



**Mi Universidad**

*Nombre del Alumno: Veronica Mariana Hernández Rincón*

*Nombre del tema: Unidad uno, estadística descriptiva*

*Parcial: Primer parcial*

*Nombre de la Materia: Bioestadística*

*Nombre del profesor: Rosario Gómez Lujano*

*Nombre de la Licenciatura: Enfermería*

*Cuatrimestre: Cuarto cuatrimestre*

Estadística descriptiva

Estadística en enfermería

Nos ayuda 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.

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.

En salud pública los componentes aleatorios se deben, entre otros aspectos, al conocimiento o a la imposibilidad de medir algunos determinantes de los estados de salud y enfermedad, así como a la variabilidad en las respuestas por los pacientes, similares entre sí, que son sometidos al mismo tratamiento.

Introducción histórica

El primer médico que utilizó métodos matemáticos para cuantificar variables de pacientes y sus enfermedades fue el francés Pierre Charles-Alexandre Louis, en su clásico estudio de la tuberculosis.

En Francia Louis René Villermé y en Inglaterra William Farr primeros mapas epidemiológicos usando métodos cuantitativos y análisis epidemiológicos.

William Heaton Hamer, propuso un modelo temporal discreto en un intento de explicar la ocurrencia regular de las epidemias de sarampión;

Pierre Simón Laplace, publicó en 1812 un tratado sobre la teoría analítica de las probabilidades.

John Brownlee , primer director del British Research Council, luchó durante veinte años con problemas de cuantificación de la infectividad epidemiológica.

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 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.

No todas las variables estadísticas son iguales y, por supuesto, no todas se pueden (en principio) expresar en forma de número.

Otra variable que podríamos encontrarnos es el color de ojos de una persona. Por ejemplo, Juan tiene los ojos verdes y Andrés los tiene azules. La variable sería el color de ojos y sería una variable cualitativa, no se expresa con número.

Bibliografía:

Estadística descriptiva

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.

**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.

Representaciones gráficas

Una gráfica o una representación gráfica o un gráfico, es un tipo de representación de datos, generalmente cuantitativos, mediante recursos visuales.

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.

Tipos de representación de graficas

**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.

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

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 presentan numéricamente las características de la distribución de un conjunto de datos o muestra.

Tabla de frecuencias con datos no agrupados.

Tabla de frecuencias con datos agrupados.

Bibliografía:

Estadística descriptiva

Tipos de frecuencia

**Frecuencia absoluta**

$$\sum_{i=1}^k ni = n1 + n2 + \dots + nk = N$$

**Frecuencia absoluta acumulada**

$$Ni = n1 + n2 + \dots + ni$$

**Frecuencia relativa**

$$fi = \frac{ni}{N}$$

**Frecuencia relativa acumulada**

$$Fi = \frac{Ni}{N}$$

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. Así, existen tres cuartiles. Los valores inferiores de la distribución se sitúan por debajo del primero (Q1).

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 D5 se corresponde con la mediana.

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

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.

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

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.

Bibliografía:

Estadística descriptiva

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.

- Coefficiente de asimetría de Fisher
- Asimétrica por la derecha.
- Asimétrica por la izquierda

- Coefficiente de curtosis o apuntamiento de Fisher
- 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.

Descripción numérica de una variable estadística bidimensional

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. Si de una cierta población se estudian dos caracteres simultáneamente se obtienen dos series de datos.

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).

Distribuciones marginales y condicionadas

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

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.

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.

Incorrelación  
Es el grado de dispersión entre los puntos de una variable. Características numéricas

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, etc.

Bibliografía:

Estadística descriptiva

Regresión y correlación

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.

El análisis de regresión ayuda a entender cómo el valor de la variable dependiente varía al cambiar el valor de una de las variables independientes, manteniendo el valor de las otras variables independientes fijas.

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.

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.

**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.

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.

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

Otros tipos de regresión

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

**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**  
**El coeficiente de determinación múltiple**

Bibliografía:

Análisis de atributos

Objetivo

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

- 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 graficas de atributos

Defectivos

- np - número de unidades no-conformes
- p - proporción de unidades no-conformes

Defectos

- c - número de defectos
- u - proporción de defectos

Bibliografía:

**Antología de bioestadística universidad del sur**

## Problemas

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.

$$\text{Media aritmética} = \frac{52+60+58+54+72+65+55+76}{8} = \frac{492}{8} = 61.5$$

$$\text{Mediana} = 52, 54, 55, 58, 60, 65, 72, 76 \quad \frac{58+60}{2} = \frac{118}{2} = 59$$

Moda= No tiene moda

$$\text{Rango} = 76-52 = 24$$

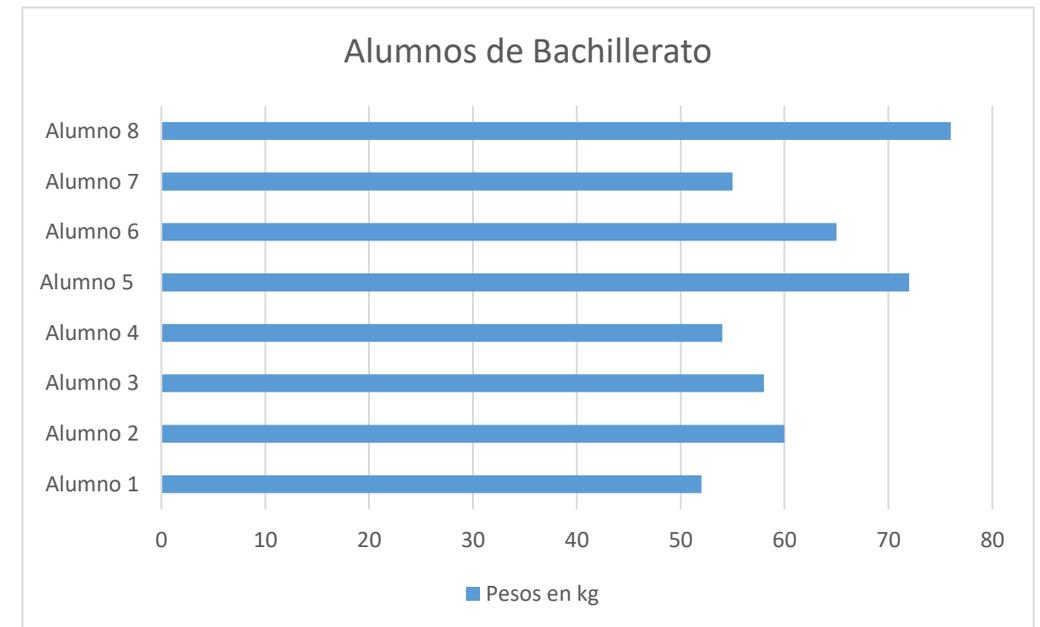
$$\text{Varianza} = \frac{(52-61.5)^2 + (54-61.5)^2 + (55-61.5)^2 + (58-61.5)^2 + (60-61.5)^2 + (65-61.5)^2 + (72-61.5)^2 + (76-61.5)^2}{7}$$

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

$$\text{Varianza} = \frac{90.25 + 56.25 + 42.25 + 12.25 + 2.25 + 12.25 + 110.25 + 210.25}{7}$$

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

$$\text{Desviación estándar} = \sqrt{76.57} = 8.7$$



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.

Datos	F	Fa	Fr	Fra	Fr%
106	3	3	0.15	0.15	15%
109	5	8	0.25	0.4	25%
110	7	15	0.35	0.75	35%
119	2	17	0.1	0.85	10%
124	3	20	0.15	1	15%
<b>Total</b>	<b>20</b>		<b>1</b>		<b>100%</b>

