

Clase 8

Tema: Distribuciones de frecuencias según el tipo de variable

Sumario:

- Concepto de clases o intervalo
 - Límite inferior y límite superior de una clase
 - Punto medio de la clase
 - Rango. Amplitud de la clase
 - Frecuencia absoluta, relativa y acumulada.
-
- **Objetivo:** Distribuir los datos en tabla de frecuencia según el tipo de variable.

Tipo de clase: CTP

Bibliografía: Bioestadística. Tomo II

Bayarre H, Hersford R, Oliva M: Estadística Descriptiva y Estadística de Salud, Ed Ciencias Médicas. 2005. La Habana,

ANP y motivación

Revisió.

-n de la tarea 3

Se tienen los datos del tipo de sangre de 12 pacientes

A,B,A,AB,O,A,B,O,O,O,A,AB

- a. Identifique la variable. **Grupo sanguíneo**
- b. Clasifíquela según su naturaleza. **Cualitativa nominal**
- c. ¿Cómo agruparías la información para la variable? Los de tipo A, B, AB y así según el grupo sanguíneo.

GRUPO SANGUÍNEO	FA	FR	FP
+	4	4/12=0.333	33.3
A			
B	2	2/12=0.167	16.7
AB	2	2/12=0.167	16.7
O	4	4/12=0.333	33.3
Total(n)	12	1	100

Los datos para que sean útiles se deben organizar de manera que faciliten su análisis, por lo que se debe conocer los métodos de organización y representación, la finalidad de estos métodos es permitir ver rápidamente todas las características posibles de estos datos que se han recolectado.

Organización de la información

Para **ORGANIZAR** los datos es necesario definir la variable de estudio, clasificarla y agrupar los datos.

OHO y asunto

Distribuciones de frecuencia: Es cuando se conoce la cantidad de elementos correspondientes a cada valor(clases) de una variable. Una forma cómoda de

organizar los datos en una distribución de frecuencia es mediante la confección de tablas.

Se utiliza para resumir grandes grupos de datos para así poder interpretarlos. Una forma de organizar los datos en estos casos es agruparlos en diferentes clases o categorías, **según una escala única**, es decir, construir una distribución de frecuencia que será cualitativa o cuantitativa discreta o cuantitativa continua en dependencia de la escala empleada.

Clase: Grupo de valores que describen una característica. Deben incluir todas las observaciones y ser excluyentes.

EJEMPLOS

De las siguientes variables ¿cuáles serían las clases o categorías que puede tomar la variable?

1. Sexo, estado civil, cantidad de hijos

Se denominan tabla de distribución de frecuencia (TDF) al conjunto de clases asociadas a una variable y sus correspondientes frecuencias.

Frecuencia absoluta (Fa): Es el número de ocurrencias en cada clase.

La frecuencia relativa(fr): Se define como el cociente de la frecuencia absoluta (fa) de la clase y el total(n) **fr = fa /Total**

Frecuencia porcentual (FP)=Es la frecuencia relativa de la clase multiplicada por 100 **FP=(fr x 100)**

Para confeccionar la distribución de frecuencias es necesario:

1. Definir las clases en que serán agrupados los elementos en una escala única
2. clasificarlos y,
3. calcular cuántos pertenecen a cada clase.

A continuación estudiaremos los procedimientos para construir tablas de distribuciones de frecuencias para los dos casos siguientes:

- Variables cualitativas y cuantitativas discretas
- Variables cuantitativas continuas

Distribución de frecuencias para datos cualitativos y cuantitativos discretos

Cuando la variable bajo estudio es cualitativa, las clases corresponderán a las categorías que están presentes dentro del conjunto de posibles valores a tomar por la variable.

Ejemplo:

Suponga que un médico de familia quiere conocer la composición por raza, de los 15 pacientes con hipertensión arterial que él atiende, para ello utilizará una distribución de frecuencias. Los datos recogidos son: B-N-M-A-B-B-N-N-M-M-M-M-A-B-N

1. Primero tendrá que considerar cuáles son las **distintas clases**; pudieran ser: blanca B, negra N, mestiza M y amarilla A.

En esto hay que observar un detalle muy importante y es que una buena escala debe ser exhaustiva y constar de clases mutuamente excluyentes, o sea, debe

permitir la clasificación de cualquiera de los elementos, pero además estos deben ser clasificados en una clase y solamente en esa.

En el caso del ejemplo, si no existe ningún individuo de la raza amarilla, se puede obviar esta categoría; pero de existir, sería un error omitirla, ya que el individuo quedaría fuera de la clasificación (la escala no sería exhaustiva).

2. Después de definidas las clases o categoría se procede a la ubicación de cada elemento en la clase que le corresponda y,

3. Por último, al **conteo**, para conocer el número de ocurrencias en cada una de ellas, valor conocido como **frecuencia absoluta** de la clase. En el ejemplo citado sería como sigue:

raza	Frecuencia		
	FA	Relativa	Por ciento
Blanca	4	0,27	27
Negra	4	0,27	27
Mestiza	5	0,33	33
Amarilla	2	0,13	13
Total	15	1	100

Esta sería la distribución de frecuencias, las clases vendrían dadas por las diferentes categorías que aporta la raza, los valores de frecuencia absoluta significarían que 4 de los pacientes hipertensos son de la raza blanca, 4 de la raza negra y así sucesivamente.

Vale destacar en los totales que la suma de las frecuencias absolutas tiene que coincidir con el total de elementos bajo estudio y la suma de las frecuencias relativas debe ser igual a la unidad (1) y el total del porcentaje debe ser 100%. Cuando la variable es nominal no tiene sentido calcular las frecuencias acumuladas.

Del ejercicio de la motivación, realice una tabla de frecuencia para la variable grupo sanguíneo

Se tienen los datos del tipo de sangre de 12 pacientes

A,B,A,AB,O,A,B,O,O,O,A,AB

Clases	Frecuencia		
	Absoluta	Relativa	Por ciento
A			
B			
AB			
O			
Total			

Distribución de frecuencias para datos cuantitativos continuos

Para la distribución por intervalos se debe tener en cuenta una serie de conceptos:

Rango o recorrido de la clase: Es la diferencia entre la puntuación mayor y la menor.

Límite de clases: Es el mayor y menor valor en un intervalo de clase, los cuales se denotan por Li, Ls para indicar el límite inferior y superior respectivamente.

Amplitud de clases: Es la amplitud del intervalo de clase la cual se obtiene por $a=Ls- Li +1$ o $a=R/k$ (número de clases)

Marca de clase: Es el punto medio de cada intervalo de clase, el cual es un valor representativo en esta de la variable que se estudia, se calcula: $Mc= \frac{Li+Ls}{2}$

Los siguientes datos corresponden a la estatura en cm de un grupo de mujeres embarazadas que padecen de hipertensión arterial.

148 -149 -159 – 162 – 165 – 165- 169- 171 – 184 – 184-150-160-170-180-160

Como la variable es cuantitativa continua, se realiza una escala de intervalos, para ello se establece un conjunto de valores o agrupaciones en los que se moverá la variable, denominado intervalo de clase, así como determinar el número de clases en que se agruparán los datos y de donde a donde irá cada una.

Paso 1

Hallar el rango o Recorrido de la clase

Cuanto más grande sea el rango mayor será la dispersión de los datos. No dice nada acerca de la dispersión de los datos dentro del intervalo.

Rango= MAYOR VALOR – MENOR VALOR
184-148 = 36

Paso 2: Hallar la amplitud de la clase para 5 clases

Para esto se debe tener en cuenta el número de clases que se va a asumir para la variable. Algunos autores plantean que no debe ser menos de 5 ni más de 15.

148 -149 -159 – 162 – 165 – 165- 169- 171 – 184 – 184-150-160-170-180-172

$$a = \frac{R}{\text{clases}} = \frac{36}{5} = 7.2$$

Paso 3: Formar los intervalos de clase y determinar las frecuencias.

estatura	CONTEO	MC	FA	FR	Porcentaje
148	153	111	3		
154	159	1	1		
160	165	1111	4		
166	171	111	3		
172	177	1	1		
178	184	111	3		
Total			15		

Qué interpretación podemos hacer de la tabla?

Estudio individual

1. Se tienen las calificaciones de un grupo de estudiantes.

3-3-3-2-2-2-4-4-5-5-5-3-2-3-2-4-4-5-4-4-3-4-3-4-3-3-3-5-5-5

- a. ¿Cuántos estudiantes tiene el grupo?
- b. Realice una TDF
- c. ¿Cuántos estudiantes están suspensos?
- d. ¿Qué por ciento representan los estudiantes con 5 puntos?
- e. ¿Cuál es la calificación más frecuente?
- f. ¿Cuál es el promedio de las calificaciones del grupo?