

TITULO: Análisis Estadístico de enfermedades infecto contagiosas: el caso de la Tuberculosis en la provincia del Guayas

AUTORES: Eduardo Moyano Baquero¹, Franciso Vera Alcívar²

¹Ingeniero en Estadística e Informática

²Director de Tesis, Ingeniero en Estadística e Informática, Escuela Superior Politécnica del Litoral, 1999. Profesor de ESPOL desde 2000

RESUMEN

Este trabajo ha analizado la eficiencia del programa de control de la tuberculosis en el transcurso a partir del año 93 hasta septiembre del año 2000. También se ha analizado como la resistencia ha ido aumentando y además se pronosticó los posibles valores futuros de algunas variables. Para lograr estos objetivos se utilizó varias herramientas estadísticas como por ejemplo: la matriz de correlación, series temporales, gráficos de secuencia y de proporciones, etc.

1. Introducción

La finalidad principal de realizar este artículo es aportar y ayudar a las personas que trabajan en el programa de la tuberculosis, proporcionando conclusiones y recomendaciones basadas en un análisis estadístico, ya que por falta de profesionales en este campo de la ciencia en nuestro país, no se ha podido realizar adecuadamente un análisis estadístico serio, necesario y requerido por parte de las organizaciones internacionales que ayudan a solventar los gastos del programa.

Los objetivos principales de este artículo son:

Mostrar la utilización correcta de las herramientas estadísticas.

Analizar como el programa de Tuberculosis ha ido mejorando con el transcurso del tiempo.

Tratar de pronosticar la cantidad de pacientes a ser analizados, cantidad total de muestras, cantidad total de enfermos.

2. La enfermedad de la Tuberculosis

En la Edad Media, a la enfermedad de la Tuberculosis se la conocía como la “peste blanca”, además se la consideraba poco contagiosa. Con el avance de la ciencia se ha descubierto que es una enfermedad contagiosa, que se da exclusivamente en lugares cerrados; es decir, en los lugares donde el cambio de aire y los rayos ultravioletas son pocos.

Esta enfermedad se da principalmente en las áreas urbanas más pobres de los países desarrollados, debido a la falta de una alimentación correcta y en los países del “tercer mundo”. Con la aparición del VIH esta enfermedad ha resurgido, hasta el punto de que ha provocado un tercio de las muertes en pacientes infectados.

Debido ha este resurgimiento de la enfermedad de la tuberculosis, la OMS está apoyando los programas de control, detección y prevención de la tuberculosis en todos los países del mundo. Presionando a los directores de los programas a llevar estadísticas y medidas de control eficientes, para poder realizar las respectivas correcciones a los programas y proporcionando información a los

investigadores para desarrollar nuevos medicamentos debido a la resistencia que está adquiriendo la enfermedad, por diversos motivos.

La tuberculosis es una infección aerógena que se propaga mediante núcleos de gotitas; es decir, se transmite por gérmenes que se encuentran en el aire que llegan a los pulmones. Estos núcleos de gotitas se generan al hablar, toser, estornudar, etc. Las personas infectadas por el SIDA tienen mayor riesgo de contraer esta infección. Esta infección es provocada por el microorganismo *Mycobacterium tuberculosis*.

3. Información de las variables utilizadas en el estudio

Las variables a utilizar en el análisis, fueron seleccionadas y recopiladas a partir del informe mensual de cultivos, elaborado por el departamento de Tuberculosis del Instituto de Higiene. Se ha seleccionado este informe, debido al grado de confiabilidad y al mismo tiempo, el Instituto de Higiene es la única red de laboratorios del país interconectados, donde se recopila y se controla toda la información concerniente a esta enfermedad.

Esta información es recogida a partir de los diferentes centros de salud donde se encuentra un laboratorio de análisis y el jefe de laboratorio tiene que realizar un informe de sus labores mensuales. Además los hospitales, los laboratorios particulares, envían sus informes mensuales, ya que deben cumplir ciertos requisitos sanitarios para seguir funcionando. Esta información es además confiable, porque nadie más puede alterarla o manipularla, ya que existen registros individuales donde se pueden rastrear y confirmar la información.

Este informe es elaborado por dos motivos: para llevar una documentación de las labores mensuales y de los gastos incurridos, por el departamento de tuberculosis. Para poder así explicar sus labores anuales al Ministerio de Salud y poder controlar alguna epidemia a nivel nacional.

La unidad de investigación es el programa de tuberculosis implementado por el gobierno nacional asesorado por los organismos internacionales, a partir de la información proporcionada por el Instituto de Higiene.

Las variables que vamos a utilizar en nuestra investigación son:

Centros de Salud.- Esta variable es de tipo numérico, concierne a la cantidad de pacientes que han sido atendidos por los centros de salud para averiguar si son o no portadores de la enfermedad.

Hospitales.- Esta variable es de tipo numérico, concierne a la cantidad de pacientes que han sido atendidos por los hospitales para averiguar si son o no portadores de la enfermedad.

Particulares.- Esta variable es de tipo numérico, concierne a la cantidad de pacientes que han sido atendidos por los consultorios particulares para averiguar si son o no portadores de la enfermedad.

Ciudad.- Esta variable es de tipo numérico, se refiere a la cantidad de pacientes que pertenecen a la ciudad.

Afuera de la Ciudad.- Esta variable es de tipo numérico, se refiere a la cantidad de pacientes que provienen de fuera de la ciudad.

Isoniacida.- Esta variable es de tipo numérico, nos indica la cantidad de pacientes que son resistentes a este medicamento.

Estreptomycin.- Esta variable es de tipo numérico, nos indica la cantidad de pacientes que son resistentes a este medicamento.

Etambutol.- Esta variable es de tipo numérico, nos indica la cantidad de pacientes que son resistentes a este medicamento.

Rifampicina.- Esta variable es de tipo numérico, nos indica la cantidad de pacientes que son resistentes a este medicamento.

Pirazinamida.- Esta variable es de tipo numérico, nos indica la cantidad de pacientes que son resistentes a este medicamento.

Resistencia a dos o más medicamentos.- Esta variable es de tipo numérico, nos indica la cantidad de pacientes resistentes a dos o más medicamentos.

4. Análisis Estadístico Univariado

4.1 Introducción

Esta sección está constituido en dos puntos. El primer punto concierne a la explicación de las herramientas

4.2 Explicación de las herramientas estadísticas

La media es el resultado del promedio de los datos en conjunto. Su fórmula matemática es:

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

donde N es el tamaño de la población.

La mediana es el dato que ocupa la posición central de los datos.

Las variables artificiales han sido creadas a partir de los datos recolectados. Estas variables son:

Total.- Esta variable es de tipo numérico y está conformada por la suma de las variables: centro de salud, hospital y particular.

Pulmonar.- Esta variable es de tipo numérico y está conformada por la suma de las variables: Espustos, H. Laríngeo, Gástricos y Bronquios.

Extrapulmonar.- Esta variable es de tipo numérico y está conformada por la suma de las variables: Orina, PERC, LCR, Fístula, Biopsias, Ganglios, Piel, Perianal, Orofaringeo, Heces, Médula ósea, Mama, Granulo, Otica, Líquido sinovial, Sec. Traqueal, Osteoarticular, Tejido ganglio y Genital.

Total de Enfermos.- Esta variable es de tipo numérico y está conformada por la suma de las variables Pulmonar y Extrapulmonar.

estadísticas. El segundo punto trata acerca del análisis estadístico de los resultados obtenidos de las diferentes variables.

La varianza es un valor que nos indica la extensión o magnitud de la separación entre los elementos de una población. Su fórmula matemática es:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

La desviación estándar se basa en las desviaciones con respecto a la media. Es igual a la raíz cuadrada de la varianza. Su fórmula matemática es:

$$\sigma = \sqrt{\sigma^2}$$

Intervalo de confianza para la media al 95%: Inferior < θ < Superior. Se denomina intervalo de confianza de 100 (1- α) por ciento para el parámetro θ , en este caso es la media. Para interpretar este intervalo debe tomarse en cuenta que si en muestras aleatorias repetidas, se consideran un gran número de estos intervalos, el 100 (1- α) por ciento de ellos contendrá el valor verdadero de θ .

Amplitud intercuartil es la diferencia entre el tercero y primer cuartiles Su fórmula matemática es:

$$\text{Amplitud Intercuartil} = Q_3 - Q_1$$

El valor mediano, Q_2 , separa el 50% superior de un conjunto de observaciones, del 50% inferior. De manera semejante, el primer cuartil, Q_1 , es el valor que corresponde al punto por debajo del cual se encuentra el 25% de las observaciones. El tercer cuartil, Q_3 , es el valor que corresponde al punto por encima del cual se encuentra el 25% de las observaciones. Por tanto, el 50% central de las observaciones se localiza entre el Q_3 y Q_1 .

Para calcular la asimetría y la curtosis debemos utilizar las funciones generatrices de momentos. La función generatriz de momentos de la variable aleatoria x , donde exista, está dada por:

$$M_x(t) = E(e^{tx}) = \sum_x e^{tx} * f(x)$$

cuando x es discreta y

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} * f(x)$$

cuando x es continua.

Asimetría es un valor que nos permitirá reconocer si la distribución tiene una asimetría hacia la izquierda o hacia la

derecha o es nula. Su fórmula matemática es:

$$\alpha_3 = \frac{\mu_3}{\sigma^3}$$

Curtosis mide el grado de agudeza de una distribución; es decir, mide la elevación o achatamiento de la distribución, comparada con la distribución normal. Su fórmula matemática es:

$$\alpha_4 = \frac{\mu_4}{\sigma^4}$$

Para utilizar la mediana como mejor medida de estimación en vez de la media, debemos fijarnos primero el valor de la asimetría y de la curtosis, si ambos valores son mayores que uno, es aconsejable utilizarla caso contrario utilizar la media.

La distribución normal es una de las distribuciones continuas más ampliamente usada en la teoría estadística porque explica muchos fenómenos de la naturaleza. Las variables aleatorias continuas pueden asumir un número infinito de valores sobre un rango finito o infinito. Su función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

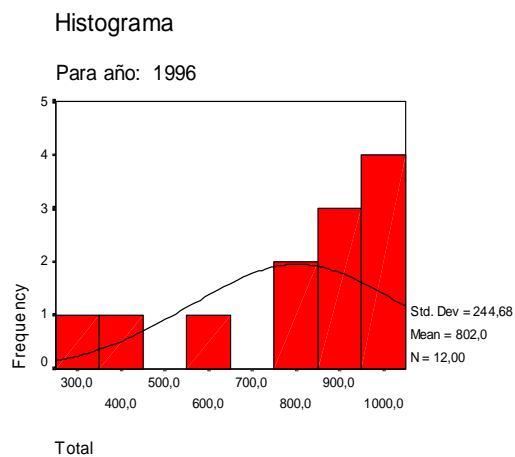
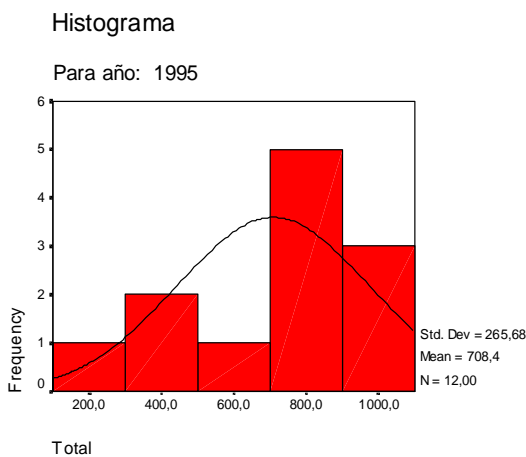
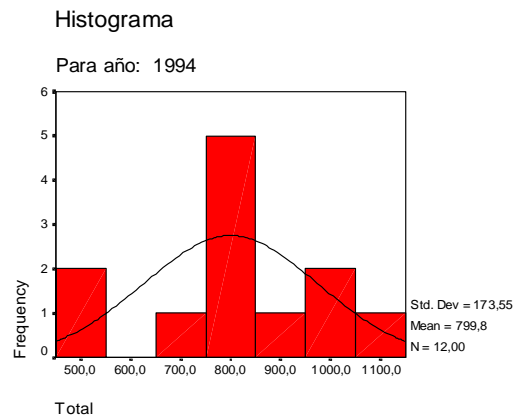
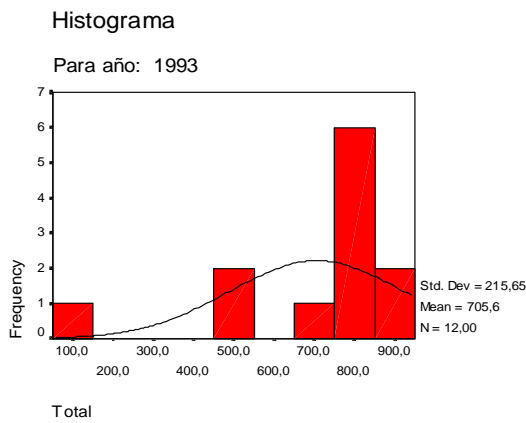
para $-\infty < x < \infty$. Como la mayoría de las variables pertenecen a la escala de razón y los datos son cantidades numéricas mayores a cero y la mayor parte de los datos normales deberían estar en el intervalo $3\sigma \pm \mu$ (95%-99%), debemos compararlo con su media, si la media es mayor a este resultado significa que los datos de ese año son aceptables, caso contrario debemos suponer con evidencia estadística que algo sucedió en ese año y los datos no son aceptables.

Variable: Total

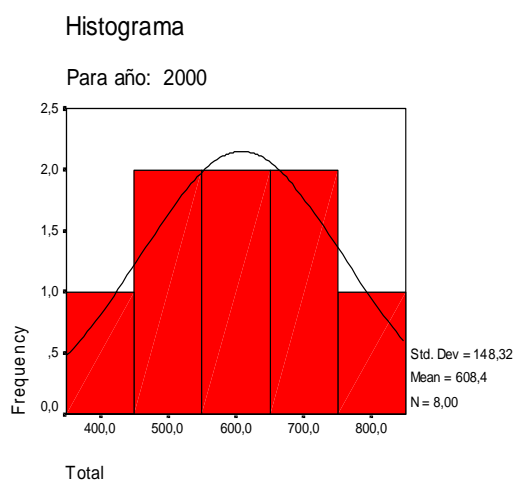
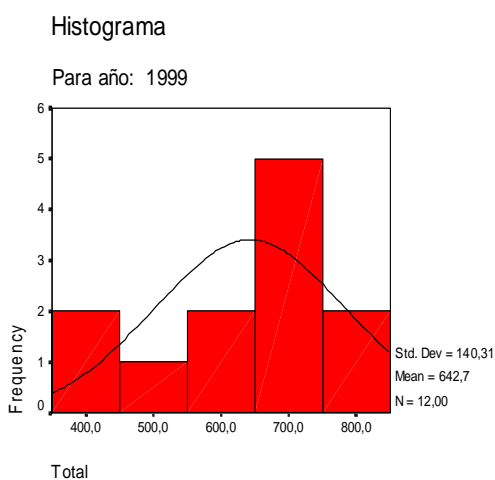
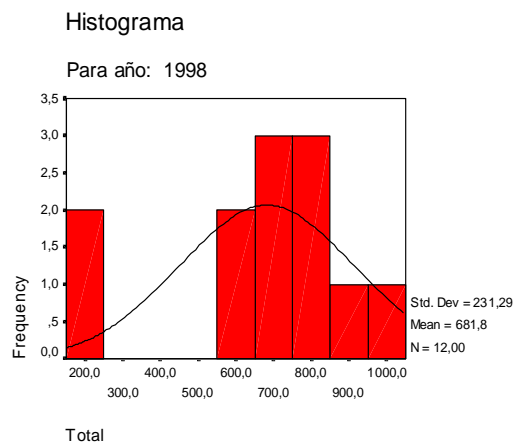
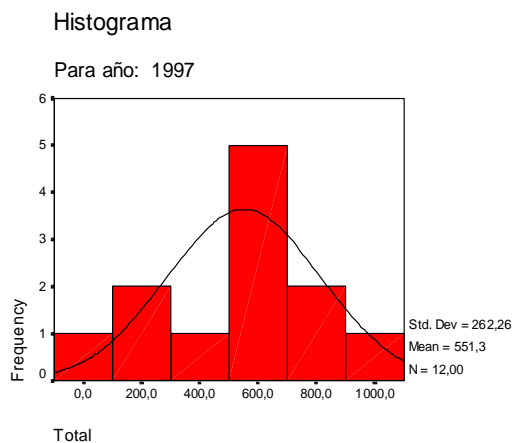
Estadística Básicas de la Variable Total

Estadístico		Año							
		1993	1994	1995	1996	1997	1998	1999	2000
Media		705.58	799.75	708.42	802.00	551.25	681.83	642.67	608.38
Intervalo De confianza	L. I.	568.57	689.48	539.61	646.54	384.62	534.88	553.52	484.38
	L. S.	842.60	910.02	877.22	957.46	717.88	828.79	731.82	732.37
Mediana		769.5	805.5	790.5	871	594.5	740.5	701.5	612.5
Varianza		46,503.36	30,118.93	70,587.36	59,869.64	68,779.30	53,497.24	19,687.33	21,998.55
Desviación Típica		215.65	173.55	265.68	244.68	262.26	231.29	140.31	148.32
A. Inter.		240.5	242.75	429.5	361	349.5	188.75	199	227.75
Asimetría		-1.90	-0.44	-1.12	-1.13	-0.47	-1.18	-0.70	-0.11
Curtosis		4.06	-0.14	0.69	0.12	0.71	0.98	-0.32	-0.70

Histogramas de Frecuencias de la Variable Total para los años: 1993, 1994, 1995, 1996



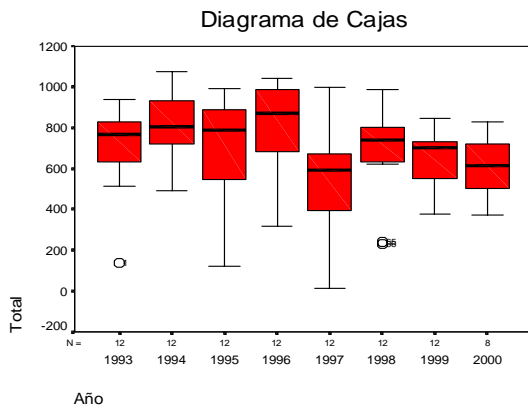
Histogramas de Frecuencias de la Variable Total para los años: 1997, 1998, 1999, 2000



En este cuadro y gráficos, podemos observar lo siguiente:

1. Existe demasiada dispersión de la información en todos los años.
2. Para los años 1993,1997 los datos son no aceptables.
3. Para los años 1993, 1995, 1996, 1998 la información se encuentra sesgada hacia la izquierda.
4. La curtosis nos indica: que para el año 1993 existe una acumulación de la información. Para el año 1993, hay evidencia estadística de utilizar la mediana como estimador en vez de la media.

Diagrama de Caja de la Variable Total



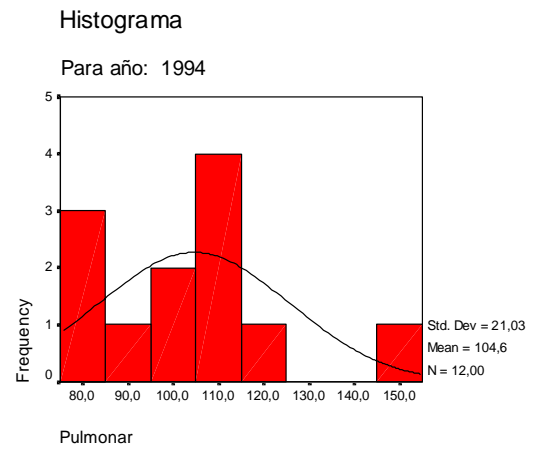
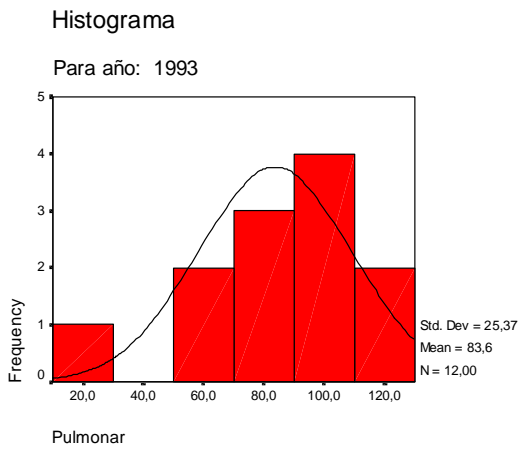
En el gráfico de diagrama de Caja, nos indica lo siguiente:
 La mediana nos indica como ha afectado las correcciones al programa de control de la TB con el transcurso del tiempo y podemos ver la existencia de valores extremos para los años 1993, 1998. Además, se puede observar de manera global la cantidad total de pacientes han sido atendidos por todos los lugares, con sospechas de tuberculosis.

Variable: Pulmonar

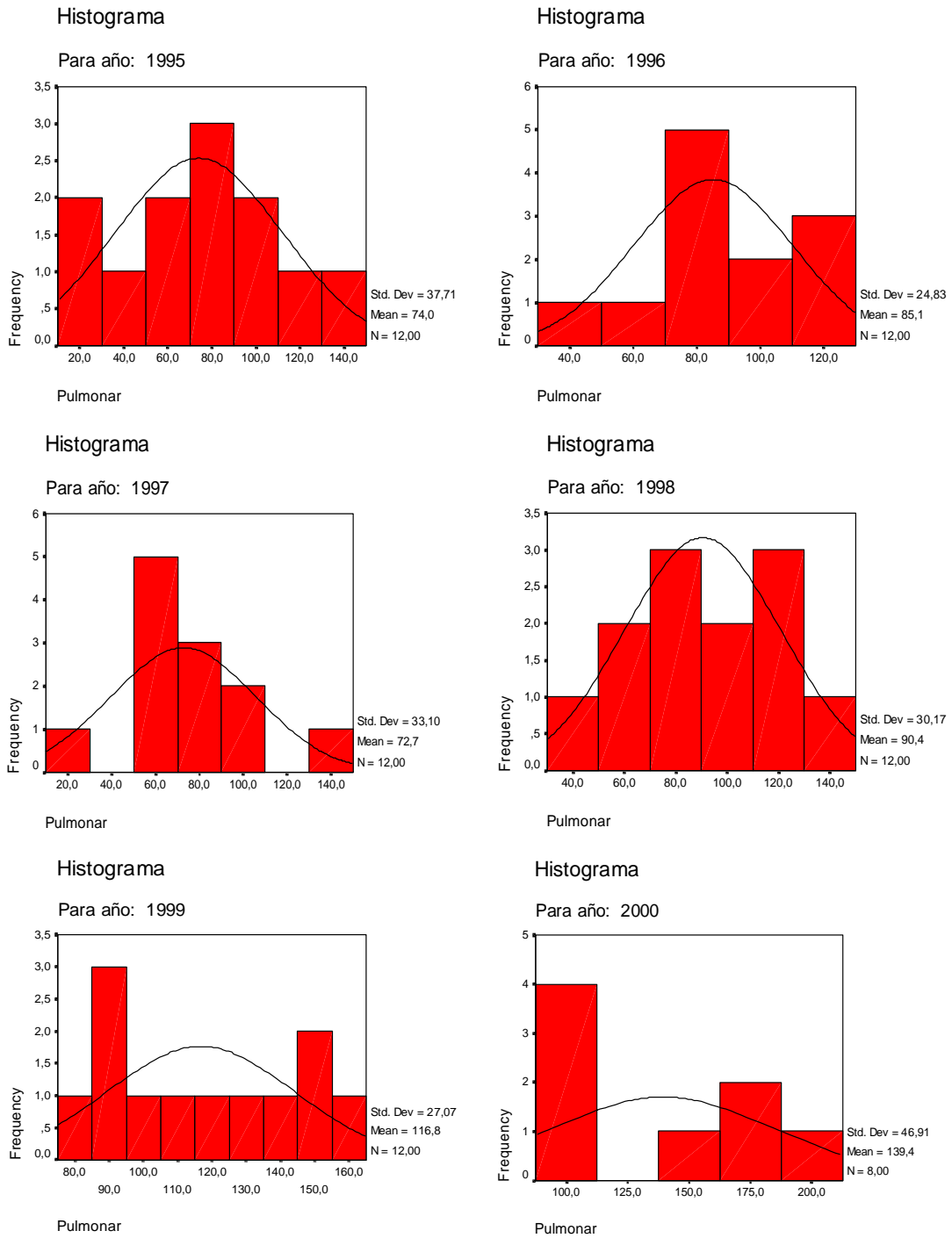
Estadística Básicas de la Variable Pulmonar

Estadístico		Año							
		1993	1994	1995	1996	1997	1998	1999	2000
Media		83.58	104.58	74.00	85.08	72.67	90.42	116.75	139.38
Intervalo	L. I.	67.46	91.22	50.04	69.31	51.63	71.25	99.55	100.16
De confianza	L. S.	99.70	117.95	97.96	100.86	93.70	109.58	133.95	178.59
Mediana		90	104.5	80	80.5	68	88.5	115.50	129.50
Varianza		643.72	442.45	1,422.1	616.63	1,095.8	910.08	732.75	2,200.5
Desviación Típica		25.37	21.03	37.71	24.83	33.10	30.17	27.07	46.91
Amp. Intercuartil		33	28.25	58	36.75	40.25	48.5	50.75	89.75
Asimetría		-0.90	0.85	-0.32	-0.41	0.86	0.17	0.11	0.29
Curtosis		1.07	1.49	-0.50	0.40	1.55	-0.72	-1.37	-1.98

Histogramas de Frecuencias de la Variable Pulmonar para los años: 1993, 1994



Histogramas de Frecuencias de la Variable Pulmonar para los años: 1995, 1996, 1997, 1998, 1999, 2000

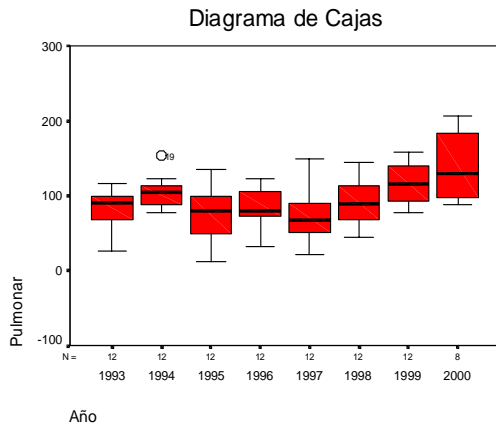


En este cuadro y gráficos, podemos observar lo siguiente:

1. Existe demasiada dispersión de la información en cada año.
2. Para los años 1995, 1997, los datos son no aceptables.
3. Para el 1993 la información se encuentra sesgada hacia la izquierda y para el año 1994 se encuentra sesgada hacia la derecha.

4. La curtosis nos indica: que para los años 1993, 1994, 1997 existe una acumulación de la información y para los años 1999, 2000 existe una expansión de la información.

Diagrama de Caja para la Variable Pulmonar



En el gráfico de diagrama de Caja, nos indica lo siguiente:

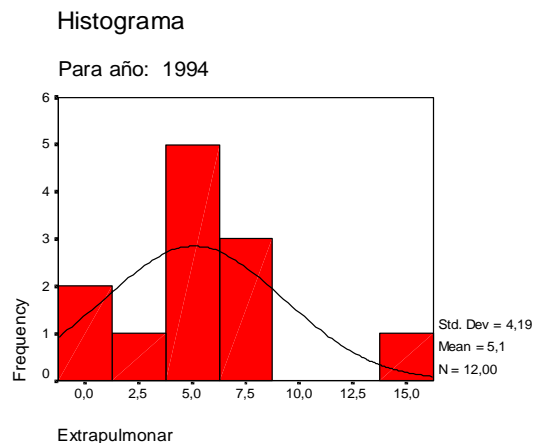
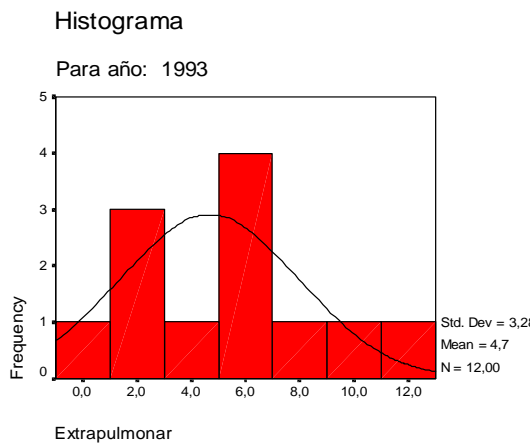
La mediana nos indica como ha afectado las correcciones al programa de control de la TB con el transcurso del tiempo y podemos ver la existencia de valores extremos para el año 1994. Además, se puede observar de manera global la cantidad total de enfermos pulmonares y esta tendencia nos indica mucha fluctuación en el tiempo, a pesar que en los últimos años existe una tendencia creciente.

Variable: Extrapulmonar

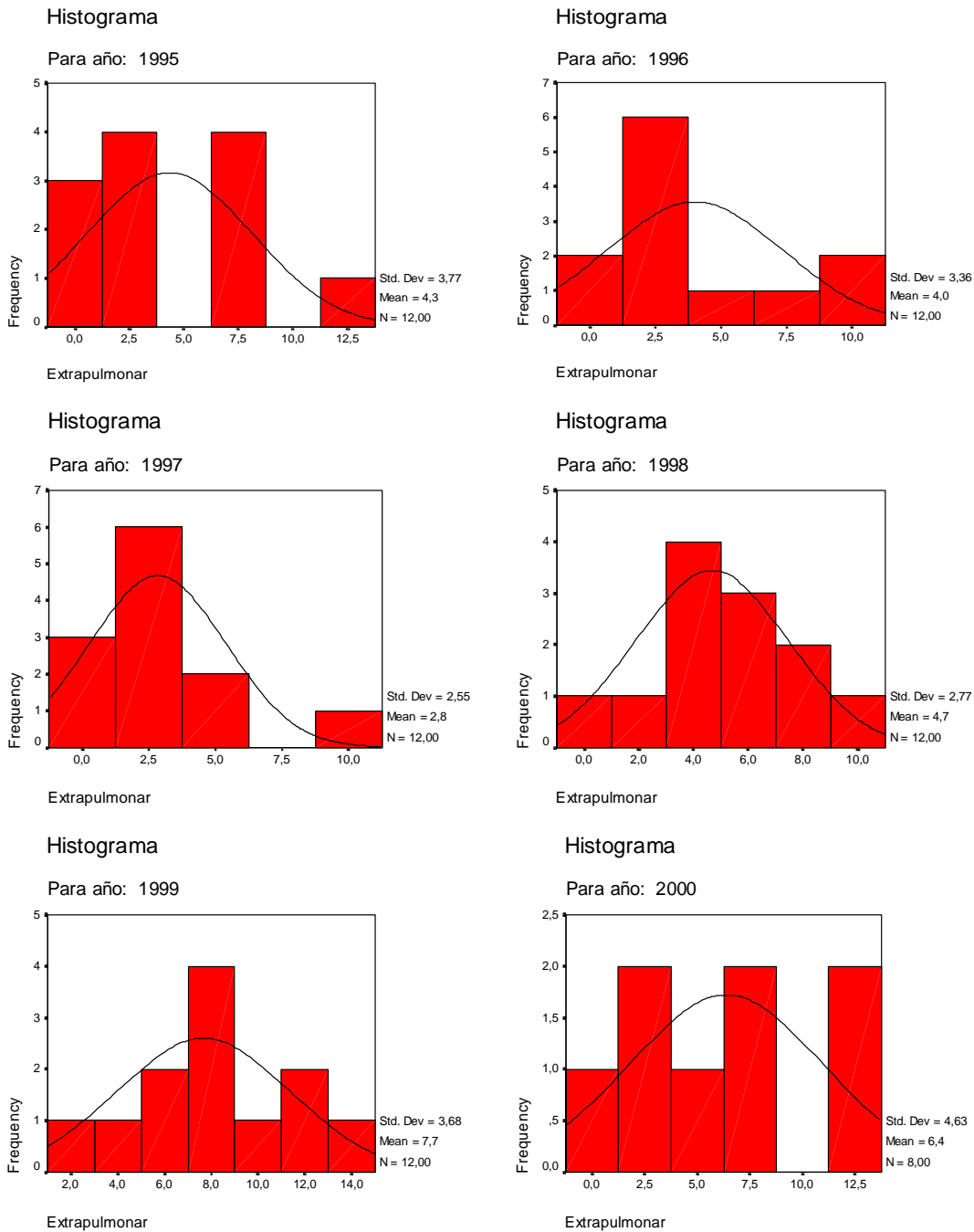
Estadísticas Básicas de la Variable Extrapulmonar

Estadístico	Año								
	1993	1994	1995	1996	1997	1998	1999	2000	
Media	4.67	5.08	4.33	4.00	2.83	4.67	7.67	6.38	
Intervalo de confianza	L. I.	2.58	2.42	1.94	1.87	1.21	2.90	5.33	2.51
	L. S.	6.75	7.74	6.73	6.13	4.46	6.43	10.00	10.24
Mediana	5	4	3	2	2	4.5	7.50	6.50	
Varianza	10.79	17.54	14.24	11.27	6.52	7.70	13.52	21.41	
Desviación Típica	3.28	4.19	3.77	3.36	2.55	2.77	3.68	4.63	
A. Intercuartil	4.75	4.5	5.75	4.75	3	4.5	5.75	8.75	
Asimetría	0.48	1.64	0.66	1.07	1.56	-0.01	0.00	0.18	
Curtosis	-0.30	3.83	-0.40	-0.40	2.12	-0.65	-0.26	-1.10	

Histogramas de Frecuencias de la Variable Extrapulmonar para los años: 1993, 1994



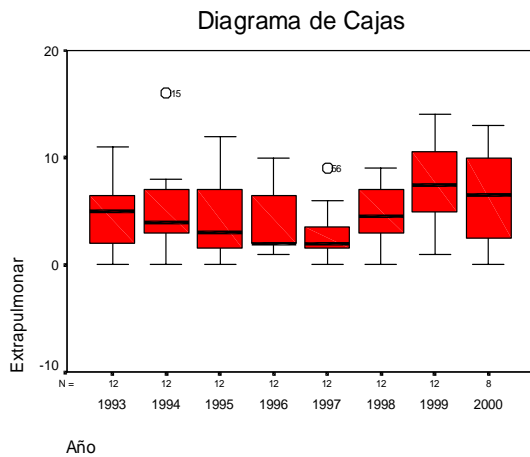
Histogramas de Frecuencias de la Variable Extrapulmonar para los años: 1995, 1996, 1997, 1998, 1999, 2000



En este cuadro y gráficos, podemos observar lo siguiente:

1. Existe demasiada dispersión de la información para todos los años.
2. Para todos los años los datos son no aceptables.
3. Para los años 1994, 1996 y 1997 los datos se encuentran sesgados hacia la derecha.
4. La curtosis nos indica: que para los años 1994, 1997 existe una acumulación de la información y para el año 2000 existe una expansión de la información. Para el año 1994 y 1997, hay evidencia estadística de utilizar la mediana como estimador en vez de la media.

Diagrama de Caja de la Variable Extrapulmonar



En el gráfico de diagrama de Caja, nos indica lo siguiente:

La mediana nos indica como ha afectado las correcciones al programa de control de la TB con el transcurso del tiempo y podemos ver la existencia de valores extremos para los años 1994, 1997. Además, se puede observar de manera global la cantidad total de enfermos extrapulmonares.

5. Análisis de los datos por medio de gráficos

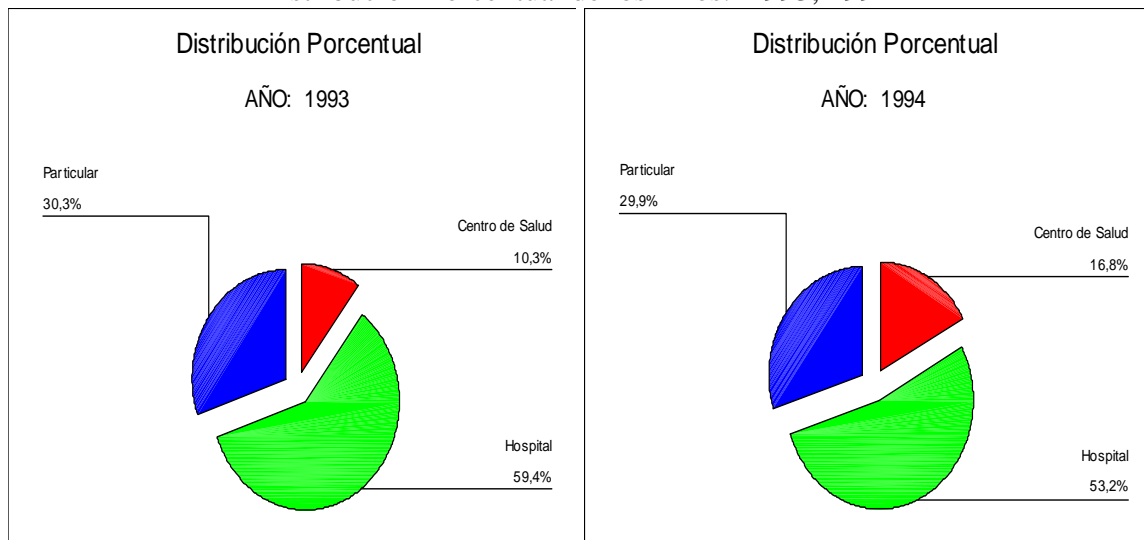
5.1 Introducción

Esta sección esta conformado en 2 puntos. El primer punto abarca los gráficos de distribución porcentual anual de pacientes analizados: Centro de Salud, Hospital, Particular. El segundo punto concierne a los gráficos de secuencia.

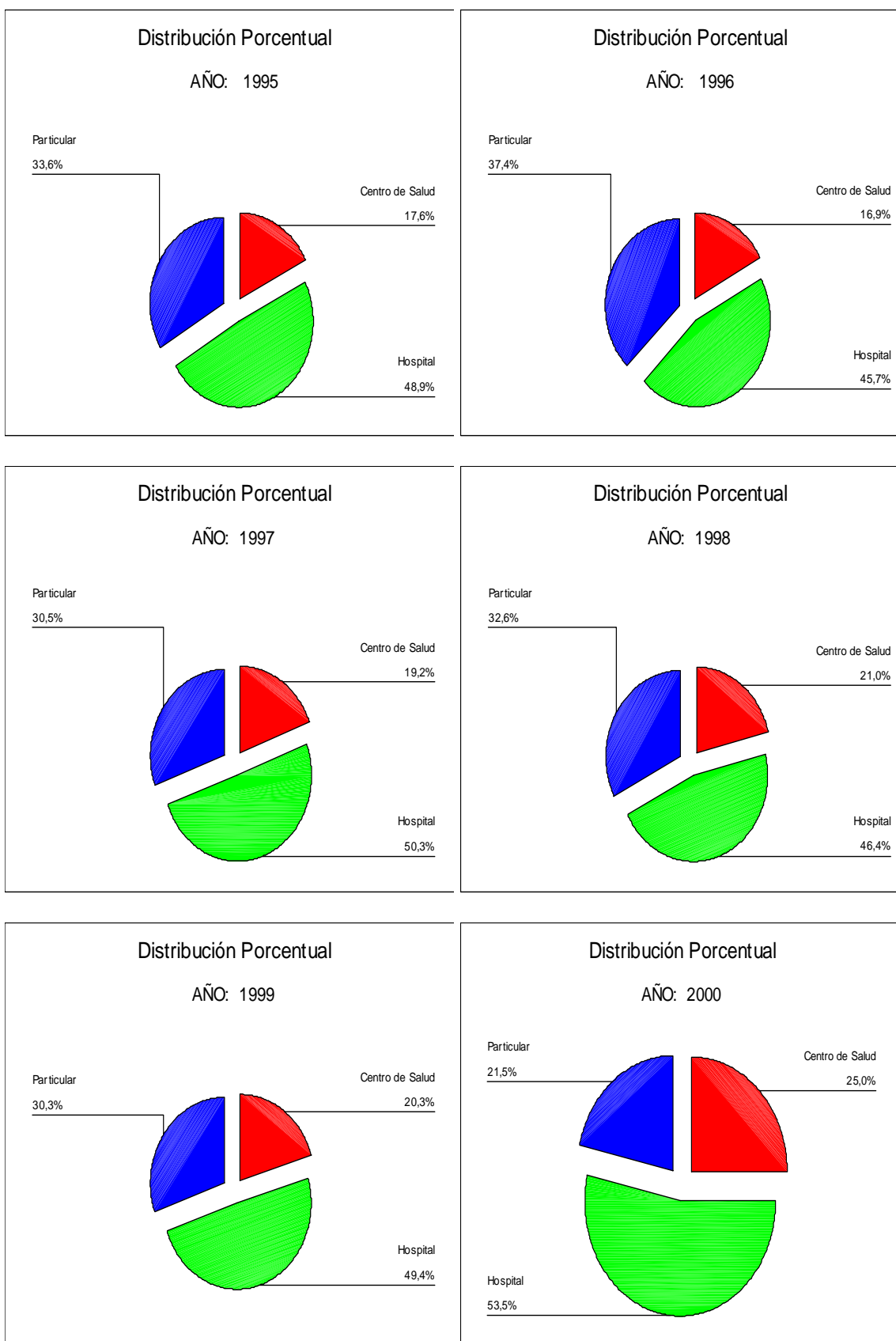
Los gráficos de distribución porcentual anual de pacientes analizados, nos permitirán tener una idea de cual fue la participación de ellos en cada uno de los centros de atención: Centro de Salud, Hospital y Particular. Estos gráficos nos servirán para conocer como los diversos centros de atención han aportado en brindar sus servicios a la comunidad en la aplicación del tratamiento del programa de tuberculosis.

5.2 Distribución Porcentual anual de pacientes analizados: Centro de Salud, Hospital, Particular

Distribución Porcentual de los Años: 1993, 1994



Distribución Porcentual de los Años: 1995, 1996, 1997, 1998, 1999, 2000

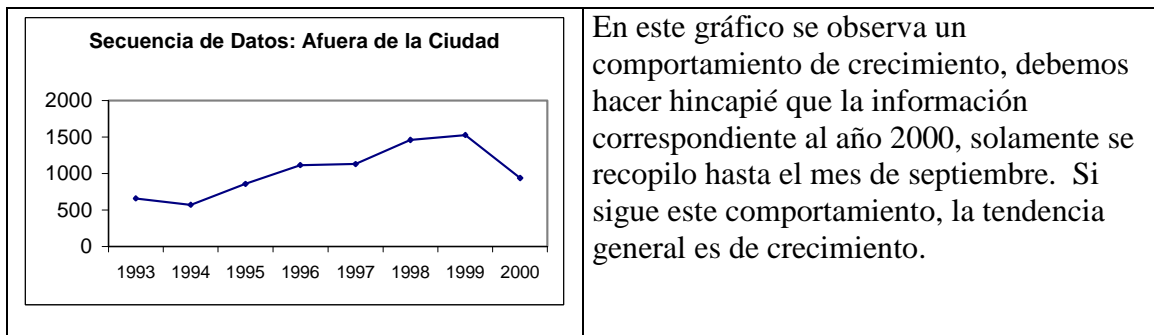
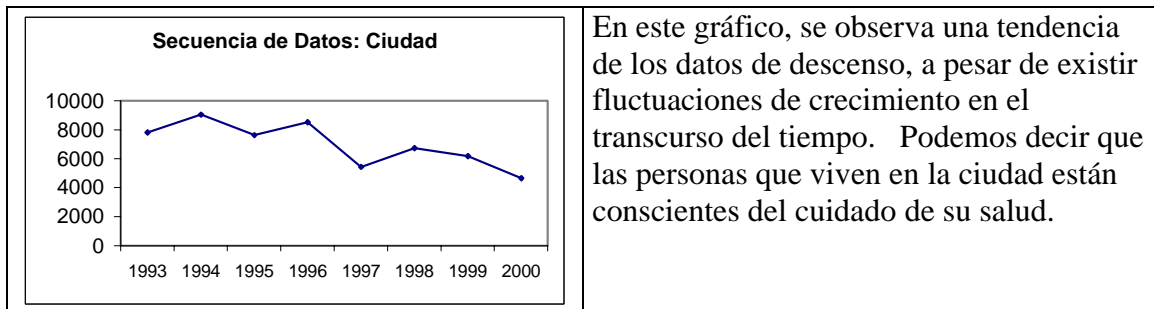
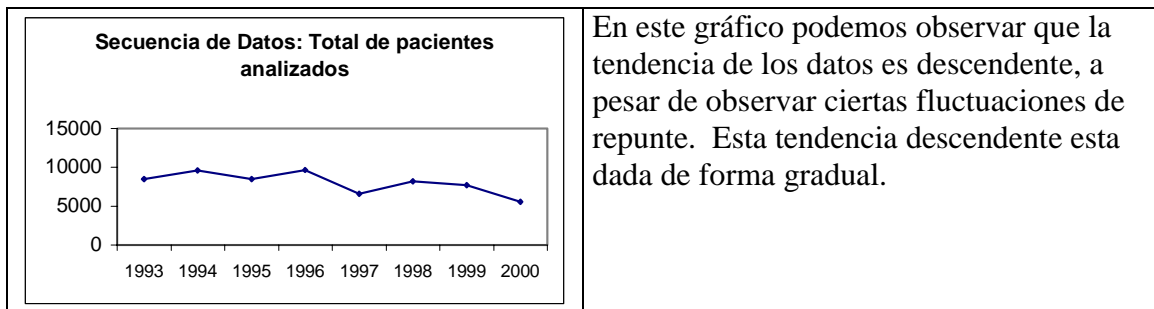


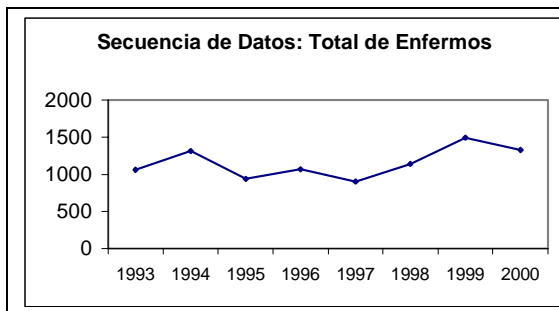
Como podemos observar, con el transcurso del tiempo, la participación de los centros de salud ha ido aumentando, mientras que la participación de particulares y hospitales ha ido disminuyendo, podemos alegar que se debe a la mejor participación de los centros de salud, en hacer conscientizar a sus pacientes la importancia de seguir un tratamiento de salud. También, podemos suponer a la mejor preparación del personal del centro de salud en brindar los servicios

requeridos, según los procedimientos del programa. Debemos señalar que los datos del año 2000, solo se ha procesado hasta el mes de septiembre.

5.3 Análisis de Secuencia

Los gráficos de análisis de secuencia nos permite observar la tendencia de la información en el transcurso del tiempo. También nos permite conocer las fluctuaciones que han ocurrido.





En este gráfico se observa un comportamiento de crecimiento, debemos hacer hincapié que la información correspondiente al año 2000, solamente se recopiló hasta el mes de septiembre. Si sigue este comportamiento, la tendencia general es de crecimiento.

6. Análisis de la Matriz de Correlación

6.1 Introducción

Esta sección está constituida en 2 puntos: El primer punto da la explicación de las herramientas estadísticas y el segundo punto corresponde a los resultados obtenidos.

6.2 Explicación de las Herramientas Estadísticas

Es importante estudiar la relación o comportamiento que puede existir entre dos variables, por este motivo vamos a explicar que es la covarianza y coeficiente simple de correlación lineal.

La covarianza entre dos variables: y_1 y y_2 , se denota por la siguiente esperanza:

$$Cov(y_1, y_2) = E[(y_1 - \mu_1)(y_2 - \mu_2)]$$

Desafortunadamente, el uso de la covarianza como una medida absoluta de dependencia es difícil porque su valor depende de la escala de medición. Es por eso que a primera vista sea difícil saber si una covarianza es grande. Esta dificultad se puede eliminar estandarizando su valor por medio del coeficiente simple de correlación lineal. Su fórmula es:

$$p = \frac{Cov(y_1, y_2)}{\sigma_1 \sigma_2}$$

está relacionado a la covarianza y toma valores en el intervalo $-1 \leq p \leq 1$.

La matriz de correlación es una matriz conformada por n filas y por n columnas. Además es una matriz simétrica; es decir, que los valores de los elementos a_{ij} de la matriz, es el mismo valor en los elementos a_{ji} de la matriz.

La matriz de correlación nos explica como se encuentran relacionadas cada una de las variables con otra variable. Su diagonal siempre contendrá el valor de 1. Si tiene un valor 0, nos indicará que no tiene ninguna relación con esa variable, por lo menos no lineal; es decir, pueda que tenga una relación cuadrática o de otro grado.

Cuando la correlación es positiva, esto nos indica que su proyección de la regresión lineal va a tender a crecer conjuntamente con la contra variable.

Cuando la correlación es negativa, esto nos indica que su proyección de la regresión lineal va a tender a decrecer conjuntamente con la contra variable.

6.3 Matriz de Correlación

En el cuadro siguiente, los valores de las correlaciones más altas, las vamos a distinguir por medio de otro color.

Matriz de Correlación

Centro de Salud	1,00																						
Hospital	0,53	1,00																					
Particular	0,54	0,68	1,00																				
Ciudad	0,65	0,92	0,88	1,00																			
Afuera de la Ciudad	0,58	0,24	0,30	0,20	1,00																		
Muestra de Espujo	0,74	0,84	0,91	0,95	0,40	1,00																	
Muestra de Orina	0,35	0,86	0,60	0,82	-0,03	0,64	1,00																
Muestra de H. L.	0,61	0,45	0,65	0,60	0,39	0,65	0,34	1,00															
Muestra Varios	0,52	0,20	0,09	0,13	0,71	0,26	-0,12	0,23	1,00														
Pulmonar	0,20	-0,02	-0,12	-0,05	0,22	0,05	-0,26	-0,04	0,39	1,00													
Extrapulmonar	0,16	0,04	0,03	0,03	0,21	0,10	-0,09	-0,08	0,27	0,46	1,00												
Total Cultivos	0,11	-0,07	-0,13	-0,07	0,03	-0,01	-0,20	0,02	0,16	0,70	0,30	1,00											
Pruebas de S.	0,16	0,03	0,02	0,02	0,22	0,10	-0,13	0,01	0,27	0,36	0,22	0,26	1,00										
Isoniacida	0,08	-0,05	0,05	-0,01	0,11	0,06	-0,15	-0,05	0,19	0,27	0,13	0,19	0,89	1,00									
Estreptomicona	-0,01	-0,18	0,00	-0,10	0,01	-0,07	-0,19	-0,09	0,14	0,05	0,00	0,04	0,51	0,66	1,00								
Etambutol	-0,02	-0,10	0,10	-0,02	-0,02	0,00	-0,10	-0,01	0,07	-0,03	-0,11	-0,07	0,29	0,47	0,91	1,00							
Rifampicina	0,07	-0,13	0,00	-0,08	0,12	0,00	-0,23	-0,07	0,20	0,25	0,08	0,16	0,85	0,95	0,77	0,58	1,00						
Pirazinamida	-0,01	-0,10	0,12	0,00	-0,05	0,03	-0,09	0,01	-0,04	-0,01	-0,04	-0,01	0,39	0,54	0,84	0,89	0,65	1,00					
R. a dos o más med.	0,14	-0,13	-0,04	-0,08	0,15	0,01	-0,25	0,00	0,21	0,24	0,12	0,18	0,84	0,81	0,51	0,29	0,81	0,41	1				

6.3.1 Observaciones:

A partir de la matriz de correlación podemos decir las siguientes conclusiones:

La variable centro de salud está correlacionada positivamente con las variables muestra de esputo, muestra de hisopados laringeos. Esta relación nos indica si crece la variable centro de salud también crecerá la variable muestra de esputo, la misma relación sucede con la variable muestra de hisopados laringeos. Hay evidencia estadística que los centros de salud piden a los pacientes que se realicen más estos tipos de muestras.

La variable hospital está correlacionada positivamente con las variables ciudad, muestra de esputo y muestra de orina. Esta relación de crecimiento entre la variable hospital y las variables de muestra de esputo y muestra de orina se debe a la realización de los exámenes que se debe realizar un paciente. La relación entre la variable hospital y ciudad hay evidencia de creer que se debe que el origen de los pacientes pertenecen a la ciudad. Además hay evidencia estadística de suponer que los hospitales piden a los pacientes que se realicen más estos tipos de muestras.

La variable particular está correlacionada positivamente con las variables ciudad, muestra de esputo, muestra de orina y muestra de hisopados laringeos. Esta relación nos indica que la mayoría de los pacientes de los centros particulares viven en la ciudad y piden que se realicen más estos tipos de muestras

La variable Ciudad está correlacionada positivamente con las variables muestra de esputo, muestra de orina y muestra de hisopados laringeos. Esta relación nos indica que la mayoría de los pacientes que viven en la ciudad se realizan estos tipos de exámenes.

La variable afuera de la ciudad está correlacionada positivamente con la variable muestra varios. Esta relación nos indica que a los pacientes se les realizan otros tipos de muestras.

La variable prueba de sensibilidad está correlacionada positivamente con las variables isoniacida, rifampicina y resistencia a dos o más medicamentos. Esta relación nos indica que la mayoría de los pacientes están adquiriendo resistencia a estos medicamentos utilizados en el esquema de tratamiento, mucho más rápido de lo normal.

La variable isoniacida está correlacionada positivamente con las variables estreptomycin, rifampicina y resistencia a dos o más medicamentos. Esta relación se debe a que el medicamento es administrado conjuntamente con otros medicamentos en diferentes dosis, provocando una resistencia a uno de ellos.

La variable estreptomycin está correlacionada positivamente con las variables etambutol, rifampicina y pirazinamida. Esta relación nos indica que si el paciente tiene resistencia a este medicamento hay evidencia estadística de que también será resistente a los otros medicamentos.

La variable etambutol está correlacionada positivamente con la variable pirazinamida. Esta relación nos indica que si el paciente tiene resistencia a este medicamento hay evidencia estadística de que también será resistente al medicamento pirazinamida.

La variable rifampicina está correlacionada positivamente con la variable pirazinamida y resistencia a dos o más medicamentos. Esta relación nos indica que si el paciente

tiene resistencia a este medicamento hay evidencia estadística de que también será resistente al medicamento pirazinamida y a otros medicamentos.

7 Análisis para pronóstico de Series de Tiempo

7.1 Introducción

Esta sección está constituido en dos puntos: El primer punto da una breve explicación de las herramientas matemáticas utilizadas. El segundo punto contiene los resultados del pronóstico.

7.2 Explicación de las herramientas

Una Serie de Tiempo es una secuencia de observaciones de una variable aleatoria indexada en el tiempo.

Generalmente la variable independiente es el tiempo y en este caso, el análisis consiste en identificar el patrón de comportamiento de la variable dependiente, en el pasado y proyectarlo al futuro.

Típicamente, para analizar una serie de tiempo, esta puede ser considerado como la superposición de 4 componentes:

Tendencia ($Z_n(t)$).- Debido al comportamiento promedio a largo plazo.

Componente cíclicas ($Z_n(c)$).- Debido a movimientos periódicos.

Componente oscilatorio ($Z_n(o)$).- Debido a movimientos irregulares periódicos.

Componente aleatorio (Y_n).- Debido a efectos aleatorios.

$$X_n = Z_n(t) + Z_n(c) + Z_n(o) + Y_n$$

La componente tendencia se la separa haciendo un ajuste generalmente lineal, luego la componente cíclica se la separa haciendo un filtrado del espectro de frecuencias; y por último la componente oscilatoria se separa de la componente aleatoria haciendo mediante un filtrado en bloque de las frecuencias.

Debido a esta descomposición, solamente a la componente aleatoria se le aplica los modelos ARIMA y SARIMA

Los valores aberrantes de los datos originales, se eliminan considerando un preajuste por regresión simple, de la componente tendencia.

Para pronosticar tendremos que sumar la tendencia, la componente Cíclica tiene período igual al número de datos entonces se la repite. La componente oscilatoria se pone los valores calculados al momento de hacer el filtrado. El error absoluto se puede pronosticar mediante un modelo ARIMA.

Para aceptar si un modelo es aceptado o rechazado, se considerará que el promedio de los errores de las variables a predecir, deberá estar comprendido en un intervalo entre -6% a 6%. Debemos recordar que ningún modelo matemático se ajusta al 100% a la realidad.

7.3 Pronóstico

La tabla siguiente contiene los valores pronosticados con sus respectivas variables:

Año	Mes	Total	Ciudad	Afuera de la Ciudad	Total de Enfermos	Total de Cultivos
2000	Octubre	659	495	198	122	684
2000	Noviembre	593	450	154	134	562
2000	Diciembre	593	444	162	118	585
2001	Enero	595	387	195	139	558
2001	Febrero	606	416	164	118	580
2001	Marzo	593	410	155	96	610
2001	Abril	653	481	166	153	662

Observaciones:

A partir del cuadro de predicción elaborado se puede inducir que:

Los promedios de errores relativos son aceptables, que se puede observar en la siguiente tabla:

Promedio del Error Relativo

Variable	Promedio del Error Relativo
Total	-0,53%
Ciudad	3,31%
Afuera de la Ciudad	-4,46%
Total de Enfermos	5,35%
Total de Cultivos	2,24%

En conclusión podríamos decir que si la enfermedad de la TB, se encuentra en un buen nivel de control, las predicciones hechas podrán ayudar al personal que labora en el programa de control de la TB; caso contrario estas predicciones fallarán.

CONCLUSIONES

- La eficiencia del programa de control de la enfermedad de tuberculosis en la provincia del Guayas, entre los años 1993 al 2000 (solo hasta el mes de septiembre), hay evidencia estadística de concluir que ha sido regular.
- Con el transcurso del tiempo, la participación de los centros de salud ha ido aumentando, mientras que la participación de particulares y hospitales ha ido disminuyendo, podemos alegar que se debe a la mejor participación de los centros de salud, en hacer conscientizar a sus pacientes la importancia de seguir un tratamiento de salud. También, podemos suponer a la mejor preparación del personal del centro de salud en brindar los servicios requeridos, según los procedimientos del programa.
- El porcentaje de pacientes a ser analizados de la variable afuera de la ciudad, ha ido creciendo con el transcurso del tiempo.
- La variable Total de pacientes analizados tiene una tendencia descendente gradual, a pesar de existir fluctuaciones de repunte.
- La variable Ciudad tiene una tendencia descendente, con fluctuaciones de repunte.
- La variable Afuera de la Ciudad tiene una tendencia de crecimiento lineal, a pesar que todo comportamiento de vida humana, tiene un crecimiento exponencial.
- La variable Total de Enfermos tiene una tendencia de crecimiento, dando evidencia del repunte de está enfermedad.

BIBLIOGRAFÍA

1. De March P. En busca de la oportunidad perdida: prioridades en el control de la tuberculosis. *Med Clin (Barc)* 1997; 109: 55-57.
2. Groups at Risk. Global tuberculosis programme. WHO Report on the tuberculosis epidemic. Ginebra: WHO, 1996.
3. Vázquez E, Blanco Aparicio M, Fernández E, Anibarro L, Lema R, Penas A et al. Study of contacts of persons with newly diagnosed tuberculosis in Galicia, Spain, in 1995 and 1996. *Int J Tuberc Lung Dis* 1997; 1 (Supl): 101.
4. John E. Freund. *Estadística Matemática con Aplicaciones*. Editorial Prentice-Hall Hispanoamericana, Cuarta Edición, 1990
5. Hair, Anderson, Tatham, Black. *Multivariate Data Analysis* Editorial Prentice-Hall, Quinta Edición, 1998.
6. U. Narayan Bhat. *Elements of applied stochastic processes*. Editorial John Wiley&Sons. Primera Edición, 1976.