

**MATERIA:** BIOESTADISTICA

**DOCENTE:** Dr. RODRIGO M. BRAVO LOPEZ

**ALUMNA:** NOEMI CONCEPCION PUERTO ANTONIO

**CARRERA:** ENFERMERIA GENERAL

**RESUMEN** UNIDAD I

## **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 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.

## **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. Asimismo, sus métodos son aplicados en estudios relacionados con la ecología y la genómica.

## 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.

## 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. También es el nombre de un conjunto de puntos que se plasman en coordenadas cartesianas y sirven para analizar el comportamiento de un proceso o un conjunto de elementos o signos que permiten la interpretación de un fenómeno. La representación gráfica permite establecer valores que no se han obtenido experimentalmente sino mediante la interpolación (lectura entre puntos) y la extrapolación (valores fuera del intervalo experimental).

## **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.
- **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. Estos gráficos se usan en representación de caracteres cualitativos y cuantitativos discretos. En variables cuantitativas continuas, se emplea una variante de los mismos llamada histograma.

## **Polígonos de frecuencias**

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. Con ello se obtiene una forma de línea poligonal abierta. Los polígonos de frecuencias se utilizan preferentemente en la presentación de caracteres cuantitativos, y tienen especial interés cuando se indican frecuencias acumulativas. Se usan en la expresión de fenómenos que varían con el tiempo, como la densidad de población, el precio o la temperatura.

## **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. Por ello, resultan útiles particularmente para mostrar comparaciones entre datos, sobre todo en forma de frecuencias relativas de las variables expresadas en forma de porcentaje. Pictogramas y cartogramas Para aligerar la presentación de datos estadísticos, con frecuencia se recurre a imágenes pictóricas representativas del valor de las variables.

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.
- La ojiva: Esta gráfica consiste en la representación de las frecuencias acumuladas de una distribución de frecuencias. Puede construirse de dos maneras diferentes; sobre la base "menor que" o sobre la

base "o más". Puede determinar el valor de la mediana de la distribución.

## 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

- 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.

### Tipos de frecuencias

- Frecuencia absoluta
- Frecuencia absoluta acumulada
- Frecuencia relativa
- Frecuencia relativa acumulada

## 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. De esta forma, reflejan los valores superiores, medios e inferiores. Los más habituales son:

- **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 (Q1). La mitad o mediana son los valores menores iguales al cuartil dos (Q2) y los superiores son representados por el cuartil tres (Q3).
- **El quintil:** En este caso, divide la distribución en cinco partes. Por tanto, hay cuatro quintiles. Además, no existe ningún valor que divida la distribución en dos partes iguales. Es menos frecuente que el anterior.
- **El decil:** Estamos ante un cuartil que divide los datos en diez partes iguales. Existen nueve deciles, de D1 a D9. El D5 se corresponde con la mediana. Por su lado, los valores superiores e inferiores (equivalentes a los diferentes cuartiles) se sitúan en puntos intermedios entre estos.
- **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

- **La media aritmética, geométrica o armónica:** Son tres medidas centrales que nos indican un promedio ponderado de los datos. La primera es la más utilizada y la más conocida de las tres. La geométrica se aplica en series que muestran crecimientos porcentuales. Por su parte, la armónica es útil en el análisis de inversiones en bolsa.
- **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. Es muy útil en variables como los ingresos o salarios, a la vez que está muy relacionada con la media y algunos de los cuantiles vistos.
- **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.

## 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.

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.**

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. Esto contrasta con la distribución condicional, que proporciona probabilidades contingentes sobre el valor conocido de otras variables.

Se denomina distribución de la variable y condicionada a un valor dado  $x_i$  de la variable  $X$  a la distribución unidimensional definida por el conjunto de valores tomados por  $Y$  y de las frecuencias de dichos valores de  $Y$  condicionadas a que  $X$  tome el valor  $x_i$ .

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.

## **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.

## **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.

## **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. Por eso la variable Y recibe el nombre de variable dependiente y la variable X el de variable independiente.

## **Coeficiente 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.

## **Regresión y correlación lineal**

### 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.

La regresión puede ser utilizada de diversas formas. Se emplean en situaciones en la que las dos variables miden aproximadamente lo mismo, pero en las que una variable es relativamente costosa, o, por el contrario, es poco interesante trabajar con ella, mientras que con la otra variable no ocurre lo mismo.

## Otros tipos de regresión

**Regresión Múltiple:** Este tipo se presenta cuando dos o más variables independientes influyen sobre una variable dependiente.

## 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.

### 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