Nombre del Alumno: Josue Alfredo Díaz Sanchez

Nombre del tema: Datos agrupados y no agrupados

Parcial: I

Nombre de la Materia: Estadística

Nombre del profesor: Víctor Tadeo Cruz

Nombre de la Licenciatura: Administración de Empresas

Cuatrimestre: 1er

Grupo: "A"

Frontera Comalapa, Chiapas 13 de Octubre del 2021

DATOS AGRUPADOS Y NO AGRUPADOS

FRECUENCIA DE CLASE

punto que divide a la clase en dos partes iguales. Es el promedio entre los límites superior e inferior de la clase.

Intervalo de clase

para una distribución de frecuencias que tiene clases del mismo tamaño, el intervalo de clase se obtiene restando el límite inferior de una clase del límite inferior de la siguiente.

FRECUENCIA RELATIVA

Es la relación o cociente entre la frecuencia absoluta y el número total de observaciones. Es la proporción entre la frecuencia de un intervalo y el número total de datos.

PUNTO MEDIO

es único y equidista de los extremos del segmento. Cumpliendo esta última condición, pertenece a la mediatriz del segmento

La fórmula para determinar el punto medio de un segmento en el plano, con coordenadas

(x1,y1) y (x2,y2) es: [(x1 + x2) / 2] + [(y1 + y2) / 2]

LIMITES

Son los valores extremos que tiene el intervalo de clase, inferior y superior, entre los cuales van a estar los valores de los datos agrupados en ese intervalo de clase

Medidas de tendencia centra

observar la distribución de los datos, estas tienden a estar localizadas generalmente en su parte central.

Media aritmética (□x)

También se le conoce como promedio ya que es el promedio de las lecturas o mediciones individuales que se tienen en la muestra

Media geométrica (G)

Es la raíz en enésima del producto de los valores de los elementos de la muestra, es usada cuando los valores de los datos de la muestra no son lineales, es decir que su valor depende de varios factores a la vez

Media aritmética ponderada (xw)

Esta media se usa cuando el peso que tiene cada uno de los datos de la muestra es diferente

Media armónica (H)

se define como el recíproco del promedio de los recíprocos de cada uno de los datos que se tienen en la muestra

Mediana (xmed)

La mediana es aquel valor que se encuentra en la parte central de los datos que se tienen en la muestra una vez que estos han sido ordenados según su valor o magnitud

Cuando el número de datos en la muestra es impar

En este caso después de ordenar los datos de la muestra en cuanto a su magnitud, es decir de mayor a menor valor o de menor a mayor valor, se procede a localizar aquel dato que se encuentra justo en el centro de los datos o en la parte central de los mismos, el valor de este dato será el que dé valor a la mediana

Cuando el número de datos en la muestra es par

En este caso después de ordenar los datos en cuanto a su magnitud, observamos que en la parte central de los datos no se encuentra dato alguno, en este caso, la mediana tomará el valor del promedio de dos datos; el que se encuentra antes de la parte central y el que se encuentra después de la parte central

Moda (xmod)

se define como aquel valor o valores que más se repiten o que tienen mayor frecuencia entre los datos que se han obtenido en una muestra, la muestra de una población nos genera la distribución de los datos una vez que estos se han graficado y en esta gráfica es posible observar la moda o modas de la misma, es por esto que una distribución de datos puede ser amodal (carece de moda), unimodal (tiene una sola moda), bimodal (tiene dos modas) o polimodal (tiene más de dos modas)

Medidas de Dispersión

Cuando se tiene una muestra de datos obtenida de una población cualquiera, es importante determinar sus medidas de tendencia central así como también es básico el determinar que tan dispersos están los datos en la muestra, por lo que se hace necesario determinar su rango, la varianza, la desviación estándar

Rango o recorrido

es la diferencia entre el valor mayor y el valor menor encontrados en la muestra, también se le denomina recorrido ya que nos dice entre que valores hace su recorrido la variable de interés

Desviación absoluta media ( )

Esta medida de dispersión nos representa la diferencia absoluta promedio que existe entre cada dato que se encuentra en la muestra y la media de los datos

¿Por qué sacar el valor absoluto de las diferencias entre cada dato y la media aritmética?

Si solo se hicieran diferencias entre cada dato y la media aritmética, estas tendrían signos positivos y negativos ya que algunos datos son menores que la media y otros son mayores que la media, luego al sumar las diferencias, con sus signos correspondientes, éstas se irían anulando unas con otras y no sería posible medir leal grado de alejamiento promedio de los datos

Varianza o variancia (s2)

Es el promedio de las diferencias elevadas al cuadrado entre cada valor que se tiene en la muestra (xi) y la media aritmética ( ) de los datos

Desviación estándar (s)

Es la desviación o diferencia promedio que existe entre cada dato de la muestra y la media aritmética de la muestra. Y se obtiene a partir de la varianza, sacándole raíz cuadrada.

Media aritmética para datos agrupados

Se calcula sumando todos los productos de marca clase con la frecuencia absoluta respectiva y su resultado dividirlo por el número total de datos

Moda

Es el valor que representa la mayor frecuencia absoluta. En tablas de frecuencias con datos agrupados, hablaremos de intervalo modal. La moda se representa por Mo.

Mediana

Es el valor que ocupa el lugar central de todos los datos cuando éstos están ordenados de menor a mayor. La mediana se representa por Me. La mediana se puede hallar sólo para variables cuantitativas.

Cálculo de la mediana para datos agrupados

La mediana se encuentra en el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas. Es decir tenemos que buscar el intervalo en el que se encuentre. N / 2

CUANTILES

son medidas de posición que se determinan mediante un método que determina la ubicación de los valores que dividen un conjunto de observaciones en partes iguales

El primer cuartil Q1,

es el menor valor que es mayor que una cuarta parte de los datos; es decir, aquel valor de la variable que supera 25% de las observaciones y es superado por el 75% de las observaciones

El segundo cuartil Q2,

(coincide, es idéntico o similar a la mediana, Q2 = Md), es el menor valor que es mayor que la mitad de los datos, es decir el 50% de las observaciones son mayores que la mediana y el 50% son menores.

El tercer cuartil Q3

es el menor valor que es mayor que tres cuartas partes de los datos, es decir aquel valor de la variable que supera al 75% y es superado por el 25% de las observaciones.

DECILES

son ciertos números que dividen la sucesión de datos ordenados en diez partes porcentualmente iguales. Son los nueve valores que dividen al conjunto de datos ordenados en diez partes iguales, son también un caso particular de los percentiles. Los deciles se denotan D1, D2,..., D9, que se leen primer decil, segundo decil, etc.

CENTILES O PERCENTILES

son, tal vez, las medidas más utilizadas para propósitos de ubicación o clasificación de las personas cuando atienden características tales como peso, estatura, etc.

El fundamento del coeficiente de Pearson es el siguiente

Cuanto más intensa sea la concordancia (en sentido directo o inverso) de las posiciones relativas de los datos en las dos variables, el producto del numerador toma mayor valor (en sentido absoluto)