



Nombre del Alumno : Cecilia Guadalupe Gómez Morales

Nombre del tema : Estadística Descriptiva

Parcial : 1ª

Nombre de la Materia : Bioestadística

Nombre del profesor : Roberto Gómez Lujano

Nombre de la Licenciatura: lic . Enfermería

Cuatrimestre : 4ª

Lugar y Fecha de elaboración: Pichucalco Chiapas

A 10 de Noviembre del 2024

UNIDAD I. ESTADISTUCA DESCRIPTIVA

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, 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.

La Bioestadística ha debido ampliar su campo para, de esta manera, incluir cualquier modelo cuantitativo,

La principal ventaja del pensamiento estadístico interviniendo en las ciencias de la vida es que no solo resuelve, sino que también comprende una compleja metodología para dar respuesta a las hipótesis, además de agilizar la cuestión de organización del sistema de investigación

La extensión de los conocimientos y aptitudes de carácter estadístico que necesitan adquirir los profesionales de la salud pública son importantes, porque el conocimiento de los principios y métodos estadísticos

INTRODUCCION HISTORICA

El primer médico que utilizó métodos matemáticos para cuantificar variables de pacientes y sus enfermedades fue el francés Pierre Charles-Alexandre Louis (1787-1872).

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.

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.

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.

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.

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

Algunas de las aportaciones más importantes de la bioestadística se han dado en el estudio de las enfermedades.

De ese modo, ha permitido resolver y optimizar la metodología para dar respuesta a las diversas hipótesis que se manejan en el mundo de las ciencias de la vida.

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 (en principio) expresar en forma de número.

La variable sería el color de ojos y sería una variable cualitativa. Es decir, no se expresa con número.

Definiciones Básicas

Variable estadística: Una variable estadística es una característica de una muestra o población de datos que puede adoptar diferentes valores.

Además, cada una de estas variables podría tener más subtipos, ya que tenemos variables de tipo económico, categóricas, dicotómicas, dependientes, independientes. Es decir, como ya hemos dicho, muchos tipos de variables estadísticas. Por ejemplo, podríamos tener una variable estadística de tipo cuantitativo, discreta y dependiente.

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.

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:

- *Diagramas de barras
- *Histogramas
- *Polígonos de frecuencia
- *Pictograma
- *Cartogramas

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.

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.

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.

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.

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 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.

Tipos de Frecuencias

Existen cuatro tipos de frecuencias:

Frecuencia absoluta La frecuencia absoluta (n_i) de un valor X_i es el número de veces que el valor está en el conjunto (X_1, X_2, \dots, X_N). La suma de las frecuencias absolutas de todos los elementos diferentes del conjunto debe ser el número total de sujetos N . Si el conjunto tiene k números (o categorías) diferentes, entonces:

$$\sum_{i=1}^k n_i = n_1 + n_2 + \dots + n_k = N$$

Frecuencia absoluta acumulada La frecuencia absoluta acumulada (N_i) de un valor X_i del conjunto (X_1, X_2, \dots, X_N) es la suma de las frecuencias absolutas de los valores menores o iguales a X_i , es decir:

$$N_i = n_1 + n_2 + \dots + n_i$$

Frecuencia relativa La frecuencia relativa (f_i) de un valor X_i es la proporción de valores iguales a X_i en el conjunto de datos (X_1, X_2, \dots, X_N). Es decir, la frecuencia relativa es la frecuencia absoluta dividida por el número total de elementos N :

$$f_i = \frac{n_i}{N}$$

siendo (X_1, X_2, \dots, X_N) el conjunto de datos y n_i el total de valores igual a X_i

Frecuencia relativa acumulada Definimos la frecuencia relativa acumulada (F_i) de un valor X_i como la proporción de valores iguales o menores a X_i en el conjunto de datos (X_1, X_2, \dots, X_N). Es decir, la frecuencia relativa acumulada es la frecuencia absoluta acumulada dividida por el número total de sujetos N : La frecuencia relativa acumulada de cada valor siempre es mayor que la frecuencia relativa.

$$F_i = \frac{N_i}{N}$$

siendo (X_1, X_2, \dots, X_N) el conjunto de datos y N_i el total de valores igual o menor a X_i

Características 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 (Q_1). La mitad o mediana son los valores menores iguales al cuartil dos (Q_2) y los superiores son representados por el cuartil tres (Q_3).

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 D_1 a D_9 . El D_5 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.

<i>Cuatriles</i>	<i>Deciles</i>	<i>Percentiles</i>
<i>k.N</i>	<i>k.N</i>	<i>k.N</i>
4	10	100

Descripción numérica de una variable estadística bidimensional.

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

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.

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.

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.

En teoría de probabilidades, la distribución marginal es la distribución de probabilidad de un subconjunto de variables aleatorias de un conjunto de variables aleatorias.

Esto contrasta con la distribución condicional, que proporciona probabilidades contingentes sobre el valor conocido de otras variables. El término variable marginal se usa para referirse a una variable del subconjunto de retenido y cuyos valores pueden ser conocidos.

Partiendo de una distribución bidimensional de frecuencias (xi, yj; nij), se denomina distribución condicionada de la variable X a un valor dado yj 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 yj.

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.

Ejemplo :

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

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

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

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

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.

$P(B/A) = P(B)$ 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. Ejemplo: la probabilidad de que al tirar una moneda salga cara (suceso B), condicionada a que haga buen tiempo (suceso A), es igual a la propia probabilidad del suceso B.

$P(A/B) = P(A)$ es decir, que la probabilidad de que se dé el suceso A, condicionada a que previamente se haya dado el suceso B, es exactamente igual a la probabilidad de A.

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.

b = valor de la base del sistema

n = número del dígito o posición del mismo

A = dígito.

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

Más comúnmente, el análisis de regresión estima la esperanza condicional de la variable dependiente dadas las variables independientes - es decir, el valor promedio de la variable dependiente cuando se fijan las variables independientes. Con menor frecuencia, la atención se centra en un cuantil, u otro parámetro de localización de la distribución condicional de la variable dependiente dadas las 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.

Otros tipos de Regresión

Regresión Múltiple: Este tipo se presenta cuando dos o más variables independientes influyen sobre una variable dependiente. Ejemplo: $Y = f(x, w, z)$.

Para poder resolver se puede utilizar programas como AD+, SPSS y Minitab y Excel. 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 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.

En los resultados de Excel se llama error típico y para explicar la relación del aprendizaje de PHP que se viene desarrollando es de 0.861

Conclusiones El 69.70% del aprendizaje del Curso de PHP puede ser explicado mediante las notas obtenidas por las asignaturas de Algoritmos, Base de Datos y Programación.

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

.Obtener .Media aritmética, mediana, moda, rango, varianza y desviación estándar de los siguientes datos no agrupados. 6, 7, 8, 9, 10, 8

Medida aritmética: $6+7+8+9+10+8= 48/6 =8$

Mediana: 6,7,8,8,9,10= $8+8/2 =8$

Moda: 8

Rango: 6,7,8,8,9,10= $10-6=4$

Varianza: $(6-8)^2+(7-8)^2+(8-8)^2+(9-8)^2+(10-8)^2/6-14+1+0+0+1+4/5=10/5=2$

Desviación estándar: $\sqrt{2} = 1.41421$

Media aritmética

Es el promedio de un conjunto de valores, calculado sumando todos los valores y dividiendo por el número de observaciones.

La Mediana

Es el valor que se encuentra justo en el medio cuando los datos se ordenan de menor a mayor.

La Moda se define como el valor que tiene mayor frecuencia absoluta.

Rango en estadística

El rango es una herramienta que nos ayuda a entender cuánto varían los datos dentro de un conjunto.

Se calcula restandole al valor más grande del conjunto el valor más pequeño.

Varianza

La varianza es una forma de entender cuánto se dispersan o varían los datos alrededor de su media.

Cuanto más se desvíe un valor de su media, más alta será la varianza, y viceversa.

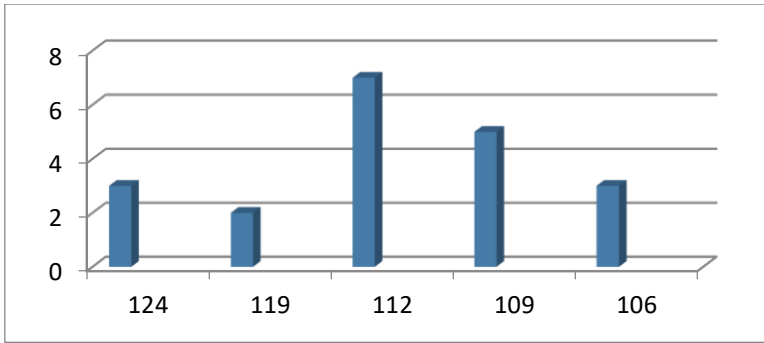
DESVIACIÓN ESTÁNDAR

La desviación estándar es un índice que muestra cuánto se diferencian una puntuación común y la media de un grupo de puntuaciones.

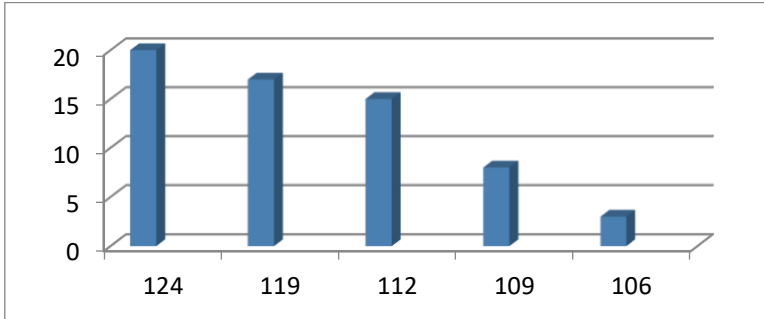
Para obtenerla, calculamos la desviación de las puntuaciones respecto a la media y luego la elevamos al cuadrado. Las sumamos y promediamos; y, finalmente, le sacamos la raíz cuadrada.

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, 112, 109, 112, 124, 109, 109, 109, 106, 124, 112, 112, 106.

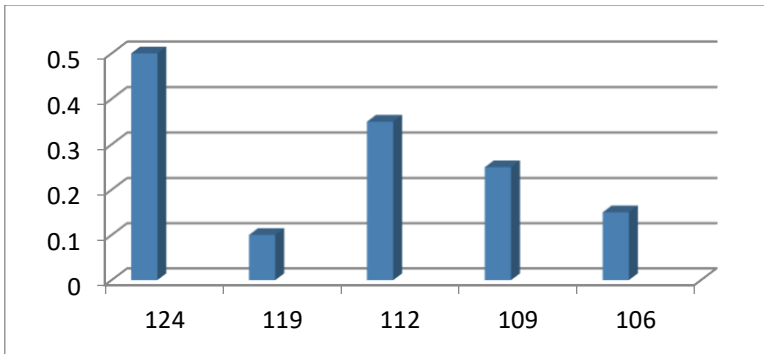
Valor	Frecuencia .Absoluta	Frecuencia .Absoluta acumulada	Frecuencia. Relativa	Frecuencia .Relativa acumulada
106	4	3	0.15	0.15
109	5	8	0.25	0.4
112	7	15	0.35	0.75
119	2	17	0.1	0.85
124	3	20	0.15	1



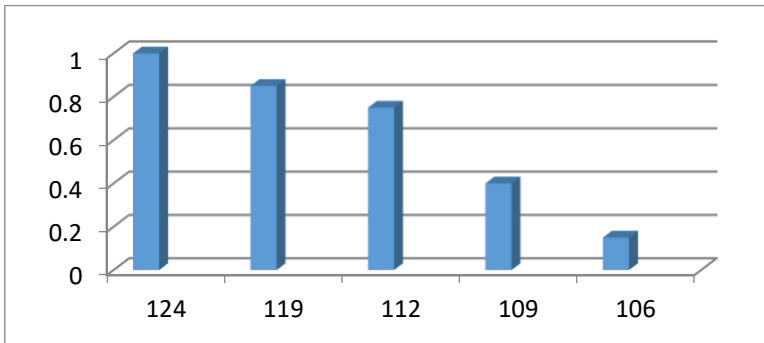
F. Absolutas de Valor



F. Absoluta Acumulada



F. Relativa



F. Relativa Acumulada

Escriba aquí la ecuación.