

UNIDAD 11: ESTADÍSTICA. TABLAS Y GRÁFICOS

Contenido

| | | |
|---|--|---|
| 1 | ESTADÍSTICA: CLASES Y CONCEPTOS BÁSICOS..... | 2 |
| 2 | VARIABLES O CARACTERES ESTADÍSTICOS..... | 2 |
| 3 | TABLAS ESTADÍSTICAS: RECUENTO | 3 |
| 4 | TABLAS ESTADÍSTICAS: FRECUENCIAS | 4 |
| 5 | GRÁFICOS PARA VARIABLES ESTADÍSTICAS CUALITATIVAS | 5 |
| 6 | GRÁFICOS PARA VARIABLES ESTADÍSTICAS CUANTITATIVAS | 8 |

1 ESTADÍSTICA: CLASES Y CONCEPTOS BÁSICOS

En sus orígenes históricos, la Estadística estuvo ligada a cuestiones de Estado (recuentos, censos, etc.) y de ahí proviene su nombre. Hoy en día está presente en todos los ámbitos humanos, tanto individuales como colectivos.

La Estadística surge ante la necesidad de poder tratar y comprender conjuntos numerosos de datos.

Definición: La Estadística es la ciencia que se ocupa de la recogida de datos, su organización y análisis; así como de las predicciones que, a partir de estos datos, pueden hacerse.

Fases de un estudio estadístico:

- Recogida de datos
- Recuento de datos: tablas y gráficos estadísticos.
- Análisis de los datos: parámetros estadísticos.
- Extracción de conclusiones de los datos.
- Toma de decisiones.

Los conceptos básicos que aparecen en cualquier estudio estadístico son:

- Población: Es el conjunto formado por todos los elementos que existen para el estudio de un determinado fenómeno.
- Individuo u objeto: Es cada elemento de la población.
- Muestra: Es el subconjunto que tomamos de la población para determinar el estudio del fenómeno.
- Tamaño de la muestra: Es el número de individuos que componen la muestra.

Podemos distinguir entre dos clases de Estadística:

Definición: La **Estadística descriptiva** se ocupa de tomar los datos de un conjunto, organizarlos en tablas o en representaciones gráficas y del cálculo de unos números que nos informen de manera global del conjunto estudiado.

Definición: La **Estadística inferencial** trata sobre la elaboración de conclusiones para la población, partiendo de los resultados de una muestra y del grado de fiabilidad de estas conclusiones.

2 VARIABLES O CARACTERES ESTADÍSTICOS

Variable o carácter estadístico: Es la cualidad o propiedad de la población que se analiza en el estudio estadístico.

Tipos de variables:

- Variables cualitativas: No se pueden medir y se describen con palabras.
Ejemplos: razas de perro, estado civil de una persona, color de ojos...
- Variables cuantitativas: Se pueden medir y expresar con números.
 Pueden ser de dos tipos:
 - Discretas: Sólo pueden tomar un número finito de valores numéricos:
Ejemplos: número de hermanos, número de parados de una ciudad, número de habitantes menores de edad...
 - Continuas: pueden tomar cualquier valor en un intervalo dado.
Ejemplos: estatura de una persona, peso de una persona...

3 TABLAS ESTADÍSTICAS: RECuento

El primer paso de cualquier estudio estadístico es recoger los datos. Normalmente se suele llevar a cabo a través de encuestas o entrevistas, según la población a estudiar, su tamaño, el tiempo de que dispongamos, ...

Una vez que tenemos los datos recogidos, pasamos a hacer el recuento: contando el número de veces que aparece cada valor de la variable a estudiar.

Ejemplo: Preguntamos a 20 alumnos el número de miembros de su familia, y sus respuestas fueron:

3, 5, 4, 3, 5, 6, 8, 3, 3, 5, 7, 5, 6, 5, 4, 4, 7, 4, 5, 3

| Miembros por familia x_i | Frecuencia |
|-------------------------------|------------|
| 3 | 5 |
| 4 | 4 |
| 5 | 6 |
| 6 | 2 |
| 7 | 2 |
| 8 | 1 |

Los valores de las variables estadísticas continuas se agrupan por **intervalos o clases**. Además, si la variable es discreta y toma muchos valores, también se suele agrupar por intervalos o clases.

El valor medio de cada clase o intervalo se llama **marca de clase** y se calcula como la semisuma de los extremos del intervalo.

Para construir los intervalos debemos tener en cuenta que:

- Es conveniente que el número de intervalos que debemos considerar en cualquier estudio esté entre 5 y 10.
- Usualmente tomamos los intervalos con igual amplitud o longitud.
- El recorrido de la variable es la diferencia entre el valor más grande y el más pequeño.
- La amplitud de cada intervalo se calcula dividiendo el recorrido de la variable entre el número total de intervalos que vayamos a considerar.

Ejemplo: A los 100 empleados de una empresa de piezas de precisión, se les ha realizado una prueba de habilidad manual. En una escala de 0 a 100 se han obtenido las siguientes puntuaciones:

27, 66, 32, 55, 46, 37, 75, 81, 18, 33, 47, 74, 37, 52, 47, 66, 80, 87, 37, 29,

46, 15, 29, 90, 76, 67, 23, 35, 94, 23, 25, 56, 73, 78, 17, 28, 76, 58, 45, 36,

55, 60, 17, 56, 23, 82, 64, 50, 51, 45, 37, 65, 62, 26, 69, 36, 54, 42, 40, 54,

27, 62, 28, 65, 46, 92, 36, 33, 23, 66, 18, 82, 47, 49, 59, 45, 73, 43, 47, 83,

78, 65, 39, 36, 53, 91, 38, 35, 68, 78, 91, 23, 34, 43, 55, 56, 74, 56, 62, 38.

Observamos que los valores extremos son 15 y 94. La amplitud total entre los datos es de 80 puntos, ya que ambas puntuaciones están incluidas.

Agruparemos los datos en 8 intervalos de amplitud 10:

$$(14,24], (24,34], \dots, (84,94].$$

Realizando el recuento con atención, se obtiene la tabla que sigue:

| Habilidad manual | Marca de clase x_i | Frecuencias f_i |
|------------------|-------------------------|----------------------|
| $(14,24]$ | 19 | 10 |
| $(24,34]$ | 29 | 12 |
| $(34,44]$ | 39 | 17 |
| $(44,54]$ | 49 | 18 |
| $(54,64]$ | 59 | 13 |
| $(64,74]$ | 69 | 13 |
| $(74,84]$ | 79 | 11 |
| $(84,94]$ | 89 | 6 |

4 TABLAS ESTADÍSTICAS: FRECUENCIAS

Para construir una tabla estadística completa tenemos que calcular:

- Frecuencia absoluta de cada valor: Es el número total de veces que aparece el dato (x_i). Se representa por f_i
- Frecuencia absoluta acumulada de cada valor: Es la suma de todas las frecuencias absolutas correspondientes a los valores anteriores a x_i y a la suya propia. Se representa por F_i . **“No tiene sentido para variables cualitativas”**.

$$F_i = f_1 + f_2 + \dots + f_i$$

- Frecuencia relativa de cada valor: Se calcula dividiendo la frecuencia absoluta correspondiente f_i al valor (x_i) entre el número total de datos N . Se representa por h_i y se obtiene por:

$$h_i = \frac{f_i}{N}$$

La suma de todas las frecuencias relativas es la unidad:

$$h_1 + h_2 + h_3 + \dots + h_{n-1} + h_n = 1$$

- Frecuencia relativa acumulada de cada valor: Es la suma de todas las frecuencias relativas correspondientes a los valores anteriores a x_i y a la suya propia. Se representa por H_i . No tiene sentido para variables cualitativas.

$$H_i = h_1 + h_2 + \dots + h_i$$

Ejemplo: Preguntamos a 20 alumnos el número de miembros de su familia, y sus respuestas fueron:

3, 5, 4, 3, 5, 6, 8, 3, 3, 5, 7, 5, 6, 5, 4, 4, 7, 4, 5, 3

| Miembros por familia x_i | Frecuencia absoluta f_i | Frecuencia absoluta acumulada F_i | Frecuencia relativa h_i | Frecuencia relativa acumulada H_i |
|-------------------------------|------------------------------|--|------------------------------|--|
| 3 | 5 | 5 | 0,25 | 0,25 |
| 4 | 4 | 9 | 0,2 | 0,45 |
| 5 | 6 | 15 | 0,3 | 0,75 |
| 6 | 2 | 17 | 0,1 | 0,85 |
| 7 | 2 | 19 | 0,1 | 0,95 |
| 8 | 1 | 20 | 0,05 | 1 |

5 GRÁFICOS PARA VARIABLES ESTADÍSTICAS CUALITATIVAS

Las tablas estadísticas muestran la información de forma esquemática y están preparadas para cálculos posteriores. La misma información estadística puede mostrarse de forma global y más expresiva, utilizando los gráficos estadísticos. Los gráficos poseen un fuerte poder de comunicación de los resultados de un estudio estadístico.

Detallamos, a continuación, los principales gráficos que permiten describir variables cualitativas:

a) Diagrama de barras

Consiste en dibujar un rectángulo por cada uno de los valores de la variable (x_i), de modo que las bases sean todas iguales, y la altura de cada rectángulo puede ser la frecuencia absoluta f_i o la frecuencia relativa h_i .

Ejemplo: En una empresa se desea conocer el color de ojos de sus empleados, se observa a los 50 empleados y se obtienen los siguientes resultados:

| Color ojos | Empleados |
|------------|-----------|
| Negros | 14 |
| Marrones | 24 |
| Verdes | 4 |
| Azules | 8 |

El diagrama de barras asociado es:

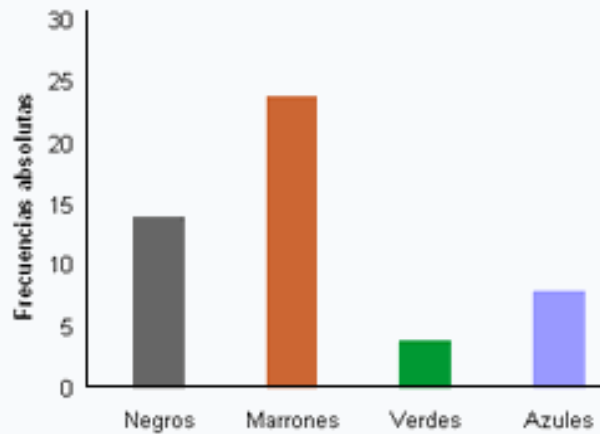


Diagrama de sectores

Consiste en dividir un círculo en sectores circulares, uno para cada x_i . El ángulo de cada sector será proporcional a la frecuencia y se calcula con una regla de tres simple.

Ejemplo: En una clase de 30 alumnos, 12 juegan a baloncesto, 3 practican la natación, 9 juegan al fútbol y el resto no practica ningún deporte.

Calculamos los ángulos correspondientes a cada variable con su frecuencia absoluta

$$\alpha_1 = \frac{360^\circ}{30} \cdot 12 = 144^\circ$$

$$\alpha_2 = \frac{360^\circ}{30} \cdot 3 = 36^\circ$$

$$\alpha_3 = \frac{360^\circ}{30} \cdot 9 = 108^\circ$$

$$\alpha_4 = \frac{360^\circ}{30} \cdot 6 = 72^\circ$$

| | Alumnos | Ángulo |
|-------------|---------|--------|
| Baloncesto | 12 | 144° |
| Natación | 3 | 36° |
| Fútbol | 9 | 108° |
| Sin deporte | 6 | 72° |
| Total | 30 | 360° |

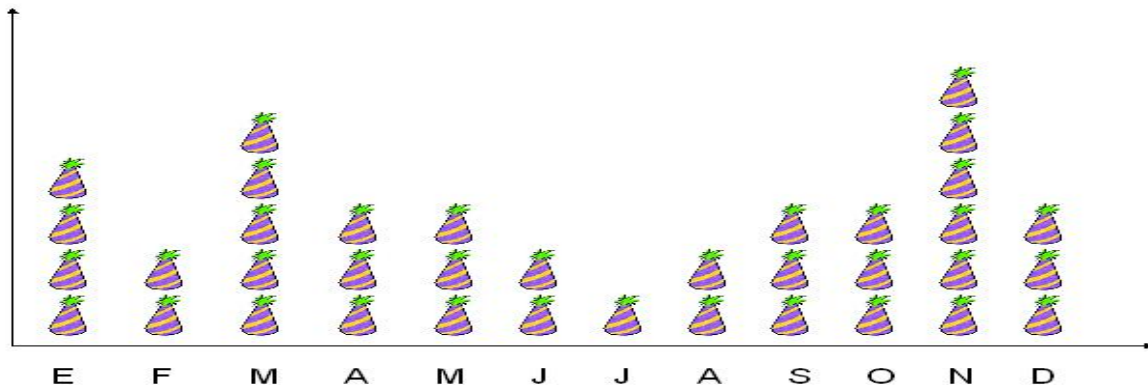
Diagrama de sectores




b) Pictograma

Consiste en realizar dibujos alusivos a la distribución que se desea representar. En muchas ocasiones son gráficos poco precisos, aunque fáciles de interpretar a simple vista.

Ejemplo: De un grupo de 1º bachillerato obtenemos los meses de sus cumpleaños y lo representamos en un pictograma

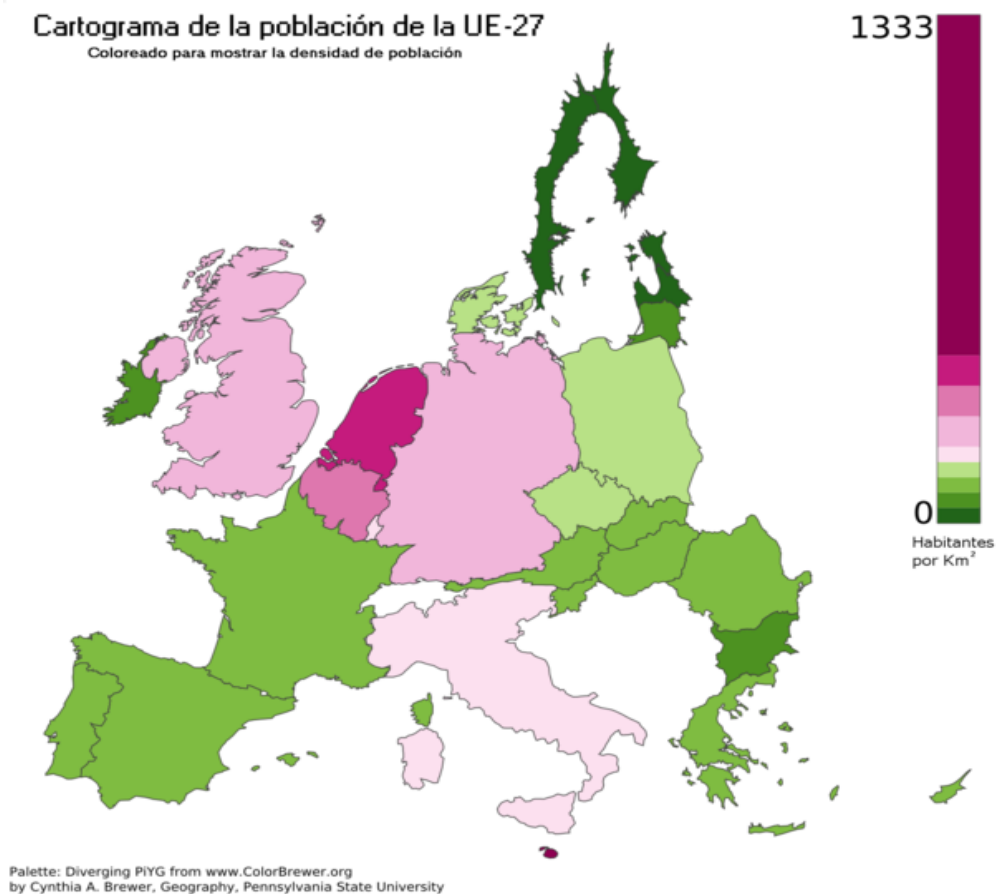


Cada  representa a un niño o niña que está de cumpleaños ese mes. Así entonces en marzo, 5 niños o niñas están de cumpleaños.

c) Cartogramas

Consiste en representar sobre un mapa cualquier tipo de datos relacionados con un área geográfica.

Ejemplo:



6 GRÁFICOS PARA VARIABLES ESTADÍSTICAS CUANTITATIVAS

Los gráficos más utilizados para representar distribuciones de variable cuantitativas, tanto discretas como continuas, son los que se describen a continuación.

a) Diagrama de barras o de columnas

Representan distribuciones de variables discretas por medio de barras o de columnas independientes, situadas encima de la variable representada. En muchas ocasiones se superponen dos o más diagramas con el fin de comparar los datos de diferentes situaciones.

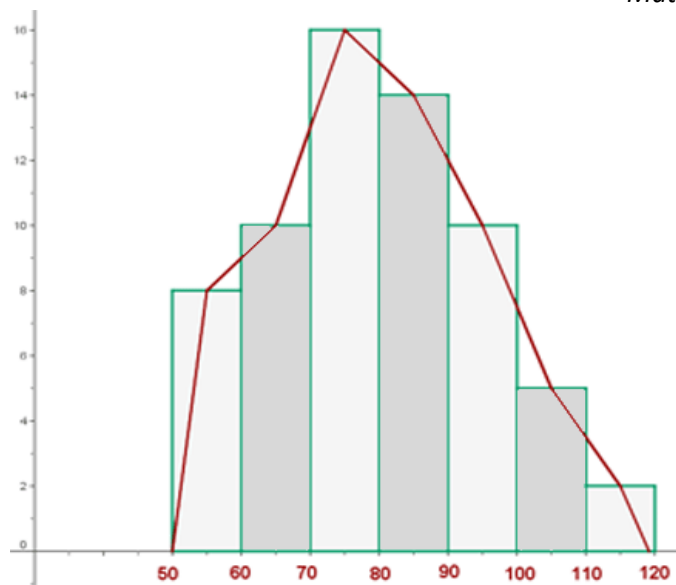
Ya lo hemos visto en el apartado anterior.

b) Diagrama de frecuencias (o polígono de frecuencias)

Se obtiene uniendo los extremos más altos de las barras o columnas mediante una línea quebrada.

Ejemplo: El peso de 65 personas adultas viene dado por la siguiente tabla:

| | X_i | f_i | F_i |
|------------|-------|-------|-------|
| [50, 60) | 55 | 8 | 8 |
| [60, 70) | 65 | 10 | 18 |
| [70, 80) | 75 | 16 | 34 |
| [80, 90) | 85 | 14 | 48 |
| [90, 100) | 95 | 10 | 58 |
| [100, 110) | 110 | 5 | 63 |
| [110, 120) | 115 | 2 | 65 |
| | | 65 | |



c) Histogramas

Son análogos a los diagramas de barras o columnas, pero para variables cuantitativas continuas

Consisten en rectángulos cuyas bases son cada uno de los intervalos y la altura es la frecuencia absoluta correspondiente a dicho intervalo.

Ejemplo: Se ha medido la talla de 200 niños de 6 años, y que se han obtenido valores entre 100 y 130 cm, con la siguiente distribución de frecuencias en intervalos de amplitud 2 cm:

| Intervalo | Frecuencia absoluta (n) | Frecuencia relativa (%) | Intervalo | Frecuencia absoluta (n) | Frecuencia relativa (%) |
|------------|-------------------------|-------------------------|--------------|-------------------------|-------------------------|
| [100, 102) | 2 | 1,0 | [116, 118) | 28 | 14,0 |
| [102, 104) | 2 | 1,0 | [118, 120) | 26 | 13,0 |
| [104, 106) | 2 | 1,0 | [120, 122) | 21 | 10,5 |
| [106, 108) | 6 | 3,0 | [122, 124) | 14 | 7,0 |
| [108, 110) | 11 | 5,5 | [124, 126) | 13 | 6,5 |
| [110, 112) | 14 | 7,0 | [126, 128) | 8 | 4,0 |
| [112, 114) | 25 | 12,5 | [128, 130] | 3 | 1,5 |
| [114, 116) | 25 | 12,5 | Total | 200 | 100,0 |

El histograma correspondiente es:

