



UNIVERSIDAD DEL SUERESTE

**Materia:**

Bioestadística

**Docente:**

Enrique Eduardo Arreola

**Alumno:**

J. Carlos De Los Santos De La Cruz

# INTRODUCCION

La estadística es uno de los pilares del método científico una vez alcanzada la fase de análisis de los datos.

La estadística descriptiva permite organizar y presentar los datos en tablas o gráficos, así como resumirlos con medidas de centralización y de dispersión, simplificando la interpretación de los mismos. La estadística inferencial estudia las variables o características que presentan los individuos, generalizando los datos obtenidos a partir de una muestra a un número mayor de individuos (población).

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 en el método estadístico. En las recomendaciones de Louis para evaluar diferentes métodos de tratamiento están las bases de los ensayos clínicos que se hicieron un siglo después. 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

## La estadística en enfermería

Aunque aparentemente la bioestadística parece una ciencia fundamentalmente teórica, es utilizada en la práctica médica a diario. Cuando hablamos de la dosis media de eritropoyetina administrada en diálisis o el tiempo medio de duración de una sesión de hemodiálisis estamos utilizando la estadística. O cuando decidimos tratar a un paciente con unas cifras de colesterol o de presión arterial elevadas, previamente se ha demostrado estadísticamente que existe un riesgo elevado cuando esas cifras están por encima de un determinado valor. O, por ejemplo, cuando queremos comparar si existe diferencias entre un fármaco inmunosupresor u otro a la hora de prevenir el rechazo de un trasplante renal. El objetivo es iniciar y familiarizar a la enfermería con el método científico, y más concretamente con las nociones básicas del análisis estadístico necesario para cualquier estudio de investigación

La bioestadística es una rama de la estadística que se ocupa de los problemas planteados dentro de las ciencias de la vida, como la biología, la medicina, la enfermería, entre otras.

## Elementos del análisis estadístico en enfermería

La estadística descriptiva comprende la presentación, organización y resumen de los datos de una manera científica. Incluye diversos métodos de organizar y representar gráficamente los datos, para dar una idea de lo que nos muestran. Las tablas, los diagramas de barras o los gráficos sectoriales o "tartas" son algunos de los elementos de estadística descriptiva. También incluye varios parámetros numéricos (como la media aritmética) que resumen los datos con muy pocos números clave. Por otra parte, la estadística inferencial o inductiva permite generalizar los datos obtenidos a partir de una muestra a un número mayor de individuos (población). La estadística inferencial se basa en la teoría de las probabilidades y trabaja con los datos que le proporciona la estadística descriptiva.

## La estadística como herramienta de trabajo en enfermería

El conocimiento de la estadística favorece el desarrollo personal pues fomenta un razonamiento crítico, aumenta la capacidad de usar datos cuantitativos para controlar nuestros juicios e interpretar los ajenos y transformarlos para resolver problemas de decisión y efectuar predicciones. (Ottaviani, 1998). En Enfermería el estudio de la Estadística aporta los conceptos fundamentales y necesarios con el dominio adecuado del instrumental para aproximarse al estudio y conocimiento de los fenómenos de competencia de la Enfermería. La práctica de la investigación y la transferencia de conocimientos producidos al ejercicio profesional, constituye la actividad básica para el desarrollo de la Enfermería a través del cual se aspira a la meta social de dar respuesta a los problemas y necesidades de la comunidad. En el campo de la Salud, las prioridades de investigación exigen que el personal que se forma y trabaja en el sector incorpore la investigación como una actividad permanente en su ámbito de acción. La Estadística desempeña un papel importante en la toma de decisiones en todas las áreas, entre ellas la salud pública. Las medidas relativas a diferentes programas sanitarios confían en parte, en las predicciones sobre la longevidad de la población, o cómo invertir recursos para reducir la mortalidad infantil, disminuir la probabilidad de muerte en accidentes vehiculares con el uso del cinturón de seguridad, cuáles factores incrementan el riesgo de que un individuo desarrolle una enfermedad coronaria

## Descripción de una variable estadística

Una variable estadística es una característica que puede fluctuar y cuya variación es susceptible de adoptar diferentes valores, los cuales pueden medirse u observarse. Las variables adquieren valor cuando se relacionan con otras variables, es decir, si forman parte de una hipótesis o de una teoría. En este caso se las denomina constructos o construcciones hipotéticas

# Definiciones básicas

**Variables cualitativas** Son el tipo de variables que como su nombre lo indica expresan distintas cualidades, características o modalidad. Cada modalidad que se presenta se denomina atributo o categoría, y la medición consiste en una clasificación de dichos atributos. Las variables cualitativas pueden ser dicotómicas cuando sólo pueden tomar dos valores posibles, como sí y no, hombre y mujer o ser politómicas cuando pueden adquirir tres o más valores. Dentro de ellas podemos distinguir:

- **Variable cualitativa ordinal o variable cuasi cuantitativa:** La variable puede tomar distintos valores ordenados siguiendo una escala establecida, 14 aunque no es necesario que el intervalo entre mediciones sea uniforme, por ejemplo: leve, moderado, fuerte.

- **Variable cualitativa nominal:** En esta variable los valores no pueden ser sometidos a un criterio de orden, como por ejemplo los colores o el lugar de registro. **Variables cuantitativas** Son las variables que toman como argumento cantidades numéricas, son variables matemáticas. Las variables cuantitativas además pueden ser:

- **Variable discreta:** Es la variable que presenta separaciones o interrupciones en la escala de valores que puede tomar. Estas separaciones o interrupciones indican la ausencia de valores entre los distintos valores específicos que la variable pueda asumir. Ejemplo: El número de hijos (1, 2, 3, 4, 5).

- **Variable continua:** Es la variable que puede adquirir cualquier valor dentro de un intervalo especificado de valores. Por ejemplo, la masa (2,3 kg, 2,4 kg, 2,5 kg...) o la altura (1,64 m, 1,65 m, 1,66 m..), o el salario. Solamente se está limitado por la precisión del aparato medidor, en teoría permiten que exista indefinidos valores entre dos variables.

**Variables independientes** Una variable independiente es aquella cuyo valor no depende de otra variable. Es aquella característica o propiedad que se supone es la causa del fenómeno estudiado. En investigación experimental se llama así a la variable que el investigador manipula. 15 Las variables independientes son las que el investigador escoge para establecer agrupaciones en el estudio, clasificando intrínsecamente a los casos del mismo. Un tipo especial son las variables de control, que modifican al resto de las variables independientes y que de no tenerse en cuenta adecuadamente pueden alterar los resultados por medio de un sesgo. La variable independiente se suele representar en el eje de abscisas. La variable independiente es la que se le asignan valores arbitrarios **Variables dependientes** Una variable dependiente es aquella cuyos valores dependen de los que tomen otra variable. La variable dependiente es una función que se suele representar por la  $y$ .

La variable dependiente se representa en el eje ordenadas. Son las variables de respuesta que se observan en el estudio, y que podrían estar influidas por los valores de las variables independientes. La variable dependiente es el factor que es observado y medido para determinar el efecto de la variable independiente.

## Representaciones gráficas

gráfica o representación gráfica es un tipo de representación de datos, generalmente numéricos, 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).



# Representación numérica

La presentación de datos estadísticos constituye en sus diferentes modalidades uno de los aspectos de más uso en la estadística descriptiva

1-Presentación escrita: Esta forma de presentación de informaciones se usa cuando una serie de datos incluye pocos valores, por lo cual resulta más apropiada la palabra escrita como forma de escribir el comportamiento de los datos; mediante la forma escrita, se resalta la importancia de las informaciones principales.

2-Presentación tabular: Cuando los datos estadísticos se presentan a través de un conjunto de filas y de columnas que responden a un ordenamiento lógico; es de gran uso e importancia para el usuario ya que constituye la forma más exacta de presentar las informaciones.

Título: Es la parte más importante del cuadro y sirve para describir todo el contenido de este.

Encabezados: Son los diferentes subtítulos que se colocan en la parte superior de cada columna.

Columna matriz: Es la columna principal del cuadro.

Cuerpo: El cuerpo contiene todas las informaciones numéricas que aparecen en la tabla.

Fuente: La fuente de los datos contenidos en la tabla indica la procedencia de estos.

Notas al pie: Son usadas para hacer algunas aclaraciones sobre aspectos que aparecen en la tabla o cuadro y que no han sido explicados en otras partes

# Características de posición, dispersión y forma

Las medidas de posición proporcionan información resumida de la variable objeto de estudio. Medidas de posición centrales

Medidas de posición centrales

- Media (aritmética, geométrica y armónica)
- Mediana
- Moda

Medidas de posición no centrales

- Cuantiles (cuartiles, deciles y percentiles)

Las medidas de dispersión estudian la separación existente entre los valores que toma la variable. Medidas de dispersión absolutas

- Rango
- Recorrido intercuartílico
- Desviación absoluta media respecto a la media
- Varianza
- Desviación típica

Medidas de dispersión relativas

- Coeficiente de apertura
- Recorrido relativo
- Recorrido semi-intercuartílico
- Coeficiente de variación
- Variable tipificada

## 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.

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: a) Simples. b) 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> La distribución de las variables marginales, la distribución marginal, se obtiene marginalizando sobre la distribución de variables descartadas y las variables descartadas se llaman a veces variables marginalizadas

Partiendo de una distribución bidimensional de frecuencias  $(x_i, y_j; n_{ij})$ , se denomina distribución condicionada de la variable X a un valor dado  $y_j$  de la variable Y a la distribución unidimensional definida por el conjunto de valores tomados por X y de las frecuencias condicionadas de dichos valores de X a qué Y tome el valor  $y_j$ . Análogamente, 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  $x_i$

## 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. <sup>20</sup> 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

## Características numéricas

Los sistemas de numeración que poseen una base tienen la característica de cumplir con la notación posicional, es decir, la posición de cada número le da un valor o peso, así el primer dígito de derecha a izquierda después del punto decimal, tiene un valor igual a  $b$  veces el valor del dígito, y así el dígito tiene en la posición  $n$  un valor igual a:  $(b^n) * A$  donde:

$b$  = valor de la base del sistema

$n$  = número del dígito o posición del mismo

$A$  = dígito.

# Regresión y correlación

El análisis de regresión es una herramienta de frecuente uso en Estadística que permite estudiar y valorar las relaciones entre diferentes variables cuantitativas tenidas en cuenta mediante la construcción de una ecuación

La correlación estadística determina la relación o dependencia que existe entre las dos variables que intervienen en una distribución bidimensional. Es decir, determinar si los cambios en una de las variables influyen en los cambios de la otra. En caso de que suceda, diremos que las variables están correlacionadas o que hay correlación entre ellas. En probabilidad y estadística, la correlación indica la fuerza y la dirección de una relación lineal y proporcionalidad entre dos variables estadísticas.

En probabilidad y estadística, la correlación indica la fuerza y la dirección de una relación lineal y proporcionalidad entre dos variables estadísticas. Se considera que dos variables cuantitativas están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto a los valores homónimos de la otra: si tenemos dos variables (A y B) existe correlación entre ellas si al disminuir los valores de A lo hacen también los de B y viceversa

# Definiciones

En estadística, el análisis de la regresión es un proceso estadístico para estimar las relaciones entre variables

En el análisis de regresión, también es de interés caracterizar la variación de la variable dependiente en torno a la función de regresión, la cual puede ser descrita por una distribución de probabilidad. Se denomina correlación al vínculo recíproco o correspondiente que existe entre dos o más elementos. El concepto se emplea de diferentes maneras de acuerdo al contexto. Correlación en el ámbito de las matemáticas y las estadísticas, la correlación alude a la proporcionalidad y la relación lineal que existe entre distintas variables.



## Curva de regresión y coeficiente de determinación

En estadística la regresión lineal o ajuste lineal es un modelo matemático usado para aproximar la relación de dependencia entre una variable dependiente  $Y$ , las variables independientes  $X_i$  y un término aleatorio.

En estadística, el coeficiente de determinación, denominado  $R^2$  y pronunciado  $R$  cuadrado, es un estadístico usado en el contexto de un modelo estadístico cuyo principal propósito es predecir futuros resultados o probar una hipótesis

En este caso, el  $R^2$  es simplemente el cuadrado del coeficiente de correlación de Pearson, lo cual es sólo cierto para la regresión lineal simple. Si existen varios resultados para una única variable, es decir, para una  $X$  existe una  $Y$ ,  $Z$ .... el coeficiente de determinación resulta del cuadrado del coeficiente de determinación múltiple. En ambos casos el  $R^2$  adquiere valores entre 0 y 1. Existen casos dentro de la definición computacional de  $R^2$  donde este valor puede tomar valores negativos.

# Regresión y correlación lineal

La correlación lineal y la regresión lineal simple son métodos estadísticos que estudian la relación lineal existente entre dos variables

- La correlación cuantifica como de relacionadas están dos variables, mientras que la regresión lineal consiste en generar una ecuación (modelo) que, basándose en la relación existente entre ambas variables, permita predecir el valor de una a partir de la otra
- El cálculo de la correlación entre dos variables es independiente del orden o asignación de cada variable a X e Y, mide únicamente la relación entre ambas sin considerar dependencias. En el caso de la regresión lineal, el modelo varía según qué variable se considere dependiente de la otra (lo cual no implica causa-efecto)

## Otros tipos de regresión

- Regresión simple: Cuando la variable Y depende únicamente de una única variable X.
- Regresión múltiple: Cuando la variable Y depende de varias variables ( $X_1, X_2, \dots, X_r$ ) En segundo lugar, en función del tipo de función  $f(X)$ :
- Regresión lineal: Cuando  $f(X)$  es una función lineal.
- Regresión no lineal: Cuando  $f(X)$  no es una función lineal. En tercer lugar, en función de la naturaleza de la relación que exista entre las dos variables:
  - La variable X puede ser la causa del valor de la variable Y. Por ejemplo, en toxicología, si  $X =$  Dosis de la droga e  $Y =$  Mortalidad, la mortalidad se atribuye a la dosis administrada y no a otras causas.

# 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.

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