



Nombre de alumnos:

María Elena Hernández Ballona

Nombre del profesor:

Rosario Gómez Lujano

Nombre del trabajo:

Ensayo “Estadística”

Materia:

Estadística

Grado:

1er Cuatrimestre

Grupo: “A”

ESTADÍSTICA

La psicología es la ciencia que estudia la conducta humana y los procesos mentales. Como todas las ciencias, la psicología se basa también de las matemáticas, hace uso de ellas para lograr sus objetivos de predecir y controlar la conducta humana. Para este fin es importante el uso de la probabilidad, que permite a los científicos tener una mayor certeza a la hora de prever cómo actúan las personas.

Como parte de las matemáticas, la estadística tiene mucho que ver con la psicología, pero ¿qué es la estadística?, la estadística es la ciencia que se encarga de recopilar, organizar, procesar, analizar e interpretar datos con el fin de deducir las características de un grupo población o objetivo. Dentro de esta rama de la matemática, se encuentran las gráficas y encuestas.

La estadística se clasifica en 4 tipos: la descriptiva o deductiva, que se encarga de mostrar el resultado de los datos estudiados de forma específica, es decir, sin generalizaciones. La inferencial o inductiva, estadística que, a diferencia de la descriptiva, sí ofrece resultados junto con datos generales de investigación amplia. La estadística aplicada proporciona resultados específicos y generalizados sobre la investigación.

La estadística matemática utilizara el álgebra y ciertos análisis más profundos para ofrecer un punto de vista enfocado y formal.

Por otra parte, se entiende por población al conjunto de seres humanos que hacen vida en un espacio determinado geográfico o territorio. Es decir, suele referirse a la población humana. Por ejemplo, un conjunto de animales dentro de un espacio geográfico.

Posteriormente, la muestra consiste en la porción que se extrae de una población estadística para un determinado estudio, con el fin de representar, conocer y determinar los aspectos de dicha población.

El muestreo es el proceso de seleccionar un conjunto de individuos de una población con el fin de estudiarlo y poder caracterizar el total de la población. Por ejemplo, si queremos saber qué porcentaje de los habitantes de México fuma habitualmente. Una forma de obtener este dato sería contactar con todos los habitantes de México y preguntarles si fuman. La otra forma sería un subconjunto de individuos, preguntarles si fuman y usar esta información como una aproximación de información buscada, pues bien este grupo de 1.000 personas que me periten conocer mejor cómo se comporta en total de mexicanos es una muestra, y la forma en que los seleccionamos es la técnica de muestreo.

Continuando con el tema, el censo es el reencuentro de datos estadísticos sobre toda una población específica. Es decir, la recopilación de información respecto a un grupo de personas. Por ejemplo, todos los ciudadanos de una nación. Dentro de los censos, la información que obtenemos con: edad, sexo, nivel de educación,

accesos a servicios básicos, ingresos mensuales, características de la vivienda habitada, entre otros.

La variable estadística es cada una de las características o cualidades que poseen los individuos de una población. Las variables se dividen en dos y son: las variables cualitativas y las variables cuantitativas. La primera se refiere a características o cualidades que no pueden ser medidas con números y la segunda, es la que se expresa mediante un número, por tanto se pueden realizar operaciones aritméticas con ella.

Por otro lado, las medidas de tendencia central son parámetros estadísticos que informan sobre el centro de distribución de la muestra o población estadística.

Entre las medidas de tendencia central encontramos: la media, que es el valor promedio de un conjunto de datos numéricos, calculada como la suma del conjunto de valores dividida entre el número total de valores. La mediana es un estadístico de posición central que parte la distribución en dos, es decir, deja la misma cantidad de valores a un lado que a otro. Por tanto, la moda es el valor que más se repite en una muestra estadística o población. Cada una de estas medidas es necesaria cuando se hará algún dato estadístico.

Las medidas de dispersión tratan, a través del cálculo de diferentes fórmulas, de arrojar un valor numérico que ofrezca información sobre el grado de variabilidad de una variable. Dentro de estas medidas se encuentran el rango, la varianza, la desviación típica y el coeficiente de variación. Estas son importantes porque nos hablan de la variabilidad que encontramos en una determinada muestra o población.

En una distribución de datos, las medidas de dispersión tienen un papel muy importante. Estas medidas complementan a las de posición central, caracterizando la variabilidad de los datos. Además son esenciales en una distribución de datos. La percepción de la variabilidad de los datos nos da información sobre la dispersión de los datos con respecto a un promedio o media.

En un estudio estadístico, a la hora de generalizar los datos de una muestra de una población las medidas de dispersión son muy importantes ya que condicionan de manera directa el error con el que trabajemos. Así, cuanto más dispersión recojamos en una muestra, más tamaño necesitaremos para trabajar con el mismo error. Por otro lado, estas medidas nos ayudan a determinar si nuestros datos se alejan mucho del valor central. Estas medidas son muy útiles para comparar distribuciones y comprender los riesgos en la toma de decisiones. A mayor dispersión, menos representativo es el valor central. Cada una de estas medidas cumplen con una función, la función de un rango es para una comparación primaria, es decir, se define como la diferencia entre el último valor de la variable y el primero.

Por su parte, la desviación media indica dónde estarían concentrados los datos si todos estuvieran a la misma distancia de la media aritmética. Consideramos la desviación de un valor de la variable como la diferencia en valor absoluto entre ese valor de la variable y la media aritmética de la serie. Así pues, se considera como la media aritmética de las desviaciones.

La varianza es una función algebraica de todos los valores, apropiada para tareas de estadística inferencial. Se puede definir como las desviaciones al cuadrado. La desviación estándar se trata de la raíz cuadrada de la varianza.

El coeficiente de variación se trata de una medida utilizada principalmente para comparar la variación entre dos conjuntos de datos medidos en diferentes unidades.

Permite comparar las dispersiones de dos distribuciones distintas siempre que sus medias sean positivas la mayor dispersión corresponderá al valor del coeficiente de variación mayor.

DISTRIBUCIÓN DE FRECUENCIAS PARA DATOS AGRUPADOS

Las tablas de distribución de frecuencias se utilizan cuando se recolectan datos, con ellas se pueden representar los datos de manera que es más fácil analizarlos.

Cuando hay muchos datos se agrupan en clases. Esto consiste en agrupar los datos en una distribución de frecuencias, que puede definirse como una ordenación o arreglo de datos en clases o categorías que muestran para cada una de ellas, el número de elementos que contiene, denominada frecuencia.

Clase es cada uno de los grupos en que se dividen los datos. Para determinar cuántas clases crear, se puede utilizar la siguiente fórmula:

Número de clases = $1 + 3,322 \log n$ donde n es el número total de datos.

El intervalo de clase o el ancho de la clase (tamaño de la clase) es el espacio que hay entre el límite superior y el límite inferior de la clase, los cuales corresponden a los valores extremos de la clase. Para obtener el ancho de clase se utiliza la siguiente fórmula:

Ancho de clase = (dato superior – dato inferior)/ número de clases.

La frecuencia absoluta es el número de veces que se repite cada dato. Cuando se agrupan los datos, es el número de datos que tiene cada clase. Se simboliza con f_j .

La marca de clase es el punto medio de la clase. Se obtiene dividiendo entre dos la suma de los valores extremos de cada clase.

El rango es la diferencia entre el valor mayor y el valor menor en estudio de una distribución de datos.

La frecuencia absoluta acumulada es la frecuencia total hasta el límite superior de cada clase. Se simboliza con F_i .

La frecuencia relativa de un dato da información sobre qué parte de la población o de la muestra en estudio corresponde a la característica analizada. Se obtiene dividiendo la frecuencia absoluta entre el número total de datos y se puede expresar como una fracción, como un decimal o como un porcentaje. Se simboliza con f_i / n donde n es el número de datos.

La frecuencia acumulada relativa es la frecuencia relativa total hasta el límite superior de cada clase. Se simboliza con F_j / n donde n es el número total de datos.

DISTRIBUCION DE FRECUENCIAS PARA DATOS NO AGRUPADOS

Para construir la tabla de datos no agrupados se debe calcular primero lo siguiente:

Número de clases, rango, amplitud de clase

Punto medio: m_i es el valor central de la clase. Se obtiene calculando el promedio de los límites reales, sumando al límite real inferior el límite real superior y dividiendo por dos.

Frecuencia absoluta. Se define como el número de elementos u observaciones pertenecientes a una misma clase.

Frecuencia relativa: Se obtiene dividiendo la frecuencia absoluta por el número total de observaciones. Indica la importancia relativa de la clase.

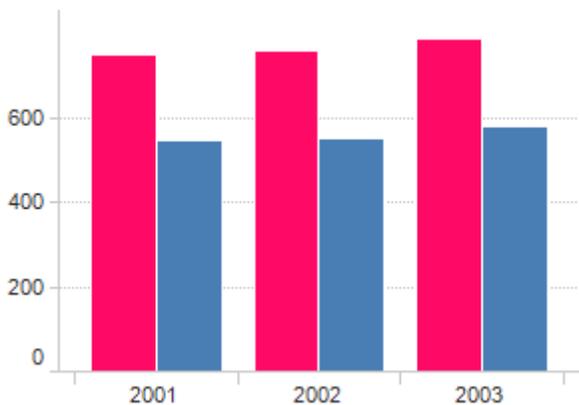
Frecuencias acumuladas: Es la suma de las frecuencias absolutas o relativas en sentido ascendente o descendente según se quieran acumular “hacia arriba” o “hacia abajo”

	Punto medio	Frecuencia absoluta	Frecuencia absoluta acumulada	Frecuencia relativa	Frecuencia relativa acumulada
Pesos (Kg)	m_i	f_i	F_i	fr_i	Fr_i
53 - 55	54	2	2	4,00%	4,00%
56 - 58	57	5	7	10,00%	14,00%
59 - 61	60	9	16	18,00%	32,00%
62 - 64	63	15	31	30,00%	62,00%
65 - 67	66	12	43	24,00%	86,00%
68 - 70	69	5	48	10,00%	96,00%
71 - 73	72	2	50	4,00%	100,00%
		50		100,00%	

GRAFICAS

GRAFICA DE BARRAS

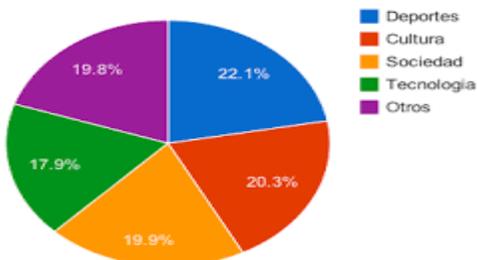
Un gráfico de barras es una forma de resumir un conjunto de datos por categorías. Muestra los datos usando varias barras de la misma anchura, cada una de las cuales representa una categoría concreta. La altura de cada barra es proporcional a una agregación específica (por ejemplo, la suma de los valores de la categoría que representa). Las categorías podrían ser desde grupos de edad a ubicaciones geográficas.



GRAFICAS DE PASTEL

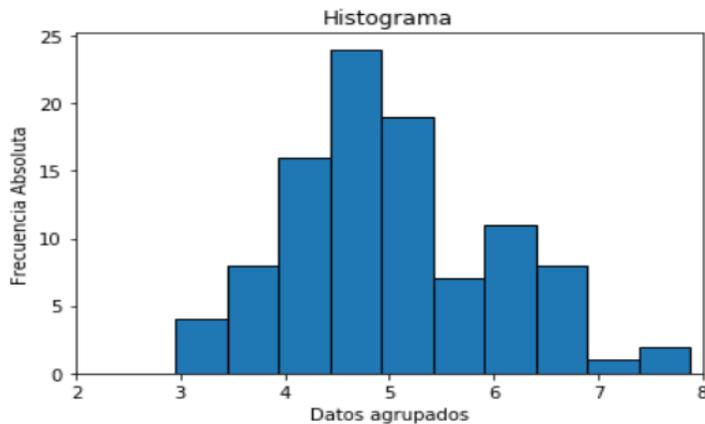
Un gráfico de pastel es un gráfico circular, el cual está dividido de tal manera que se asemeja a las tajadas de un pastel. Este gráfico divide tus datos en sectores, ilustrando cada proporción numérica con el fin de mostrar la composición como un todo.

Visitas a contenidos



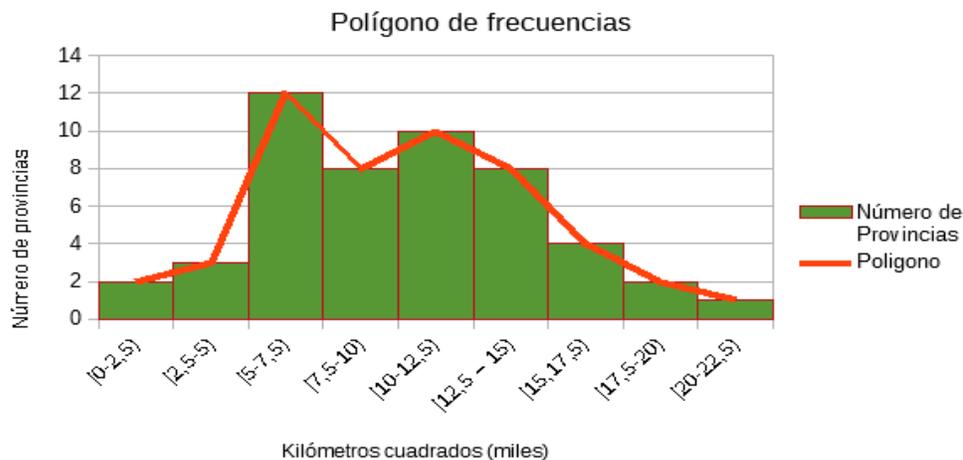
HISTOGRAMAS

Un histograma es la representación gráfica en forma de barras, que simboliza la distribución de un conjunto de datos. Sirven para obtener una "primera vista" general, o panorama, de la distribución de la población, o de la muestra, respecto a una característica, cuantitativa y continua.



POLIGONOS DE FRECUENCIA

Polígono de frecuencia es el nombre que recibe una clase de gráfico que se crea a partir de un histograma de frecuencia. Estos histogramas emplean columnas verticales para reflejar frecuencias: el polígono de frecuencia es realizado uniendo los puntos de mayor altura de estas columnas.



Calcula la media aritmética o promedio, mediana, moda, rango, desviación estándar y varianza para las siguientes calificaciones que obtuvo un alumno LPSI de la UDS:
6, 7, 8, 9, 9, 8, 7, 6, 10.

Media: 7.7

$$6+7+8+9+9+8+7+6+10=70/9=7.7$$

Mediana: 8

6, 6, 7, 7, 8, 8, 9, 9, 10

Moda: 6, 7, 8, 9

6, 6, 7, 7, 8, 8, 9, 9, 10

Rango: 4 Numero mayor menos el numero menor

6, 7, 8, 9, 9, 8, 7, 6, **10.** $10-6=4$

Desviación estándar: 1.39687

Varianza: 1.95125