



Universidad del sureste

súper nota

Estadística

Bryant reyes robles

22/ de mayo / 24



I.- LA ESTADÍSTICA DESCRIPTIVA

La estadística descriptiva es, junto con la inferencia estadística o estadística inferencial, una de las dos grandes ramas de la estadística. Su propio nombre lo indica, trata de describir algo. Pero no describirlo de cualquiera forma, sino de manera cuantitativa. Pensemos en el peso de una caja de verduras, en la altura de una persona o en la cantidad de dinero que gana una empresa. De estas variables podríamos decir muchas cosas. Por ejemplo, podríamos indicar que esta o aquella caja de tomates pesan mucho o pesan menos que otras. Siguiendo con otro ejemplo, podríamos decir que el ingreso de una empresa varía mucho a lo largo del tiempo o que una persona tiene una altura promedio.



1.1.- Tipos de variables

Dentro de la estadística descriptiva, podemos describir los datos de manera cualitativa o cuantitativa.

Variable cualitativa: Hace referencia a una cualidad. Ejemplos: el color de ojos de una persona o el color de pelo.

Variable cuantitativa: Hace referencia a una medida cuantitativa. Ejemplos: la altura de una persona en centímetros o el peso de una persona en kilogramos.

Así pues, sobre estas variables se pueden calcular ciertos parámetros.

Especialmente sobre las variables cuantitativas. Ya que, por ejemplo, ¿cuál es el valor promedio del color de ojos? Si hay cinco personas con color de ojos azul y cinco con color de ojos verde, el promedio no será que tienen un color de ojos promedio de azul-verde. Por tanto, en ese caso no sería posible calcular algunos de los parámetros que veremos a continuación.

Variable estadística

Parámetros estadísticos básicos

Con el objetivo de resumir la información, se idearon diversas fórmulas que ofrecían medidas de un determinado tipo. Así, están aquellas que nos ofrecen información sobre el centro, otras sobre la dispersión o variabilidad y otras sobre la posición de un valor.

Medidas de tendencia central: Denominadas así porque ofrecen información sobre el centro de conjunto de datos. Por ejemplo, la media es una medida de tendencia o posición central ya que el promedio nos ofrece un valor centrado del conjunto de datos. ¿Dónde podríamos decir que se encuentra el punto medio? En el centro, en la mitad aproximadamente. Otro ejemplo de medida de tendencia central es la mediana.

Medidas de dispersión: También son conocidas como medidas de variabilidad. Por ejemplo, la desviación típica es una medida de variabilidad ya que nos dice si los valores de un conjunto de datos son muy dispares o no. Dos ejemplos más sobre medidas de dispersión podrían ser la varianza y el rango estadístico.



1.2.- CONCEPTOS BÁSICOS ESTADÍSTICA

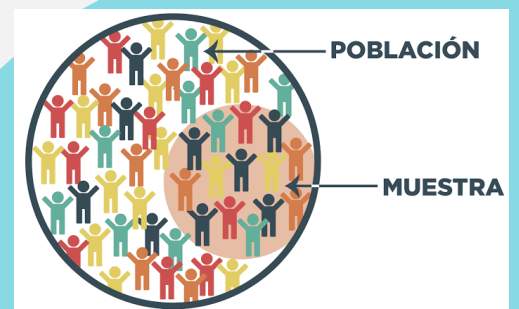
Universo:

En estadística es el nombre específico que recibe particularmente en la investigación social la operación dentro de la delimitación del campo de investigación que tienen por objeto la determinación del conjunto de unidades de observaciones del conjunto de unidades de observación que van a ser investigadas. Para muchos investigadores el término universo y población son sinónimos. En general, el universo es la totalidad de elementos o características que conforman el ámbito de un estudio o investigación.



POBLACION

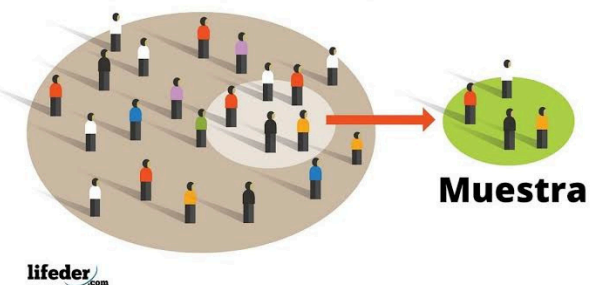
En estadística el concepto de población va más allá de lo que comúnmente se conoce como tal. En términos estadísticos, población es un conjunto finito o infinito de personas, animales o cosas que presentan características comunes, sobre los cuales se quiere efectuar un estudio determinado. En otras palabras, la población se define como la totalidad de los valores posibles (mediciones o conteos) de una característica particular de un grupo especificado de personas, animales o cosas que se desean estudiar en un momento determinado. Así, se puede hablar de la población de habitantes de un país, de la población de estudiantes universitarios de la zona sur del Estado Anzoátegui, de la población de casas de la Urbanización Los Ríos de la ciudad de El Tigre, el rendimiento académico de los estudiantes del IUTJAA, el número de carros marca Corola de la ciudad de El Tigre, la estatura de un grupo de alumnos del IUTJAA, la talla, etc.



MUESTRA

La muestra es un subconjunto de la población, seleccionado de tal forma, que sea representativo de la población en estudio, obteniéndose con el fin de investigar alguna o algunas de las propiedades de la población de la cual procede. En otras palabras es una parte de la población que sirve para representarla. Según el DRAE, es una parte o porción extraída de un conjunto por métodos que permiten considerarla como representativa del mismo.

Población estadística



lifeder.com



MUESTREO

Es el procedimiento mediante el cual se obtiene una o más muestras de una población determinada. Existen dos tipos de muestreos a saber:



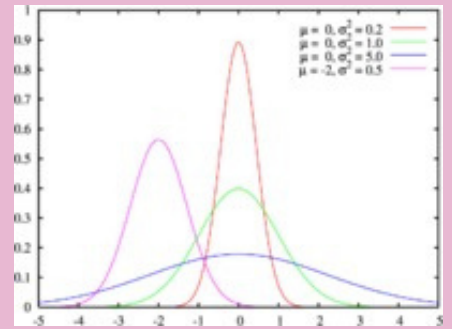
Muestreo Probabilístico

QuestionPro

LOS PARAMETROS

Son cualquiera característica que se pueda medir y cuya medición se lleve a cabo sobre todos los elementos que integran una población determinada, los mismos suelen representarse con letras griegas. El valor de un parámetro poblacional es un valor fijo en un momento dado.

Ejemplo: La media Aritmética = μ (miu), La desviación Típica = σ , (Sigma) etcétera.



DATO ESTADISTICO

Es un conjunto de valores numéricos que tienen relación significativa entre sí. Los mismos pueden ser comparados, analizados e interpretados en una investigación cualquiera. Se puede afirmar que son las expresiones numéricas obtenidas como consecuencia de observar un individuo de la población; por lo tanto, son las características que se han tomado en cuenta de cualquiera población para una investigación determinada.



FRECUENCIA

La frecuencia es el número de veces que se repite (aparece) el mismo dato estadístico en un conjunto de observaciones de una investigación determinada, las frecuencias se les designan con las letras f_i , y por lo general se les llaman frecuencias absolutas.



DISTRIBUCION DE FRECUENCIAS

En estadística existe una relación con cantidades, números agrupados o no, los cuales poseen entre sí características similares. Existen investigaciones relacionadas con los precios de los productos de la dieta diaria, la estatura y el peso de un grupo de individuos, los salarios de los empleados, los grados de temperatura del medio ambiente, las calificaciones de los estudiantes, etc., que pueden adquirir diferentes valores gracias a una unidad apropiada, que recibe el nombre de variable. La representación numérica de las variables se denomina dato estadístico. La distribución de frecuencia es una disposición tabular de datos estadísticos, ordenados ascendente o descendientemente, con la frecuencia (f_i) de cada dato. Las distribuciones de frecuencias pueden ser para datos no agrupados y para datos agrupados o de intervalos de clase.



DISTRIBUCION DE FRECUENCIAS PARA DATOS NO AGRUPADOS

Es aquella distribución que indica las frecuencias con que aparecen los datos estadísticos, desde el menor de ellos hasta el mayor de ese conjunto sin que se haya hecho ninguna modificación al tamaño de las unidades originales. En estas distribuciones cada dato mantiene su propia identidad después que la distribución de frecuencia se ha elaborado. En estas distribuciones los valores de cada variable han sido solamente reagrupados, siguiendo un orden lógico con sus respectivas frecuencias.



shutterstock.com · 261177431

DISTRIBUCIÓN DE FRECUENCIA DE CLASE O DE DATOS AGRUPADOS:

Es aquella distribución en la que las disposiciones tabulares de los datos estadísticos se encuentran ordenados en clases y con la frecuencia de cada clase; es decir, los datos originales de varios valores adyacentes del conjunto se combinan para formar un intervalo de clase.



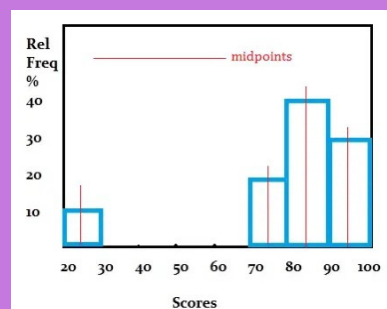
1. - AMPLITUD DE CLASE, LONGITUD O ANCHO DE UNA CLASE

1. La amplitud o longitud de una clase es el número de valores o variables que concurren a una clase determinada. La amplitud de clase se designa con las letras l_c . Existen diversos criterios para determinar la amplitud de clases, ante esa diversidad de criterios, se ha considerado que lo más importante es dar un ancho o longitud de clase a todos los intervalos de tal manera que respondan a la naturaleza de los datos y al objetivo que se persigue y esto se logra con la práctica.



1. - PUNTO MEDIO O MARCA DE CLASE

El centro de la clase, es el valor de los datos que se ubica en la posición central de la clase y representa todos los demás valores de esa clase. Este valor se utiliza para el cálculo de la media aritmética.



5.- FRECUENCIA DE CLASE

La frecuencia de clase se le denomina frecuencia absoluta y se le designa con las letras f_i . Es el número total de valores de las variables que se encuentran presente en una clase determinada, de una distribución de frecuencia de clase.



6.- FRECUENCIA RELATIVA

La frecuencia relativa es aquella que resulta de dividir cada uno de los f_i de las clases de una distribución de frecuencia de clase entre el número total de datos (N) de la serie de valores. Estas frecuencias se designan con las letras fr ; si cada fr se multiplica por 100 se obtiene la frecuencia relativa porcentual ($fr \%$).

X_i	Frecuencia absoluta (n_i)	Frecuencia relativa ($f_i = n_i/N$)	Frecuencia relativa ($f_i = n_i/N$) en %
3	2	0,07	7%
4	4	0,13	13%
5	6	0,20	20%
6	7	0,23	23%
7	5	0,17	17%
8	3	0,10	10%
9	2	0,07	7%
10	1	0,03	3%

7.- FRECUENCIAS ACUMULADAS

Las frecuencias acumuladas de una distribución de frecuencias son aquellas que se obtienen de las sumas sucesivas de las f_i que integran cada una de las clases de una distribución de frecuencia de clase, esto se logra cuando la acumulación de las frecuencias se realiza tomando en cuenta la primera clase hasta alcanzar la última.

X_i	Frecuencia absoluta (n_i)	Frecuencia absoluta acumulada (N_i)
3	2	2
4	4	6
5	6	12
6	7	19
7	5	24
8	3	27
9	2	29
10	1	30
Total	30	30

8.- FRECUENCIA ACUMULADA RELATIVA

La frecuencia acumulada relativa es aquella que resulta de dividir cada una de las fa de las diferentes clases que integran una distribución de frecuencia de clase entre el número total de datos (N) de la serie de valores, estas frecuencias se designan con las letras far. Si las frace multiplican por 100 se obtienen las frecuencias acumuladas relativas porcentuales y las mismas se designan así: far %.

X_i	Frecuencia absoluta (n_i)	Frecuencia relativa ($f_i = n_i/N$)	Frecuencia relativa ($f_i = n_i/N$) en %
3	2	0,07	7%
4	4	0,13	13%
5	6	0,20	20%
6	7	0,23	23%
7	5	0,17	17%
8	3	0,10	10%
9	2	0,07	7%
10	1	0,03	3%
Total	30	1	100%

LA MEDIANA

La mediana (Md) es una medida de posición que divide a la serie de valores en dos partes iguales, un cincuenta por ciento que es mayor o igual a esta y otro cincuenta por ciento que es menor o igual que ella. Es por lo tanto, un parámetro que esta en el medio del ordenamiento o arreglo de los datos organizados, entonces, la mediana divide la distribución en una forma tal que a cada lado de la misma queda un número igual de datos.

La Media	La Mediana	La Moda
Es el valor promedio de un conjunto de datos. Cómo se calcula: Suma todos los valores y divide por la cantidad de datos.	Es el valor que se encuentra justo en el medio cuando los datos se ordenan de menor a mayor. Cómo se Calcula: Ordena los datos y selecciona el valor central.	Es el valor que aparece con mayor frecuencia en un conjunto de datos. Cómo se Calcula: Identifica el número que se repite más veces.
Ejemplo $(10 + 15 + 20 + 25) / 4 = 17,5$	Ejemplo Para 5, 8, 10, 12, 15, la mediana es 10.	Ejemplo En 2, 3, 3, 4, 5, 5, 5, la moda es 5.

LA MODA

La moda es la medida de posición que indica la magnitud del valor que se presenta con más frecuencia en una serie de datos; es pues, el valor de la variable que más se repite en un conjunto de datos. De las medias de posición la moda es la que se determina con mayor facilidad, ya que se puede obtener por una simple observación de los datos en estudio, puesto que la moda es el dato que se observa con mayor frecuencia. La moda se designa con las letras Mo.

La Media	La Mediana	La Moda
Es el valor promedio de un conjunto de datos. Cómo se calcula: Suma todos los valores y divide por la cantidad de datos.	Es el valor que se encuentra justo en el medio cuando los datos se ordenan de menor a mayor. Cómo se Calcula: Ordena los datos y selecciona el valor central.	Es el valor que aparece con mayor frecuencia en un conjunto de datos. Cómo se Calcula: Identifica el número que se repite más veces.
Ejemplo $(10 + 15 + 20 + 25) / 4 = 17,5$	Ejemplo Para 5, 8, 10, 12, 15, la mediana es 10.	Ejemplo En 2, 3, 3, 4, 5, 5, 5, la moda es 5.

DESVIACIÓN TÍPICA O ESTÁNDAR

Es la medida de dispersión más utilizada en las investigaciones por ser la más estable de todas, ya que para su cálculo se utilizan todos los desvíos con respecto a la media aritmética de las observaciones, y además, se toman en cuenta los signos de esos desvíos.

QuestionPro
Fórmula para calcular la desviación estándar

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

VARIANZA

Es otra de las variaciones absolutas y la misma se define como el cuadrado de la desviación típica; viene expresada con las mismas letras de la desviación típica pero elevada al cuadrado, así s^2 y s^2 . Las fórmulas para calcular la varianza son las mismas utilizadas por la desviación típica, exceptuando las respectivas raíces, las cuales desaparecen al estar elevados el primer miembro al cuadrado

QuestionPro
Varianza de una muestra (s^2)

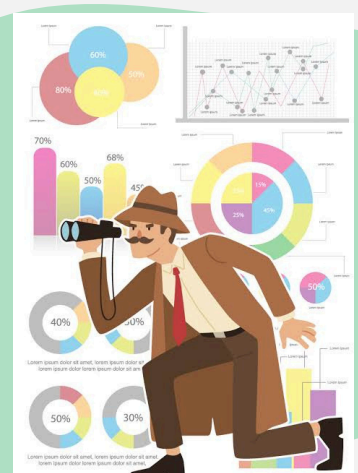
FÓRMULA:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

- s^2 = Varianza
- x_i = Término de conjunto de datos
- \bar{x} = Medida de la muestra
- \sum = Sumatoria
- n = Tamaño de la muestra

LA ESTADÍSTICA DENTRO DEL MÉTODO CIENTÍFICO

La estadística no se puede utilizar como una caja mágica para extraer certezas, donde se introducen datos y se extraen leyes. La estadística, en el contexto de probabilidades y técnicas de inferencia, es incapaz por sí misma de suplantar al Método Científico, sólo es un gran apoyo.



Bibliografía

<https://plataformaeducativauds.com.mx/assets/docs/libro/LNU/2e38faf807e4310316facdc1b7d23494-LC-LNU302%20ESTADISTICA%20DESCRIPATIVA%20EN%20NUTRICION.pdf>