



**Mi Universidad**

*Nombre del Alumno: Emilly Cruz Martínez*

*Nombre del tema: Estadística Descriptiva*

*Nombre de la Materia: Bioestadística*

*Nombre del Maestro: Rosario Gómez Lujano*

*Nombre de la Licenciatura: Enfermería*

*Cuatrimestre: 4to*

# ESTADISTICA DESCRIPTIVA

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.

## 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.

Los primeros intentos de hacer coincidir las matemáticas de la teoría estadística con los conceptos emergentes de la infección bacteriana tuvieron lugar a comienzos del siglo XX. William Heaton Hamer propuso un modelo temporal en un intento de explicar la ocurrencia regular de las epidemias de sarampión y Ronald Ross 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.

## La estadística como herramienta de trabajo en enfermería

la estadística juega un papel fundamental en la investigación en ciencias de la salud, y a través de un equipo multidisciplinar que engloba a profesionales del ámbito sanitario, académico y perfiles expertos en metodología estadística se obtienen investigaciones de mayor calidad. Esta disciplina es usada en diversos campos de la medicina y la salud pública, como la epidemiología, nutrición y salud ambiental.

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. El pensamiento estadístico ha permitido establecer un sistema organizado de investigación, desde el diseño de la misma, el muestreo, el control de calidad, el análisis y la presentación de la información.

## 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. Claro que no todas las variables estadísticas son iguales y, por supuesto, no todas se pueden expresar en forma de número.

Por norma general podemos encontrarnos dos tipos de variables: Variable cuantitativa: Son variables que se expresan numéricamente. Variable continua: Toman un valor infinito de valores entre un intervalo de datos. Variable discreta: Toman un valor finito de valores entre un intervalo de datos. Variable cualitativa: Son variables que se expresan, por norma general, en palabras. Variable ordinal: Expresa diferentes niveles y orden. Variable nominal: Expresa un nombre claramente diferenciado.

# ESTADISTICA DESCRIPTIVA

## 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.

**Tipos de representaciones gráficas:** Diagramas de barras: muestran los valores de las frecuencias absolutas sobre un sistema de ejes cartesianos, cuando la variable es discreta o cualitativa. 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. Gráficos de sectores: circulares o de tarta, dividen un círculo en porciones proporcionales según el valor de las frecuencias relativas. Pictogramas: En realidad, son diagramas de barras en los que las barras se sustituyen con dibujos alusivos a la variable. Cartogramas: expresiones gráficas a modo de mapa. Pirámides de población: para clasificaciones de grupos de población por sexo y edad.

## Representación numérica

La tabla 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. **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. Está indicado para representarlos en un histograma. Existen cuatro tipos de frecuencias: Frecuencia absoluta, Frecuencia absoluta acumulada, Frecuencia relativa y Frecuencia relativa acumulada.

## Características de posición, dispersión y forma

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

### 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, Desviación estándar). **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 • Coeficiente de asimetría de Fischer. • Coeficiente de curtosis a apuntamiento de Fisher.

# ESTADISTICA DESCRIPTIVA

## Descripción numérica de una variable estadística bidimensional

Variable estadística bidimensional es el conjunto de pares de valores de dos caracteres o variables estadísticas unidimensionales X e Y sobre una misma población. La variable estadística bidimensional se representa por el símbolo (X, Y) y cada uno de los individuos de la población viene caracterizado por la pareja (xi, yi), en el cual xi representa los datos, valores o marcas de clase x1, x2, ..., xn de la variable X; e yi representa los datos, valores o marcas de clase y1, y2, ..., ym de la variable Y.

## Distribuciones marginales y condicionadas

la distribución marginal es la distribución de probabilidad de un subconjunto de variables aleatorias de un conjunto de variables aleatorias. La distribución marginal proporciona la probabilidad de un subconjunto de valores del conjunto sin necesidad de conocer los valores de las otras variables.

La distribución marginal de X es simplemente la función de probabilidad de x, pero la palabra marginal sirve para distinguirla de la distribución conjunta de X e Y. La distribución marginal de dos variables aleatorias se puede obtener a partir de su distribución conjunta. Para una variable aleatoria se puede especificar probabilidades para dicha variable sin tener en cuenta los valores de cuales quiera otras variables aleatorias.

Por ejemplo:

$$P(H) = 610/1000 = 0.61$$

$$P(M) = 390/1000 = 0.39$$

$$P(\text{Messenger}) = 423/1000 = 0.423$$

$$P(\text{WhatsApp}) = 577/1000 = 0.577$$

De cada distribución bidimensional se pueden deducir dos distribuciones marginales: una correspondiente a la variable x, y otra correspondiente a la variable

## 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.

Ejemplo: el suceso estatura de los alumnos de una clase y el color del pelo son independientes: el que un alumno sea más o menos alto no va a influir en el color de su cabello, ni viceversa. Para que dos sucesos sean independientes tienen que verificar al menos una de las siguientes condiciones:  $P(B/A) = P(B)$  es decir, que la probabilidad de que se dé el suceso B, condicionada a que previamente se haya dado el suceso A, es exactamente igual a la probabilidad de B.

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.

# ESTADÍSTICA DESCRIPTIVA

## 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.

Curva de regresión y coeficiente de determinación. La curva de regresión de Y sobre X visualiza como cambia la media de la variable Y de aquellos grupos de observaciones caracterizados por tener un mismo valor en la otra variable X.

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. El coeficiente de determinación puede adquirir resultados que oscilan entre 0 y 1. Así, cuando adquiere resultados más cercanos a 1, mayor resultará el ajuste del modelo a la variable que se pretende aplicar para el caso en concreto.

## Otros tipos de regresión

Regresión Múltiple: Este tipo se presenta cuando dos o más variables independientes influyen sobre una variable dependiente. Ejemplo:  $Y = f(x, w, z)$ . El error estándar de la regresión múltiple Es una medida de dispersión la estimación se hace más precisa conforme el grado de dispersión alrededor del plano de regresión se hace más pequeño. El coeficiente de determinación múltiple Mide la tasa porcentual de los cambios de Y que pueden ser explicados por  $x_1$ ,  $x_2$  y  $x_3$  simultáneamente.

## Análisis de atributos

Su principal objetivo es el de evitar un error muy común consistente en tratar de encontrar la forma de mejorar un producto, servicio o proceso analizándolo como un todo.

Características para las Gráficas de Control de Atributos

- Están basadas en decisiones de pasa/no pasa.
- Se pueden aplicar en casi cualquier operación donde se recolectan datos.
- Se utilizan en características de calidad que no pueden ser medidas o que son costosas o difíciles de medir. A diferencia de las gráficas de control de datos variables, las gráficas de datos atributos se pueden establecer para una característica de calidad o para muchas.

Tipos de Gráficas de Atributos:

- Defectivos – np - número de unidades no-conformes – p - proporción de unidades no-conformes.
- Defectos – c - número de defectos – u - proporción de defectos.