



ESTADÍSTICA



"analysis-laptop-business-style-cartoon-graph-1449661" (Mohamed Hassan, CCO)



ÍNDICE

ESTADÍSTICA

1. ESTUDOS ESTADÍSTICOS.....	1
1.1 Fases dun estudo estatístico.....	1
2. POBOACIÓN E VARIABLES.....	1
<i>Exercicios</i>	2
3. ORGANIZACIÓN DOS DATOS.....	3
<i>Exercicios</i>	3
4. GRÁFICOS ESTADÍSTICOS.....	4
5. PARÁMETROS ESTADÍSTICOS.....	6
5.1 Parámetros de centralización.....	6
Media aritmética.....	6
Moda.....	6
Mediana.....	6
5.2 Parámetros de dispersión.....	6
Varianza.....	6
Desviación típica.....	6
<i>Exercicios</i>	6
SOLUCIÓNS.....	7



1. ESTUDOS ESTADÍSTICOS

A estatística é a parte das matemáticas que se encarga de estudar e analizar datos para extraer conclusións.

1.1 Fases dun estudo estatístico

Todo estudo estatístico consta dunha serie de fases e tarefas ben diferenciadas:

- 1) Definición da poboación e da característica que imos estudar: nesta fase temos que fixar a poboación a estudar, a característica, que e como imos recoller os datos etc.
- 2) Selección da mostra: necesitamos establecer o tamaño da mostra.
- 3) Recollida de datos: debemos deseñar o cuestionario para realizar a recollida dos datos.
- 4) Organización e representación gráfica: estudaremos os datos recollidos, faremos táboas de frecuencias, gráficas para visualizar mellor os datos etc.
- 5) Obtención de conclusións.

2. POBOACIÓN E VARIABLES

A poboación é o colectivo sobre o que se quere facer o estudo. Cando a poboación é demasiado grande para ser estudada, selecciónase unha mostra.

Unha mostra é un subgrupo da poboación que permite obter características de toda a poboación.

Un individuo é cada un dos elementos da poboación ou a mostra.

Sobre esta poboación ou mostra estúdanse unha serie de propiedades ou características que reciben o nome de variables estatísticas. Clasifícanse en:

- Variable cualitativa: se non se pode representar numericamente.
- Variable cuantitativa: cando se pode representar numericamente. Poden ser discretas (toman un número finito de valores numéricos) e continuas (poden tomar, teoricamente, todos os valores dentro dun intervalo).



EXERCICIOS

Exercicio 1

Clasifica as seguintes variables en cualitativas, cuantitativas discretas ou continuas: cor favorito, temperatura, número de discos vendidos, números de aprobados, estado civil, persoas que asisten a un concerto, altura dos alumnos dunha clase e profesión.

Exercicio 2

Dos 2.000 individuos dunha aldea escollemos unha mostra formada por 50 individuos aos que se lles pregunta o número de veces que van ao cine nun mes.

- Cal é a poboación de estudo?
- Que variable imos que estudar?
- Que tipo de variable é?

Exercicio 3

Queremos realizar un estudo dos homes, maiores de idade, dunha determinada cidade. Preguntouse a 32 homes cal é a súa altura.

- Cal é a poboación de estudo?
- Que variable temos que estudar?
- Que tipo de variable é?

3. ORGANIZACIÓN DOS DATOS

Unha vez recollidos os datos hai que tabulalos, é dicir, confeccionar unha táboa para organizalos. A dita táboa chamámola táboa de frecuencias.

Unha táboa de frecuencias pode conter os seguintes elementos:

- Frecuencia absoluta, f_i : número de veces que se repite un dato.
- Frecuencia relativa, h_i : cociente entre a frecuencia absoluta, f_i , e o número total de datos, N .
- Frecuencia absoluta acumulada, F_i : suma de todas as frecuencias absolutas dos valores menores ou iguais có valor escollido.
- Frecuencia relativa acumulada, H_i : suma de todas as frecuencias relativas dos valores menores ou iguais có valor escollido.

Para construír unha táboa de frecuencias, existe unha lixeira diferenza entre unha variable discreta e unha continua.

Nas variables continuas, agruparemos os datos en intervalos (ou clases) e para traballar con eles colleremos o punto medio dos extremos. A este punto chamámolo marca de clase (x_i). O número de clases non debe ser excesivo (entre 4 e 10). Recoméndase, ademais, que os intervalos teñan a mesma lonxitude.

EXERCICIOS

Exercicio 4

Dos 2.000 individuos dunha aldea escollemos unha mostra formada por 50 individuos aos que se lles pregunta o número de veces que van ao cine nun mes e as respostas foron:

0 1 2 1 0 2 0 1 1 1 1 0 0 1 0 3 1 1 1 1 0 1 0 1 1 1 1 2 1 0 2 1 1 0 1 1 1 1 1 0 1 1 1 1 2 1 1 1 1 1

- Constrúa unha táboa de frecuencias.
- Que porcentaxe de individuos van como moito unha vez ao mes ao cine?

Exercicio 5

Queremos realizar un estudo dos homes, maiores de idade, dunha determinada cidade. Preguntouse a 32 homes, que verificaban os requisitos, cal é súa altura en cm e os resultados foron os seguintes: 161, 168, 166, 171, 167, 172, 170, 167, 170, 165, 169, 170, 169, 172, 162, 169, 166, 174, 178, 169, 162, 168, 167, 178, 175, 168, 164, 179, 172, 167, 170, 173. Constrúe unha táboa de frecuencias.

4. GRÁFICOS ESTADÍSTICOS

Os gráficos estadísticos axudan a organizar e interpretar dun xeito máis visual os datos recollidos. Existen diferentes gráficos estadísticos:

- **Diagrama de barras:** utilízanse para representar frecuencias de variables cuantitativas discretas. As barras son estreitas e sitúanse sobre os valores puntuais da variable. Tamén se poden utilizar para representar variables cualitativas.
- **Histograma:** emprégase para representar variables cuantitativas continuas. En lugar de barras, utilízanse rectángulos cuxo ancho é o ancho dos intervalos da variable.
- **Polígono de frecuencias:** Constrúense unindo os puntos medios dos rectángulos dos histogramas ou das barras dos diagramas de barras.
- **Diagrama de sectores:** É un círculo dividido en tantos sectores como datos teña a variable, sendo a súa amplitude proporcional á súa frecuencia.

Diagrama de barras

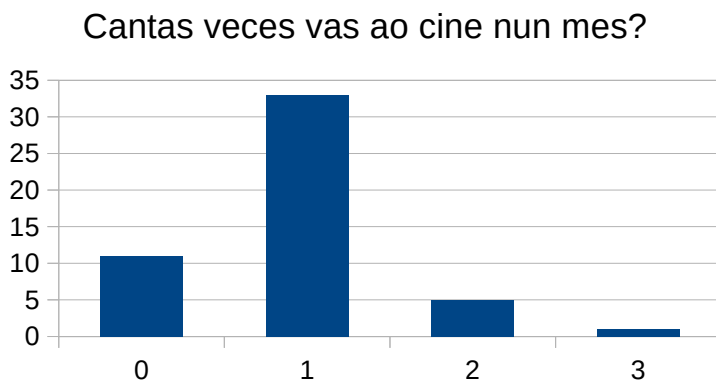
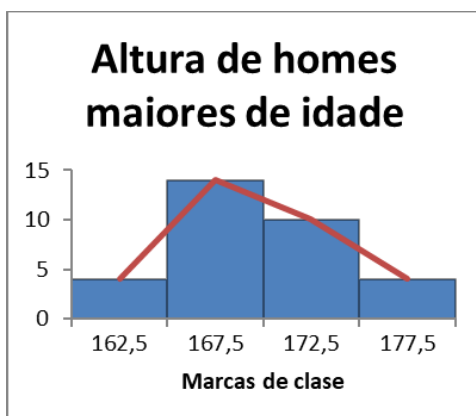


Diagrama de barras

No eixe de abscisas (o horizontal) escríbense os datos da variable e no de ordenadas (o vertical) as frecuencias absolutas.

Histograma e polígono de frecuencias



Histograma e polígono de frecuencias

Divídese o eixe de abscisas en intervalos e levántase un rectángulo en cada tramo con altura a súa frecuencia.

Diagrama de sectores

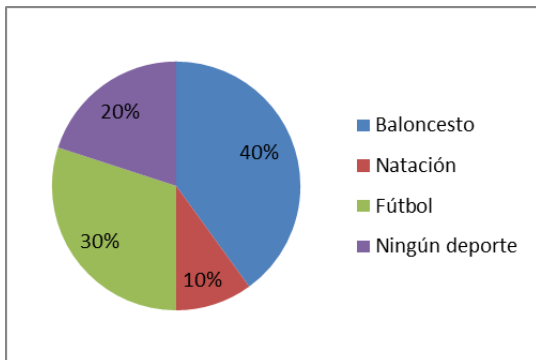


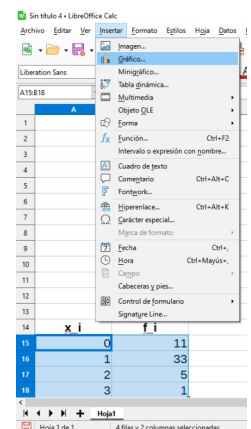
Diagrama de sectores

Dividimos o círculo en tantos sectores como datos teña a variable, sendo a súa amplitude proporcional á frecuencia relativa.

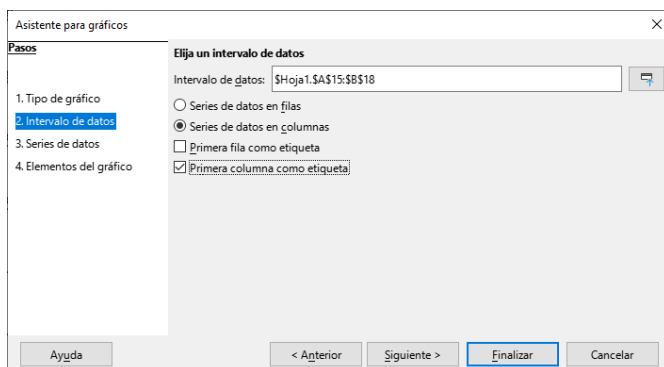
Para facilitar a súa construción, pódese engadir unha columna á táboa de frecuencias onde se multiplique a frecuencia relativa por 360°. Os resultados obtidos son as amplitudes dos sectores.

Podemos crear gráficos utilizando follas de cálculo.

- Escribimos os datos da variable e das frecuencias.
- Seleccionamos todos os datos.
- Pinchamos Inserir→Gráfico
- No tipo de gráfico seleccionamos a columna para facer o diagrama de barras ou un histograma, circular para facer un diagrama de sectores ou liña e columna para facer un diagrama de barras con polígono de frecuencias. Para un histograma, unha vez creado o gráfico de columnas, pinchamos co botón dereito sobre o gráfico, formato de series de datos e espazado 0%.
- Picamos en seguinte e marcamos series de datos en columnas e primeira columna como etiqueta.



Inserir un gráfico



Seleccionar os datos

5. PARÁMETROS ESTADÍSTICOS

Despois de obter os datos dunha distribución, necesitamos sintetizar a información para a súa posterior análise. Para iso, obteremos os parámetros estatísticos.

5.1 Parámetros de centralización

Os parámetros de centralización indícanos arredor de que valor (centro) se distribúen os datos; son a media aritmética, a moda, e a mediana.

Media aritmética

A media (\bar{x}) é o cociente de todos os datos multiplicados pola súa frecuencia entre o número total de datos.

$$\bar{x} = \frac{\sum f_i \cdot x_i}{N}$$

Moda

A moda (M_o) é o dato ou datos que aparecen con maior frecuencia. Se a variable é continua fálase de intervalo modal.

Mediana

A Mediana (M_e) é o valor que ocupa a posición central dos datos, despois de ordenalos, se o número de datos é impar, ou a media dos datos centrais, se o número é par.

5.2 Parámetros de dispersión

Varianza

A varianza (σ^2) utilízase para analizar como de dispersos están os datos con respecto á media.

$$\sigma^2 = \frac{\sum f_i \cdot (x_i - \bar{x})^2}{N} = \frac{f_i \cdot x_i^2}{N} - \bar{x}^2$$

Desviación típica

A desviación típica (σ) é a raíz cadrada da varianza. Ten as mesmas unidades que a media.

EXERCICIOS

Exercicio 6

Obtén os parámetros estatísticos das seguintes notas: 2, 4, 4, 4, 5, 7, 9, 9, 10. Fai o mesmo cos datos dos exercicios 4 e 5.



SOLUCIÓNS

Exercicio 1

Clasifica as seguintes variables en cualitativas, cuantitativas discretas ou continuas: cor favorito, temperatura, número de discos vendidos, números de aprobados, estado civil, persoas que asisten a un concerto, altura dos alumnos dunha clase e profesión.

- Variables cualitativas: cor favorito, estado civil, profesión.
- Variables cuantitativas
 - Discretas: número de discos vendidos, números de aprobados, persoas que asisten a un concerto.
 - Continuas: temperatura e altura dos alumnos dunha clase.

Exercicio 2

Dos 2000 individuos dunha aldea escollemos unha mostra formada por 50 individuos aos que se lles pregunta o número de veces que van ao cine nun mes.

- Cal é a poboación de estudo? As 2000 persoas da aldea.
- Que variable teremos que estudar? O número de veces que van ao cine por mes.
- Que tipo de variable é? Variable cuantitativa discreta.

Exercicio 3

Queremos realizar un estudo dos homes, maiores de idade, dunha determinada cidade. Preguntouse a 32 homes, que verificaban os requisitos, a súa altura.

- a) Cal é a poboación de estudo? Os homes maiores de idade dunha determinada cidade.
- b) Que variable temos que estudar? A altura.
- c) Que tipo de variable é? Variable cuantitativa continua

Exercicio 4

Dos 2000 individuos dunha aldea escollemos unha mostra formada por 50 individuos aos que se lles pregunta o número de veces que van ao cine nun mes e as respostas foron:

0 1 2 1 0 2 0 1 1 1 1 0 0 1 0 3 1 1 1 1 0 1 0 1 1 1 1 2 1 0 2 1 1 0 1 1 1 1 1 0 1 1 1 1 2 1 1 1 1 1

a) Constrúa unha táboa de frecuencias.

x_i	Reconto	f_i	F_i	h_i	H_i
0	III III I	11	11	$\frac{11}{50}=0,22$	0,22
1	III III III III III III III	33	11+33=44	$\frac{33}{50}=0,66$	0,22+0,66=0,88
2	IIII	5	44+5=49	$\frac{5}{50}=0,10$	0,88+0,1=0,98
3	I	1	49+1=50	$\frac{1}{50}=0,02$	0,98+0,02=1
		50		1	

b) Que porcentaxe de individuos van como moito unha vez ao mes ao cine?

Un 88% (Miramos o valor de 1 na columna H_i).

Exercicio 5

Queremos realizar un estudo dos homes, maiores de idade, dunha determinada cidade. Preguntouse a 32 homes, que verificaban os requisitos, cal é súa altura en cm e os resultados foron os seguintes: 161, 168, 166, 171, 167, 172, 170, 167, 170, 165, 169, 170, 169, 172, 162, 169, 166, 174, 178, 169, 162, 168, 167, 178, 175, 168, 164, 179, 172, 167, 170, 173. Constrúe unha táboa de frecuencias.

Intervalos	Marca de clase (x_i)	Reconto	f_i	F_i	h_i	H_i
[160,165)	162,5	IIII	4	4	4/32=0,125	0,125
[165,170)	167,5	III III IIII	14	18	14/32=0,4375	0,5625
[170,175)	172,5	III III	10	28	10/32=0,3125	0,875
[175,180)	177,5	IIII	4	32	4/32=0,125	1
			32		1	

Exercicio 6

Obtén os parámetros estatísticos das seguintes notas: 2, 4, 4, 4, 5, 7, 9, 9, 10. Fai o mesmo cos datos dos exercicios 4 e 5.

- Media: $\bar{x} = \frac{2+4+4+4+5+7+9+9+10}{9} = \frac{54}{9} = 6$
- Moda: $Mo = 4$
- Mediana: $Me = 5$
- Varianza: $\sigma^2 = \frac{1 \cdot 2^2 + 3 \cdot 4^2 + 1 \cdot 5^2 + 1 \cdot 7^2 + 2 \cdot 9^2 + 1 \cdot 10^2}{9} - 6^2 = \frac{388}{9} - 36 = 7,11$
- Desviación típica: $\sigma = \sqrt{7,11} = 2,67$

Exercicio 4

x_i	f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
0	11	11	$0 \cdot 11 = 0$	$0^2 \cdot 11 = 0$
1	33	44	$1 \cdot 33 = 33$	$1^2 \cdot 33 = 33$
2	5	49	$2 \cdot 5 = 10$	$2^2 \cdot 5 = 20$
3	1	50	$3 \cdot 1 = 3$	$3^2 \cdot 1 = 9$
Total	50		46	62

- Media: $\bar{x} = \frac{46}{50} = 0,92$
- Moda: $Mo = 1$
- Mediana. Os valores centrais son o 25 e o 26. Ambos valen 1 porque a frecuencia acumulada de 1 é 44 e a de 0 é 11: $Me = \frac{1+1}{2} = 1$
- Varianza: $\sigma^2 = \frac{62}{50} - 0,92^2 = 0,3936$
- Desviación típica: $\sigma = \sqrt{0,3936} = 0,6274$

Exercicio 5

Intervalos	Marca de clase (x_i)	f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[160,165)	162,5	4	4	$162,5 \cdot 4 = 650$	$162,5^2 \cdot 4 = 105625$
[165,170)	167,5	14	18	$167,5 \cdot 14 = 2345$	$167,5^2 \cdot 14 = 392787,5$
[170,175)	172,5	10	28	$172,5 \cdot 10 = 1725$	$172,5^2 \cdot 10 = 297562,5$
[175,180)	177,5	4	32	$177,5 \cdot 4 = 710$	$177,5^2 \cdot 4 = 126025$
		32		5430	922000

- Media: $\bar{x} = \frac{5430}{32} = 169,6875$
- Moda: $Mo = [165,170)$
- Mediana. Os valores centrais son o 16 e o 17. Ambos valen 167,5 porque a frecuencia acumulada de 167,5 é 18 e a de 162,5 é 4: $Me = \frac{167,5 + 167,5}{2} = 167,5$
- Varianza: $\sigma^2 = \frac{922000}{32} - 169,6875^2 = 18,6523$
- Desviación típica: $\sigma = \sqrt{18,6523} = 4,3188$