

## Obxectivos

Nesta quincena aprenderás a:

- Lembrar os conceptos de poboación, mostra, individuo e carácter.
- Valorar a importancia do concepto de variable estatística e distinguir entre os diferentes tipos.
- Resumir mediante unha táboa de frecuencias calquera serie de datos.
- Asociar e interpretar gráficos estatísticos valorando a súa utilización en diferentes áreas de coñecemento.
- Calcular, valorar e interpretar a media, mediana e moda en variable discreta.

### Antes de empezar

1. Vocabulario estatístico ..... páx. 4  
Poboación, mostra, individuo e carácter
2. Carácter. Variable estatística ..... páx. 6  
Carácter cualitativo. Atributos  
Variables discretas  
Variables continuas
3. Ordenación de datos. Tabulación .. páx. 10  
Para variable discreta  
Para variable cualitativa
4. Gráficos para variable cualitativa .. páx. 12  
Diagrama de barras  
Diagrama de sectores
5. Gráficos para variable discreta .... páx. 14  
Diagrama de barras  
Polígonos de frecuencias  
Diagrama de sectores
6. Medidas de centralización ..... páx. 17  
Media  
Mediana  
Moda
7. Medidas de dispersión ..... páx. 20  
Rango  
Desviación media

RESUMO

Autoavaliación



## Antes de empezar

A Estadística penetrou en múltiples aspectos da vida cotiá facendo familiares termos como poboación, mostra, media, mediana, moda...

Pode asegurarse que calquera persoa informada de hoxe en día posúe un vocabulario básico de estadística, enténdeo, utilízao e valora.

Practicamente todas as ciencias, tanto científico tecnolóxicas como sociais utilizan, en aspectos fundamentais destas, á estadística.

O deporte non é unha excepción. En todos eles e en particular no baloncesto o manexo dos datos estadísticos constitúe un aspecto a estudar e manexar tan importante ás veces como as tácticas e a técnica implícitas do propio xogo.

No exemplo seguinte simula un saque de fondo en baloncesto, represéntase con puntos vermellos os xogadores atacantes e os verdes como os defensores.

O estudo que realizan os corpos técnicos dos equipos encárgase de calcular que estatística de tiro ten cada xogador, deste xeito se se deixa desmarcado o xogador que teña peor estatística; o balón irá cara a el.



### 1. Vocabulario estatístico

#### Poboación, mostra, individuo e carácter.

As primeiras definicións necesarias para o inicio de calquera estudo estatístico son poboación, individuo, mostra e carácter.

Comecemos cun exemplo que nos faga intuír os devanditos conceptos.

##### Exemplo

##### Estudo sobre a posible existencia de vida noutras estrelas.

Existen sistemas planetarios semellantes ao noso que, quizais poidan albergar algún tipo de vida?

Ata fai moi pouco tempo, os astrónomos non tiñan probas da existencia de planetas fóra do Sistema Solar.

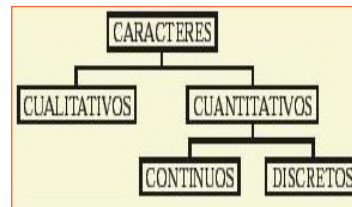
Na actualidade descubríronse algúns centenares de planetas xigantes, que dificilmente poderían encerrar vida, pero que si serían unha especie de sinal de que nesa estrela pode existir un sistema con órbitas e tamaños mais acordes coas posibilidades de vida no sentido que coñecemos.

- A poboación está constituída por todas as estrelas do universo visible.
- A mostra está constituída por todas as estrelas escollidas e observadas no proxecto.
- O individuo é cada estrela do universo observable.
- O carácter é a presenza ou non de perturbacións que indiquen a existencia de planetas xigantes.

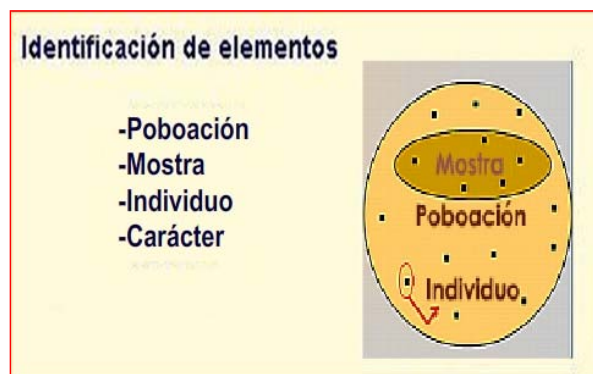
Podemos definir os conceptos anteriores do seguinte xeito:

- **Poboación:** Conxunto de todos os elementos que verifican unha característica que será obxecto de estudo.
- **Individuo:** Cada un dos elementos da poboación.
- **Mostra:** Calquera subconxunto da poboación. Este subconxunto é moi importante que sexa *representativo* da poboación.
- **Carácter:** Cada unha das propiedades que posúen os individuos da poboación e que poden ser obxecto de estudo.

A definición de carácter debe ir acompañada da seguinte clasificación:



Recordemos entón que ante calquera estudo estatístico debemos ter en conta a identificación dos elementos, desta forma evitaremos erros nas conclusións finais.



## Exemplo

### Estudo sobre a evolución da talla na xuventude española.

Os españois igualan en estatura á maioría dos europeos, pero evolucionan cara á obesidade norteamericana.

Un estudo antropométrico conxunto entre varios hospitais españois, revela que a estatura dos españois igualouse nos últimos trinta anos respecto á maioría dos países europeos.

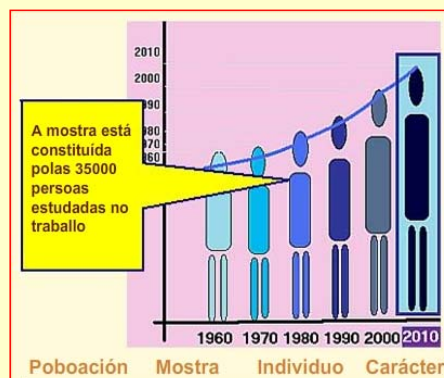
O mesmo estudo tamén alerta sobre a preocupante tendencia cara á obesidade en niveis similares a poboación norteamericana.

O traballo, levado a cabo mediante a medición de case 35000 suxeitos entre os anos 2000 e 2004, tamén demostrou que as diferenzas entre as distintas comunidades autónomas dentro de España son case inexistentes.

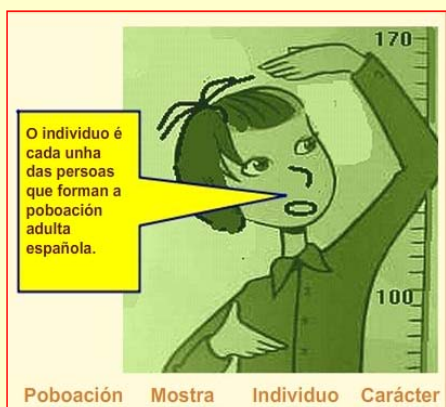
*Poboación*



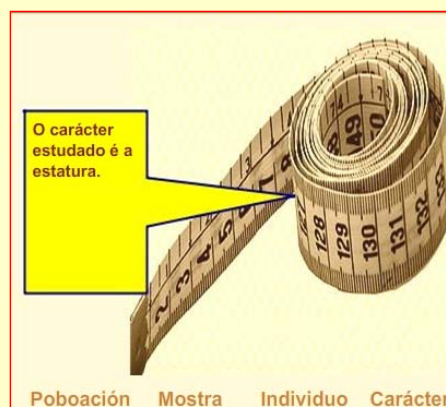
*Mostra*



*Individuo*



*Carácter*



## 2. Carácter. Variable estatística.

### Traballo de campo.

O traballo de campo é a etapa da investigación na que se establece contacto directo coa poboación ou mostra para solicitar os datos que se necesiten.

A planificación é fundamental e o seu desenvolvemento depende do método de obtención da información que se utilice.

O emprego do ordenador permite unha simulación de situacións que fai que realicemos un traballo de campo virtual sen desgaste físico.

### Carácter cualitativo. Atributos.

Comecemos novamente cun exemplo que nos ilustre.

#### Exemplo

##### Afección ao fútbol.

Preguntamos a unha serie de persoas sobre as súas preferencias en canto a afección futbolística.

A mostra que consideramos será de 9 persoas de distintas cidades españolas.

Os datos son F.C. Barcelona, Sevilla C.F., At. Bilbao, R. Madrid, R. Madrid, At. Madrid, Valencia C.F., F.C. Barcelona e Deportivo de A Coruña.

As características destes valores son:

- Non son medibles con números.
- Non ten sentido a ordenación.
- As distintos valores identifícanse co nome do equipo elixido.

Todos os individuos da poboación que imos estudar teñen unha serie de propiedades ou calidades que en estatística reciben o nome de caracteres.

Os caracteres poden ser de dous grandes tipos:

- a) CUALITATIVOS
- b) CUANTITATIVOS

Un carácter cualitativo caracterízase porque as súas diferentes modalidades non poden expresarse con números.

#### Exemplo

##### A túa cor preferida.

Preguntamos a unha serie de persoas sobre as súas preferencias en canto a cores.

Neste caso a simulación da poboación e cor elixida pódese realizar mediante o ordenador, existen programas que permiten xerar mostras aleatorias que simulan o traballo de campo.

A mostra sobre a que actuamos será de 10 persoas dunha cidade calquera.

DATOS OBTIDOS	
	branco
	verde
	amarelo
	verde
	laranja
	amarelo
	vermello
	vermello

#### CARACTERÍSTICAS:

- Os valores que toma non son medibles numericamente.
- Non ten sentido a ordenación.
- Non ten sentido falar de valores CONSECUTIVOS
- As distintas modalidades do carácter non representables numericamente, se identifican mediante o nome da cor elixida



## VARIABLES DISCRETAS. CARÁCTER CUANTITATIVO DISCRETO.

Denomínase así ao carácter cuxas modalidades se poden representar con números.

Dentro dos caracteres cuantitativos distínguense dous tipos: Discreto e continuo.

É discreto se toma valores illados, de maneira que entre dous consecutivos non existe outro intermedio.

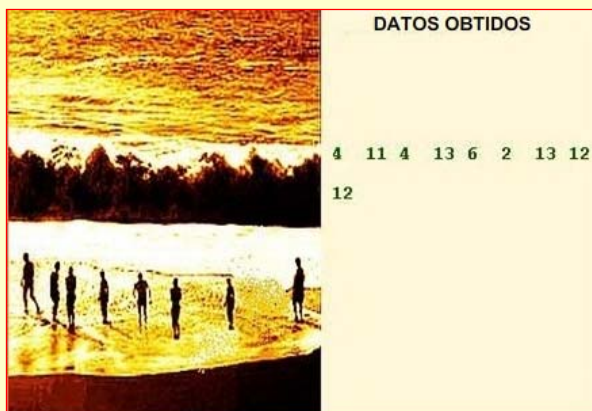
### Exemplo

#### Canta xente hai na praia?

Realizamos unha fotografía dunha determinada zona de praia a distintas horas do día e anotamos as persoas que aparecen nela.

Neste caso dispoñemos dun banco virtual de fotos e un procedemento totalmente aleatorio que simula as distintas situacións.

A mostra sobre a que actuamos é de 9 fotografías.



#### CARACTERÍSTICAS:

- Os valores que toma son illados: entre 1 e 13.
- Os valores pódense ordenar e contar
- Entre dous valores CONSECUTIVOS non existen valores intermedios.

#### -ESQUEMA



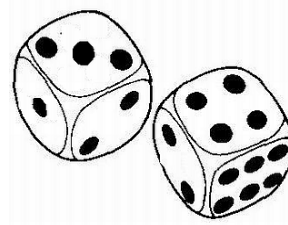
**Non hai valores intermedios**

### Exemplo

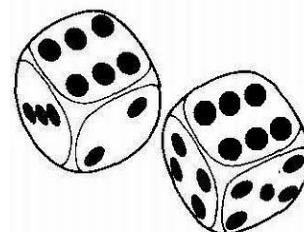
#### ¿Canto suman as caras superiores de dous dados previamente lanzados?

Lanzamos dous dados perfectos anotando a suma dos resultados das caras superiores.

A mostra que consideramos será a suma de 8 pares de lanzamentos.



$$3 + 4 = 7$$



$$6 + 6 = 12$$

Os datos obtidos son: 7, 6, 9, 2, 8, 1, 8 e 7.

As características destes valores son:

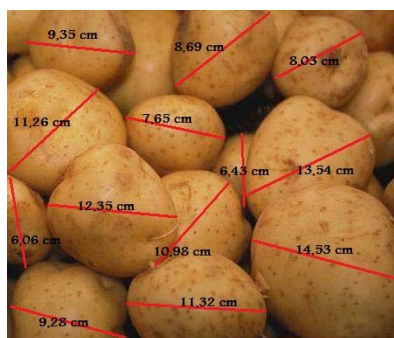
- Os valores que toman son illados; entre 2 e 12.
- Os valores pódense ordenar e contar.
- Entre dous valores consecutivos non existen valores intermedios.

## Variáveis contínuas. Carácter quantitativo contínuo.

Cando as modalidades dun carácter quantitativo poden tomar valores dun conxunto de números reais ou un intervalo, polo menos teoricamente, dise que estamos ante un carácter quantitativo.

### Exemplo

Medindo patacas. Diámetro medio.



Para unha posterior clasificación de calidade realízase un estudo sobre o diámetro medio en distintas producións de patacas. A mostra que consideramos será de 8 producións.

Todo o, neste caso, custoso procedemento de recollida das mostras substituímos por unha simulación por ordenador.

Os datos obtidos son: 14,6 ; 6,7 ; 9,8 ; 13,2 ; 8,1 ; 9,3; 13,8 e 10,1.

As características que se teñen sobre os valores son:

- Entre dous valores sempre existe a posibilidade doutro.
- Non ten sentido falar de valores consecutivos.
- Toma valores dentro dun intervalo.

### Exemplo

#### Pesando neonatos.

Imos preguntar o peso dos neonatos nunha determinada cidade.

Existen programas informáticos que permiten xerar mostras aleatorias que simulan o traballo de campo.

A mostra sobre a que actuamos é de 40 bebés.

#### DATOS OBTIDOS

2,28	3,22	3,95	2,43	3,59
2,47	3,41	3,50	2,07	2,13
3,55	2,61	3,21	2,29	4,07
2,33	2,51	3,38	3,67	4,10
3,84	2,31	2,63	4,03	2,95
3,86	4,03	3,02	2,99	4,00
3,72	2,98	3,88	3,07	3,28
3,84	3,37	3,55	4,07	2,94

#### CARACTERÍSTICAS:

- Entre dous valores sempre existe a posibilidade de outro.
- Non ten sentido falar de valores CONSECUTIVOS
- Toma valores dentro do intervalo [2,07 , 4,10]
- INTERVALO DE VALORES





## Exercicios resoltos

1. Clasifica as seguintes variables: cualitativas, discreta ou continua, escribindo unha X no recadro correspondente.

Nº de fillos varóns, tipo de música preferida, nº de fillos, peso de neonatos, páxinas dun libro, estatura.

SOLUCIÓN:

	CUALITATIVA	DISCRETA	CONTINUA
Nº de fillos varóns		X	
Tipo de música preferida	X		
Nº de fillos		X	
Peso de neonatos			X
Páxinas dun libro		X	
Estatura			X

2. Clasifica as seguintes variables: cualitativas, discreta o continua, escribindo unha X no recadro correspondente.

Raza de cans, nº de fillos, lonxitude do pé, materias pendentes, perímetro cranial, cantante favorito.

SOLUCIÓN:

	CUALITATIVA	DISCRETA	CONTINUA
Raza de cans	X		
Nº de fillos		X	
Lonxitude do pé			X
Materias pendentes	X		
Perímetro cranial			X
Cantante favorito	X		

Observa:

A variable "materias pendentes" fai referencia ao nome das materia e por iso é variable cualitativa, mentres a variable "nº de materias pendentes" sería discreta.

## 3. Ordenación de datos.

Ao final poderás comprobar que...

### PROPIEDADES INTERESANTES DAS TÁBOAS ESTADÍSTICAS

- A suma de todas as frecuencias absolutas é igual ao tamaño da poboación ou da mostra.
- A suma das frecuencias relativas é sempre igual a 1.

*Se se realizou algún redondeo nas frecuencias relativas é usual que a súa suma non sexa exactamente igual a un debido aos erros cometidos..*

### Exemplo

#### Idade dos estudantes.

As idades de 30 estudantes dun instituto de ensino secundario dan os valores que posteriormente tabulamos como segue:

$X_i$	$f_i$	$h_i$	$F_i$	$H_i$
12	2	$\frac{2}{30}$	2	$\frac{2}{30}$
13	6	$\frac{6}{30}$	8	$\frac{8}{30}$
14	5	$\frac{5}{30}$	13	$\frac{13}{30}$
15	3	$\frac{3}{30}$	16	$\frac{16}{30}$
16	10	$\frac{10}{30}$	26	$\frac{26}{30}$
17	4	$\frac{4}{30}$	30	$\frac{30}{30}$

## Tabulación para variable discreta.

O paso seguinte ao traballo de campo é a disposición dos datos de xeito ordenado, conciso e visualmente atractivo.

En estatística, este proceso recibe o nome de tabulación.

Os valores obtidos ordénanse, especifican e agrupan de tal forma que sexa doado a información e busca.

As primeiras columnas que deben aparecer serán:

- Valores de la variable,  $X_i$ .
- Frecuencias absolutas,  $f_i$ .
- Frecuencias relativas,  $h_i$ .
- Frecuencias absolutas, acumuladas,  $F_i$ .
- Frecuencias relativas acumuladas,  $H_i$ .

Nalgúns casos pódese utilizar a porcentaxe en lugar das frecuencias relativas ou ademais das frecuencias relativas.

## Tabulación para variable cualitativa.

Nos casos de carácter cualitativo, a tabulación dos datos é moi simple. As tres columnas que teñen sentido fan referencia a:

- O valor dos atributos.
- A frecuencia absoluta
- A frecuencia relativa

### Exemplo

#### A práctica de deporte.

Recollida de datos sobre deportes practicados, tabulada:

fútbol, tenis, balonmán, tenis, voleibol, atletismo, baloncesto, fútbol, fútbol, balonmán, fútbol, voleibol, balonmán, fútbol, balonmán, fútbol, fútbol, tenis, atletismo.

$X = \text{deporte}$	$f$	$h$
baloncesto	2	2/20
balonmán	4	4/20
voleibol	2	2/20
tenis	3	3/20
atletismo	2	2/20
fútbol	7	7/20
	20	

### Exercicios resoltos

3. Para un estudo de accesibilidade, durante 30 días anotamos o número de prazas libres de aparcamento ás 5 da tarde.

1 2 1 2 0 1 3 2 1 5 0 2 2 1 3  
 3 2 1 1 5 0 5 3 0 3 3 2 2 3 1

Realiza unha tabulación dos datos na que aparezan as columnas correspondentes ás frecuencias absolutas, relativas, acumuladas absolutas e relativas.

SOLUCIÓN:

Nº de prazas de aparcamento	f	h	F	H
0	4	0,14	4	0,14
1	8	0,28	12	0,41
2	8	0,28	20	0,69
3	7	0,24	27	0,93
4	0	0	27	0,93
5	3	0,1	30	1,03
	30			

4. Preguntamos a 20 estudantes elixidos aleatoriamente polo tipo de música que prefiren escoitar.

Os resultados son: disco, rock, rock, clásica, rock, latina, pop, rock, latina, rock, flamenco, flamenco, flamenco, latina, rock, clásica, disco, disco, latina, rock.

Realiza unha tabulación dos datos na que aparezan as columnas correspondentes ás frecuencias absolutas e relativas.

SOLUCIÓN:

Tipo de música	F	h
Disco	3	0,15
Rock	7	0,35
Latina	4	0,20
Clásica	2	0,10
Flamenco	3	0,15
Pop	1	0,05
	20	

## 4. Gráficos para unha variable cualitativa.

### Diagrama de barras.

O diagrama de barras é xunto ao de sectores o gráfico máis utilizado para variable cualitativa.

Utilízase como complemento á táboa de frecuencias ou mesmo en substitución desta.

No eixe de abscisas sitúanse a igual distancia os distintos atributos.

A partir de cada atributo levántanse barras de igual grosor e a altura das cales sexa a da correspondente frecuencia absoluta.

### Diagrama de sectores.

O diagrama de sectores en variables cualitativas é un dos recursos estatísticos máis utilizados.

É especialmente útil nos casos en que existen poucas modalidades do carácter. Adóitase utilizar xunto á táboa de frecuencias ou substituíndo a esta.

Para calcular o ángulo do sector que corresponde a cada valor da frecuencia:

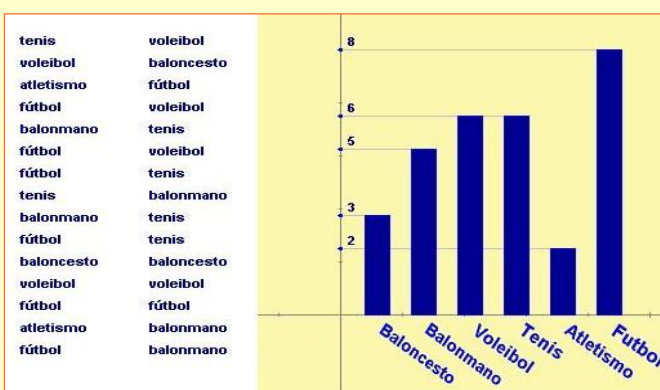
$$x = \frac{f_i \cdot 360}{\text{Total Datos}}$$

### Exemplos

#### Deporte practicados. (Diagrama de barras)

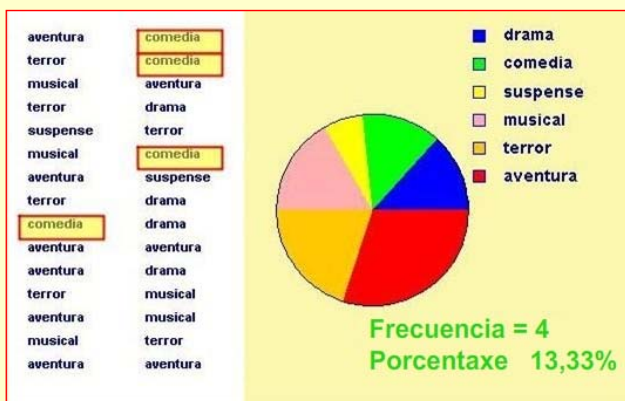
Os datos corresponden ás contestacións realizadas por 30 estudantes sobre o deporte que practicaban con maior frecuencia no instituto.

Se queremos ter unha rápida visión dos datos, unha forma de organizalos é a través dunha representación de diagrama de barras. Neste exemplo podes ver a diferenza entre facer unha análise sobre o listado ou sobre a gráfica. Cal che resulta máis doado?.



#### Tipo de película. (Diagrama de sectores)

Volvemos preguntar aos nosos estudantes sobre o tipo de película que lles gusta ver. Outra forma de organizalos de forma máis doada de ver é o diagrama de sectores. Serías capaz de recordar algunha outro exemplo? (Axuda: acontece cada catro anos).



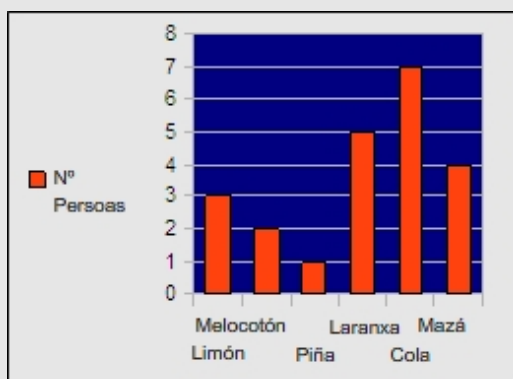
## Exercicios resoltos

5. Os datos corresponden ás respostas realizadas por 22 persoas elixidas aleatoriamente, acerca do sabor preferido nos refrescos dunha determinada marca.

Laranxa, mazá, cola, laranxa, limón, cola, melocotón, cola, limón, cola, cola, mazá, limón, laranxa, cola, piña, mazá, laranxa, cola, laranxa, mazá e melocotón.

Debuxa o diagrama de barras que representa os datos anteriores.

SOLUCIÓN:

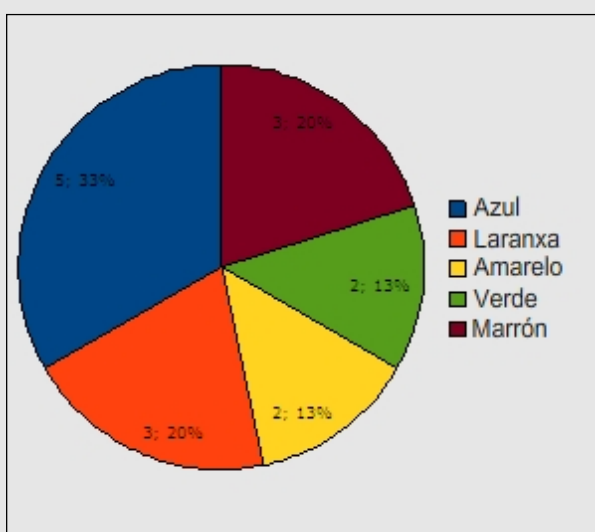


6. Os resultados corresponde ás contestacións realizadas por 15 estudantes acerca de cal é a súa cor preferida.

As respostas que deron son: azul, marrón, laranxa, amarelo, azul, laranxa, verde, verde, azul, marrón, azul, laranxa, amarelo, marrón, e azul.

Debuxa o diagrama de sectores que representa os datos anteriores.

SOLUCIÓN:



## 5. Gráficos para unha variable discreta.

### Diagrama de barras.

É o gráfico estatístico máis utilizado para variables discretas. Para elaborar o diagrama, sitúanse no eixe de abscisas os valores correspondentes da variable.

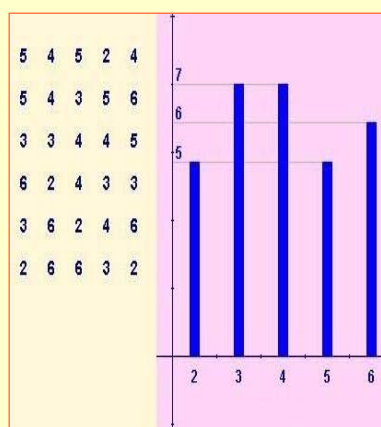
A partir de cada valor levántanse barras do mesmo grosor e a altura das cales será a correspondente a cada frecuencia.

#### Exemplo

##### Faltas de ortografía.

A profesora anotou o número de faltas de ortografía dos seus estudantes.

Quere unha representación que lle permita ver os datos rapidamente, sabendo cuántos estudantes comenten un número determinado de faltas de ortografía.



### Polígonos de frecuencias.

O polígono de frecuencias constrúese a partir do diagrama de barras, unindo os puntos medios da base superior dos rectángulos que constitúen as barras.

Se se constrúe un diagrama de barras considerando en lugar das frecuencias as frecuencias acumuladas e unimos os puntos medios das bases superiores mediante segmentos, obtemos unha poligonal crecente que denominamos *polígono de frecuencias acumuladas*.

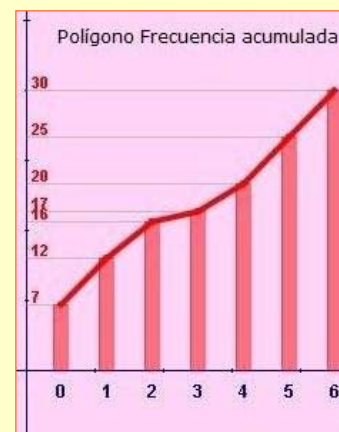
#### Exemplo

##### Número de chamadas.

Unha empresa de telecomunicacións quere facer un estudo sobre os seus clientes, vendo o número de chamadas que recibe un grupo destes.

O estudo realízase sobre 30 persoas, anotando o número de chamadas recibidas nun día.

$X_i$	$f_i$	$F_i$
0	7	7
1	5	12
2	4	16
3	1	17
4	3	20
5	5	25
6	5	30





## Diagrama de sectores.

Vexamos primeiro un exemplo.



Como complemento á tabulación e ás veces substituíndo a esta, en Estadística é moi habitual recorrer a gráficos, o efecto visual directo dos cales amosa as primeiras características dunha distribución estatística.

Para variables cuantitativas discretas, así como para as cualitativas, os gráficos que se utilizan con mais frecuencia son:

1. O diagrama de barras
2. O diagrama de sectores

Non obstante, depende do tipo de información que queiramos obter, ás veces, resulta útil realizar o polígono de frecuencias, e o polígono de frecuencia acumuladas.

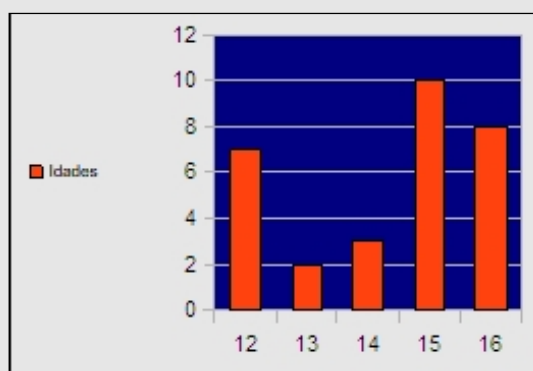
## Exercicio resolto

7. As idades de 30 estudantes dun instituto de ensino secundario son as seguintes:

15 15 16 15 16 16 16 16 16 12 13 12 15 16 14  
12 14 12 15 13 14 16 15 15 12 15 12 15 15 12

Representa o diagrama de barras correspondente:

SOLUCIÓN:



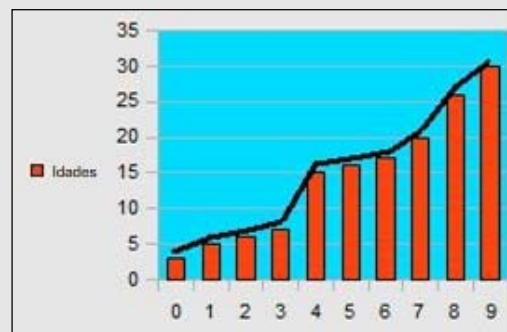
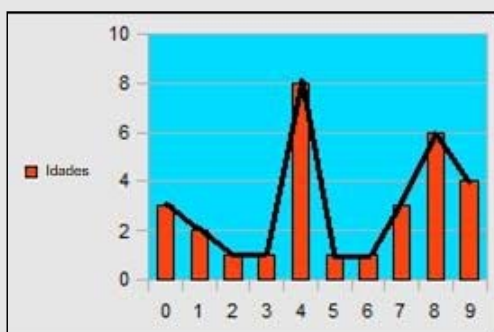
## Exercicios resoltos

8. Os datos corresponden ao número de chamadas telefónicas que reciben ao día 30 persoas.

0 8 8 8 3 9 0 4 4 7 9 7 2 7 4  
4 9 1 4 1 4 5 6 4 9 8 8 1 8 8

Debuxa o diagrama os polígonos de frecuencia e de frecuencia acumuladas que representa os datos anteriores.

SOLUCIÓN:



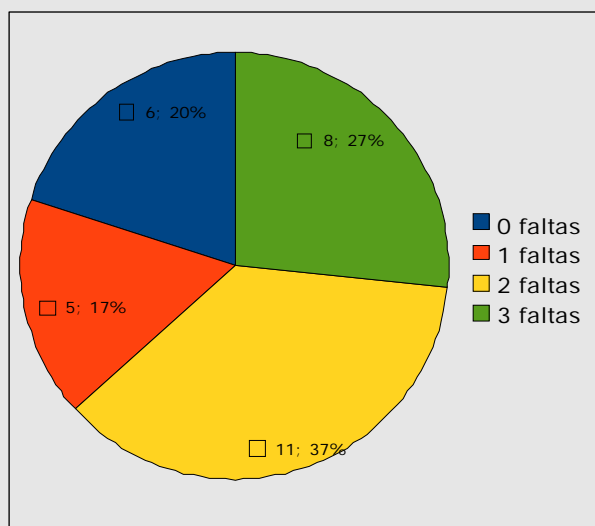
9. Os datos corresponden ao número de faltas de ortografía no mesmo texto de 30

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

estudantes.

Representa o diagrama de sectores correspondente.

SOLUCIÓN:



## 6. Medidas de centralización

### Media aritmética.

Aos parámetros ou medidas estatísticas que informan sobre a tendencia habitual ou central dos datos dunha distribución denomínaselles en estatística *medidas de tendencia central*. A máis utilizada é a media aritmética.

A media aritmética defínese como a suma de todos os datos dividida entre o número total destes. Como habitualmente disporemos dunha táboa de datos coas súas frecuencias, aplicaremos:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \dots + x_n \cdot f_n}{N}$$

1. A media non ten porqué ser un valor propio da variable.
2. É moi sensible a valores extremos nos datos.
3. Compórtase de forma natural en relación ás operacións aritméticas.

#### Exemplo

##### Faltas de asistencia. (Moitos datos)

Cando temos moitos datos, para evitar realizar unha conta con grande cantidade de números, primeiro organizamos unha táboa.

Vexamos o exemplo en que se teñen anotados as faltas de asistencia dun grupo de 27 estudantes.

Hai 6 estudantes que faltaron 0 veces, 4 que faltaron 1 vez, ...

$X_i$	$f_i$	$x_i \cdot f_i$
0	6	0
1	4	4
2	4	8
3	0	0
4	6	24
5	3	15
6	4	24
<b>27</b>	<b>75</b>	

Logo de tabular os datos, construímos a columna correspondente aos produtos  $x_i \cdot f_i$

Na última casa, calculamos a suma total da columna  $\sum_{i=1}^n x_i \cdot f_i$

A media obtense dividindo  $\sum_{i=1}^n x_i \cdot f_i = 75$  entre o total de datos 27,

**media =  $\frac{75}{27} = 2,77$**

#### Exemplo

##### Faltas de asistencia.

(Poucos datos)

As faltas de asistencia de 4 estudantes nun mes veñen recollida polos seguintes valores: 0, 3, 2 e 1.

A media aritmética calcúlase:

$$\frac{0+3+2+1}{4} = 1,5$$

### Mediana.

A mediana é aquel valor da variable estatística que deixa o 50% de observacións inferiores a el; así pois, a mediana divide en dous partes iguais á distribución estatística.

Dentro das propiedades da mediana pódense destacar:

1. Como medida descritiva non se ve tan afectada como a media pola presenza de valores extremos.
2. É de cálculo rápido e de doada interpretación.
3. Ten propiedades matemáticas complicadas que fan que se utilice pouco en inferencia estatística.

#### Caso de poucos datos e en número impar.

Neste caso procédese a ordenar os datos de menor a maior, considérase o valor da mediana o que corresponde ao lugar central.

## Caso de poucos datos e en número par.

Neste caso procédese a ordenar os datos de menor a maior, considérase o valor da mediana o correspondente á semisuma dos dous lugares centrais.

### Exemplo

#### A mediana do número de suspensos. (Moitos datos)

Entramos nunha clase de 25 estudantes e preguntamos o número de suspensos na última avaliación, hai 4 estudantes con 0 suspensos, 2 con 1 suspensos,...

Como temos moitos datos, organizámoslos na seguinte táboa para calcular a mediana.

Para localizar a mediana, en primeiro lugar calculamos a metade dos datos:

$$\frac{N}{2} = 12,5$$

Agora buscamos na columna de frecuencias acumuladas a primeira vez que se supera a metade dos datos.

O valor correspondente de  $X_i$  é a mediana da distribución estatística. Neste caso:

$$Me = 3$$

$X_i$	$f_i$	$F_i$
0	4	4
1	2	6
2	3	9
3	7	16
4	4	20
5	5	25
	25	25

## Moda

Defínese a moda como o valor da variable estatística que ten a frecuencia absoluta máis alta.

Se existen varios valores con esta característica, entón dise que a distribución ten varias modas (*plurimodal*).

Esta medida de centralización é sen dúbida a de máis doado cálculo. Adóitase utilizar como complemento á media aritmética e mediana xa que por si soa non achega unha información determinante da distribución.

Non é tan sensible como a media aritmética a valores extremos.

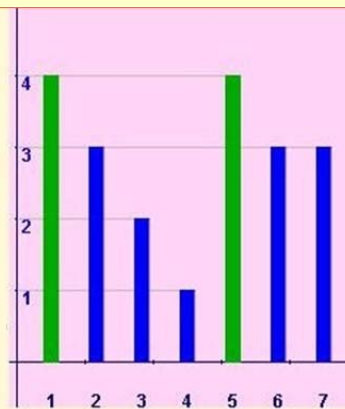
### Exemplo

#### Número de chamadas.

Nun grupo de 20 persoas recóllese o número de chamadas que realizan durante un día.

Resultando os seguintes valores: 4 persoas fan 1 chamada, 3 persoas fan 2 chamadas, 2 persoas fan 3 chamadas...

$X_i$	$f_i$	$F_i$
1	4	4
2	3	7
3	2	9
4	1	10
5	4	14
6	3	17
7	3	20



A moda é o maior valor de frecuencia absoluta

Observa que neste exemplo temos que a distribución é bimodal, xa que  $X_1 = 1$  e  $X_5 = 5$  corresponden con  $f_1 = 4 = f_5$ . Sendo ambas o máximo número de chamadas.

Compara o devandito dato co xa aprendido da media aritmética e a mediana.

### Exercicios resoltos

10. As idades dun grupo de 9 amigas son: 12, 14, 13, 16, 13, 15, 15, 17 e 13. Calcula a media, mediana e moda.

SOLUCIÓN:

X = Idade	f	F	X·f
12	1	1	12
13	3	4	39
14	1	5	14
15	2	7	30
16	1	8	16
17	1	9	17
			128

Media:  $\frac{128}{9} = 14,22$

Mediana 14 (se ordenamos os datos, aparece na posición 5).

Moda: 13 (aparece 3 veces).

11. O número de chamadas telefónicas que reciben ao día os 9 integrantes dunha familia son:

7, 8, 15, 12, 13, 5, 10, 4, 8

Calcula a media, mediana e moda.

SOLUCIÓN:

X = Nº de chamadas	f	F	X·f
4	1	1	4
5	1	2	5
7	1	3	7
8	2	5	16
10	1	6	10
12	1	7	12
13	1	8	13
15	1	9	15
			73

Media:  $\frac{73}{9} = 8,11$

Mediana 8 (se ordenamos os datos, aparece na posición 5).

Moda: 8 (aparece 2 veces).

A media aritmética infórmanos que se ti comiches dous polos e eu ningún, de promedio ambos temos comido un. Pero é obvio que ti te deches un atracón e eu pasei fame.

As medidas de centralización aportan unha información, pero é necesario completalas con outras medidas que denominaremos de dispersión. Así pois, no exemplo anterior:

- O rango de valores vai de 0 a 2 polos, hai unha amplitude de  $2-0=2$ .
- Ti tomaches un polo máis do promedio e eu un menos. Ambos nos temos desviado unha unidade respecto á media.

### Exemplo

$x_i$	$f_i$	$x_i \cdot f_i$	$ \bar{x} - x_i  \cdot f_i$
5	2	10	29,44
10	1	10	9,72
15	3	45	14,16
20	4	80	1,11
25	6	150	31,66
30	2	60	20,55
<b>18</b>	<b>355</b>	<b>106,66</b>	

$$\text{Rango} = 30 - 5 = 25$$

$$\text{DM} = \frac{106,66}{18} = 5,92$$

## 7. Medidas de dispersión

- As medidas de **dispersión** indican se os datos están máis ou menos agrupados respecto das medidas de centralización. En particular respecto á media aritmética.

### Rango

O rango ou percorrido é a diferenza entre o maior e o menor valor da variable. Indica a lonxitude do intervalo no que se achan todos os datos.

### Desviación

A desviación respecto á media é a diferenza en valor absoluto entre cada valor da variable estatística e a media aritmética.

$$D_i = |x_i - \bar{x}|$$

### Desviación media

É a media aritmética das desviacións dos diferentes datos.

$$D_{\bar{x}} = \frac{|x_1 - \bar{x}|f_1 + |x_2 - \bar{x}|f_2 + \dots + |x_n - \bar{x}|f_n}{N}$$

## EXERCICIOS resoltos

12. Calcula o rango e la desviación media dos datos:

	8	8	6	10	9	6	7	8	9	7
	7	6	6	7	9	5	5	7	10	7
$x_i$	$f_i$	$x_i \cdot f_i$	$ x_i - \bar{x}  \cdot f_i$							
5	2	10	4,8							
6	4	24	5,6							
7	6	42	2,4							
8	3	24	1,8							
9	3	27	4,8							
10	2	20	5,2							
<b>20</b>	<b>147</b>	<b>24,6</b>								

#### Solución:

O rango oscila entre 5 e 10 cunha amplitude de 5.

Facemos o reconto.

$$\text{A media: } \bar{x} = \frac{147}{20} = 7,4$$

Calculamos a desviación de cada dato respecto da media, en valor absoluto. A media das desviacións:

$$\text{DM} = \frac{24,6}{20} = 1,23$$

13. Calcula a desviación media dos datos tabulados seguintes:

$x_i$	$f_i$	$x_i \cdot f_i$	$ x_i - \bar{x}  \cdot f_i$
100	7	700	2831,82
300	8	2400	1636,36
500	13	6500	59,09
700	9	6300	1759,09
900	7	6300	2768,18
<b>Total:</b>	<b>44</b>	<b>22200</b>	<b>9054,55</b>

#### Solución:

Calculamos a media:

$$\bar{x} = \frac{22200}{44} = 504,55$$

Completamos a última columna:

$$\text{DM} = \frac{9054,55}{44} = 205,59$$



## Para practicar



### A. Clasifica segundo o carácter da variable as seguintes situacións:

#### 1. Situacións:

Cantante favorito
Lonxitude de espárragos
Marca de refresco favorita
Tipo de música preferida
Raza de cans
Nº de días soleados ao mes

#### 2. Situacións:

Nº de días de vacacións
Autor literario favorito
Nº irmáns
Nota media en selectividade
Temperatura media cidade
Nº días faltas a clase

#### 3. Situacións:

Nº de días chuviosos ao mes
Tempo de espera autobús
Nº faltas nun ditado
Cor de ollos
Películas vistas ao mes
Nota media en selectividade

### B. Realiza unha tabulación que inclúa a frecuencia absoluta, relativa e as súas acumuladas, cando sexa necesario aproxima ata as centésimas, dos datos que se corresponden coas situacións seguintes:

4. O número de veces que cambiaron de domicilio 23 persoas.  
2, 2, 0, 2, 4, 2, 4, 4, 3, 4, 3, 3, 3, 3, 0, 1, 0, 4, 0, 3, 0, 3 e 5.

5. O número de irmáns que teñen 20 estudantes dun centro.  
1, 4, 0, 2, 3, 1, 0, 3, 4, 1, 1, 3, 3, 3, 4, 1, 1, 2, 1 e 1.

6. O número de dormitorios de 28 vivendas dunha cidade.  
3, 5, 0, 4, 2, 3, 0, 0, 1, 1, 3, 0, 2, 4, 1, 3, 3, 3, 1, 4, 4, 0, 3, 3, 1, 4, 3 e 1.

7. O número de faltas de ortografía no mesmo texto de 30 estudantes son:  
0, 0, 2, 1, 4, 6, 6, 5, 0, 4, 6, 5, 5, 1, 0, 0, 3, 5, 1, 2, 5, 1, 0, 5, 2, 0, 4, 3, 6 e 4.

### C. Efectúa unha tabulación dos datos na que aparezan as columnas de frecuencias absolutas e relativas. Cando sexa necesario aproxima ata as centésimas.

8. O sabor preferido nos refrescos dunha determinada marca de 22 persoas.  
Laranxa, cola, laranxa, limón, cola, melocotón, cola, limón, cola, cola, mazá, limón, laranxa, cola, piña, cola, laranxa, mazá, laranxa, cola, laranxa e mazá.
9. As actividades realizadas por 20 estudantes no seu tempo libre.  
Deporte, amigos, idiomas, música, idiomas, idiomas, amigos, música, deportes, baile, baile, música, deportes, idiomas, cine, amigos, deportes, amigos, música, e cine.

10. O tipo de programa de televisión que prefiren **ver** no seu tempo libre. Ficción, infantís, deportivos, espectáculo, documentais, infantís, ficción, culturais, espectáculo, infantís, ficción, deportivos, deportivos, espectáculo, ficción, documentais, culturais, ficción, deportivos e espectáculo.

**D. Debuxa o diagrama de barras correspondente ás situacións que aparecen.**

11. Preguntamos a 25 estudantes elixidos aleatoriamente polo tipo de música que prefiren escoitar. Os resultados son: disco, disco, rock, clásica, rock, latina, pop, rock, pop, latina, rock, flamenco, flamenco, latina, flamenco, latina, rock, clásica, disco, disco, latina, rock, disco, latina e rock.

12. Os datos corresponden ás contestacións realizadas por 25 persoas elixidas aleatoriamente, acerca do tipo de película que prefiren ver. Os datos son os seguintes: comedia, terror, suspense, comedia, aventura, drama, aventura, aventura, comedia, musical, terror, musical, suspense, aventura, comedia, terror, musical, terror, terror, comedia, suspense, suspense, comedia, aventura e aventura.

13. Os resultados seguintes corresponden ás contestacións realizadas por 25 estudantes acerca das actividades realizadas no seu tempo libre. Deporte, amigos, amigos, idiomas, música, idiomas, deporte, música, idiomas, amigos, música, deportes, baile, música, baile, música, deportes, idiomas, cine, amigos, deportes, cine, amigos, música, e cine.

14. As idades de 30 estudantes dun instituto de ensino secundario son as seguintes:  
12, 13, 12, 15, 12, 15, 13, 14, 15, 12, 12, 12, 15, 15, 13, 14, 14, 16, 13, 12, 13, 14, 15, 16, 15, 13, 14, 15, 15 e 12.

15. Número de materias suspensas de 30 estudantes son:  
2, 0, 3, 2, 4, 0, 1, 3, 4, 2, 5, 0, 3, 2, 5, 4, 1, 1, 0, 1, 1, 1, 3, 4, 2, 0, 5, 5, 3 e 2.

16. O número de chamadas telefónicas que reciben un día un grupo de 20 amigos son:  
4, 5, 1, 9, 5, 3, 6, 3, 7, 8, 3, 4, 1, 0, 9, 7, 6, 2, 1 e 5.

17. Para un estudo de accesibilidade, durante 30 días anotamos o número de prazas libres de aparcamento ás 5 da tarde.  
1, 1, 3, 5, 4, 0, 1, 3, 4, 2, 5, 0, 3, 2, 5, 4, 3, 1, 0, 1, 4, 1, 3, 4, 2, 3, 5, 4, 3 e 0.

**E. Debuxa o diagrama de sectores correspondente ás situacións que aparecen nos exercicios D.11, D.12, D.16 e D.17**

**F. Realiza o polígono de frecuencias e o de frecuencias acumuladas dos exercicios dos apartados D.14 e D.15**

**G. Acha as medidas de centralización dos exercicios dos apartados B.6 e B.7**

**H. Acha as medidas de dispersión dos exercicios dos apartados B.6 e B.7**

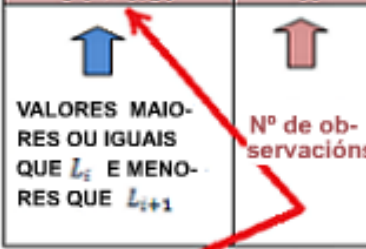
## Para saber máis

En **estadística** en moitas ocasións unha variable discreta toma tal variedade de valores que para que a tabulación sexa efectiva, debe realizarse mediante intervalos. A variable queda deste xeito dividido en clases (intervalos, xeralmente da mesma amplitude). Esta é a técnica que se utiliza para variables continuas.

Ademais do rango e desviación media outras medidas de dispersión dos datos que se utilizan normalmente son:

- Varianza
- Desviación típica

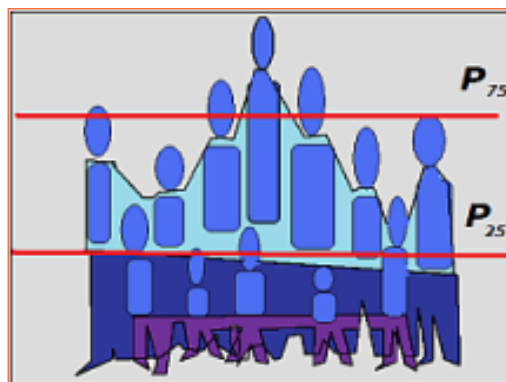
Intervalos	frecuencia
...	...
...	...
$[L_i, L_{i+1})$	$f_i$



VALORES MAIORES OU IGUAIS QUE  $L_i$  E MENORES QUE  $L_{i+1}$

Nº de observacións

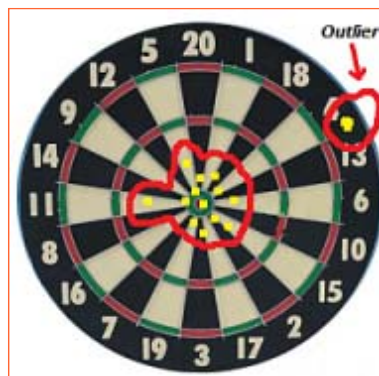
**Como representante de todo o intervalo considérase o valor central (MARCA DE CLASE)**



O principal parámetro estatístico e o máis utilizado é a *media aritmética*, non obstante unha característica importante da media é que vese moi afectada por valores extremos na distribución.

A variación dun simple dato na distribución afecta ás medidas de tendencia central aínda que non da mesma forma. En estadística a un valor especialmente anómalo denomínaselle **Outlier**. Decidir se nun estudo estatístico se depuran estes valores extremos, é unha das primeiras accións que debe realizar un investigador.

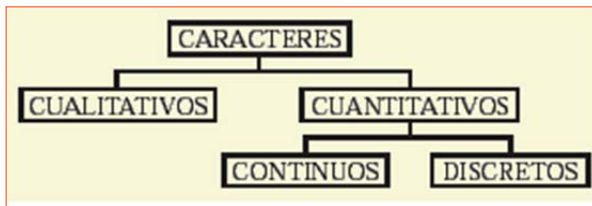
A *mediana* é a medida de posición que máis se utiliza, non obstante é moi habitual na maioría dos estudos estatísticos facer referencia a outras medidas de posición como os **cuartiles**, **deciles** o **percentiles**.





## Lembra o máis importante

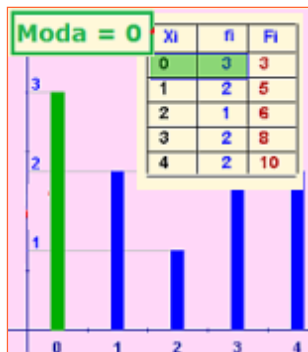
As primeiras definicións necesarias para o inicio de calquera estudo estatístico son: **Poboación, Individuo, Mostra e Carácter.**



CUALITATIVOS: *Non expresables numericamente.*  
 CUANTITATIVOS: *Pódese expresar mediante número.*

### Medidas de tendencia central

**Moda:** Valor que ten a frecuencia absoluta máis alta.



É a única que pode calcularse para variable cualitativa.

Non é tan sensible como a media aritmética a valores extremos.

**Media aritmética:** suma de todos os datos dividida entre o número total destes.

$X_i$	$f_i$	$x_i \cdot f_i$
1	5	5
2	2	4
3	2	6
4	5	20
5	0	0
6	0	0
7	5	35
8	5	40
9	0	0
	<b>24</b>	<b>110</b>

Moi sensible a valores extremos nos datos  
 Non ten por que ser un valor propio da variable

$$\text{media} = \frac{\sum_{i=1}^n x_i \cdot f_i}{\sum_{i=1}^n f_i} = \frac{110}{24} = 4,58$$

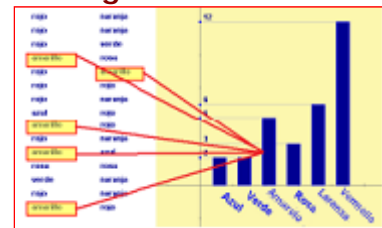
**Mediana:** divide en dous partes iguais á distribución estatística.

$X_i$	$f_i$	$F_i$
0	8	8
1	7	15
2	5	20
3	16	36
4	9	45
	<b>45</b>	<b>45</b>

$Me = 3$   
 $\frac{22,50}{2} = 11,25$   
 $\frac{45}{2} = 22,50$

De cálculo rápido e de doada interpretación. Non é tan sensible como a media aritmética a valores extremos.

### Diagrama de barras



### Polígono de frecuencias



### Diagrama de sectores



## Autoavaliación



1. Dados os datos: 7, 5, 7, 5, 6, e 8.  
Calcula a media aritmética con dúas cifras decimais.
2. A nota media obtida en cinco exames foi 6,8. Se catro das notas foron 4,7; 9,5; 8,3 e 9,2. Cal é a quinta?
3. A nota media de catro notas é 4,2. Se saquei agora un 8,0. Que nota media terei agora?
4. Nunha proba de ximnasia a puntuación de cada atleta calcúlase eliminando a peor e a mellor nota dos xuíces. Se as puntuacións obtidas foron: 8,1; 9,0; 9,3; 9,6; 8,2; 8,7 e 9,5. Que nota corresponde?
5. Calcula a mediana destes datos: 9, 15, 19, 22, 31, 38 e 43.
6. Calcula a desviación media destes datos: 18, 15, 28, 22 e 36.
7. Nunha distribución de 63 datos, a frecuencia absoluta dun valor da variable é 21. Cantos graos corresponderían a ese valor nun diagrama de sectores?
8. Para obter a nota final de curso danos a elixir entre a media, a mediana e a moda das nove notas obtidas. Cal elixirías? As notas son: 6, 3, 3, 4, 6, 8, 7, 9 e 3.
9. Calcula a mediana destes datos: 1, 17, 26, 5, 11 e 24.
10. Indica se a variable é discreta, continua ou cualitativa: Perímetro cranial.

## Soluciones dos exercicios para practicar

**A.1** Cualitativa, continua, cualitativa, cualitativa, cualitativa, discreta.

**A.2** Discreta, cualitativa, discreta, continua, continua, discreta.

**A.3** Discreta, continua, discreta, cualitativa, discreta, continua.

### B.4

Nº de cambios de domicilio	f	h	F	H
0	5	0,22	5	0,22
1	1	0,04	6	0,26
2	4	0,17	10	0,43
3	7	0,30	17	0,73
4	5	0,22	22	0,95
5	1	0,04	23	0,99
	23			

### B.5

Nº de irmáns	f	h	F	H
0	2	0,10	2	0,10
1	8	0,40	10	0,50
2	2	0,10	12	0,60
3	5	0,25	17	0,85
4	3	0,15	20	1,00
	20			

### B.6

Nº de dormitorios	f	h	F	H
0	5	0,18	5	0,18
1	6	0,21	11	0,39
2	2	0,07	13	0,45
3	9	0,32	22	0,77
4	5	0,18	27	0,95
5	1	0,04	28	0,99
	28			

### B.7

Nº de faltas de ortografía	f	h	F	H
0	7	0,23	7	0,23
1	4	0,13	11	0,36
2	3	0,10	14	0,46
3	2	0,07	16	0,53
4	4	0,13	20	0,66
5	6	0,20	26	0,86
6	4	0,13	30	0,99
	30			

### C.8

Sabor preferido	f	h
Laranxa	6	0,27
Cola	8	0,36
Limón	3	0,14
Piña	1	0,05
Melocotón	1	0,05
Mazá	3	0,14
	22	

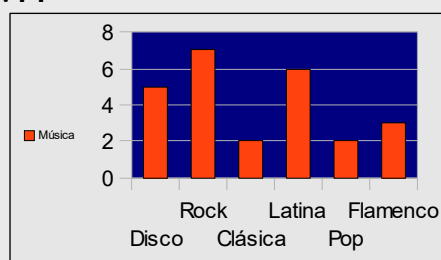
### C.9

Actividades tempo libre	F	h
Deporte	4	0,20
Amigos	4	0,20
Idiomas	4	0,20
Baile	2	0,10
Cine	2	0,10
Música	4	0,20
	20	

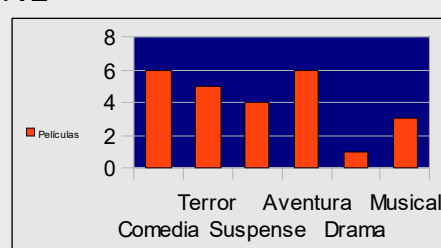
### C.10

Programas de televisión	f	h
Ficción	5	0,25
Infantís	3	0,15
Deportivos	4	0,20
Espectáculo	4	0,20
Documentais	2	0,10
Culturais	2	0,10
	20	

### D.11

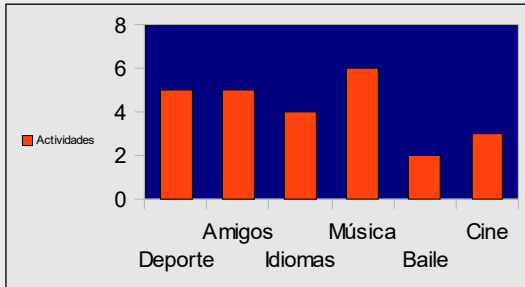


### D.12

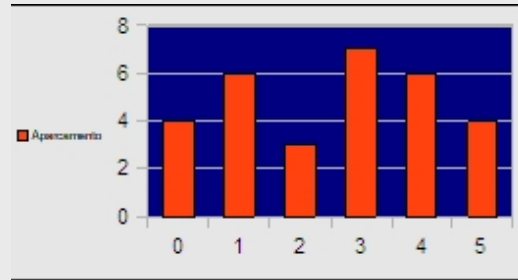




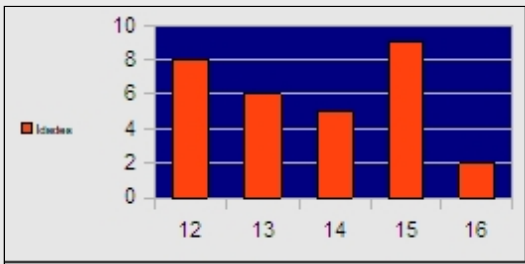
**D.13**



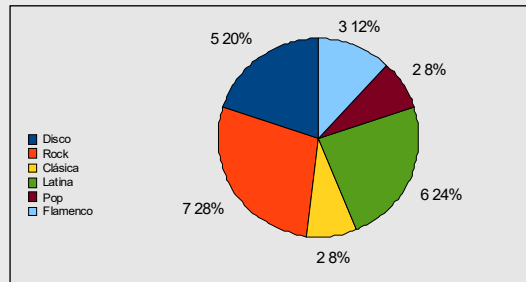
**D.17**



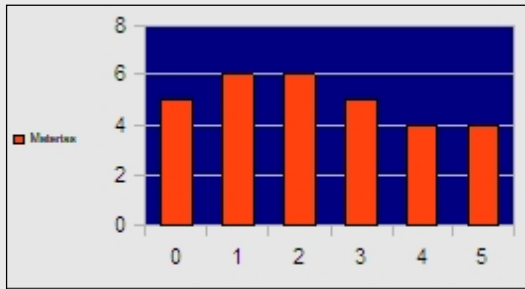
**D.14**



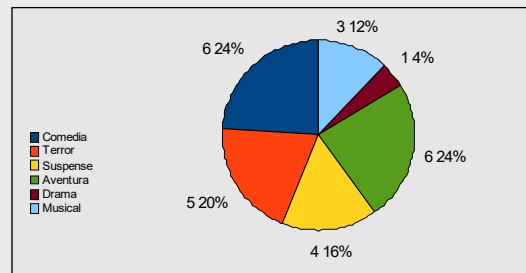
**E.11**



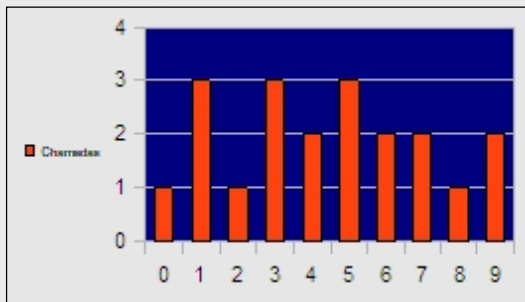
**D.15**



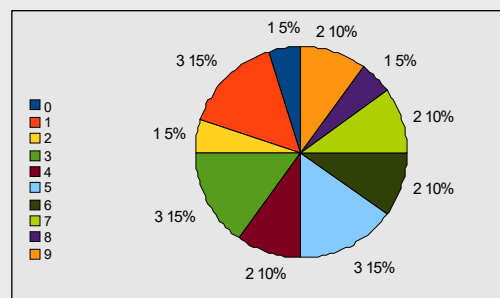
**E.12**



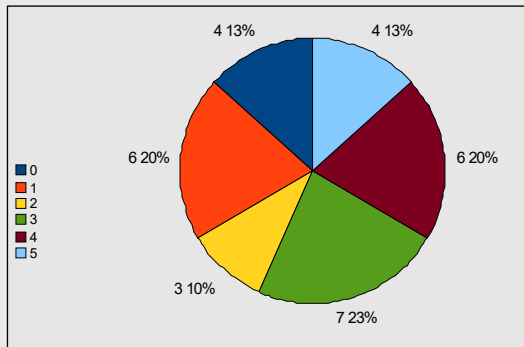
**D.16**



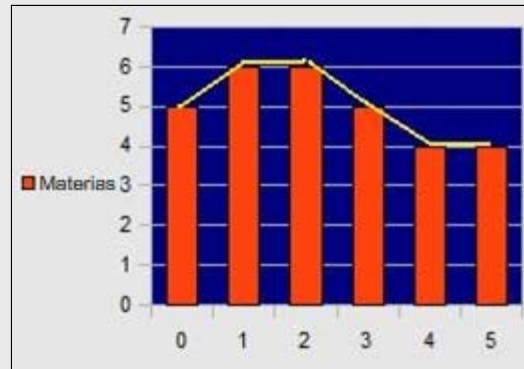
**E.16**



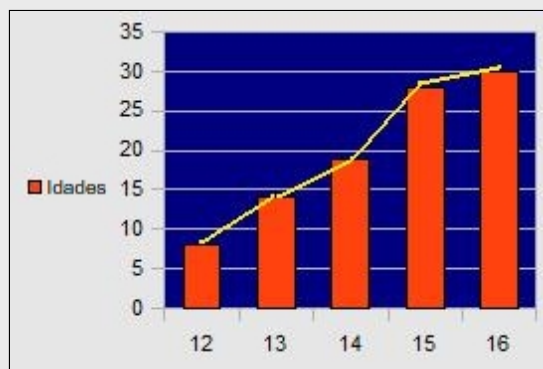
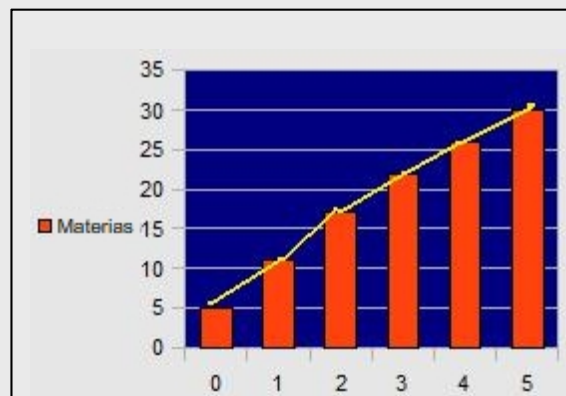
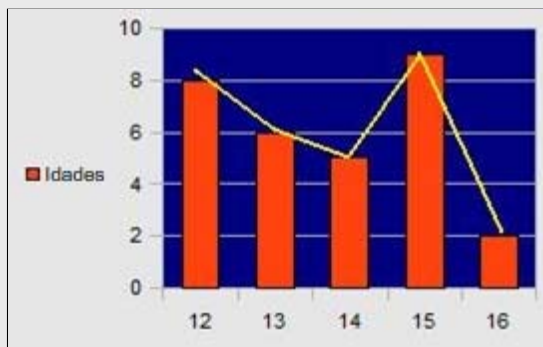
**E.17**



**G.15**



**G.14**



**G.6**

Media = 2,21  
Mediana = 3  
Moda = 3

**G.7**

Media = 2,86  
Mediana = 3  
Moda = 0

**H.6**

Rango = 5  
Desviación media = 1,34

**H.7**

Rango = 6  
Desviación media = 2,01

## Soluciones AUTOAVALIACIÓN

1. 6'33
2. 2'3
3. 5'8
4. 8'94
5. 22
6. 6,56
7. 120°
8. A mediana
9. 14
10. Cualitativa