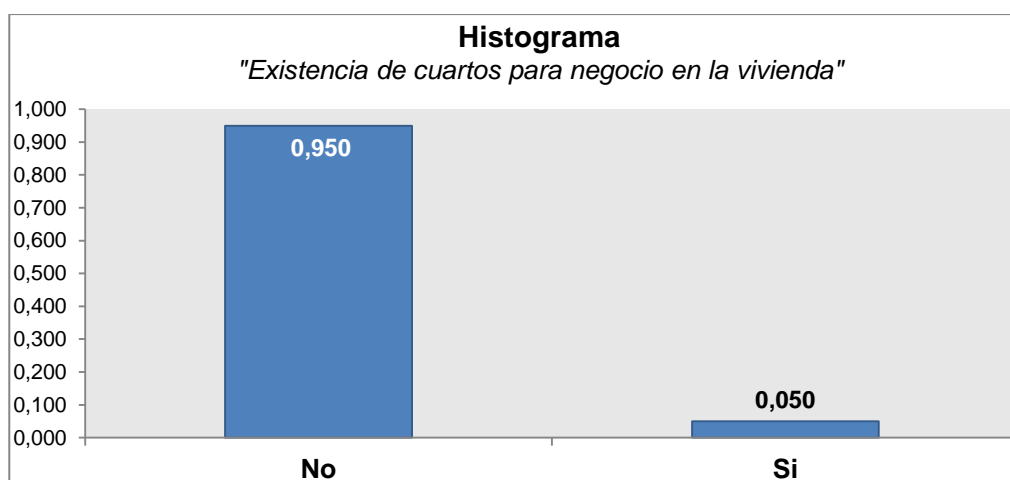


Gráfico 3.7. Existencia de cuartos para negocio en la vivienda

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Los hogares entrevistados informaron que el 95% de ellos no tienen al menos un cuarto para negocio en la vivienda, y solo el 5% si dispone de este sitio en su vivienda.

Tabla 3.9:

Número de cuartos para negocio

Tipo de categorías	Frecuencia	Porcentaje
Un sólo cuarto	113	0,050
Dos cuartos	1	0,000
Total General	144	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

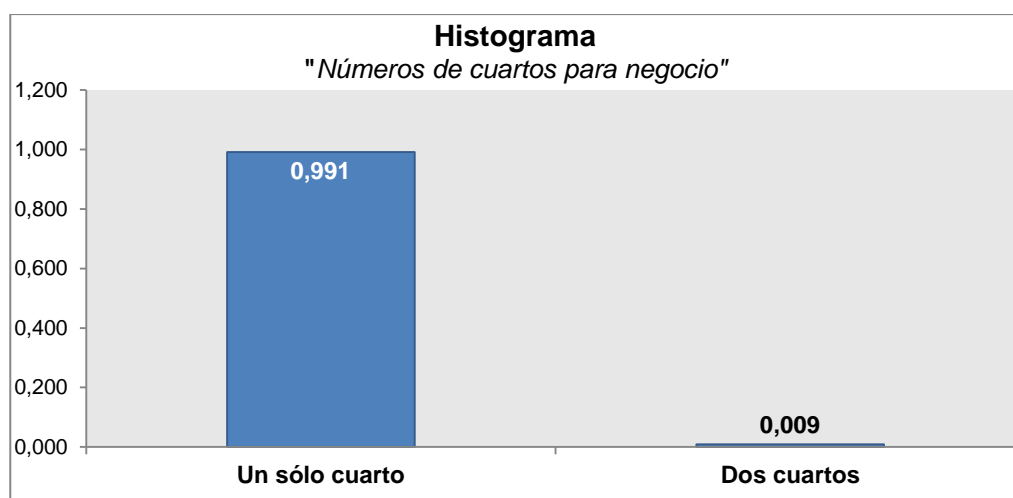


Gráfico 3.8. Números de cuartos para negocio

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

De los 144 hogares del cantón Guayaquil que respondieron que sí tienen cuartos para negocio, el 99% tienen un sólo cuarto para negocio en su vivienda, mientras que el restante 1% poseen más de uno.

Tabla 3.10:

Sitio de la vivienda donde se cocinan los alimentos

Tipo de categorías	Frecuencia	Porcentaje
En cuarto exclusivo	1491	0,658
En sala comedor	410	0,181
En cuarto utilizado también para dormir	268	0,118
No cocinan	77	0,034
En el patio, corredor u otro sitio	19	0,008
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

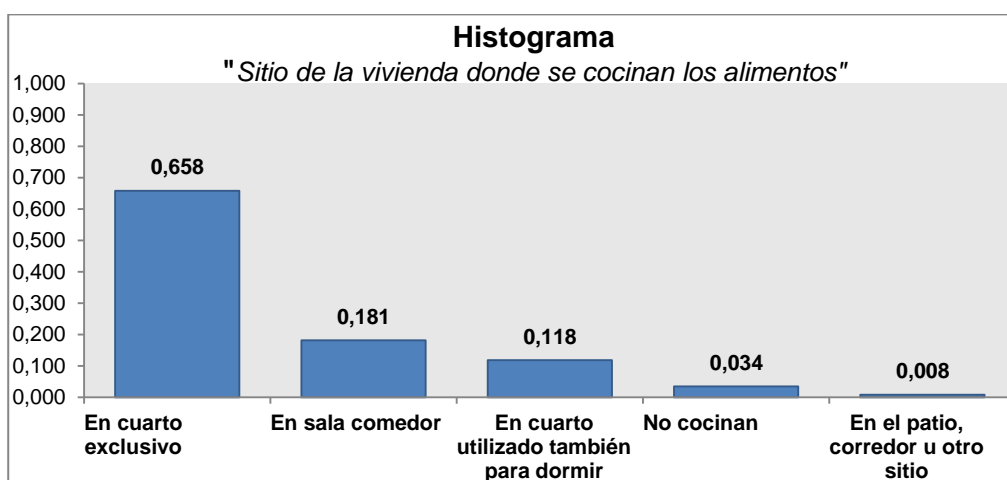


Gráfico 3.9. Sitio de la vivienda donde se cocinan los alimentos

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

El 66% de las viviendas informaron tener un cuarto exclusivo para cocinar, otro 18% de hogares cocinan sus alimentos en la sala comedor mientras que

el restante 16% cocinan en los dormitorios, o utilizan el patio, el corredor u otros sitios, y en casos especiales no cocinan.

Tabla 3.11:

Combustible con que principalmente cocinan los alimentos

Tipos de categoría	Frecuencia	Porcentaje
Gas de tanque o cilindro	2158	0,953
Electricidad	20	0,009
Leña o carbón	7	0,003
Gas centralizado	3	0,001
No Informa	77	0,034
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

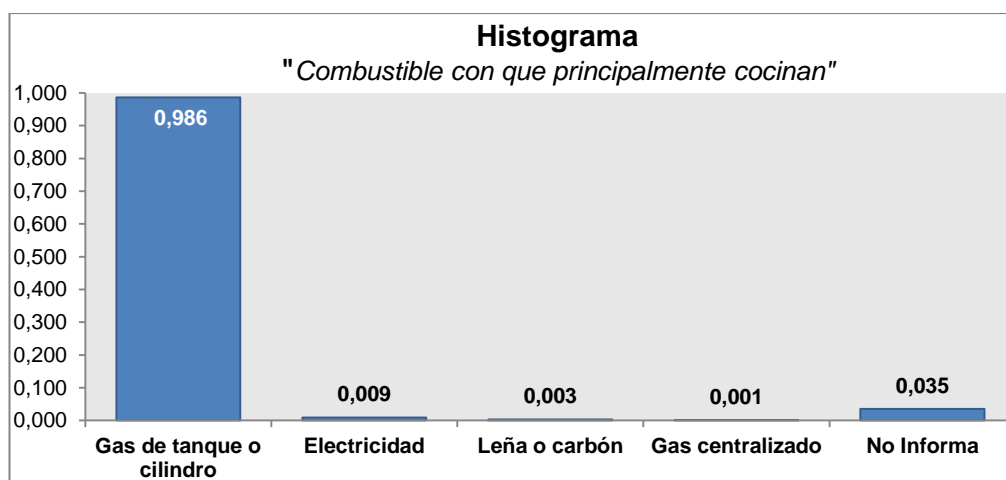


Gráfico 3.10. Combustible con que principalmente cocinan

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

De los 2,188 hogares que informaron que si preparan sus alimentos en casa el 99% lo hacen utilizando como combustible de principal uso el gas de

tanque o cilindro, otro grupo del 1% utiliza leña o carbón mientras que los demás lo hacen por medio de electricidad y gas centralizado.

Tabla 3.12:

Tipo de servicio higiénico del hogar

Tipos de categoría	Frecuencia	Porcentaje
Inodoro y alcantarillado	1515	0,669
Inodoro y pozo séptico	511	0,226
Inodoro y pozo ciego	163	0,072
No tiene	41	0,018
Letrina	35	0,015
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

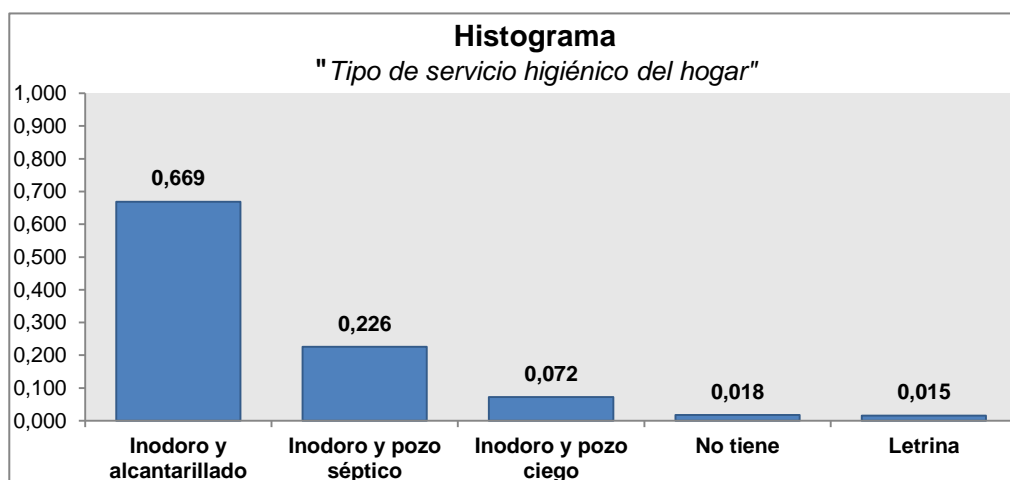


Gráfico 3.11. Tipo de servicio higiénico del hogar

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Como tipo de servicio higiénico predominante de los hogares de Guayaquil son en un 67% inodoros y alcantarillados, seguido de un 23% de los inodoros

y pozo séptico, sin embargo los demás hogares usan los inodoros y pozo ciego, letrina o simplemente no tienen.

Tabla 3.13:

Vía de acceso principal al agua de uso diario

Tipos de categoría	Frecuencia	Porcentaje
Red pública	1939	0,856
Carro repartidor, triciclo	300	0,132
Otro cual	22	0,010
Pila, pileta o llave pública	4	0,002
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

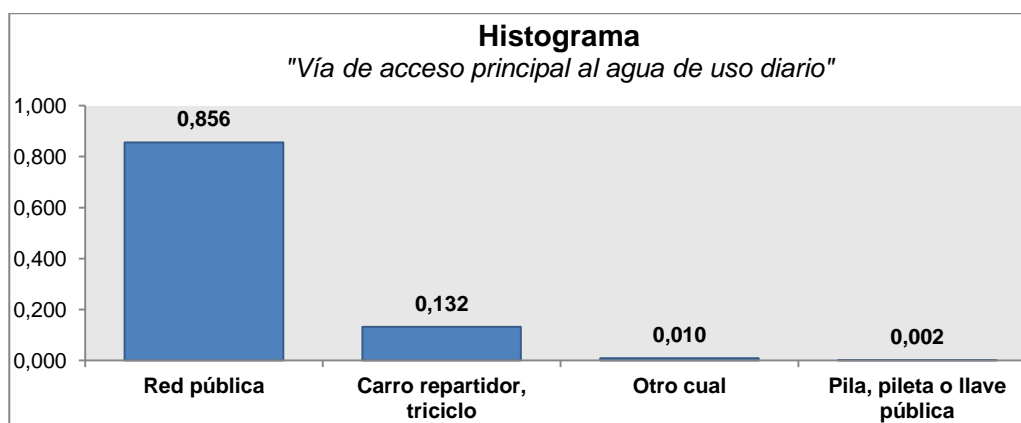


Gráfico 3.12. Vía de acceso principal al agua de uso diario

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

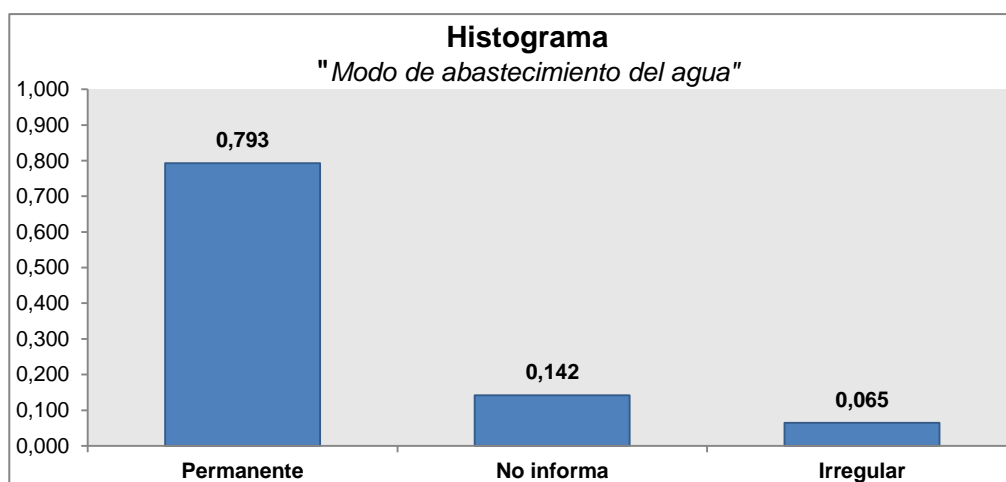
El 86% hogares utilizan agua que llega de la red pública para cocinar sus alimentos como se observa en el gráfico 3.12, mientras que un 23% recogen agua del carro repartidor, triciclo, y una pequeña minoría utilizan agua que pila, pileta o llave pública y otras fuentes de agua.

Tabla 3.14:

Modo de abastecimiento de agua

Tipos de categoría	Frecuencia	Porcentaje
Permanente	1796	0,793
No informa	322	0,142
Irregular	147	0,065
Total General	2265	1,000

Nota Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

**Gráfico 3.13. Modo de abastecimiento del agua**

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

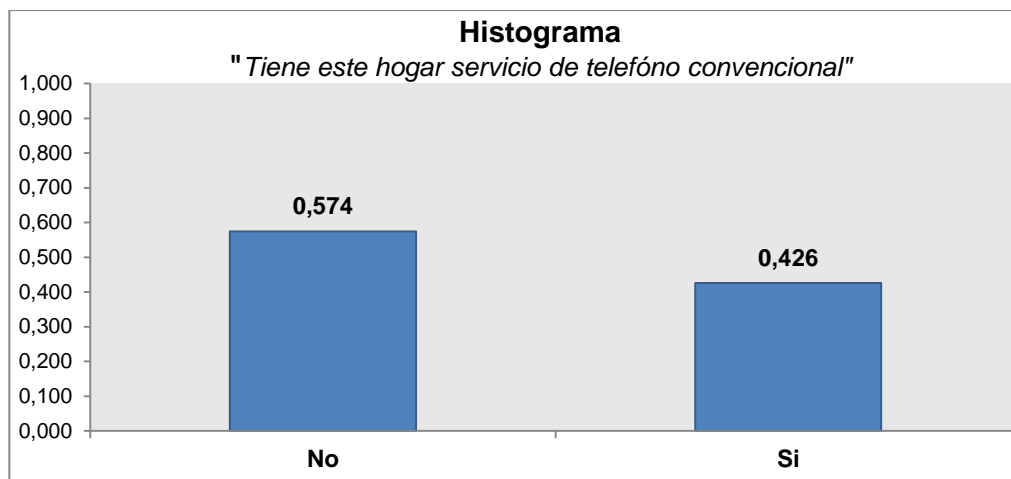
Considerando un frecuencia permanente del abastecimiento del agua, poco más de la mitad el 79% de los hogares tienen esta frecuencia y apenas el 6% indica que el abastecimiento es irregular, mientras que los demás deciden en no informar.

Tabla 3.15:

Tiene este hogar servicio de telefonía convencional

Tipos de categoría	Frecuencia	Porcentaje
No	1301	0,574
Si	964	0,426
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

**Gráfico 3.14. Tiene este hogar servicio de telefonía convencional**

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

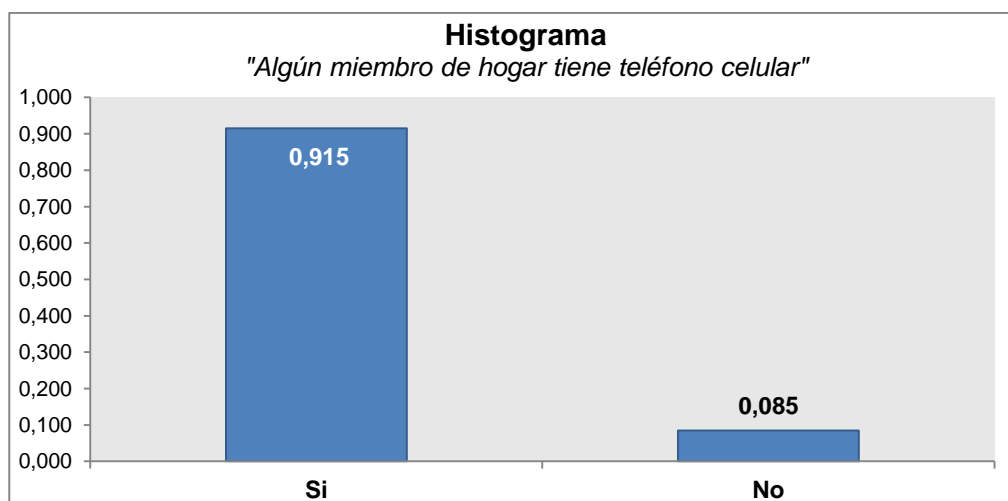
El 57% los hogares disponen de servicio de teléfono convencional mientras que el restante 43% no posee. Los hogares que informaron no poseer telefonía convencional sustituyen este servicio por la telefonía móvil.

Tabla 3.16:

Algún miembro del hogar tiene teléfono celular

Tipos de categoría	Frecuencia	Porcentaje
Si	2073	0,915
No	192	0,085
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

**Gráfico 3.15. Algún miembro del hogar tiene teléfono celular**

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Dentro de los hogares analizados el 92% tienen al menos un miembro del hogar que posee un teléfono celular, el restante 8% que informaron no tener telefonía móvil.

Tabla 3.17:

Distribución de hogares por posesión de telefonía

Hogares por servicio de telefonía	Convencional	Celular
Si	964	2073
No	1301	192
Totales	2265	2265

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

De los 2,265 hogares entrevistados 964 hogares se comunican por medio de teléfonos convencionales, los demás no utilizan este tipo de telefonía o se comunican por algún otro medio, también se observa que 2,073 fueron el número de hogares que utilizan celulares para comunicarse y apenas 192 no utilizan este tipo de comunicación.

Tabla 3.18:

Número de líneas de celular

Número de líneas celular	Frecuencia	Porcentaje
1	740	0,327
2	841	0,371
3	312	0,138
4	110	0,049
5	44	0,019
6	20	0,009
7	4	0,002
10	2	0,001
No informa	192	0,085
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

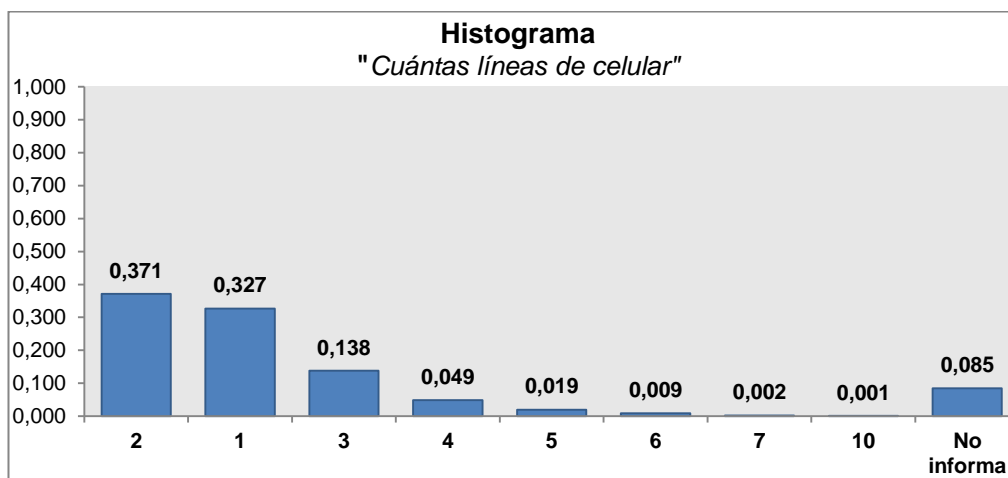


Gráfico 3.16. Número de líneas de celular

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

De los 2265 hogares entrevistados solo 2073 respondieron que tienen líneas de telefonía móvil, el 82% utilizan de una a tres líneas, un 6% superan las tres líneas telefónicas y el faltante 12% no usan teléfonos celulares.

Tabla 3.19:

Tipo de alumbrado con que cuenta el hogar

Tipos de categorías	Frecuencia	Porcentaje
Empresa eléctrica pública	2261	0,998
Planta eléctrica privada	1	0,000
Vela, candil, mechero, gas	3	0,001
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

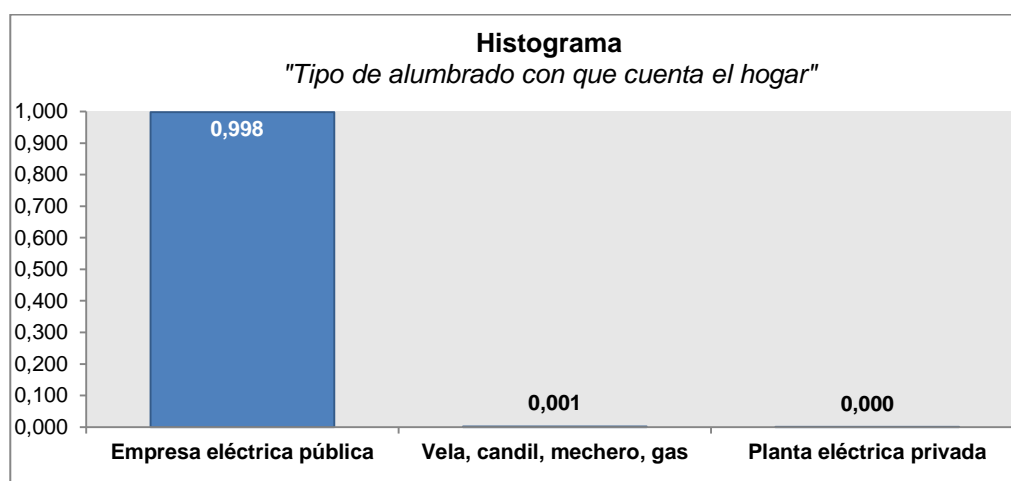


Gráfico 3.17. Tipo de alumbrado con que cuenta el hogar

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí LozaG

Dando una respuesta casi unificada el 99% de los hogares informaron que como principal vía de alumbrado que se utiliza en el hogar proviene de la empresa eléctrica pública, mientras que los demás hogares se alumbran mediante vela, candil, mechero, gas, planta eléctrica privada paneles solares.

Tabla 3.20:

Percepción del nivel de vida mediante los ingresos

Tipo de categoría	Frecuencia	Porcentaje
Viven más o menos bien	1791	0,791
Viven mal	314	0,139
Viven bien	160	0,071
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

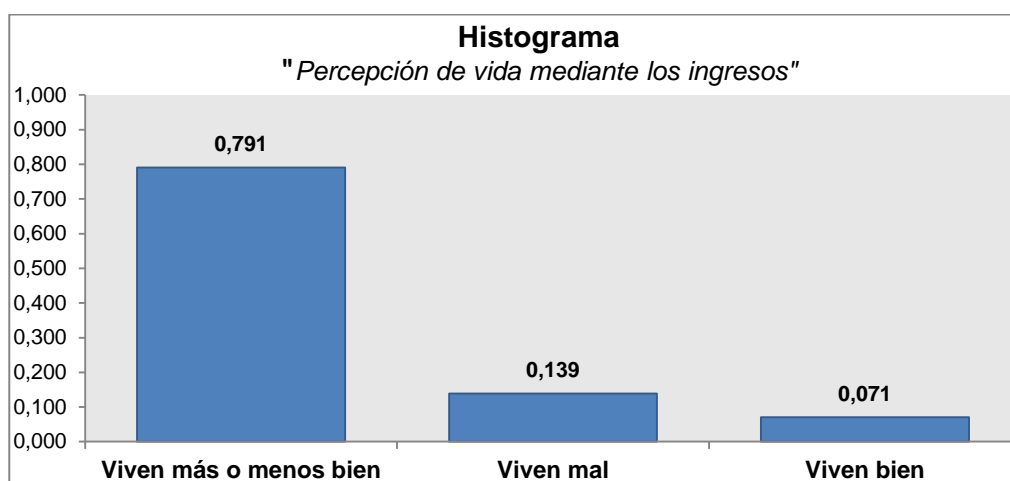


Gráfico 3.18. Percepción del nivel de vida mediante los ingresos

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

Tomando en consideración sus ingresos, los hogares respondieron sólo el 7% que vive bien, el 79% vive más o menos bien y los que tienen una percepción que viven mal son el 14% de este grupo.

Tabla 3.22:

Situación económica de su hogar

Tipos de categoría	Frecuencia	Porcentaje
Apenas logran equilibrar sus ingresos y gastos	1378	0,608
Se ven obligados a endeudarse	602	0,266
Logran ahorrar dinero	185	0,082
Se ven obligados a gastar los ahorros	100	0,044
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

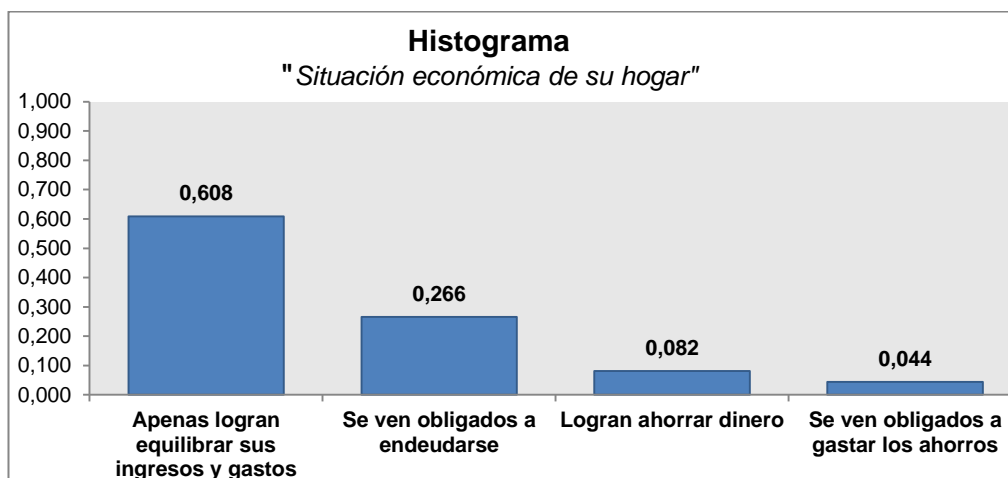


Gráfico 3.20. Situación económica de su hogar

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

Consultando a los entrevistados sobre la actual situación económica de sus hogares, ellos informaron que el 61% apenas logran equilibrar sus ingresos, los hogares que se ven obligados a endeudarse son el 31% y apenas el 8% logran ahorrar dinero, mientras que los demás se ven obligados a gastar los ahorros.

Tabla 3.23:

Nivel de vida del hogar

Tipos de categoría	Frecuencia	Porcentaje
Está igual	1404	0,620
Empeoró	556	0,245
Mejóro	305	0,135
Total General	2265	1,000

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

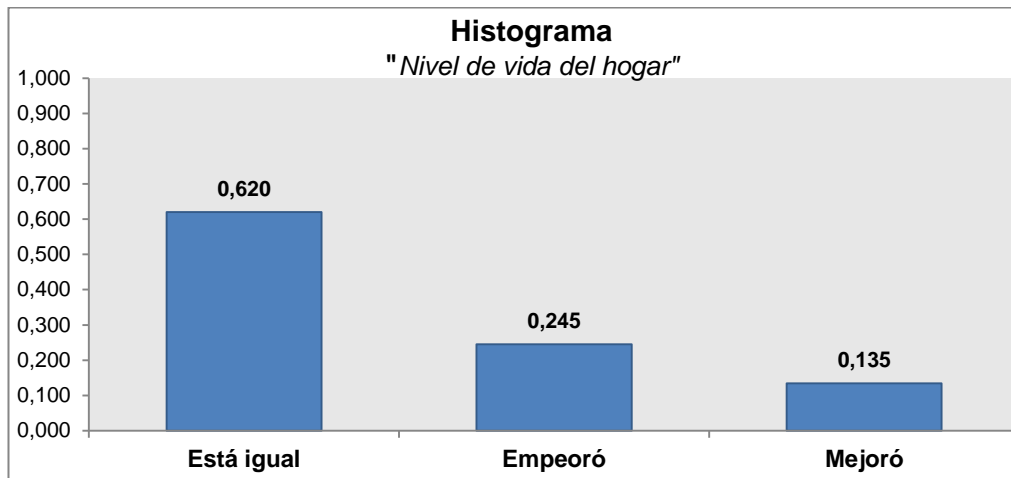


Gráfico 3.21. Nivel de vida del hogar

Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

En la información obtenida de la pregunta, ¿existió algún cambio en el nivel de vida en su hogar?, el 62% manifestó estar igual, el 25% indicaron que empeoraron y sólo el 14% mostró mejora.

CAPÍTULO 4

ANÁLISIS EXPLORATORIO MULTIVARIADO PARA LA PERCEPCIÓN DE RELACIONES ENTRE VARIABLES Y CATEGORÍAS.

Este capítulo fue desarrollado para identificar la estructura de componentes sobre la situación socio económica de los hogares del cantón Guayaquil mediante un análisis de las principales variables de la ENIGHUR 2011-2012. Se aplicó la técnica Análisis de componentes principales categóricos a cada una variables pertenecientes a las diferentes secciones de los tres formularios que dispone esta encuesta.

La ventaja de usar esta técnica no lineal es que se puede utilizar variables categóricas de tipo ordinal y no es necesario que estas posean supuestos de normalidad o linealidad

El propósito de realizar este estudio fue de encontrar patrones que permitan relacionar y distinguir a los hogares de Guayaquil por los tipos de variables con las cuales fueron entrevistados. Obteniendo gran éxito dado que los individuos se aglomeraron dependiendo del tipo de atributo que posee su hogar, como por ejemplo en la sección que describe la vivienda y el hogar, se encontró que los entes formaron tres grupos por tipo de característica que posee su vivienda mientras que en las otras secciones se concentraron tomando una forma similar a una pirámide, la cual se la bautizó como la pirámide socio económica.

También se desarrolló un modelo de predicción logística, que simula en base un grupo de predictores (variables independientes) el evento de interés de los hogares del cantón Guayaquil posean ingresos inferiores al percentil 30 del total de ingresos.

IDENTIFICACIÓN DE UNA ESTRUCTURA USANDO LAS PRINCIPALES VARIABLES DE LA ENIGHUR 2011-2012 MEDIANTE EL ANÁLISIS DE COMPONENTES PRINCIPALES CATEGÓRICO

El principal objetivo de esta tesis como fue ya mencionado en el primer capítulo es disminuir el número de variables analizadas que fueron seleccionadas de la ENIGHUR realizada por el INEC durante el 2011 y el 2012. Ahondamos en un estudio exploratorio en el que esperamos definir un conjunto de pequeñas dimensiones, tratando de concentrar la mayor variabilidad y explicación de los datos.

Entre las variables seleccionadas encontramos diferentes tipos de medición, están las de tipo nominal como por ejemplo el “Sexo” de la persona entrevistada, de tipo ordinal como “El nivel de instrucción” y por último las variables escalares donde tenemos “Ingresos del hogar” y “Gastos del Hogar”, Etc...

Una parte de las variables fueron sometidas a una recodificación de sus categorías, mientras otras que no poseían categorías se transformaron en variables categóricas tomando en consideración la naturaleza la medición.

4.1. REDUCCIÓN DE VARIABLES RELACIONADAS A LA VIVIENDA Y EL HOGAR

Primero se analizan las variables que describen las características principales de la vivienda de los entrevistados.

De las variables involucradas en esta sección, gran parte son medidas en nivel ordinal, seguidas por las ordinales y unas pocas son escalares.

La Tabla 4.1 muestra la matriz de componentes, donde podemos observar a las variables que poseen cargas muy bajas en relación a las dos componentes principales, las que se detallan a continuación: “Tipo de alumbrado de la vivienda” con un valor máximo de 0.284, “Existe cuartos para negocio en la vivienda” con un máximo de 0.222 y “Metros cuadrados de construcción de la vivienda” con 0.445.

Tabla 4.1:

Matriz de Componentes Sección I. Datos de la Vivienda y Hogar (12 Variables)

	Dimensión	
	1	2
Tiene este hogar servicio telefónico convencional	.644	-.269
Tipo de servicio higiénico	.859	-.009
Manera de adquisición de vivienda	-.245	-.885
Vía de acceso principal a la vivienda	.767	.030
Material predominante del piso de la vivienda	.775	-.094
Fuente de obtención del agua	.753	-.004
Tipo de vivienda	.792	.141
Existe cuartos para negocio en la vivienda	.112	-.222
Tipo de alumbrado de la vivienda	.284	-.103
Adquisición del terreno donde está construida la vivienda	.078	.702
Metros cuadrados de construcción de la vivienda	-.445	.410
Años de construcción que tiene la vivienda	-.486	-.177

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Se observa en el Gráfico 4.1 la formación de 3 grupos de individuos que muestran indicios de las relaciones de los entes estudiados.

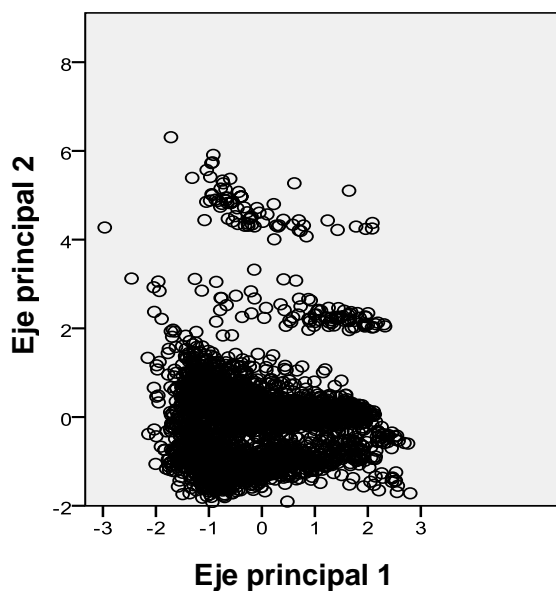


Gráfico 4.1 Individuos por características de la vivienda

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

En este conjunto de características se visualiza que los individuos forman grupos que son separados en el eje principal 2, con el propósito aclarar la distribución de los individuos se procede a continuar con el estudio mediante la matriz de componente con el objetivo de eliminar las variables que generen las menores aportaciones.

Las variables mencionadas anteriormente en la Tabla 4.1 muestran cargas muy bajas por lo que se toma la decisión de removerlas y aplicar el análisis de componentes principales categóricos una vez más.

En la tabla 4.2 se presenta a los resultados luego de aplicar por segunda vez la técnica el análisis de componentes principales categórico con las variables que poseen las cargas más altas.

Tabla 4.2:

**Matriz de Componentes Sección I. Datos de la Vivienda y Hogar
(9 Variables)**

	Dimensión	
	1	2
Tiene este hogar servicio telefónico convencional	.653	-.062
Tipo de servicio higiénico	.873	-.047
Manera de adquisición de vivienda que ocupa este hogar	-.010	-.901
Vía de acceso principal a la vivienda	.786	-.040
Años de construcción que tiene la vivienda	-.488	-.210
Material predominante del piso	.779	-.050
Adquisición del terreno donde está construida la vivienda	.023	.893
¿De dónde obtiene el agua principalmente?	.771	-.036
Tipo de vivienda	.800	.058

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

De hecho podemos observar que las variables analizadas tienen cargas altas superiores al 0.450, ya estén estas relacionadas a la componente 1 o a la componente 2. Por lo que decidimos preservar las variables para un próximo análisis donde serán incluidas con otras variables de las siguientes secciones.

En el Gráfico 4.2 se distingue una agrupación de individuos en tres conjuntos diferenciados por el segundo eje principal. Estos grupos muestran a los

individuos aglomerados por el tipo vivienda que poseen. En el primer grupo se observan a los individuos que habitan las viviendas que fueron construidas con materiales de calidad y que poseen los mejores servicios para el hogar, seguidos por el grupo de individuos cuyas viviendas fueron construidas con pocos materiales de buena calidad en donde los servicios básicos del hogar se encuentran en buenas condiciones, y por último donde se distingue mayor aglomeración de individuos encontramos a las vivienda construidas con materiales más comunes de fácil adquisición por su precio y comercialización. Estos hogares no disponen en muchos casos de servicios básicos lo que puede ser perjudicial a la salud a largo plazo de los individuos.

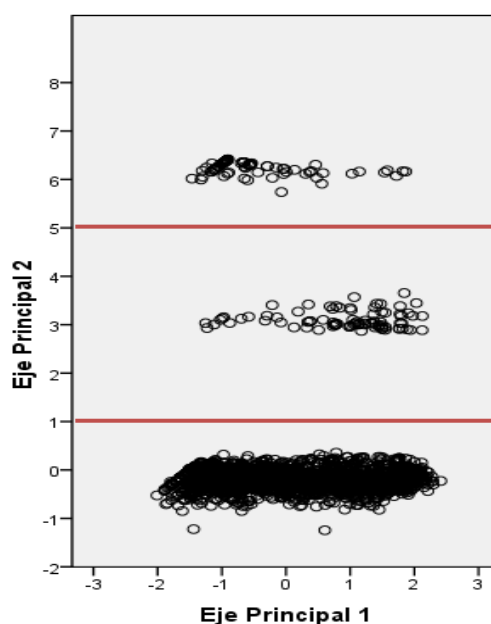


Gráfico 4.2 Individuos por características de la vivienda reducido

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

4.1.1. REDUCCIÓN DE VARIABLES RELACIONADAS AL EQUIPAMIENTO Y CARACTERÍSTICAS DEL HOGAR.

En la sección II del Formulario 1 de la Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos y Rurales 2011 – 2012, se detalla el equipamiento de los hogares en función de la posesión de bienes durables.

La información requerida con relación a estos bienes son el número de bienes y su valor medido en el tiempo del bien.

Esta sección tiene como objetivo encontrar el patrimonio de los hogares medidos en bienes durables, y saber si se encuentra en posesión y acceso del artículo, así como de terrenos, casas, departamentos y locales comerciales.

En el proceso de tabulación de la base se pudo observar la existencia de muchos datos perdidos, por lo que procedió a explorar las variables con estadística descriptiva, obteniendo los siguientes resultados; de los 28 ítems involucrados en esta sección se obtuvo que sólo la variable Nro. De Refrigeradoras constaba con pocos datos perdidos, que en comparación con el resto de variables de esta sección resulta ser la única variable a considerarse para el análisis multivariado.

4.1.2. REDUCCIÓN DE VARIABLES RELACIONADAS A LA INFORMACIÓN DE LOS MIEMBROS DEL HOGAR Y CARACTERÍSTICAS OCUPACIONALES

En la sección III, IV, V del Formulario 1 de la Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos y Rurales 2011 – 2012, se mide la pobreza subjetiva de los hogares, observado desde la visión del jefe del hogar y su percepción del nivel de vida. Además, clasifica a los integrantes del hogar por su nivel de instrucción y su acceso a programas sociales.

Tabla 4.3:

Matriz de Componentes Sección III, IV y V. Miembros del Hogar, Percepción del Nivel de Vida y Características Ocupacionales (14 Variables)

	Dimensión	
	1	2
Nivel de vida del hogar	-.329	-.061
Seguro afiliado	-.648	-.031
Sabe leer y escribir	-.536	.039
Acciones en actual situación económica de su hogar	-.260	-.187
Edad	-.375	.465
Nivel de instrucción	.774	.083
Se matriculó en un algún centro de educación	.276	.048
Monto mínimo mensual estimado para vivir bien	.485	.153
Estado civil	.294	-.838
Sexo	-.184	.795
Percepción de vida a base del hogar.	-.442	-.176
Categoría de ocupación	-.587	-.012
Autodefinición étnica	.356	.263
¿Considera que su hogar es pobre?	.557	.303

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

La Tabla 4.3 describe las cargas de las componentes por cada variable de donde muestra cuatro variables que poseen cargas muy bajas en valor absoluto; la variable “Nivel de vida del hogar” genera una carga máxima de -0.329, “Acciones en actual situación económica de su hogar” con un valor de -0.260, “Se matriculó en un algún centro de educación” con 0.276 y por último “Autodefinición Étnica” posee como valor máximo un 0.356. Se decide eliminar estas variables y se vuelve a realizar el análisis.

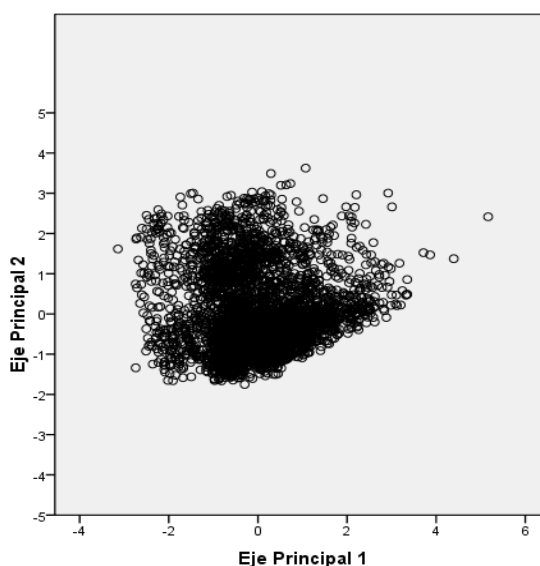


Gráfico 4.3 Individuos por percepción de nivel de vida y características ocupacionales.

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

En la Tabla 4.4 de matriz de componentes de la sección IV y V notamos que todas variables resultantes tienen un carga máxima que supera a 0.400

correspondiente a una de las dos componentes, por lo preservamos estas variables para un próximo análisis.

Tabla 4.4:

Matriz de Componentes Sección III, IV y V. Miembros del Hogar, Percepción del Nivel de Vida y Características Ocupacionales (10 V.)

	Dimensión	
	1	2
Seguro afiliado	-.668	-.149
Sabe leer y escribir	-.569	-.016
Edad	-.408	.441
Nivel de instrucción	.782	.161
Estado civil	.369	-.827
Sexo	-.256	.784
Percepción de vida a base de los ingresos del hogar	-.408	-.205
Categoría de ocupación	-.601	-.136
¿Considera que su hogar es pobre?	.522	.347
Monto mínimo mensual estimado para vivir bien	.490	.216

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Entre las variables que más información aportan a la componente 1 tenemos al “Nivel de Instrucción” con un peso de 0.782, seguido de la variable “Seguro del Afiliado” con -0.668 y a “Categoría de ocupación” con un peso de -0.601, mientras que para la componentes 2 observamos que las variables más explicativas son “Estado Civil” aportando con un peso de -0.827 y a “Sexo” con un peso de 0.784. Observamos en el Gráfico 4.4 una figura en forma de triángulo donde en la parte superior se encuentran los individuos con mayor nivel de instrucción, los mismo que perciben tener un mejor estilo de vida en

base a los ingresos que su hogar obtiene. En el centro se encuentran las personas que tiene un nivel de educación medio a básico, estas personas que en su mayoría son empleados dependientes de empresas públicas o privadas que poseen un monto mínimo de ingreso que fluctúa entre un sueldo básico y los 600 dólares. Mientras que en la base del triángulo se ubican las personas que no están afiliadas a ningún tipo de seguro médico, gran parte de ellos no saben leer ni escribir, son trabajadores informales o independientes y tiene una percepción de pobreza dado al monto necesario que ellos necesitan para vivir su día a día.

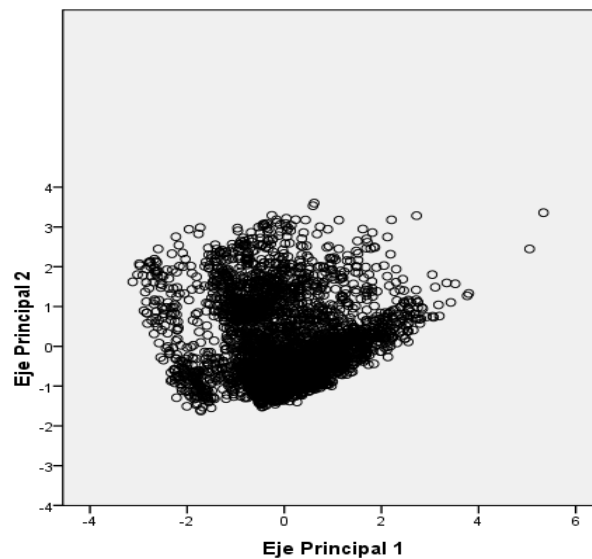


Gráfico 4.4 Individuos por percepción de nivel de vida y características ocupacionales reducido.

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

4.1.3. REDUCCIÓN DE VARIABLES RELACIONADAS A INGRESOS Y TRANSACCIONES FINANCIERAS DE LOS PERCEPTORES.

Aquí encontramos a las variables de la sección VI del formulario 1 de la ENIGHUR que trata de captar todas las formas de ingreso del hogar, sean tanto en especie y en efectivo, destinadas al consumo del hogar o a los miembros que conforman el mismo.

Entre las formas de ingreso de la principal ocupación de los perceptores del hogar tenemos: al ingreso monetario como principal forma de ingreso del perceptor, a las deducciones de sus ingresos, los ingresos en especie, otros ingresos del asalariado y el ingreso neto. El ingreso neto es la combinación lineal entre los ingresos monetarios más los ingresos en especie con la disminución de las deducciones relacionadas al ingreso monetario.

Al tener variables que son combinaciones lineales de unas de otras, procedemos a estudiar sólo a los ingresos monetarios por ser la mayor fuente relaciona con el perceptor, y esta será la que intervendrá en el análisis de componentes.

La variable seleccionada para la medición multivariada posee la menor cantidad de valores perdidos, proporcionando de esta manera mayor información de las características de los individuos con respecto al sueldo como asalariado.

4.1.4. REDUCCIÓN DE VARIABLES RELACIONADAS AL GASTO DENTRO DEL HOGAR.

Con relación a los gastos del hogar se dispone de 3.563 variables que detallan los gastos diarios, mensuales, trimestrales, semestrales y anuales del hogar; a estos gastos se los agrupó en 22 macros gastos tomando en consideración la naturaleza de cada costo incurrido.

Durante la tabulación de estas 22 macro tipologías de gastos, se observó gran cantidad de datos perdidos, por lo que se procede analizar su información descriptiva obteniendo que sólo 7 de estas nuevas variables mantienen un porcentaje de datos perdidos no más del 5% lo que las hace aceptable para ser consideradas en esta investigación, como se muestran a continuación:

Tabla 4.5:

Matriz de Componentes de Gastos Dentro del Hogar (7 Variables)

	Dimensión	
	1	2
Vehículos y transporte	.731	-.155
Productos a base de harina de trigo	.465	.661
Artículos de vestir	.760	-.002
Gastos personales	.633	-.125
Salud y medicinas	.533	-.485
Gastos en suministros de hogar	.679	-.254
Te y bebidas alcohólicas	.606	.525

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Los gastos definidos poseen altas cargas y se mantienen para el análisis.

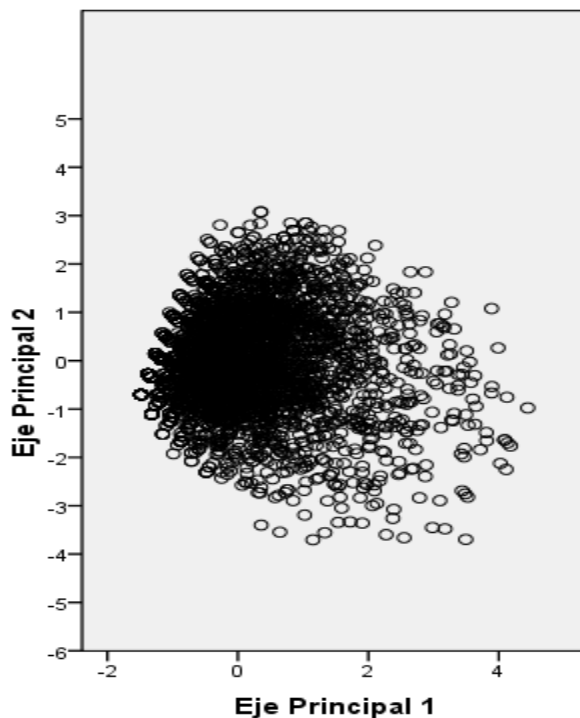


Gráfico 4.5 Individuos por monto de gasto destinado para el hogar

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

En el Gráfico 4.5 no logramos distinguir ninguna formación de los individuos analizados, por lo que se procede con las siguientes secciones del formulario para luego ejecutar la técnica con todas las variables que se conservaron en los previos análisis de las secciones anteriores.

4.1.5. REDUCCIÓN DE VARIABLES RESULTANTES EN LA DISMINUCIÓN PARCIAL EN CADA SECCIÓN

En este bloque analizamos las variables con mayores cargas resultantes de la reducción parcial en cada sección.

Agrupamos las variables que obtuvieron las mejores cargas direccionadas a una de las dos componentes en cada sub análisis, uniendo 9 variables de la sección I. 'Datos de la Vivienda y Hogar', con la variable "Nro. De Refrigeradoras" de la sección II 'Equipamiento y Propiedades del Hogar'; las 10 variables del análisis conjunto de la sección III 'Percepción del Nivel de Vida' con la sección IV. 'Información de los Miembros del Hogar' y la sección V. 'Características Ocupacionales'; procedidas por los "Ingresos Monetarios" y los 7 ítems generadores de "Gastos del Hogar."

En la Tabla 4.6 muestra a las variables sometidas en este estudio, donde se puede observar a la variable "Categoría de ocupación" con un valor de 0.416, a la variable "Manera de adquisición de vivienda que ocupa este hogar" con 0.407, "Percepción de vida a base de los ingresos del hogar" con 0.358, "Sabe leer y escribir" con 0.312, "Adquisición del terreno donde está construida la vivienda" con 0.291, "Edad" con 0.429, "Productos a base de harina" con -0.305, "Años de construcción que tiene la vivienda" con 0.389, "Te y bebidas alcohólicas" con -0.374, "Gastos de los miembros del hogar" con -0.414, las mismas que poseen los valores más bajos en la primera solución del análisis.

Tabla 4.6:

Matriz de Componentes de las Secciones I, II, III, IV, V

	Dimensión	
	1	2
Transporte y vehículos	-.604	-.352
Productos a base de harina	-.305	-.169
Artículos de Vestir	-.598	-.345
Gastos personales	-.395	-.414
Salud y medicinas	-.467	-.068
Te y bebidas alcohólicas	-.374	-.312
Gastos en suministros del hogar	-.754	-.068
Vía de acceso principal a la vivienda	.612	-.434
Tipo de vivienda	.633	-.353
Material predominante del piso de la vivienda	.734	-.207
Tipo de servicio higiénico en la vivienda	.730	-.408
¿De dónde obtiene el agua principalmente?	.585	-.413
Tiene este hogar servicio telefónico convencional	.621	-.227
Manera de adquisición de vivienda que ocupa este hogar	.407	-.148
Adquisición del terreno donde está construida la vivienda	.291	-.201
Años de construcción que tiene la vivienda	-.351	.389
Nivel de instrucción del entrevistado	-.685	-.086
Monto mínimo mensual estimado para vivir bien	-.572	-.228
¿Considera que su hogar es pobre?	-.591	.008
Sexo	.002	.603
Estado civil	-.113	-.648
Edad	-.025	.429
Percepción de vida a base de los ingresos del hogar.	.358	.146
Sabe leer y escribir	.312	.076
Categoría de ocupación	.416	.268
Seguro afiliado	.528	.274
Total ingresos monetarios	-.660	-.245

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

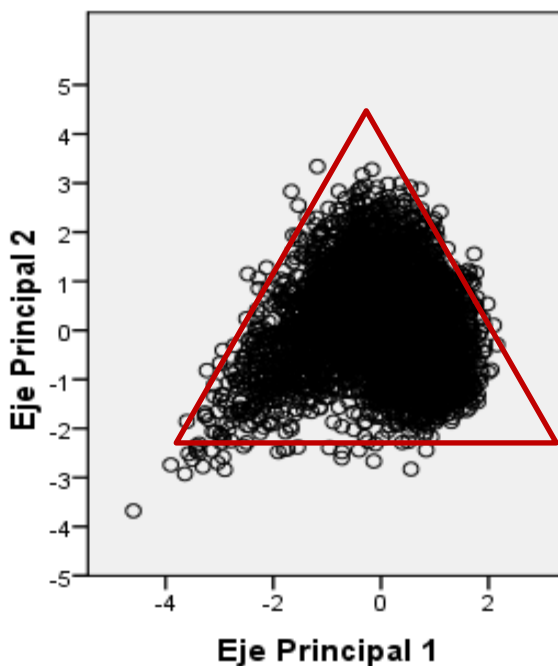


Gráfico 4.6 Individuos por nivel socio económico

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

En el Gráfico 4.6 se puede distinguir a los individuos formando una figura tipo pirámide, representando de esta manera la pirámide de económico poblacional, por lo que se analizan los pesos de las componentes para así determinar cuáles variables son las que menos aportan al estudio, para luego ser eliminadas y así aplicar una vez más la técnica tratando de acentuar la formación de los individuos en forma piramidal.

Una vez eliminadas las variables definidas en el análisis anterior se procede aplicar la técnica a las variables resultantes una vez más obteniendo los siguientes resultados.

Tabla 4.7:

Matriz de Componentes de las Secciones I, II, III, IV, V

	Dimensión	
	1	2
Tiene este hogar servicio telefónico convencional	.645	-.148
Seguro afiliado	.509	.292
Monto mínimo mensual estimado para vivir bien	-.572	-.289
Estado civil	.121	.711
Vía de acceso principal a la vivienda	.634	-.403
Artículos de vestir	-.579	-.333
¿De dónde obtiene el agua principalmente?	.607	-.408
Transporte y vehículos	-.595	-.377
Total ingresos monetarios	-.646	-.279
Material predominante del piso de la vivienda	.758	-.177
Salud y medicinas	-.481	-.170
Gastos en suministros del hogar	-.732	-.134
Nivel de instrucción del entrevistado	-.680	-.069
Tipo de servicio higiénico en la vivienda	.747	-.378
Sexo	-.021	.707
¿Considera que su hogar es pobre?	-.596	-.019
Tipo de vivienda	.646	-.319

Nota Fuente: INEC (2011). *ENIGHUR 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

De la Tabla 4.7 se obtuvo que los hogares del cantón Guayaquil son descritos por el tipo de material del piso de sus viviendas, el tipo de servicio higiénico que tiene la vivienda y lo que gastan en insumos de día a día, tales como: como productos de limpieza, enseres de cocina, mejoras de la infraestructura de la vivienda, entre otros.

Otra característica que describe a los hogares del cantón Guayaquil es nivel máximo de instrucción que tiene el receptor de ingreso o el jefe del hogar.

Existe una cierta similitud entre los hogares dado por el tipo de vivienda en el que habitan y el total de los ingresos que generan, el estudio mostró que los individuos que disponen de más ingresos monetarios viven en mejor infraestructura que los más pobres. Muy similar a estas características está la disponibilidad de acceso de los individuos a tener telefonía convencional en sus hogares estando bastante relacionada al ingreso y al tipo de vivienda.

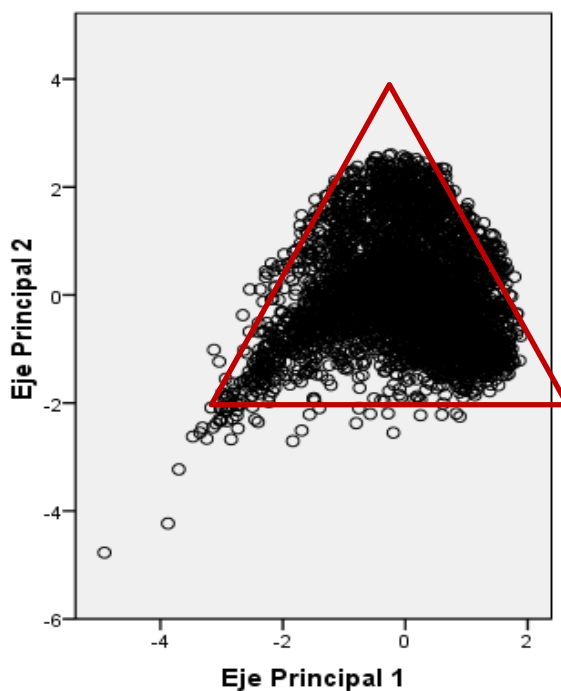


Gráfico 4.7 Individuos por nivel socio económico resultante

Fuente: INEC (2011). ENIGHUR 2011 - 2012. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

4.2. MODELO DE REGRESIÓN LOGÍSTICA PARA EVALUAR EL DESARROLLO DEL NIVEL SOCIOECONÓMICO DE LOS HOGARES DEL CANTÓN GUAYAQUIL

En esta sección se trata de diseñar un modelo para evaluar la capacidad de desarrollo económico de los hogares del cantón Guayaquil que fueron entrevistados durante la ENIGHUR 2011 - 2012.

4.2.1. IMPLEMENTACIÓN DE UN MODELO DE REGRESIÓN LOGÍSTICA PARA PREDECIR EL INGRESO DE LOS HOGARES

En esta sección se desarrolló un modelo capaz de predecir “*Los ingresos netos de los hogares del cantón Guayaquil menores al percentil 30 correspondiente a \$ 334.72*” estos ingresos se obtuvieron en la ENIGHUR 2011-2012. Como variables independientes o predictoras se consideró a 10 variables entre ellas de tipos categóricos nominales, ordinales y variables continuas.

La variable de interés fue convertida a variable dicotómica asignándole el valor 1 a los hogares que respondieron tener ingresos menores o iguales a \$ 334.72, y 0 para las respuestas que indicaron una falla o ausencia de la característica de interés, es decir, para los hogares donde sus ingresos superaban los \$ 334.72.

A continuación se detallan el conjunto de predictores usadas para evaluar la variable de interés

1. *Número de dormitorios que posee la vivienda*
2. *¿Dispone la vivienda de cuartos para negocio?*
3. *¿Dispone vivienda de servicio higiénico?*
4. *¿Algún miembro de la familia tiene teléfono celular?*
5. *Metros cuadrados de construcción de la vivienda*
6. *Monto mínimo mensual estimado para vivir bien*
7. *Edad del jefe del hogar*

8. ¿El jefe del hogar asegurado?
9. Tiene tarjeta de crédito (bancaria o comercial)
10. Ingresos por transferencias corrientes, prestaciones recibidas e inversiones

Se aplicó el análisis de regresión logística a las variables seleccionadas para el modelo, utilizando el software estadístico SPSS versión 17 seleccionando en el menú analizar la opción Regresión / Logística Binaria. Usando esta opción trabajamos con la técnica de regresión logística de tipo directo donde se introducen todos los predictores en la ecuación simultáneamente tanto como la tolerancia no sea afectada. Con este método vamos a evaluar la contribución de cada una de las variables predictoras sobre y bajo las demás.

Con las vista de resultados que posee SPSS se pudo obtener una abrumadora cantidad de información proveída de la regresión logística.

Tabla 5.1:

Resumen del procesamiento de los casos

		Cantidad	Porcentaje
Casos seleccionados	Incluidos en el análisis	2,265	100.0%
	Casos perdidos	0	0%
	Total	2,265	100%
Casos no seleccionados		0	0%
Total		2,265	100%

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

Lo primero que verificamos es el detalle concerniente al tamaño de la muestra que presenta el cuadro de resumen de casos, aquí vemos el software utilizó para ejecutar la técnica regresión logística todos los 2265 hogares del cantón Guayaquil.

Verificamos en la tabla 5.2 las categorías de la variable dependiente y su respectiva codificación, con la se procede a realizar la predicción de los datos.

Para el modelo se creó una nueva variable en base a la variable dependiente seleccionada “Ingresos netos de los hogares del cantón Guayaquil” convirtiéndola de tipo continua a dicotómica.

Tabla 5.2:

Codificación de la variable dependiente

Valores Originales	Valores Internos
Ingresos menores o iguales a los \$334,72	1
Ingresos superiores a los \$334,72	0

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012*. Quito – Ecuador.

Elaborado por: Jhon Quimí Loza

La codificación consistió en calcular el percentil 30 de la variable de interés obteniendo un valor de \$334,72, donde a los valores menores o iguales a este percentil se les asignó el valor de 1 y a los casos que superaban este estadístico se les colocó un cero.

En la tabla 5.3 siguiente se detalla la nueva codificación de las variables categóricas utilizadas como predictores para este modelo.

Tabla 5.3:

Codificaciones de las variables categóricas

	Categorías	Frecuencia	Codificación de parámetros
Tiene tarjeta de crédito (bancaria o comercial)	Si	402	.000
	No	1,863	1.000
¿Dispone vivienda de servicio higiénico?	Si	2,189	.000
	No	76	1.000
¿Algún miembro de la familia tiene teléfono celular?	Si	2,073	.000
	No	192	1.000
¿El jefe del hogar asegurado?	Si	1,021	.000
	No	1,244	1.000
¿Dispone la vivienda de cuartos para negocio?	Si	114	.000
	No	2,151	1.000

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012.* Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Entre las variables del predictivas tenemos a 5 variables categóricas y 5 variables numéricas lo que hace un total de 10 variables utilizada, en la tabla superior se describen las categorías de las 5 variables categóricas del modelo.

Tabla 5.4:

Tabla de Clasificación

Observadas			Predicción		
			Nivel de Vida		Porcentaje Correcto
			Ingresos superiores a los \$334,72	Ingresos menores o iguales a los \$334,72	
Paso 1	Ingreso de neto	Ingresos superiores a los \$334,72	0	1129	100 %
		Ingresos menores o iguales a los \$334,72	0	1136	0 %
Porcentaje general					50.2%

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

En la Tabla 5.4 anterior observamos a un modelo donde sólo se considera a la constante, observamos a un modelo hipotético que no posee variables predictoras. En este modelo asumimos que todos los entes respondieron que tienen ingresos superiores a los \$334,72 por lo que ninguno de los entrevistados cumplió con el evento de interés.

Tabla 5.5:

Resumen del modelo

Paso	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2,124.776	.361	.482

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Los estadísticos Cox & Snell R Square y Nagelkerke R Square mostrados en el resumen del modelo nos proveen de un indicador cuanta cantidad de varianza de la variable dependiente es explicada por el modelo.

Estos estadísticos sugieren que entre 36.1% y el 48.2% de variabilidad es explicada por este conjunto de variables.

Tabla 5.6:

Prueba Omnibus del Modelo de Coeficientes

		Chi-square	df	Sig.
Paso 1	Paso	1015.160	10	.000
	Bloque	1015.160	10	.000
	Modelo	1015.160	10	.000

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

En la tabla anterior se prueba a las variables predictoras seleccionadas para nuestro modelo. La prueba Omnibus nos brinda un indicador general de que tan bien se ajusta el modelo que incluye a todas variables predictoras comparándolo con un modelo hipotético que sólo dispone de la constante de la regresión, esta prueba es conocida como la prueba de bondad de ajuste del modelo. El objetivo es encontrar altos valores de significancia (el valor de sig. deberá ser menor que 0.05), como se puede observar en la Tabla 5.6 se obtuvo un valor 0.000, por lo tanto, el modelo con el conjunto variables usadas como predictores predice mejor los resultados que el modelo hipotético donde se utilizaría sólo la constante de regresión.

Se concluye que al poseer altos valores de significancia los predictores considerados para predicción de nuestra variable de analizada son los correctos.

Tabla 5.6:

Tabla de Clasificación

Observadas			Predicción		
			Nivel de Vida		Porcentaje Correcto
			Ingresos superiores a los \$334,72	Ingresos menores o iguales a los \$334,72	
Paso 1	Ingreso de neto	Ingresos superiores a los \$334,72	890	239	78.8 %
		Ingresos menores o iguales a los \$334,72	211	925	81.4 %
Porcentaje general					80.1 %

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Tabla superior indica la capacidad de predicción del modelo sobre el evento de interés estudiado y como es capaz de predecir correctamente cada una de las características (modelo incluye a la constante y a todos los predictores).

Comparando la Tabla 5.6 (análisis con predictores y constante) con la Tabla 5.4 (análisis con sólo constante) se observa una mejora significativa en el modelo cuando los predictores son incluidos. El modelo clasificó correctamente un 80.1% de los casos en general, una mejora significativa sobre el 50.2% explicado en la Tabla 5.4.

El modelo logró una sensibilidad del 81.4%, por lo que fue capaz de clasificar correctamente el 81.4% de los hogares que indicaron que tener ingresos inferiores a los \$ 334.72. Mientras que en lo que se refiere a la especificidad, predijo correctamente el 78.8% de los hogares quienes superan sus ingresos a los \$ 334.72.

El modelo mostró un valor predictivo positivo del 79,5%, es decir, la técnica encontró que de la predicción de hogares que tienen ingresos menores o iguales al percentil 30 (\$334.72) el 79,5% de ellos poseen la característica de interés. Mientras que para el valor predictivo negativo se obtuvo que de la predicción de hogares que no poseían la característica de interés el 80,8% cumplen teniendo ingresos superiores al percentil 30 (\$334.72).

Tabla 5.7 variables en la ecuación nos indica que mientras mayor sea el número de dormitorios que reporte el hogar la probabilidad de que obtenga ingresos menores o iguales a los \$334,72 se reduce a un factor de 0.845. De igual manera si los hogares reportan no tener cuartos para negocio la probabilidad que cumpla con el evento de interés disminuye a un factor de 0.409.

Tabla 5.7:

Tabla de Variables en la ecuación

Variables	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Número de dormitorios que posee la vivienda	-,168	,062	7,283	1	,007	,845	,748	,955
¿Dispone la vivienda de cuartos para negocio? (1)	-,894	,287	9,735	1	,002	,409	,233	,717
¿Dispone vivienda de servicio higiénico? (1)	,638	,307	4,323	1	,038	1,892	1,037	3,452
¿Algún miembro de la familia tiene teléfono celular? (1)	1,077	,260	17,111	1	,000	2,937	1,763	4,894
Metros cuadrados de construcción de la vivienda	,004	,002	6,789	1	,009	1,004	1,001	1,007
Monto mínimo mensual estimado para vivir bien	,000	,000	20,180	1	,000	1,000	1,000	1,001
Edad del jefe del hogar	,019	,004	22,495	1	,000	1,019	1,011	1,027
¿El jefe del hogar es asegurado? (1)	1,110	,111	99,914	1	,000	3,033	2,440	3,771
Tiene tarjeta de crédito (bancaria o comercial) (1)	-,344	,166	4,309	1	,038	,709	,513	,981
Ingresos Transferencias e Inversiones	-,003	,000	294,187	1	,000	,997	,996	,997
Constante	1,163	,387	9,015	1	,003	3,199		

Nota Fuente: INEC (2011). *Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011 - 2012*. Quito – Ecuador.
Elaborado por: Jhon Quimí Loza

Generan ingresos menores a los \$334.72 los hogares que no disponen de ningún tipo de servicio higiénico dentro de su hogar, los hogares donde ningún miembro de la vivienda utiliza como medio de comunicación el

teléfono celular, además poseen estos ingresos los hogares donde el jefe de la familia no tiene ningún tipo de afiliación al seguro.

La probabilidad de que se cumpla el evento de interés es 3.033 veces mayor para un hogar cuyo jefe no es afiliado a ningún tipo de seguro que para un jefe que si posee algún tipo de afiliación.

Si la cabeza de la familia no posee tarjeta de crédito de ninguna clase se reduce la probabilidad de que tengan ingresos menores a los \$334.72 a un factor del 0.709.

Por cada año extra que cumpla el jefe del hogar aumenta en 1.019 veces más la probabilidad de que su hogar obtenga ingresos menores a los \$334.72.

Si el hogar reporta un aumento del monto por ingresos de transferencias corrientes, prestaciones recibidas e inversiones la probabilidad que obtengan ingresos menores a los \$334.72 disminuye a un factor de 0.997.

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

Luego de haber analizado y estudiado cada uno de los capítulos de este proyecto de tesis, se mostrarán las conclusiones obtenidas del análisis realizado y las recomendaciones respecto a las falencias encontradas en el campo de interés.

5.1. CONCLUSIONES

El objetivo de esta tesis como se menciona en el primer capítulo es analizar el nivel socio económico de los hogares del cantón Guayaquil mediante las principales variables categóricas de la última encuesta de ingresos y gastos de los hogares urbanos y rurales ENIGHUR 2011-2012.

Obteniendo así que los hogares del cantón Guayaquil son descritos por el tipo de material que poseen en el piso de sus viviendas, por el tipo de

servicio higiénico, el valor que gastan en insumos diariamente y el nivel de instrucción que posee el perceptor o jefe del hogar, estas son unas de las principales variables que explican la mayor variabilidad de información proporcionada por cada uno de los hogar entrevistados.

Este estudio a diferencia del aplicado por el Ing. Santamaría exploró no sólo variables cuantitativas también añadió variables cualitativas de tipo ordinal y nominal. En los dos estudios se aplicó la técnica multivariada el Análisis de componentes principales, con la diferencia el señor Santamaría utilizó sólo variables numéricas por lo que el ACP fue la técnica más adecuada, pero para este proyecto por contener variables numéricas y categóricas estudiadas al mismo tiempo era necesario utilizar un derivado de la técnica llamada el Análisis de componente principales categóricos.

El objetivo fue además de reducir el número de dimensiones, revelar patrones entre los individuos analizados como podemos observar en el capítulo 4 de esta tesis. Donde se logró diferenciar el comportamiento de los individuos por tipos de características pertenecientes a las variables sometidas en el análisis, que a diferencia del estudio anterior no hubo un análisis más profundo del comportamiento de los individuos.

En la ENIGHU 2003 estudiada por el Ing. Santamaría se utilizó un total de 17 preguntas cuantitativas las que fueron analizadas mediante la técnica ACP, y en este proyecto se utilizaron 17 variables entre variables numéricas y

categorías de tipo nominal y ordinal. El análisis socio económico realizado en el cuarto capítulo mejora al 100% la capacidad de revelar patrones entre los individuos dado que la técnica Análisis de componentes principales categóricos relaciona al individuo con cada una de las categorías que proporcionó durante la ENIGHUR 2011-2012. Este análisis mostró el comportamiento de los entes como un todo, es decir, se visualizó el comportamiento de los individuos en un gráfico de dos dimensiones donde se obtuvo la formación de 3 grupos de individuos relacionados a las características de la vivienda y hogar, además cuando se analizaron todas variables se encontró una aglomeración de individuos en forma de pirámide, la cual fue nombrada como pirámide socio-económica.

Se desarrolló un método de regresión que pueda predecir un evento de interés establecido.

El método de regresión logística directo fue realizado con el propósito de evaluar el impacto de un número de factores en la probabilidad de que los encuestados reportarán que generan ingresos menores a los \$ 334.72. El modelo contiene 10 variables independientes (Número de dormitorios que posee la vivienda, ¿Dispone la vivienda de cuartos para negocio?, ¿Dispone vivienda de servicio higiénico?, ¿Algún miembro de la familia tiene teléfono celular?, Metros cuadrados de construcción de la vivienda, Monto mínimo mensual estimado para vivir bien, Edad del jefe del hogar, ¿El jefe del hogar asegurado?, Tiene tarjeta de crédito, Ingresos por transferencias

corrientes, prestaciones recibidas e inversiones). El modelo completo contiene todos los predictores que fueron estadísticamente significativos, $p < 0.05$, indicando que el modelo fue capaz de distinguir entre las respuestas quien reportó y no reportó ingresos menores al percentil 30. El modelo explicó el 36.1 (Cox and Snell R square) y 48.2% (Nagelkerke R squared) de la varianza total de los datos, y clasificó correctamente el 81.4% de los casos, como se muestra en la tabla 1. El predictor más fuerte que mejor explica el evento de interés fue ¿El jefe del hogar es asegurado?, registrando un odds ratio de 3.033. Este indicó que quienes respondieron que el jefe de su hogar no tiene ningún tipo de afiliación a un seguro tuvieron sobre 3.033 veces más probabilidad de reportar ingresos menos a los \$334.72 que quienes reportaron que su jefe si está afiliado a un seguro.

5.2. RECOMENDACIONES

Mediante la realización este proyecto se obtuvieron diferentes resultados los cuales sirven para mejorar algún tipo de inconformidad, ya sea ésta en la captura de los datos o la implementación de procedimientos.

Unas de las recomendaciones es reducir el número de datos perdidos obtenidos de la encuesta, dado que por esta circunstancia muchas variables

que se podrían considerar importantes fueron descartadas por el alto porcentaje de información faltante.

Otra recomendación sería que para próximas encuestas de este tipo, se utilizará como principales variables a las características que resultaron más significativas en esta tesis. Utilizar este número de preguntas en la ENIGHUR reduciría significativamente los costos que involucra realizar esta encuesta, entre los ahorros tendríamos: el ahorro de material de impresión, menor contratación de personal, disminución de horas de capacitación al personal, menor tiempo de permanencia en remotos alejados de las zonas urbanas entre otros.

Existe una gran oportunidad de mejora si se utiliza las recomendaciones antes mencionadas, lo ayudaría reducir el tiempo, dinero y el esfuerzo invertido que existe en realizar la ENIGHUR la cual tiene una duración aproximadamente de casi un año. Sería ideal profundizar en este tema, realizando un estudio del monto de dinero que se ahorraría el país reduciendo considerablemente el número de variables de la encuesta.

ANEXOS

REFERENCIAS BIBLIOGRÁFICAS

- Programa de las Naciones Unidas para el Desarrollo. (2013). *Informe sobre Desarrollo Humano 2013*. Estados Unidos: Communications Development Incorporated, Washington DC.
- BERENSON, M. L., LEVINE, D. M., & GOLDSTEIN, M. (1983). *INTERMEDIATE STATISTICAL METHODS AND APPLICATIONS*. New Jersey: Prentice - Hall.
- Brown, C. V., & Jackson, P. M. (1990). *Public Sector Economics*. Massachusetts: Basil Blackwell Ltd.
- INEC. (2011). *Encuesta de Estratificación del Nivel Socioeconómico NSE 2011*. Quito.
- INEC. (2011-2012). *Manual del Encuestador (2013): Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos y Rurales ENIGHUR 2011-2012*. Quito.
- INEC. (2013). *Principales Resultados: Encuesta Nacional de Ingresos y Gastos (ENIGHUR) 2011- 2012*. Quito.
- Instituto Nacional de Estadísticas y Censos. (23 de Enero de 2012). *INECPedia*. Recuperado el 15 de Septiembre de 2014, de <http://www.ecuadorencifras.com:8080/inecpedia/index.php/IPC>
- Kline, P. (1994). *An Easy Guide to Factor Analysis*. Routledge.
- Mangin, L. J., & Mallou, V. J. (2003). *Análisis multivariable para las ciencias sociales*. Madrid: Prentice Hall.
- Manobanda, S. G. (2005). *Análisis Estadístico Multivariado De Los Datos Obtenidos En La Encuesta Nacional de Ingresos y Gastos de los Hogares Urbanos 2003*. Tesis de Grado, Escuela Superior Politécnica, ICM, Guayaquil.
- Nacional del Ecuador*. (2005). QUITO.
- Pallant, J. (2011). *Survival Manual: A step by step guide to data analysis using SPSS* (fourth ed.). Australia: Allen & Unwin.
- Rencher, A. C. (2002). *Methods of Multivariate Analysis* (Second ed.). Estados Unidos: JOHN WILEY & SONS, INC.
- Tabachnick, B. G., & Fidell , L. S. (1996). *Using Multivariate Statistics* (Fourth ed.). Northridge, California, Estados Unidos: Allyn and Bacon.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics*. Northridge: Allyn and Bacon.

Vogt, P. W. (2005). *Dictionary of Statistics & Methodology* (Third ed.). United States: Sage Publication.