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 Polictécnica 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 sé pronóstico 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 está infección. Está 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 laborares mensuales. Además los hospitales, los laboratorios particulares, envían su 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 concurridos, por el departamento de tuberculosis. Para poder así explicar sus labores anuales al Ministerios 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.

Estreptomicina.- 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.

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

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

Pulmonar.- Esta variables es de tipo numérico y esta conformada por la suma de las variables: Esputos, H. Laringeo, Gástricos y Bronquios.

Extrapulmonar.- Esta variables es de tipo numérico y esta 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 esta conformada por la suma de las variables Pulmonar y Extrapulmonar.

4. Análisis Estadístico Univariado

4.1 Introducción

Esta sección esta constituido en dos puntos. El primer punto concierne a la explicación de las herramientas estadísticas. El segundo punto trata acerca del análisis estadístico de los resultados obtenidos de las diferentes variables.

4.2 Explicación de las herramientas estadísticas

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



donde N es el tamaño de la población.

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

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:



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:



Intervalo de confianza para la media al 95%: Inferior<<Superior. Se denomina intervalo de confianza de 100 (1-) por ciento para el parámetro , 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-) por ciento de ellos contendrá el valor verdadero de .

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



El valor mediano, Q2, separa el 50% superior de un conjunto de observaciones, del 50% inferior. De manera semejante, el primer cuartil, Q1, es el valor que corresponde al punto por debajo del cual se encuentra el 25% de las observaciones. El tercer cuartil, Q3, 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 Q3 y Q1.

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:



cuando **x** es discreta y



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:

.

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:



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:



para . 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  (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.