



Mi Universidad

**Mapa Conceptual y
Cuadro Sinóptico**

Nombre del Alumno: Eleazar Lázaro Alvarado

Nombre del profesor: Rosario Gómez Lujano

Nombre de la Materia: Bioestadística

*Nombre del tema: Estadística Descriptiva y Cálculo de
Probabilidades*

Nombre de la Licenciatura: Enfermería

Parcial: 1er Módulo

Cuatrimestre: 4to Cuatrimestre

Estadística 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.

Ventaja

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.

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. 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.

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.

Diversos campos

Esta disciplina es usada en diversos campos de la medicina y la salud pública, como la epidemiología, nutrición y salud ambiental.

Aplicación de métodos

Son aplicados en estudios relacionados con la ecología y la genómica.

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.

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.

Ejemplos:

La altura de Juan es de 180 centímetros. La variable estadística es la altura y está medida en centímetros.

También podríamos, por ejemplo, decir que el beneficio de una empresa ha sido de 22.300 dólares el último año. En este caso, la variable sería el beneficio y estaría medido en dólares.

Ambas variables son del tipo cuantitativo (se expresan con un número).

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.

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.

Representación gráfica

Definición

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

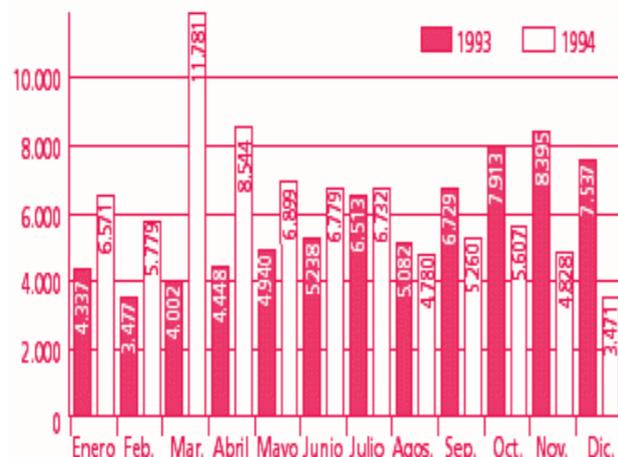
Diagramas de barras

Gráficos de sectores

Cartogramas

Pirámides de población

Ejemplo de diagrama de barras:



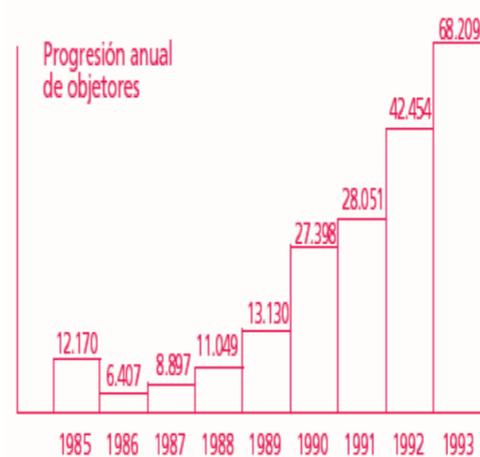
Diagramas de barras

Polígonos de frecuencia

Pictogramas

Ojiva

Ejemplo de histograma:



Representación numérica

Tabla de frecuencia

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 tablas

Tabla de frecuencias con datos no agrupados.

Tabla de frecuencias con datos agrupados.

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 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 .

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 .

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 .

Estadística Descriptiva

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

El quintil

El decil

El percentil

Medidas de posición central

Estas nos permiten resumir la distribución de los datos en un solo valor central, alrededor del cual se sitúan.

La media aritmética, geométrica o armónica

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.

La moda nos informa sobre aquellos que se repiten en más ocasiones.

Medidas de dispersión:

Varianza

Desviación estándar

Descripción numérica de una variable estadística bidimensional

Definición

Variable estadística bidimensional es el conjunto de pares de valores de dos caracteres o variables estadísticas unidimensionales X e Y sobre una misma

Distribuciones bidimensionales:

Son 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

Definición

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.

La función es:

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.

La distribución marginal de dos variables aleatorias se puede obtener a partir de su distribución conjunta.

Independencia e incorrelación

Definición

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.

Incorrelación

Es el grado de dispersión entre los puntos de una variable, es decir, el cuándo los puntos no marchan en una misma dirección si no que están dispersos por todos lados, a diferencia de la correlación que es todo lo contrario.

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.

Los cuatro primeros se caracterizan por tener una base mientras que el sistema romano no posee base y resulta más complicado su manejo tanto con números, así como en las operaciones básicas.

Estadística Descriptiva

Regresión y correlación

Definición

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.

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 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. Entre las ventajas por la que sobresale el coeficiente de correlación respecto a otras formas de medición de correlación, es la covarianza.

Otros tipos de regresión

Regresión Múltiple

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

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.

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 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.

Cálculo de Probabilidades

La teoría de la probabilidad pretende ser una herramienta para modelizar y tratar con situaciones de este tipo. El objetivo del Cálculo de Probabilidades es el estudio de métodos de análisis del comportamiento de fenómenos aleatorios.

La medida de probabilidad. Espacio Probabilístico

El conjunto muestral es un conjunto exhaustivo (contiene todas las posibles ocurrencias) y mutuamente exclusivo (no pueden darse dos ocurrencias a la vez).

Axioma 1

Un experimento se denomina aleatorio cuando puede dar resultados distintos al realizarse en las mismas condiciones (por ejemplo, lanzar un dado al aire y observar el número resultante).

Axioma 2

Es una fórmula de agregación que se usa para calcular la probabilidad de la unión de subconjuntos disjuntos. Establece que la incertidumbre de un cierto subconjunto es la suma de las incertidumbres de sus partes (disjuntas). Nótese que esta propiedad también se cumple para sucesiones finitas.

Probabilidad Condicionada

Miraremos la forma en que cambia la probabilidad de un suceso "A" cuando se sabe que otro suceso "B" ha ocurrido.

Probabilidad condicional para sucesos independientes

Dos sucesos A y B son independientes cuando la probabilidad de que suceda A no se ve afectada porque haya sucedido, o no, B. Por ejemplo, Si tiramos dos veces una moneda, el segundo resultado que obtenemos no está influenciado por el primer resultado obtenido.

Sucesos dependientes

Dos sucesos A y B son dependientes cuando la probabilidad de que suceda A se ve afectada porque haya sucedido, o no, B.

Teoremas Asociados

El teorema de Bayes es utilizado para calcular la probabilidad de un suceso, teniendo información de antemano sobre ese suceso.

Fórmula del teorema de Bayes

Para calcular la probabilidad tal como la definió Bayes en este tipo de sucesos, necesitamos una fórmula. La fórmula se define matemáticamente como:

$$P[A_n/B] = \frac{P[B/A_n] \cdot P[A_n]}{\sum P[B/A_i] \cdot P[A_i]}$$

Cálculo de Probabilidades

Variable Aleatoria

Se llama variable aleatoria a toda función que asocia a cada elemento del espacio muestral (E) un número real. Se utilizan letras mayúsculas (X, Y), para designar variables aleatorias, y las respectivas minúsculas (x, y), para designar valores concretos de las mismas.

Variable aleatoria discreta

Es discreta si los números a los que da lugar son números enteros. La forma de calcular las probabilidades de una variable aleatoria discreta es a través de la función de probabilidad.

Variable aleatoria continua

Es continua en caso de que los números a los que dé lugar no sean números enteros. Es decir, tengan decimales. La probabilidad de que se dé un suceso determinado correspondiente a una variable aleatoria continua, viene establecida por la función de densidad.

Ejemplo de variable aleatoria

Dado: No es la variable aleatoria. El dado es simplemente un objeto.

Lanzamiento de un dado: No es la variable aleatoria. El lanzamiento de un dado es el experimento aleatorio.

Resultados del lanzamiento de un dado: Sí es la variable aleatoria. Es la función que recoge los resultados del lanzamiento del dado.

Concepto de variable aleatoria

Se denomina variable aleatoria a la función que adjudica eventos posibles a números reales cuyos valores se miden en experimentos de tipo aleatorio. Estos valores posibles representan los resultados de experimentos que todavía no se llevaron a cabo o cantidades inciertas.

Permite:

Ofrecer una descripción de la probabilidad de que se adoptan ciertos valores.

Qué valor adoptará la variable cuando sea determinada o medida, pero sí se puede conocer cómo se distribuyen las probabilidades vinculadas a los valores posibles.

Función de distribución

La Función de Distribución Acumulada es una función matemática de la variable real: x (minúscula); que describe la probabilidad de que X tenga un valor menor o igual que x .

La FDA sería la función con la recta real como dominio, con imagen del área hasta aquí de la función f , siendo aquí el valor x para la variable aleatoria real X .

Variables aleatorias discretas y continuas

Una variable aleatoria es una función que asigna un valor numérico, al resultado de un experimento aleatorio. Una variable aleatoria puede ser discreta o continua.

Las variables aleatorias discretas son aquellas que presentan un número contable de valores; por ejemplo, el número de personas que viven en una casa (3, 5 o 9).

Las variables aleatorias continuas son aquellas que presentan un número incontable de valores; por ejemplo, el peso de las vacas en una granja (una vaca puede pesar 632.12 kg, otra 253.12012 kg, y nunca terminaríamos de enumerar todos los posibles valores).

En general:

Las variables aleatorias discretas representan datos que provienen del conteo del número de elementos, mientras que, las variables aleatorias continuas representan datos que provienen de mediciones, por ejemplo, tiempo, peso, longitud, etc.

Cálculo de Probabilidades

Características de una variable

- 1.- Están contenidas esencialmente en el título, el problema, el objetivo y las respectivas hipótesis de la investigación. En virtud de ello es que no se puede agregar nuevas variables de las que ya existen en los ítems mencionados.
- 2.- Son aspectos que cambian o adoptan distintos valores. Esto significa que las variables al ser medidas y observadas expresan diferencias entre los rasgos, cualidades y atributos de las unidades de análisis.

- 3.- Son enunciados que expresan rasgos característicos de los problemas medibles empíricamente. Estas variables en la práctica social pueden ser medidas y observadas con instrumentos convencionales
- 4.- Son susceptibles de descomposición empírica. Dicho de otro término, que las variables pueden desagregarse en indicadores, índices, subíndices e ítems.

Esperanza de una variable aleatoria

La esperanza matemática de una variable aleatoria es una característica numérica que proporciona una idea de la localización de la variable aleatoria sobre la recta real. Decimos que es un parámetro de centralización o de localización.

Corresponde

Con el valor medio teórico de los posibles valores que pueda tomar la variable aleatoria, o también con el centro de gravedad de los valores de la variable supuesto que cada valor tuviera una masa proporcional a la función de densidad en ellos.

Momentos de una variable aleatoria

Entre las distintas características de una distribución ocupan un importante lugar los momentos

- Momentos no centrados
- Momentos centrados en media

Los momentos centrados se calculan, como los no centrados, teniendo en cuenta la definición de esperanza de una función de una variable aleatoria.

Funciones asociadas a una variable aleatoria

Una función que asocia un número real, perfectamente definido, a cada punto muestral. A veces las variables aleatorias (v.a.) están ya implícitas en los puntos muestrales.

Funciones de densidad discreta y continua

Ambas son las funciones que acumuladas (en forma de sumatorio en el caso discreto o en forma de integral en el caso continuo) dan como resultado la función de distribución.

Diferencias

- La función de densidad discreta toma valores positivos únicamente en los puntos del recorrido y se interpreta como la probabilidad de la que la variable tome ese valor $f(x) = P(X = x)$.
- La función de densidad continua toma valores en el conjunto de números reales y no se interpreta como una probabilidad. No está acotada por 1, puede tomar cualquier valor positivo.

Resuelve el siguiente ejercicio.

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.

Construye una distribución de frecuencia que muestre frecuencia absoluta, frecuencia absoluta acumulada, frecuencia relativa y frecuencia relativa acumulada.

Construye una gráfica de barras con los datos anteriores.

Datos	Frecuencia Absoluta	Frecuencia Absoluta Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada
106	3	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
Total	20			

