

- **Desarrolle las siguientes actividades con su equipo:**
- **Investigue, estudie y documente los siguientes términos:**
- Definición de estadística.

La Estadística es una disciplina que utiliza recursos matemáticos para organizar y resumir una gran cantidad de datos obtenidos de la realidad, e inferir conclusiones respecto de ellos.

- La ciencia de recopilar, organizar, presentar, analizar e interpretar información para ayudar a tomar decisiones mas efectivas.

- Estadística descriptiva.

Puede definirse como aquellos métodos que incluyen la recolección, presentación y caracterización de un conjunto de datos con el fin de describir apropiadamente las diversas características de ese conjunto.

- Conjunto de métodos para organizar, resumir y presentar los datos de manera informativa.

- Estadística inferencial.

Puede definirse como aquellos métodos que hacen posible la estimación de una característica de una población o la toma de una decisión referente a una población basándose sólo en los resultados de una muestra

- Conjunto de métodos utilizados para determinar algún atributo medible acerca de una población con base en una muestra

- Población o universo.

es la totalidad de elementos o cosas bajo consideración.

- Variable cualitativa.

Son las variables que expresan distintas cualidades, características o modalidad. Cada modalidad que se presenta se denomina atributo o categoría y la medición consiste en una clasificación de dichos atributos. Las variables cualitativas pueden ser ordinales y nominales. Las variables cualitativas pueden ser dicotómicas cuando sólo pueden tomar dos valores posibles como Sí y no, hombre y mujer o son politémicas cuando pueden adquirir tres o más valores. Dentro de ellas podemos distinguir:

Variable cualitativa ordinal: La variable puede tomar distintos valores ordenados siguiendo una escala establecida, aunque no es necesario que el intervalo entre mediciones sea uniforme, por ejemplo, leve, moderado, grave

Variable cualitativa nominal: En esta variable los valores no pueden ser sometidos a un criterio de orden como por ejemplo los colores o el lugar de residencia.

- Cuando la característica que se estudia no es numérica

- Variable cuantitativa.

Son las variables que se expresan mediante cantidades numéricas. Las variables cuantitativas además pueden ser: Variable discreta o Variable continua

- Variable discreta.

Es la variable que presenta separaciones o interrupciones en la escala de valores que puede tomar. Estas separaciones o interrupciones indican la ausencia de valores entre los distintos valores específicos que la variable pueda asumir.

- Variable continua.

Es la variable que puede adquirir cualquier valor dentro de un intervalo especificado de valores. Por ejemplo el peso o la altura, que solamente limitado por la precisión del aparato medidor, en teoría permiten que siempre existe un valor entre dos

cualesquiera.

- Escala nominal.

sólo permite asignar un nombre al elemento medido. Esto la convierte en la menos informativa de las escalas de medición.

Los siguientes son ejemplos de variables con este tipo de escala:

Nacionalidad.

Uso de anteojos.

Número de camiseta en un equipo de fútbol.

Número de Cédula Nacional de Identidad.

- Escala ordinal.

además de las propiedades de la escala nominal, permite establecer un orden entre los elementos medidos.

Ejemplos de variables con escala ordinal:

Preferencia a productos de consumo.

Etapas de desarrollo de un ser vivo.

Clasificación de películas por una comisión especializada.

Madurez de una fruta al momento de comprarla.

- Escala de intervalo.

además de todas las propiedades de la escala ordinal, hace que tenga sentido calcular diferencias entre las mediciones.

Los siguientes son ejemplos de variables con esta escala:

Temperatura de una persona.

Ubicación en una carretera respecto de un punto de referencia (Kilómetro 85 Ruta 5).

Sobrepeso respecto de un patrón de comparación.

Nivel de aceite en el motor de un automóvil medido con una vara graduada.

- Escala de razón.

El nivel de medida más elevado es el de cocientes o razones, y se diferencia de las escalas de intervalos iguales únicamente por poseer un punto cero propio como origen; es decir que el valor cero de esta escala significa ausencia de la magnitud que estamos midiendo

- Distribución de frecuencia.

Es el ordenamiento tabular de los datos por clases conjuntamente con las frecuencias de clases. es el número de veces que se observa dicho valor, o el número de casos clasificados en la clase definida por él.

- Determinación del número de grupos o clases.

- Para determinar el número de clases es la regla de 2 a la k, en esta guía sugiere elegir el número más bajo (k) para el número de clases, de modo que 2 elevado a la k sea mayor que el número de observaciones (n).

- Determinación del intervalo o amplitud de la clase (ancho de clase).

Debe ser el mismo para todas las clases. Todas estas deben cubrir por lo menos la distancia desde el valor más bajo en los datos en bruto hasta el valor más alto. Se tiene la siguiente fórmula

$$i = (H-L)/k$$

i= intervalo de clase

H= valor observado más alto

L= valor observado más bajo

K= número de clases

- Determinación de los límites inferior y superior de cada clase.

Establezca límites de clases claros a fin de incluir cada observación en una sola categoría. Esto significa que debe evitar la superposición, así como los límites de clase inciertos.

- Cálculo de la frecuencia absoluta de cada clase.

Frecuencia absoluta

de la clase c_i es el número n_i , de observaciones que presentan una modalidad perteneciente a esa clase.

Frecuencia absoluta (n_i): Número de elementos que presentan la clase x_i .

- Cálculo de la frecuencia relativa de cada clase.

Frecuencia relativa

de la clase c_i es el cociente f_i , entre las frecuencias absolutas de dicha clase y el número total de observaciones, es decir

Obsérvese que f_i es el tanto por uno de observaciones que están en la clase c_i . Multiplicado por representa el porcentaje de la población que comprende esa clase.

Frecuencia relativa: Frecuencia relativa:

- Determinación de la marca de clase.
- punto representativo del intervalo. Si éste es acotado, tomamos como marca de clase al punto más representativo, es decir al punto medio del intervalo,

La marca de clase no es más que una forma abreviada de representar un intervalo mediante uno de sus puntos. Por ello hemos tomado como representante, el punto medio del mismo. Esto está plenamente justificado si recordamos que cuando se mide una variable continua como el peso, la cantidad con cierto número de decimales que expresa esta medición, no es el valor exacto de la variable, sino una medida que contiene cierto margen de error, y por tanto representa a todo un intervalo del cual ella es el centro.

- Cálculo de la distribución de frecuencias relativa.

f_i , se calcula sobre variables cuantitativas o cuasicuantitativas, siendo el tanto por uno de los elementos de la población que están en alguna de las clases y que presentan una modalidad inferior o igual a la c_i , es decir,

Como todas las modalidades son exhaustivas e incompatibles ha de ocurrir que

o lo que es lo mismo,

- Diseño de un diagrama de tallos y hojas.

es una forma rápida de obtener una representación visual ilustrativa del conjunto de datos, para construir un diagrama de tallo y hoja primero se debe seleccionar uno o más dígitos iniciales para los valores de tallo, el dígito o dígitos finales se convierten en hojas, luego se hace una lista de valores de tallo en una columna vertical. Prosiguiendo a registrar la hoja por cada observación junto al valor correspondiente de tallo, finalmente se indica las unidades de tallos y hojas en algún lugar del diagrama, este se usa para listas grandes y es un método resumido de mostrar los datos, posee la desventaja que no proporciona sino los datos, y no aparece por ningún lado información sobre frecuencias y demás datos importantes.

- Histograma.

Se emplea para ilustrar muestras agrupadas en intervalos. Esta formado por rectángulos unidos a otros, cuyos vértices de la base coinciden con los límites de los intervalos y el centro de cada intervalo es la marca de clase, que representamos en el eje de las abscisas. La altura de cada rectángulo es proporcional a la frecuencia del intervalo respectivo. Esta proporcionalidad se aplica por medio de la siguiente fórmula

Altura del rectángulo = frecuencia relativa/longitud de base

El histograma se usa para representar variables cuantitativas continuas que han sido agrupadas en intervalos de clase, la desventaja que presenta que no funciona

para variables discretas, de lo contrario es una forma útil y práctica de mostrar los datos estadísticos.

- Polígono de frecuencias.

es una representación gráfica de la distribución de frecuencias que resulta esencialmente equivalente al histograma y se obtiene uniendo mediante segmentos los centros de las bases superiores de los rectángulos del histograma (es decir, los puntos de las marcas de clase).

- Ojiva (polígono de frecuencia acumulada).

Es una representación gráfica con la misma idea de unir los centros de las bases superiores de los rectángulos de la distribución del histograma de frecuencias acumuladas

- Gráficas de líneas, barras, horizontales, verticales y de pastel.

Gráficas de línea.- Cuando los datos se relacionan entre sí, es decir, cuando podemos decir que existe cierta continuidad entre las observaciones (como por ejemplo el crecimiento poblacional, la evolución del peso o estatura de una persona a través del tiempo, el desempeño académico de un estudiante a lo largo de su instrucción escolar, las variaciones presentadas en la medición realizada en algún experimento cada segundo o minuto) se pueden utilizar las gráficas de líneas, que consisten en una serie de puntos trazados en las intersecciones de las marcas de clase y las frecuencias de cada una, uniéndose consecutivamente con líneas

Graficas de barras.- nombre que recibe el diagrama utilizado para representar gráficamente distribuciones discretas de frecuencias no agrupadas. Se llama así porque las frecuencias de cada categoría de la distribución se hacen figurar por trazos o columnas de longitud proporcional, separados unos de otros. Existen tres principales clases de gráficos de barras:

Barra simple: se emplean para graficar hechos únicos

Barras múltiples: es muy recomendable para comparar una serie estadística con otra, para ello emplea barras simples de distinto color o tramado en un mismo plano cartesiano, una al lado de la otra

Barras compuestas: en este método de graficación las barras de la segunda serie se colocan encima de las barras de la primera serie en forma respectiva.

El diagrama de barras proporciona información comparativa principalmente y este es su uso principal, este diagrama también muestra la información referente a las frecuencias

Gráficos de Pastel.- es un gráfico que se basa en una proporcionalidad entre la frecuencia y el ángulo central de una circunferencia, de tal manera que a la frecuencia total le corresponde el ángulo central de 360° . Para construir se aplica la siguiente fórmula:

$$X = \text{frecuencia relativa} * 360^\circ / \text{suma de frecuencias}$$

Este se usa cuando se trabaja con datos que tienen grandes frecuencias, y los valores de la variable son pocos, la ventaja que tiene este diagrama es que es fácil de hacer y es entendible fácilmente, la desventaja que posee es que cuando los valores de la variable son muchos es casi imposible o mejor dicho no informa mucho este diagrama y no es productivo, proporciona principalmente información acerca de las frecuencias de los datos de una manera entendible y sencilla.