

***NOMBRE DEL ALUMNO:***

*DIANA PATRICIA CASTILLEJOS LOPEZ*

***NOMBRE DEL TEMA:***

*ESTADISTICA DESCRIPTIVA*

***PARCIAL:***

*PRIMER PARCIAL*

***NOMBRE DE LA MATERIA:***

*BIOESTADISTICA 1*

***NOMBRE DEL PROFESOR:***

*ROSARIO GOMEZ LUJANO*

***NOMBRE DE LA LICENCIATURA:***

*LIC.ENFERMERIA*

***CUATRIMESTRE:***

*CUARTO CUTRIMESTRE*

# UNIDAD I: ESTADÍSTICA DESCRIPTIVA

## 1.1 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. Como los objetos de estudio de las ciencias de la vida son muy variados, la Bioestadística ha debido ampliar su campo para, de esta manera, incluir cualquier modelo cuantitativo, 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. En Salud Pública la estadística permite analizar situaciones en las que los componentes aleatorios contribuyen de forma importante en la variabilidad de los datos obtenidos. Se necesitan para el ejercicio eficaz de la salud pública y, adicionalmente, para la comprensión e interpretación de los datos sanitarios.

### 1.1.1 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. 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. 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. 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.

## 1.2 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. Los análisis estadísticos empleados en un estudio dependen en gran medida del tipo de estudio, del objetivo que se pretende abordar y del tamaño de la muestra, así como del grado de conocimiento por parte de los investigadores de las técnicas estadísticas y del software para su implementación. Es por ello que 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. Algunas de las aportaciones más importantes de la bioestadística se han dado en el estudio de las enfermedades. Sin lugar a dudas, 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.

### 1.3 Descripción de una variable estadística.

Cuando hablamos de variable estadística estamos hablando de una cualidad que, generalmente adopta forma numérica.

Claro que no todas las variables estadísticas son iguales y, por supuesto, no todas se pueden (en principio) expresar en forma de número. Así, otra variable que podríamos encontrarnos es el color de ojos de una persona.

## UNIDAD I: ESTADÍSTICA DESCRIPTIVA

### 1.3.1. Definiciones básicas.

**Variable estadística:** Una variable estadística es una característica de una muestra o población de datos que puede adoptar diferentes valores. Aunque hay decenas de tipos de variables estadísticas, 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. 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. Además, cada una de estas variables podría tener más subtipos, ya que tenemos variables de tipo económico, categóricas, dicotómicas, dependientes, independientes.

Adicionalmente, también debemos aclarar que el hecho que las variables cualitativas se expresen con nombre no quiere decir que no puedan ser parte de un modelo matemático.

# UNIDAD I: ESTADÍSTICA DESCRIPTIVA

## 1.4 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: o representaciones visuales figurativas. 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.

**Diagramas de barras e histogramas** Los diagramas de barras 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.

**Polígonos de frecuencias** Esta gráfica se usa para representar los puntos medios de clase en una distribución de frecuencias.  
**Gráficos de sectores** En los diagramas de sectores, 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 y cartograma**

- 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.

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

## 1.5 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. La tabla de frecuencias es una herramienta que permite ordenar los datos de manera que se presenten numéricamente las características de la distribución de un conjunto de datos o muestra.

**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.

**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.

Los pasos iniciales para formar una tabla de frecuencias con datos agrupados están encaminados a determinar el número de intervalos y definirlos (siempre que no se conozcan de antemano). Los pasos son: 1. Obtener el rango R de los datos. Es la diferencia entre el dato mayor y el menor del conjunto de valores que toma la variable a tabular.

2. Fijar cuántos intervalos o clases se desea. Se tiende a que el número de clases sea impar y que esté entre 5 y 15. Hay dos maneras de hacerlo: ♣ A criterio del investigador. ♣ Mediante el método de Sturges, que emplea la siguiente fórmula

3. Determinar la amplitud del intervalo o clase I: Es el resultado de dividir el rango R o amplitud total por el número de clases o intervalos nint que se han fijado

4. Formar los diferentes intervalos o clases, partiendo del valor mínimo del nuevo rango R'. Cada intervalo tendrá unos extremos a y b separados por la amplitud de clase o intervalo I.

# UNIDAD I: ESTADÍSTICA DESCRIPTIVA

v

## 1.6 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.

- El cuartil: Es uno de los más utilizados y divide la distribución en cuatro partes iguales.
- 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 D1 a D9.

- El percentil: Por último, este cuartil divide la distribución en cien partes.

Medidas de posición central

- 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.
- La moda: Estamos ante una medida central de los valores más frecuentes.

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.

## 1.7 Descripción numérica de una variable estadística bidimensional.

En numerosas ocasiones interesa estudiar simultáneamente dos (o más) caracteres de una población.

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  $(x_i, y_i)$ , en el cual  $x_i$  representa los datos, valores o marcas de clase  $x_1, x_2, \dots, x_n$  de la variable X; e  $y_i$  representa los datos, valores o marcas de clase  $y_1, y_2, \dots, y_m$  de la variable Y.

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.

## 1.8 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.

La función de probabilidad marginal es usada para hallar las diferentes distribuciones de probabilidad estadística de las variables individuales, con esta función podemos asignar diferentes valores a las variables conjuntas sin tener que relacionarlas, por ello se amplía las probabilidades de cada una de las variables.

# UNIDAD I: ESTADÍSTICA DESCRIPTIVA

## 1.9 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.

Esta definición puede hacerse más operativa, a través de la caracterización siguiente: Dos variables son estadísticamente independientes cuando para todos los pares de valores se cumple que la frecuencia relativa conjunta es igual al producto de las frecuencias relativas marginales.

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.

Características numéricas Los sistemas de numeración son conjuntos de dígitos usados para representar cantidades, así se tienen en los sistemas de numeración decimal, binario, octal, hexadecimal, romano, etc.

## 1.10 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 sí dos o más variables en una población.

El análisis de correlación generalmente resulta útil para un trabajo de exploración cuando un investigador o analista trata de determinar que variables son potenciales importantes, el interés radica básicamente en la fuerza de la relación.

Definiciones En estadística, el análisis de la regresión es un proceso estadístico para estimar las relaciones entre variables. Incluye muchas técnicas para el modelado y análisis de diversas variables, cuando la atención se centra en la relación entre una variable dependiente y una o más variables independientes (o predictoras).

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. Es decir, como varía, por término medio, la variable Y en función de los valores de X.

Regresión Lineal La regresión lineal simple comprende el intento de desarrollar una línea recta o ecuación matemática lineal que describe la reacción entre dos variables.

Correlación Lineal El coeficiente de correlación permite la medición de la correlación entre dos variables.

## 1.11 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.

El Error Estándar de Regresión Múltiple Mediante esta medida de dispersión se hace más preciso el grado de dispersión alrededor del plano de regresión, se hace más pequeño.

El coeficiente de determinación múltiple Utilizaremos para determinar la tasa porcentual de Y para ser explicados las variables múltiples

# UNIDAD I: ESTADÍSTICA DESCRIPTIVA

## 1.12 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. Muchas veces, la búsqueda de una idea global, salvadora, que mejore el todo, impide descubrir la característica específica que, por sí sola, podría producir el resultado deseado. 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.

UNIVERSIDAD DEL SURESTE 48 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

1.-Los pesos en kilogramos de ocho alumnos de bachillerato son los siguientes: **52, 60, 58, 54, 72, 65, 55 y 76**. Obtener: **Media aritmética, mediana, moda, rango, varianza y desviación estándar.**

**52,60,58,54,72,65,55,76**

**MEDIA ARITMETICA:61.5**

**50,54,55,58,60,65,72,76**

**MEDIANA: 59**       $\frac{58+60}{2} = \frac{118}{2} = 59$

**MODA: no tiene moda**

**RANGO: 76 – 52 = 24**

**VARIANZA**

$$s^2 = (52 - 61.5)^2 + (54 - 61.5)^2 + (55 - 61.5)^2 + (58 - 61.5)^2 + (60 - 61.5)^2 + (65 - 61.5)^2 + (72 - 61.5)^2 + (76 - 61.5)^2$$

$$s^2 = \frac{(-9.5)^2 + (-7.5)^2 + (6.5)^2 + (-3.5)^2 + (-1.5)^2 + (3.5)^2 + (10.5)^2 + (14.5)^2}{7}$$

$$s = \frac{(90.25) + (56.25) + (42.25) + (12.25) + (2.25) + (12.52) + (110.25) + (210.25)}{7}$$

$$\frac{536}{7} = 76.57$$

**DESVIACION ESTANDAR:**  $\sqrt{76} \cdot 57 = 8.75$

2.- Cierta universidad realizó un experimento sobre el coeficiente intelectual (C.I.) de sus alumnos, para lo cual aplicó un examen de C.I. a un grupo de 20 alumnos escogidos al azar, obteniendo los siguientes resultados: 119, 109, 124, 119, 106, 112, 112, 112, 109, 112, 124, 109, 109, 109, 106, 124, 112, 112, 106.

COEFICIENTE INTELECTUAL (C.I.)

<i>DATOS</i>	<i>F</i>	<i>FA</i>	<i>FR</i>	<i>FRA</i>	<i>FR%</i>
<i>106</i>	<i>3</i>	<i>3</i>	<i>0.15</i>	<i>0.15</i>	<i>15%</i>
<i>109</i>	<i>5</i>	<i>8</i>	<i>0.25</i>	<i>0.4</i>	<i>25%</i>
<i>112</i>	<i>7</i>	<i>15</i>	<i>0.35</i>	<i>0.75</i>	<i>35%</i>
<i>119</i>	<i>2</i>	<i>17</i>	<i>0.1</i>	<i>0.85</i>	<i>10%</i>
<i>124</i>	<i>3</i>	<i>20</i>	<i>0.15</i>	<i>1</i>	<i>15%</i>
	<i>20</i>		<i>1</i>		<i>100%</i>

