



**ALUMNA: KARINA MENDOZA CRUZ**

**MATERIA: BIOESTADISTICA**

**LIC. RODRIGO MANUEL BRAVO LOPEZ**

## **1. La estadística en enfermería.**

En las ciencias de la salud, la estadística tiene una gran importancia ya que posee numerosas ventajas, por ejemplo, nos puede ayudar a conocer las problemáticas presentes en una comunidad, los factores de riesgo o predisposición a ciertas patologías y puede ser muy útil a la hora de buscar una respuesta a esta o al tratar de educar para evitarlas en futuras ocasiones.

La principal ventaja del pensamiento estadístico interviniendo en las ciencias de la vida es que no solo resuelve, sino que también comprende una compleja metodología para dar respuesta a las hipótesis, además de agilizar la cuestión de organización del sistema de investigación, desde el diseño general, el de muestreo, el control de la calidad de información y la presentación de los resultados.

## **2. Introducción histórica.**

El primer médico que utilizó métodos matemáticos para cuantificar variables de pacientes y sus enfermedades fue el francés Pierre Charles-Alexandre Louis (1787-1872). La primera aplicación del Método numérico (que es como tituló a su obra y llamó a su método) en su clásico estudio de la tuberculosis, que influyó en toda una generación de estudiantes. Sus discípulos, a su vez, reforzaron la nueva ciencia de la epidemiología con base en el método estadístico.

En Francia Louis René Villermé (1782-1863) y en Inglaterra William Farr (1807-1883) que había estudiado estadística médica con Louis hicieron los primeros mapas epidemiológicos usando métodos cuantitativos y análisis epidemiológicos. Francis Galton (1822-1911), basado en el darwinismo social, fundó la biometría estadística.

Pierre Simón Laplace (1749-1827), astrónomo y matemático francés, publicó en 1812 un tratado sobre la teoría analítica de las probabilidades, *Théorie analytique des probabilités*, sugiriendo que tal análisis podría ser una herramienta valiosa para resolver problemas médicos.

Tres diferentes problemas cuantitativos fueron estudiados por otros tantos autores. William Heaton Hamer (1862-1936) propuso un modelo temporal discreto en un intento de explicar la ocurrencia regular de las epidemias de sarampión; John Brownlee (1868-1927), primer director del British Research Council, luchó durante veinte años con problemas de cuantificación de la infectividad epidemiológica, y Ronald Ross (1857-1932) exploró la aplicación matemática de la teoría de las probabilidades con la finalidad de determinar la relación entre el número de mosquitos y la incidencia de malaria en situaciones endémicas y epidémicas.

Los primeros trabajos bioestadísticos en enfermería los realizó, a mediados del siglo XIX la enfermera inglesa Florence Nightingale. Durante la guerra de Crimea, Florence Nightingale observó que eran mucho más numerosas las bajas producidas en el hospital que en el frente. Por lo tanto, recopiló información y dedujo que la causa de la elevada tasa de mortalidad se debía a la precariedad higiénica existente. Así, gracias a sus análisis estadísticos, se comenzó a tomar conciencia de la importancia y la necesidad de unas buenas condiciones higiénicas en los hospitales.

### 3. La estadística como herramienta de trabajo en enfermería.

El análisis y las técnicas estadísticas son un componente esencial en toda investigación biomédica, y la utilización de las técnicas estadísticas ha evolucionado considerablemente en los últimos años en las áreas de la investigación de ciencias de la salud.

Esta disciplina es usada en diversos campos de la medicina y la salud pública, como la epidemiología, nutrición y salud ambiental. Asimismo, sus métodos son aplicados en estudios relacionados con la ecología y la genómica.

Algunas de las aportaciones más importantes de la bioestadística se han dado en el estudio de las enfermedades. A raíz de los datos arrojados por esta disciplina se ha logrado un mejor entendimiento de la propagación de ciertas enfermedades y las características de males crónicos como el cáncer y el sida. Además, ha contribuido enormemente al desarrollo de nuevos fármacos.

Descripción de una variable estadística. Cuando hablamos de variable estadística estamos hablando de una cualidad que, generalmente adopta forma numérica. Por ejemplo, la altura de Juan es de 180 centímetros. La variable estadística es la altura y está medida en centímetros. También podríamos, por ejemplo, decir que el beneficio de una empresa ha sido de 22.300 dólares el último año. En este caso, la variable sería el beneficio y estaría medido en dólares. Ambas variables son del tipo cuantitativo (se expresan con un número).

Por ejemplo, Juan tiene los ojos verdes y Andrés los tiene azules. La variable sería el color de ojos y sería una variable cualitativa. Es decir, no se expresa con número.

### 4. Definiciones básicas.

**Variable estadística:** Una variable estadística es una característica de una muestra o población de datos que puede adoptar diferentes valores.

**Variable cuantitativa:** Son variables que se expresan numéricamente.

**Variable continua:** Toman un valor infinito de valores entre un intervalo de datos.

Por ejemplo, el tiempo que tarda un corredor en completar los 100 metros lisos.

**Variable discreta:** Toman un valor finito de valores entre un intervalo de datos.

Ejemplo: Número de helados vendidos.

**Variable cualitativa:** Son variables que se expresan, por norma general, en palabras.

**Variable ordinal:** Expresa diferentes niveles y orden. Por ejemplo, primero, segundo, tercero, etc.

**Variable nominal:** Expresa un nombre claramente diferenciado. Por ejemplo, el color de ojos puede ser azul, negro, castaño, verde, etc.

cada una de estas variables podría tener más subtipos, ya que tenemos variables de tipo económico, categóricas, dicotómicas, dependientes, independientes. Es decir, como ya hemos dicho, muchos tipos de variables estadísticas. Por ejemplo, podríamos tener una variable estadística de tipo cuantitativo, discreta y dependiente.

## 5. Representaciones gráficas

Una gráfica o una representación gráfica o un gráfico, es un tipo de representación de datos, generalmente cuantitativos, mediante recursos visuales (líneas, vectores, superficies o símbolos), para que se manifieste visualmente la relación matemática o correlación estadística que guardan entre sí.

La representación gráfica permite establecer valores que se han obtenido experimentalmente sino mediante la interpolación (lectura entre puntos) y la extrapolación (valores fuera del intervalo experimental).

**6. Tipos de representaciones gráficas:** Cuando se muestran los datos estadísticos a través de representaciones gráficas, se ha de adaptar el contenido a la información visual que se pretende transmitir.

Para ello, se barajan múltiples formas de representación:

**Diagramas de barras:** muestran los valores de las frecuencias absolutas sobre un sistema de ejes cartesianos, cuando la variable es discreta o cualitativa. se usan para representar gráficamente series estadísticas de valores en un sistema de ejes cartesianos, de manera que en las abscisas se indica el valor de la variable estadística y en las ordenadas se señala su frecuencia absoluta.

**Histogramas:** formas especiales de diagramas de barras para distribuciones cuantitativas continuas.

**Polígonos de frecuencias:** formados por líneas poligonales abiertas sobre un sistema de ejes cartesianos. Esta gráfica se usa para representar los puntos medios de clase en una distribución de frecuencias. Para construir polígonos de frecuencias, se trazan las frecuencias absolutas o relativas de los valores de la variable en un sistema de ejes cartesianos y se unen los puntos resultantes mediante trazos rectos.

**Gráficos de sectores:** circulares o de tarta, dividen un círculo en porciones proporcionales según el valor de las frecuencias relativas. también llamados circulares o de tarta, se muestra el valor de la frecuencia de la variable señalada como un sector circular dentro de un círculo completo.

**Pictogramas:** o representaciones visuales figurativas. En realidad, son diagramas de barras en los que las barras se sustituyen con dibujos alusivos a la variable.

Dos formas comunes de expresión gráfica de los datos son:

Los pictogramas, que muestran diagramas figurativos con figuras o motivos que aluden a la distribución estadística analizada (por ejemplo, una imagen antropomórfica para indicar tamaños, alturas u otros).

Los cartogramas, basados en mapas geográficos que utilizan distintas tramas, colores o intensidades para remarcar las diferencias entre los datos.

**Cartogramas:** expresiones gráficas a modo de mapa. Pirámides de población: para clasificaciones de grupos de población por sexo y edad.

**La ojiva:** Esta gráfica consiste en la representación de las frecuencias acumuladas de una distribución de frecuencias.

## 7. Representación numérica

La tabla de frecuencias (o distribución de frecuencias) es una tabla que muestra la distribución de los datos mediante sus frecuencias. Se utiliza para variables cuantitativas o cualitativas ordinales.

Construcción de la tabla de frecuencias Cabe distinguir entre:

- Tabla de frecuencias con datos no agrupados.
- Tabla de frecuencias con datos agrupados.

### **Construcción de una tabla de frecuencias con datos no agrupados**

1. En la primera columna se ordenan de menor a mayor los diferentes valores que tiene la variable en el conjunto de datos.
2. En las siguientes columnas (segunda y tercera) se ponen las frecuencias absolutas y las frecuencias absolutas acumuladas.
3. Las columnas cuarta y quinta contienen las frecuencias relativas y las frecuencias relativas acumuladas.
4. Adicionalmente (opcional) se pueden incluir dos columnas (sexta y séptima), representando la frecuencia relativa y la frecuencia relativa acumulada como tanto por cien. Estos porcentajes se obtienen multiplicando las dos frecuencias por cien.

**Construcción de una tabla de frecuencias con datos agrupados:** Se emplea cuando hay un número alto de datos. Estos se agrupan en intervalos o clases para facilitar su tabulación y análisis.

### **Tipos de frecuencias**

#### **Existen cuatro tipos de frecuencias**

**Frecuencia absoluta acumulada:** La frecuencia absoluta ( $n_i$ ) de un valor  $X_i$  es el número de veces que el valor está en el conjunto  $(X_1, X_2, \dots, X_N)$ . La suma de las frecuencias absolutas de todos los elementos diferentes del conjunto debe ser el número total de sujetos  $N$ .

**Frecuencia relativa** La frecuencia relativa ( $f_i$ ) de un valor  $X_i$  es la proporción de valores iguales a  $X_i$  en el conjunto de datos  $(X_1, X_2, \dots, X_N)$ . Es decir, la frecuencia relativa es la frecuencia absoluta dividida por el número total de elementos  $N$ :

## 8. Características de posición, dispersión y forma

**Medidas de posición:** Las medidas de posición son indicadores estadísticos que permiten resumir los datos en uno solo, o dividir su distribución en intervalos del mismo tamaño.

Las medidas de posición se suelen dividir en dos grandes grupos: la de tendencia no central y las centrales. Las medidas de posición no centrales son los cuantiles. Estos realizan una serie de divisiones iguales en la distribución ordenada de los datos.

**El cuartil:** Es uno de los más utilizados y divide la distribución en cuatro partes iguales. Así, existen tres cuartiles. Los valores inferiores de la distribución se sitúan por debajo del primero ( $Q_1$ ). La mitad o mediana son los valores menores iguales al cuartil dos ( $Q_2$ ) y los superiores son representados por el cuartil tres ( $Q_3$ ).

**El quintil:** En este caso, divide la distribución en cinco partes. Por tanto, hay cuatro quintiles.

**El decil:** Estamos ante un cuartil que divide los datos en diez partes iguales. Existen nueve deciles, de  $D_1$  a  $D_9$ . El  $D_5$  se corresponde con la mediana.

**El percentil:** Por último, este cuartil divide la distribución en cien partes. Hay 99 percentiles. Tiene, a su vez, una equivalencia con los deciles y cuartiles.

**Medidas de posición central** Estas nos permiten resumir la distribución de los datos en un solo valor central, alrededor del cual se sitúan; mientras que las segundas dividen la distribución en partes iguales.

**La media aritmética, geométrica o armónica:** Son tres medidas centrales que nos indican un promedio ponderado de los datos.

**La mediana:** En este caso, esta es la medida de posición central más reconocible. Divide la distribución en dos partes iguales. De esta forma, expresa el valor mediano, que no medio.

**La moda:** Estamos ante una medida central de los valores más frecuentes. Por tanto, la moda nos informa sobre aquellos que se repiten en más ocasiones.

**Medidas de dispersión:** Las medidas de dispersión, o de variabilidad, expresan cómo se distribuyen los datos en torno a alguna de las medidas de centralización definidas antes, y son un complemento a estas últimas para describir más fielmente un conjunto de datos.

**Varianza:** La Varianza es una medida de dispersión que se utiliza para representar la variabilidad de un conjunto de datos respecto de la media aritmética de los mismo.

**Desviación estándar:** La desviación estándar o desviación típica es una medida que ofrece información sobre la dispersión media de una variable. La desviación estándar es siempre mayor o igual que cero. Se obtiene al sacar la raíz cuadrada a la varianza.

**Medidas de forma:** Las medidas de forma son aquellas que nos muestran si una distribución de frecuencia tiene características especiales como simetría, asimetría, nivel de concentración de datos y nivel de apuntamiento que la clasifiquen en un tipo particular de distribución.

**dos tipos de medida:**

- Coeficiente de asimetría de Fischer.
- Coeficiente de curtosis a apuntamiento de Fisher.

En caso contrario, dicha distribución será asimétrica o diremos que presenta asimetría. La asimetría puede ser de dos tipos:

- Asimétrica por la derecha.
- Asimétrica por la izquierda.

La referencia citada es la distribución normal, y distinguiremos tres casos:

- Leptocúrtica, si la distribución es más picuda que la normal,
- Mesocúrtica, si la distribución es igual a la normal, y.
- Platicúrtica, si la distribución es más aplastada que la normal.

## 9. Descripción numérica de una variable estadística bidimensional

En numerosas ocasiones interesa estudiar simultáneamente dos (o más) caracteres de una población. En el caso de dos (o más) variables estudiadas conjuntamente se habla de variable bidimensional (multidimensional); si se trata de dos caracteres cualitativos, de par de atributos.

Se denominan distribuciones bidimensionales a las tablas estadísticas bidimensionales formadas por todas las frecuencias absolutas de todos los posibles valores de la variable estadística bidimensional (X, Y). Las tablas estadísticas bidimensionales pueden ser: Simples y de doble entrada.

**Distribuciones marginales y condicionadas:** En teoría de probabilidades, la distribución marginal es la distribución de probabilidad de un subconjunto de variables aleatorias de un conjunto de variables aleatorias.

El término variable marginal se usa para referirse a una variable del subconjunto de retenido y cuyos valores pueden ser conocidos.<sup>1</sup>

## 10. Independencia e incorrelación

Dos variables estadísticas son estadísticamente independientes cuando el comportamiento estadístico de una de ellas no se ve afectado por los valores que toma la otra; esto es cuando las relativas de las distribuciones condicionadas no se ven afectadas por la condición, y coinciden en todos los casos con las frecuencias relativas marginales.

Se dice que dos variables X e Y son independientes estadísticamente cuando la frecuencia relativa conjunta es igual al producto de las frecuencias relativas marginales en todos los casos, es decir:  $i, j$  Si esto no se cumple para todos los valores se dice que hay dependencia estadística.

$P(A/B) = P(A)$  es decir, que la probabilidad de que se dé el suceso A, condicionada a que previamente se haya dado el suceso B, es exactamente igual a la probabilidad de A.

**Incorrelación** Es el grado de dispersión entre los puntos de una variable, es decir, el cuándo los puntos no marchan en una misma dirección si no que están dispersos por todos lados, a diferencia de la correlación que es todo lo contrario. Características numéricas Los sistemas de numeración son conjuntos de dígitos usados para representar cantidades, así se tienen los sistemas de numeración decimal, binario, octal, hexadecimal, romano.

Los cuatro primeros se caracterizan por tener una base (número de dígitos diferentes: diez, dos, ocho, dieciséis respectivamente) mientras que el sistema romano no posee base y resulta más complicado su manejo tanto con números, así como en las operaciones básicas.

## 11. Regresión y correlación

En forma más específica el análisis de correlación y regresión comprende el análisis de los datos muestrales para saber qué es y cómo se relacionan entre si dos o más variables en una población. El análisis de correlación produce un número que resume el grado de la correlación entre dos variables; y el análisis de regresión da lugar a una ecuación matemática que describe dicha relación.

**Coefficiente de determinación** El coeficiente de determinación es la proporción de la varianza total de la variable explicada por la regresión. Es también denominado R cuadrado y sirve para reflejar la bondad del ajuste de un modelo a la variable que se pretende explicar.

BIBLIOGRAFIA:  
Antología