XUNTA DE GALICIA
CONSELLERÍA DE EDUCACIÓN
E ORDENACIÓN UNIVERSITARIA
Dirección Xeral de Educación, Formación
Profesional e Innovación Educativa

Ámbito científico tecnolóxico

Educación a distancia semipresencial

Módulo 3

Unidade didáctica 4

Estatística

Índice

1.1 Descrición da unidade didáctica....................3
1.2 Coñecementos previos....................3
1.3 Criterios de avaliación....................3
2.1 A estatística. Fases e tarefas dun estudo estatístico....................4
2.2 Poboación e mostra....................6
2.2.1 Conceptos básicos....................6
2.2.2 Métodos de selección de mostras. Representatividade....................6
2.3 Variables estatísticas. Frecuencias absolutas e relativas. Representacións gráficas...10
2.3.1 Variables estatísticas....................10
2.3.2 Frecuencias absolutas e relativas....................11
2.3.3 Frecuencias absolutas e relativas acumuladas....................14
2.3.4 Representacións gráficas....................17
2.4 Aplicacións informáticas para a representación gráfica de datos estatísticos....................22
2.4.1 Diagramas de barras con LibreOffice Calc....................22
2.4.2 Diagramas de sectores con LibreOffice Calc....................24
2.5 Parámetros estatísticos....................24
2.5.1 Parámetros de centralización....................24
2.5.2 Parámetros de dispersión....................28
2.5.3 Parámetros de posición. Os cuartís....................33
2.5.4 Diagrama de caixa e bigotes....................34
2.5.5 Interpretación conxunta da media e da desviación típica....................36

3. Actividades finais....................37

4. Solucionario....................40

4.1 Solucións das actividades propostas....................40
4.2 Solucións das actividades finais....................47

5. Glosario....................53

1. Introdución

1.1 Descrición da unidade didáctica

Nesta unidade estudaremos a estatística. Na estatística trabállase cunha grande, grandísima, cantidade de datos. Nesta unidade veremos como organizar eses datos para proceder ao seu posterior estudo e sermos quen de sacar conclusións sobre os datos cos que traballamos.

Veremos tamén as representacións gráficas máis usuais en estatística e a súa interpretación.

1.2 Coñecementos previos

Para traballar con esta unidade é necesario lembrar os conceptos sobre estatística estudados no módulo anterior, en especial:

- Conceptos de poboación, mostra e individuo.
- Concepto e tipos de variables estatísticas.
- Frecuencia absoluta e frecuencia relativa. Organización dos datos en táboas.
- Representacións gráficas: diagramas de barras, diagramas de sectores e polígonos de frecuencias.
- Medidas de centralización: media, moda e mediana.

1.3 Criterios de avaliación

- Elaborar informacións estatísticas para describir un conxunto de datos mediante táboas e gráficas adecuadas á situación analizada e xustificar se as conclusións son representativas para a poboación estudada.
- Calcular e interpretar os parámetros de posición e de dispersión dunha variable estatística para resumir os datos e comparar distribucións estatísticas.

2. Secuencia de contidos e actividades

2.1 A estatística. Fases e tarefas dun estudo estatístico

A necesidade de recontar datos sobre as diferentes poboacións xurdiu ao apareceren sociedades organizadas como, por exemplo, en China, Exipto ou Roma.

Co transcorrer do tempo creáronse novas necesidades que non se limitaban só ao campo social. Aparece a conveniencia de facer estudos de tipo estatístico nas distintas ramas do saber como poden ser a economía ou a bioloxía.

Deste xeito, a partir do século XIX aplicouse a estatística a distintos campos, dando lugar á aparición de novas especialidades e subdivisións delas: a econometría, a bioestatística (tamén chamada biometría)...

Na actualidade as técnicas estatísticas utilízanse en case todos os aspectos da vida. Deséñanse enquisas para recoller información previa ás consultas electorais. Realízanse mostraxes sobre as características de calidade dun produto. Sóndase aos consumidores para indagar se un produto que pode saír á venda terá éxito ou non. Realízanse experimentos e medicións para determinar o modo máis axeitado de evitar o deterioro do medio ambiente etcétera.

Hoxe en día, a estatística é una ciencia que se encarga de estudar unha determinada característica dun conxunto de individuos recollendo, analizando e interpretando datos.

A estatística divídese en dúas ramas principais:

- **Estatística descritiva ou dedutiva**: trata do reconto, a ordenación e a clasificación dos datos obtidos nas observacións. Constrúense táboas e represéntanse en gráficos, que permiten simplificar en gran medida a complexidade dos datos que interveñen na distribución. A partir dos datos obtéñense os parámetros estatísticos que caracterizan a distribución. Esta parte da estatística limitase a realizar deducións directamente a partir dos datos e os parámetros obtidos.
- **Estatística inferencial ou indutiva**: formula e resolve o problema de establecer previsións e deducións xerais sobre unha poboación a partir resultados obtidos dunha mostra. Utiliza resultados obtidos mediante a estatística descritiva e apóiase fortemente no cálculo de probabilidades.

Nun típico traballo estatístico, procédese á recollida de datos que se representan graficamente para facilitar a súa visión global e calcúlanse certos parámetros para caracterizar a súa tendencia central (media, moda, mediana...) e outros para indicar a súa dispersión (varianza, desvío típico...). Posteriormente inténtase ver se os datos se poden vincular con algún modelo estatístico para poder obter máis información sobre o obxecto do estudo.

A información estatística chega a nós mediante gráficas ou táboas moi ben construídas, coas que resulta doado entendermos a información dada. Pero para chegar a elas, cómpre realizarmos un longo proceso, que se inicia agora.

- *Que queremos estudar?* Necesitamos saber o que pretendemos estudar, por exemplo que afeccións deportivas teñen os alumnos e as alumnas dun centro.
- *Selección das variables que se van a analizar*. Debe ser evidente cal é a variable e cales os seus posibles valores.
- *Colleita de datos*. Efectúanse as medidas ou realízanse as enquisas.
- *Organización de datos*. Ordénanse, pásanse a papel, ou mellor, introdúcense no computador.
- Os pasos seguintes son a elaboración de táboas e gráficas e o cálculo de parámetros, aos que dedicaremos o resto da unidade.

Actividades propostas

S1. Quérese facer unha enquisa para estudar a afección da xente nova á lectura. Diga, xustificadamente cales das preguntas seguintes lle parecen razoables e cales non.

- Di cales son as túas lecturas favoritas.
- Dos xéneros literarios seguintes, sinala os que liches máis dunha hora o último mes.

☐ Novela ☐ Historia ☐ Poesía ☐ Teatro ☐ Filosofía ☐ Biografía ☐ Cómic

- Les xornais? Se é así, de que tipo?
- A cal ou cales das seguintes publicacións periódicas dedicas máis de dúas horas semanais:

☐ Xornais de información xeral ☐ Xornais deportivos ☐ Revistas científicas

☐ Revistas do corazón ☐ Outros (indíquese cales)

2.2 Poboación e mostra

2.2.1 Conceptos básicos

Como acabamos de ver no apartado anterior, a estatística estuda unha ou máis características dun grupo de individuos, ao que se chama poboación, para obter resultados e conclusións e facer predicións sobre o comportamento da poboación con respecto ás características estudadas.

Chámase **poboación** ao conxunto de persoas, obxectos ou individuos sobre os que se quere estudar unha determinada característica. Cada membro da poboación chámase **individuo**.

Se, por exemplo, queremos estudar o número de pezas defectuosas fabricadas por un obradoiro nun día, a poboación sería as pezas fabricadas nese obradoiro. Cada unha das pezas fabricadas sería un individuo.

Ás veces resulta practicamente imposible observar todos os individuos da poboación, en especial cando son moi numerosos. No canto de examinar toda a poboación o que se fai é examinar unha pequena parte dela á que chamamos **mostra**.

As mostras deben ser representativas da poboación, polo que teñen que cumprir dous requisitos: teñen que ser aleatorias e teñen que ser representativas da poboación que representan, isto é, teñen que presentar as mesmas características.

2.2.2 Métodos de selección de mostras. Representatividade

Ás veces, cando queremos estudar unha característica dunha poboación, resulta moi custoso ou non é posible estudar todos os membros desa poboación. Así, por exemplo, se queremos coñecer o peso medio das pescadas do Cantábrico, non parece recomendable a estratexia de capturalas todas e pesalas. En situacións deste tipo e doutros semellantes, a mellor solución é facer o estudo sobre unha parte dos individuos da poboación, é dicir, seleccionar unha mostra.

Analicemos a seguinte situación. Nunha vila hai 50 000 vivendas habitadas e o concello quere saber cal é a porcentaxe de vivendas que teñen lavalouzas. Para iso, escolle ao chou 400 vivendas e un empregado do concello vai casa por casa preguntando se hai ou non lavalouzas na casa. O entrevistador establece que 252 das vivendas escollidas tiñan lavalouzas e 148 non a tiñan.
Neste suposto, a poboación está formada por 50 000 individuos que son as vivendas habitadas. A mostra é o conxunto das vivendas escollidas, ten un tamaño de 400 individuos.

Como na mostra atoparon 252 vivendas con lavalouzas sobre 400, para obtermos a proporción, en tanto por cento, das vivendas que teñen lavalouzas, facemos simplemente:

$$\frac{252}{400} \times 100 = 63\ \%$$

Polo que o concello ten un indicio razoable de que a porcentaxe de vivendas do concello que dispón de lavalouzas será, aproximadamente, dun 63 %.

Pero resulta evidente que estes pequenos cálculos non resolven de xeito inapelable o problema; xorden preguntas do tipo de: "ata que punto podemos supoñer que unha porcentaxe obtida na mostra é xeneralizable a toda a poboación?"

A busca de respostas a este tipo de preguntas leva a analizar os distintos tipos de mostras posibles e a súa representatividade.

Agora, imos describir resumidamente os tipos de mostraxe máis usados:

Mostraxe aleatoria simple. Todas as mostras, do mesmo tamaño, que se poden obter sobre a poboación teñen a mesma probabilidade de saír. Para levar a cabo unha mostraxe aleatoria simple, o máis habitual é elaborar un listado con todos os individuos da poboación, asignando un número a cada un delas. A continuación escóllese a mostra por sorteo ou usando unhas táboas especiais nas que os números aparecen ao azar, sen seguir ningunha tendencia, chamadas táboas de números aleatorios.

Exemplo. Unha empresa de 500 traballadores desexa estimar cantos deles acoden camiñando a traballar. Despois de elaborar unha lista numerada dos traballadores, introduce nun bombo 500 bólas numeradas do 1 ao 500 e extrae 25. A continuación entrevista aos 25 traballadores escollidos para saber se van camiñando ao traballo. Isto é unha mostraxe aleatoria simple.

Mostraxe aleatoria con reposición. Consiste en obter unha mostra aleatoria simple, pero aínda que un individuo fora escollido unha vez para formar parte da mostra, pode ser escollido outras veces. O resultado final será unha mostra na que cada individuo pode aparecer máis dunha vez na mostra. Para poboacións grandes, moito máis grandes có tamaño da mostra final, non existen diferenzas significativas entre este método e o anterior, porque a probabilidade de obter un individuo repetido é practicamente nula.

Exemplo. Para saber a proporción de machos e femias na poboación de ras dunha lagoa, un grupo de biólogos segue o seguinte procedemento: captura unha ra, comproba o seu sexo e devólvea á lagoa; realiza este experimento 100 veces. Isto é unha mostraxe aleatoria con reposición.

Mostraxe estratificada. Divídese a poboación nuns poucos grupos grandes (estratos) e faise mostraxe aleatoria simple en cada un deles.

Exemplo:. Nun concello hai tres institutos e o número de alumnos en cada un deles é similar. O concello estuda a posibilidade de organizar unha actividade conxunta para os tres centros e quere coñecer o grao de aceptación desta iniciativa entre o alumnado. Para iso, escolle ao azar 50 alumnos en cada un dos institutos e entrevista aos 150 estudantes. Isto é unha mostraxe estratificada.

Mostraxe por conglomerados. Divídese a poboación en moitos grupos miúdos (conglomerados). A continuación, elabórase un listado deses grupos e, con eles, faise mostraxe aleatoria simple para obter unha mostra de conglomerados e estúdanse todas as unidades dos conglomerados escollidos.

Exemplo. Supoñamos que, no exemplo anterior, o concello adopta unha estratexia de sondaxe totalmente distinta: elabora un mapa de distritos do seu territorio de xeito que en cada distrito haxa, aproximadamente, 25 estudantes cursando estudos nalgún instituto. A continuación, escolle por sorteo 6 deses distritos e entrevista a todos os alumnos que se atopan nas zonas escollidas. Isto é unha mostraxe por conglomerados.

Mostraxe sistemática. Escóllese unha unidade en particular, pode ser ao azar, e todas as demais unidades da mostra son obtidas mediante un método preestablecido.

Exemplo. Deséxase saber cantos usuarios dunha autoestrada son turistas. O procedemento de sondaxe, desenvolvido nun día calquera, consiste en escoller o primeiro dos usuarios que pase pola peaxe despois das oito da mañá e preguntarlle se é turista. Despois, fáiselle a mesma pregunta a un de cada 10 usuarios ata as oito da mañá do día seguinte. Isto é unha mostraxe sistemática.

Mostraxe oportunista ou intencional. É un tipo de mostraxe non científica e de nula fiabilidade, na que a mostra é confeccionada de xeito que os resultados presentan unha tendencia definida ou nesgo. Por exemplo, non é fiable unha enquisa sobre calidade da ensinanza nun determinado nivel cando os alumnos que integran a mostra foron escollidos entre o colectivo dos que aprobaron todas as áreas dese nivel.

O tipo de mostraxe que dá maior fiabilidade á hora de xeneralizar á poboación os resultados obtidos na mostra é a mostraxe aleatoria, simple ou con reposición. É a mellor forma para conservar na mostra as características da poboación.

Á hora de xeneralizar os resultados á poboación, sempre veñen especificadas unhas marxes de erro. Esas marxes de erro pódense observar, por exemplo, en calquera sondaxe electoral que publican os medios de comunicación en vésperas de eleccións.

Actividades propostas

S2. Un profesor de matemáticas fai unha proba de control, sen influencia na nota, nunha clase de 30 estudantes, pero non examina a todos. O día do control, o profesor leva á clase un bombo con dez bólas numeradas do 0 ao 9 e extrae unha bóla por sorteo; se sae o 0 deben facer o control as persoas 10, 20 e 30; se sae o 3 deben facelo as que teñen os números 3, 13 e 23 etcétera. Que tipo de mostraxe usa o profesor para escoller as persoas que han de examinarse? Cal é a poboación? Cales son os tamaños da poboación e da mostra?

S3. Un banco decide sortear 5 coches entre todos os clientes que teñan a nómina domiciliada nalgunha das súas oficinas. Interprétese o proceso de sorteo coma unha mostraxe indicando cal é a poboación e cal é o tamaño da mostra. Que tipo de mostraxe debe levar a cabo o banco para garantir a equidade no sorteo?

S4. Nun mercado hai 15 postos de venda de pementos de Padrón. Os vendedores desexan saber que porcentaxe de pementos ten sabor picante e deciden coller 10 pementos de cada posto e probalos. Que tipo de mostraxe están a facer?

S5. Unha compañía aérea pretende coñecer a opinión dos seus clientes sobre a puntualidade dos voos. Para iso, contrata entrevistadores que preguntan aos viaxeiros dos avións que aterraron á hora prevista. Considera adecuado este método de sondaxe? Como cualificaría a técnica de mostraxe utilizada?

S6. Unha granxa experimental estuda unha determinada planta. O invernadoiro no que se leva a cabo o cultivo ten 1 000 plantas e os científicos queren estimar o número de froitos que é esperable obter nunha plantación destas características. Para iso, escollen ao azar 100 plantas e contan o número de froitos de cada unha delas, despois multiplican o número total de froitos contados por 10. Que tipo de mostraxe fixeron?

S7. Cinco depósitos cobren o subministro de auga dunha cidade. O control semanal de salubridade faise do seguinte xeito: cada día da semana é analizada, por sorteo, a auga dun destes depósitos. Que tipo de mostraxe se usa no control semanal da auga dos depósitos?

S8. Un mes antes das eleccións municipais, un partido político encarga unha sondaxe sobre intención de voto a unha empresa. No municipio viven 60 000 persoas con dereito a voto que se distribúen do seguinte xeito:

	Homes	Mulleres
Menores de 40 anos	13 500	16 500
De 40 anos ou máis	13 800	16 200

A empresa vai seleccionar unha mostra de 2 000 persoas para facer o estudo. Cantas persoas serán menores de 40 anos? Cantas serán mulleres? Cantos serán homes de 40 anos ou máis?

S9. De cada un dos seguintes estudos estatísticos, indique cal é a poboación e se considera necesario elixir unha mostra.

Horas diarias de sono dos habitantes dunha provincia.	Preferencias literarias das persoas maiores de idade que viven nun edificio.

2.3 Variables estatísticas. Frecuencias absolutas e relativas. Representacións gráficas

2.3.1 Variables estatísticas

Chámase **variable estatística** a cada unha das propiedades que se pretendan estudar.

As variables estatísticas poden ser cualitativas ou cuantitativas. As **variables estatísticas cualitativas** son as que non se poden expresar como valores numéricos e as **variables estatísticas cuantitativas** son as que poden tomar valores numéricos.

As variables estatísticas cuantitativas poden ser **discretas**, cando no son posibles todos os valores numéricos, e **continuas** cando son posibles todos os valores numéricos.

Actividades propostas

S10. Indique se as seguintes variables son cualitativas ou cuantitativas: a música preferida, o peso, a cor do pelo, a idade, as cualificacións dun exame.

S11. Está vostede a estudar a idade (en anos) dos 25 estudantes dunha clase. Esta variable estatística, é discreta ou é continua?

S12. Está vostede a estudar o peso das mazás dunha maceira da súa propiedade. Esta variable estatística, é discreta ou é continua?

2.3.2 Frecuencias absolutas e relativas

Cando se recollen datos sobre un estudo estatístico, adoitan ser moitos. Se nos limitamos a presentalos como unha listaxe de valores, serán pouco explicativos e de difícil comprensión, por iso se empregan métodos gráficos e numéricos para describir unha variable estatística e os seus valores.

Para representarmos os datos recollidos, o primeiro é construírmos unha táboa de frecuencias. Nela aparecen todos os valores que pode tomar a variable e o número de veces que se obtén cada valor da mesma.

Para calquera valor da variable que esteamos estudando, chámase **frecuencia absoluta** de cada valor ó número de veces que aparece dito valor. A frecuencia absoluta desígnase por f_i.

Actividade resolta

Preguntámoslle a 50 persoas o número de veces que van ao cine nun mes e obtemos as seguintes respostas:

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

Calcule as frecuencias absolutas desta variable.

Efectuaremos un reconto dos datos ordenándoos nunha táboa que amose a frecuencia absoluta (número de veces que aparece ese dato), que chamaremos f_i.

Fixémonos que, neste caso, os valores máximos e mínimo que toma a nosa variable estatística (número de veces que asiste ao cine nun mes) son 3 e 0 respectivamente. Así, na táboa de recollida de datos debemos poñer os valores da variable que hai entre 0 e 3.

x_i = número de veces que asiste ao cine nun mes.	Reconto	f_i = frecuencia absoluta
0	||||| ||||| |	11
1	||||| ||||| ||||| ||||| ||||| ||||| |||	33
2	|||||	5
3	|	1
TOTAL		N = 50

É dicir, hai 11 persoas que non van ao cine, 33 que van unha vez por mes, 5 que van dúas veces e 1 que vai tres veces ao mes. Como preguntamos a 50 persoas, a suma das frecuencias absolutas ten que ser 50.

Ás veces, as variables non son discretas, senón continuas, polo que poden tomar calquera valor. Nestes casos, establécense intervalos que chamaremos **intervalos de clase** e a cada un deles asociarémoslle como valor o punto medio do intervalo, que se chama **marca de clase**.

Tamén é posible que a variable sexa discreta, pero, ante a dificultade de facer un reconto de cada valor da variable ou que haxa un número moi grande de valores diferentes, pódese facer un reconto dos datos agrupándoos en intervalos de clase.

As clases non teñen por que ser todas da mesma amplitude, pero é aconsellable facelas da mesma amplitude para evitar problemas nalgunha das representacións gráficas que veremos nesta unidade.

Unha vez feito isto, procédese do mesmo xeito que no exemplo anterior.

Actividade resolta

Quérese realizar un estudo sobre a lonxitude dun tipo de parafusos que se fan nunha fábrica. Elíxese ao chou unha mostra de 32 e obtéñense os seguintes resultados en milímetros.

161	171	167	172	170	170	165	169	170	169	172	162	169	166	174	178
167	169	168	176	169	162	168	167	175	168	164	179	172	167	170	173

Agrupe os datos en clases, estableza as marcas de clase e calcule cadansúa frecuencia absoluta.

Efectuaremos un reconto dos datos ordenándoos nunha táboa que amose a frecuencia absoluta (número de veces que aparece ese dato), que chamaremos f_i.

Reparemos en que, neste caso, os valores máximo e mínimo que toma a nosa variable estatística (lonxitude dos parafusos) son 179 e 161, respectivamente. Así, na táboa de recollida de datos debemos poñer clases que inclúan estes valores.

Comezaremos no 160 e remataremos no 180 facendo clases de lonxitude 5.

Intervalos de clase	x_i = marca de clase	Reconto	f_i = frecuencia absoluta
$160 \leq x < 165$	162,5	||||	4
$165 \leq x < 170$	167,5	||||| ||||| ||||	14
$170 \leq x < 175$	172,5	||||| |||||	10
$175 \leq x < 180$	177,5	||||	4
TOTAL			N = 32

Como analizamos 32 parafusos, a suma das frecuencias absolutas ten que ser 32.

Pero a frecuencia absoluta, como só indica o número de veces que apareceu un valor, non dá moita información. Vexamos un exemplo: o primeiro día de clase, o profesor do ámbito científico tecnolóxico do módulo 3 na presentación da materia indica que o curso pasado suspenderon 5 persoas. Esta información por si mesma non permite ás persoas matriculadas facerse unha idea da posible dificultade da materia, pois se as persoas matriculadas eran 10, suspenderon a metade (o 50 %); pero se as persoas matriculadas eran 50, entón suspende unha de cada 10 (un 10 %). Polo tanto, para poder ter unha idea mellor, debemos poñer en relación a frecuencia absoluta co número total de datos.

O número total de datos represéntase habitualmente por **N**, como podes ver nas actividades resoltas anteriores.

Chámase **frecuencia relativa**, dun valor calquera da variable estatística, a frecuencia absoluta dese valor dividida entre o número total de datos. A frecuencia relativa desígnase por h_i.

Actividade resolta

Calcule as frecuencias relativas nas actividades resoltas anteriores.

Na actividade do número de veces que van ao cine nun mes, xa tiñamos feito o reconto e calculadas as frecuencias absolutas.

Agora, na táboa que tiñamos, engadimos unha columna para calcular as frecuencias relativas.

x_i	Reconto	f_i	h_i = frecuencia relativa
0	||||| |||||		11	$\frac{11}{50}=0{,}22$
1	||||| ||||| ||||| ||||| ||||| ||||| |||	33	$\frac{33}{50}=0{,}66$
2	|||||	5	$\frac{5}{50}=0{,}1$
3			1	$\frac{1}{50}=0{,}02$
TOTAL		N=50	1,00

Cómpre sinalar que a suma das frecuencias relativas sempre vale 1.

Na actividade da lonxitude dos parafusos, xa tiñamos feito o reconto e calculadas as frecuencias absolutas.

Agora, na táboa que tiñamos, engadimos unha columna para calcular as frecuencias relativas.

Clases	x_i = marca de clase	Reconto	f_i	h_i = frecuencia relativa
160 ≤ x < 165	162,5	||||	4	$\frac{4}{32} = 0{,}125$
165 ≤ x < 170	167,5	||||| ||||| ||||	14	$\frac{14}{32} = 0{,}4375$
170 ≤ x < 175	172,5	||||| |||||	10	$\frac{10}{32} = 0{,}3125$
175 ≤ x < 180	177,5	||||	4	$\frac{4}{32} = 0{,}125$
TOTAL			N = 32	1,00

Cómpre sinalar que a suma das frecuencias relativas sempre vale 1.

A frecuencia relativa é o "*tanto por un*" de veces que aparece cada valor da variable. É dicir, se multiplicamos por 100 a frecuencia relativa obtemos o tanto por cento de veces que aparece cada un dos valores.

Así, no primeiro dos exemplos, un 22 % das persoas ás que lles preguntamos non vai ao cine, un 66 % vai unha vez ao mes, un 10 % vai dúas veces ao mes e un 2 % vai ao cine tres veces ao mes.

2.3.3 Frecuencias absolutas e relativas acumuladas

As frecuencias absolutas e as relativas estudadas no punto anterior dannos unha visión concreta de cada un dos valores da variable estatística, pero, ás veces, queremos ver tamén a evolución da variable desde o primeiro dos valores ata un valor concreto. Para poder ver isto, defínense as frecuencias acumuladas, tanto as absolutas como as relativas.

As frecuencias acumuladas das actividades resoltas anteriores permitirán responder, con só mirar a táboa, preguntas do tipo *"cantas persoas van ao cine menos de dúas veces ao mes?"* ou *"cantos parafusos miden menos de 170 milímetros?"*

A **frecuencia absoluta acumulada** dun valor x_i (representarémola por F_i) é o número de veces que aparece o valor x_i, ou calquera dos anteriores a el. É dicir, para calcular a frecuencia absoluta do valor x_i, hai que sumar as frecuencias absolutas de todos os valores anteriores a x_i máis a frecuencia absoluta de x_i.

Chámase **frecuencia relativa acumulada**, dun valor calquera da variable estatística, a frecuencia absoluta acumulada dese valor dividida entre o número total de datos. A frecuencia relativa acumulada desígnase por H_i.

Actividade resolta

Calcule as frecuencias absolutas acumuladas e as frecuencias relativas acumuladas nas actividades resoltas anteriores.

Na actividade do número de veces que van ao cine nun mes, xa tiñamos feito o reconto e calculadas as frecuencias absolutas e as relativas.

Agora, na táboa que tiñamos, engadimos dúas columnas, unha para as frecuencias absolutas acumuladas e outra para as frecuencias relativas acumuladas. Por cuestións de espazo, suprimiremos a columna do reconto.

x_i	f_i	h_i	F_i = frecuencia absoluta acumulada	H_i = frecuencia relativa acumulada
0	11	$\frac{11}{50} = 0{,}22$	11	$\frac{11}{50} = 0{,}22$
1	33	$\frac{33}{50} = 0{,}66$	44	$\frac{44}{50} = 0{,}88$
2	5	$\frac{5}{50} = 0{,}1$	49	$\frac{49}{50} = 0{,}98$
3	1	$\frac{1}{50} = 0{,}02$	50	$\frac{50}{50} = 1$
TOTAL	N = 50	1,00		

É importante decatarse de que a frecuencia absoluta acumulada do maior valor da variable ten que ser igual ao número de datos (N), pois todos os valores obtidos son menores ou iguais có maior valor da variable.

Do mesmo xeito, e pola mesma razón, a frecuencia relativa acumulada do maior valor da variable ten que ser igual a 1.

Na actividade da lonxitude dos parafusos, xa tiñamos feito o reconto e calculadas as frecuencias absolutas e as relativas.

Agora, na táboa que tiñamos, engadimos dúas columnas, unha para as frecuencias absolutas acumuladas e outra para as frecuencias relativas acumuladas. Por cuestións de espazo suprimiremos a columna do reconto.

Clases	x_i	f_i	h_i	F_i = frecuencia absoluta acumulada	H_i = frecuencia relativa acumulada
$160 \leq x < 165$	162,5	4	$\frac{4}{32} = 0{,}125$	4	$\frac{4}{32} = 0{,}125$
$165 \leq x < 170$	167,5	14	$\frac{14}{32} = 0{,}4375$	18	$\frac{18}{32} = 0{,}5625$
$170 \leq x < 175$	172,5	10	$\frac{10}{32} = 0{,}3125$	28	$\frac{28}{32} = 0{,}875$
$175 \leq x < 180$	177,5	4	$\frac{4}{32} = 0{,}125$	32	$\frac{32}{32} = 1$
TOTAL		N = 32	1,00		

Repare en que a frecuencia absoluta acumulada do maior valor da variable é igual ao número de datos (32) e que a súa frecuencia relativa acumulada tamén vale 1.

De todas as formas, cómpre sinalar que para calcular as frecuencias acumuladas na táboa podemos proceder da seguinte forma:

O primeiro valor das frecuencias acumuladas coincide co primeiro valor das frecuencias.

x_i	f_i	h_i	F_i = frecuencia absoluta acumulada	H_i = frecuencia relativa acumulada
0	11	$\frac{11}{50}=0{,}22$	11	$\frac{11}{50}=0{,}22$
1	33	$\frac{33}{50}=0{,}66$	44	$\frac{44}{50}=0{,}88$
2	5	$\frac{5}{50}=0{,}1$	49	$\frac{49}{50}=0{,}98$
3	1	$\frac{1}{50}=0{,}02$	50	$\frac{50}{50}=1$
TOTAL	N = 50	1,00		

O resto dos valores poden obterse da seguinte forma:

- Na frecuencia absoluta acumulada, o valor seguinte obtense sumando a frecuencia absoluta acumulada do valor anterior coa a frecuencia absoluta deste valor.

x_i	f_i	h_i	F_i = frecuencia absoluta acumulada	H_i = frecuencia relativa acumulada
0	11	$\frac{11}{50}=0{,}22$	11	$\frac{11}{50}=0{,}22$
1	33	$\frac{33}{50}=0{,}66$	44	$\frac{44}{50}=0{,}88$
2	5	$\frac{5}{50}=0{,}1$	49	$\frac{49}{50}=0{,}98$
3	1	$\frac{1}{50}=0{,}02$	50	$\frac{50}{50}=1$
TOTAL	N = 50	1,00		

- Na frecuencia relativa acumulada, o valor seguinte obtense sumando a frecuencia relativa acumulada do valor anterior coa frecuencia relativa deste valor.

x_i	f_i	h_i	F_i = frecuencia absoluta acumulada	H_i = frecuencia relativa acumulada
0	11	$\frac{11}{50}=0{,}22$	11	$\frac{11}{50}=0{,}22$
1	33	$\frac{33}{50}=0{,}66$	44	$\frac{44}{50}=0{,}88$
2	5	$\frac{5}{50}=0{,}1$	49	$\frac{49}{50}=0{,}98$
3	1	$\frac{1}{50}=0{,}02$	50	$\frac{50}{50}=1$
TOTAL	N = 50	1,00		

Actividades propostas

S13. Un concello revisa 50 pisos para saber canta xente vive en cada un deles obtendo os seguintes resultados:

0	2	4	1	0	2	3	3	1	0	4	2	3	0	1	4	2	4	1	0	5	2	1	3	0
4	2	4	3	5	1	2	3	4	0	1	2	3	2	1	3	2	0	1	4	2	3	1	2	0

Elabore a táboa de frecuencias incluíndo as relativas, as absolutas, as relativas acumuladas e as absolutas acumuladas.

S14. As posibles respostas a unha enquisa son: MB (moi bo), B (bo), R (regular), M (malo) e MM (moi malo). As respostas de 50 persoas foron as seguintes:

R	B	MB	M	R	R	MM	MB	M	R
R	MM	R	B	R	MB	R	R	MB	R
M	R	B	R	MB	R	R	B	R	R
M	R	B	R	MB	R	R	B	R	MM
R	R	B	R	R	M	R	B	B	R

Elabore a táboa de frecuencias e responda as seguintes preguntas:

a) Cantas persoas responden mal ou moi mal?

b) Que porcentaxe de persoas responden ben ou moi ben?

S15. A seguinte táboa representa a puntuación obtida por 100 estudantes nun test que constaba de oito preguntas.

Puntos	0	1	2	3	4	5	6	7	8
Nº persoas	0	2	6	9	18	22	24	12	7

Elabore a táboa de frecuencias e responda as seguintes preguntas:

a) Cantas persoas obteñen 5 puntos? Que porcentaxe representan?

b) Cantos persoas teñen 6 ou máis puntos? Que porcentaxe representan?

2.3.4 Representacións gráficas

Para representar graficamente as frecuencias, tanto absolutas como relativas, dispoñemos de varias alternativas:

- Diagrama de barras.
- Polígono de frecuencias.
- Diagrama de sectores.
- Histogramas.

Diagrama de barras

Utilízase para representar táboas de frecuencias correspondentes a variables cuantitativas discretas. Por iso, as barras son estreitas e sitúanse sobre os valores puntuais da variable. Tamén se pode utilizar para representar variables cualitativas.

Para debuxar un diagrama de barras, representamos no eixe horizontal os diferentes valores da variable (x_i) e no eixe vertical a frecuencia que desexemos representar; sobre cada valor da variable debuxamos unha barra ata a altura da frecuencia que queremos representar.

Actividade resolta

Debuxe os diagramas de barras para as frecuencias absolutas e as absolutas acumuladas correspondentes á variable estatística “nota obtida no exame de matemáticas” nunha clase de 20 estudantes onde as notas acadadas son as seguintes:

10	1	3	9	7	5	4	5	6	2	5	5	3	4	5	9	5	4	6	8

Calculamos as frecuencias absolutas e as absolutas acumuladas que aparecen na táboa seguinte.

x_i	f_i	F_i
1	1	1
2	1	2
3	2	4
4	3	7
5	6	13
6	2	15
7	1	16
8	1	17
9	2	19
10	1	20
TOTAL	N = 20	

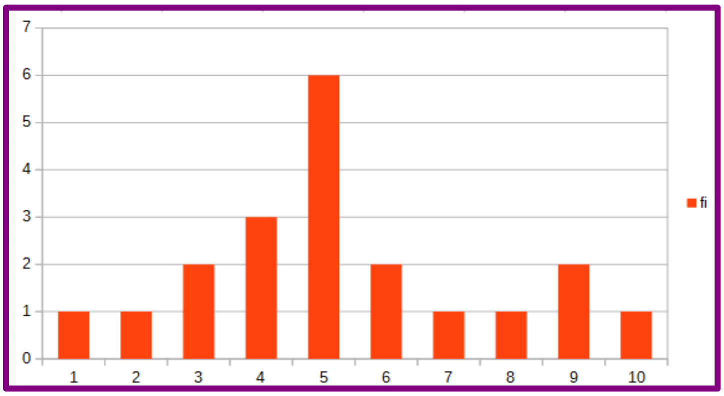


Diagrama de barras das frecuencias absolutas

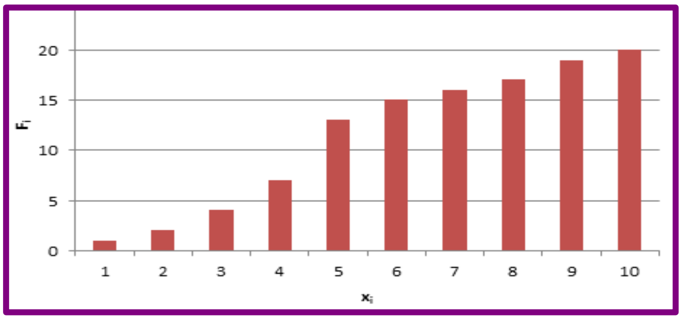


Diagrama de barras das frecuencias absolutas acumuladas

No caso de que sexa unha variable continua agrupada en clases, tomaremos como valores da variable as marcas de clase.

Polígono de frecuencias

Para debuxar un polígono de frecuencias representamos no eixe horizontal os diferentes valores da variable (x_i) e no eixe vertical a frecuencia que desexemos representar. Sobre cada valor da variable debuxamos un punto á altura da frecuencia que queremos representar e posteriormente unimos ditos puntos.

Actividade resolta

Constrúa os polígonos de frecuencias para as frecuencias relativas e as relativas acumuladas correspondentes á variable estatística da actividade resolta anterior.

10	1	3	9	7	5	4	5	6	2	5	5	3	4	5	9	5	4	6	8

Calculamos as frecuencias relativas e as relativas acumuladas que aparecen na táboa seguinte.

x_i	f_i	h_i	H_i
1	1	0,05	0,05
2	1	0,05	0,1
3	2	0,1	0,2
4	3	0,15	0,35
5	6	0,3	0,65
6	2	0,1	0,75
7	1	0,05	0,8
8	1	0,05	0,85
9	2	0,1	0,95
10	1	0,05	1
TOTAL	N = 20	1	

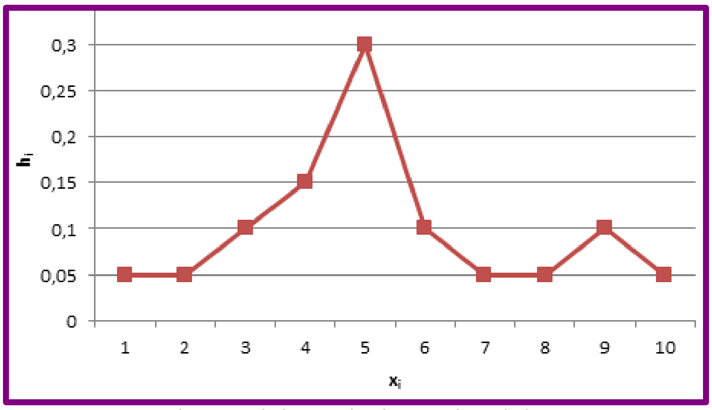


Diagrama de barras das frecuencias relativas

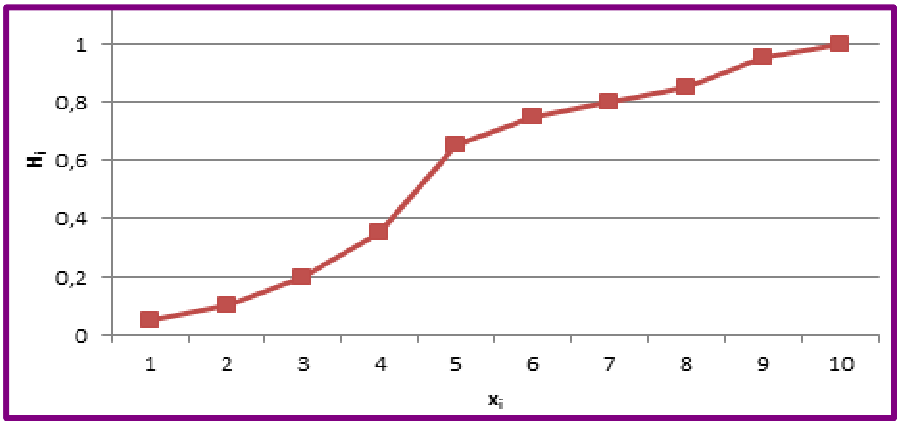


Diagrama de barras das frecuencias relativas acumuladas

Do mesmo xeito que para o diagrama de barras, no caso de que sexa unha variable continua agrupada en clases, tomaremos como valores da variable as marcas de clase.

Diagramas de sectores

Úsanse para representar frecuencias absolutas e frecuencias relativas. Non se usa para representar as frecuencias acumuladas, nin as absolutas nin as relativas.

Para debuxar un diagrama de sectores debuxamos un círculo e, a modo de tortas de cores, representamos proporcionalmente a frecuencia no ángulo de cada sector. Pódese utilizar para todo tipo de variables, pero úsase, frecuentemente, para as variables cualitativas.

Actividade resolta

Debuxe os diagramas de sectores para as frecuencias absolutas da variable estatística da actividade resolta anterior "nota obtida no exame de matemáticas nunha clase de 20 estudantes".

Xa tiñamos as frecuencias absolutas que aparecen na táboa seguinte.

x_i	f_i
1	1
2	1
3	2
4	3
5	6
6	2
7	1
8	1
9	2
10	1
TOTAL	N = 20

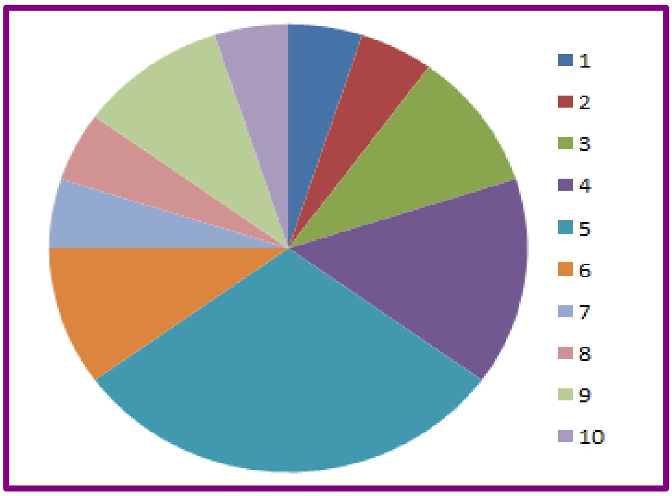


Diagrama de sectores das frecuencias absolutas

No caso de que sexa unha variable continua agrupada en clases, tomaremos como valores da variable as marcas de clase.

Histogramas

Úsanse para representas frecuencias absolutas e frecuencias relativas para variables continuas.

Para debuxar un histograma debuxamos no eixe horizontal as clases. Sobre cada clase debuxamos un rectángulo coa área proporcional á frecuencia que queremos representar. Se todas as clases son da mesma amplitude, o histograma constrúese como o diagrama de barras, pero neste caso o ancho da barra é o do intervalo de clase.

Actividade resolta

No apartado 2.3.2 traballamos nunha actividade resolta que trataba sobre a lonxitude dun tipo de parafusos que se fan nunha fábrica. Tiñamos unha mostra de 32 parafusos cos seguintes resultados en milímetros.

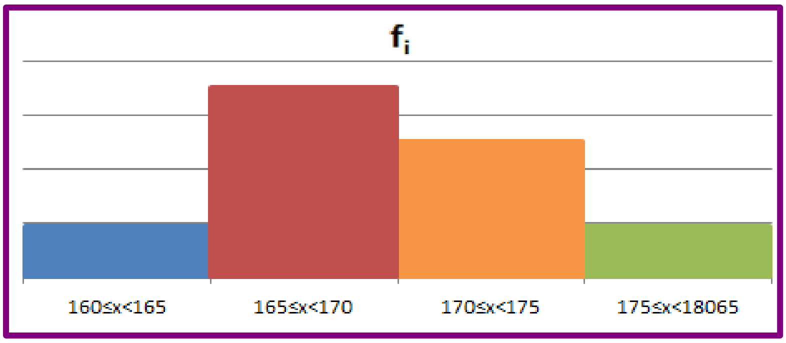
161	171	167	172	170	170	165	169	170	169	172	162	169	166	174	178
167	169	168	176	169	162	168	167	175	168	164	179	172	167	170	173

Agrupábamos os datos en clases e establecíamos as marcas de clase.

Debuxe o histograma correspondente a esta variable.

Xa tiñamos a táboa seguinte:

Intervalos	x_i	f_i
160 ≤ x < 165	162,5	4
165 ≤ x < 170	167,5	14
170 ≤ x < 175	172,5	10
175 ≤ x < 180	177,5	4
TOTAL		N = 32



Actividade resolta

Nun teatro miden a duración dos aplausos: 40 veces os aplausos duraron menos de 10 segundos, 30 veces duraron entre 10 e 15 segundos, 20 veces duraron entre 15 e 20 segundos e 36 veces duraron entre 20 e 30 segundos. Debuxe o histograma correspondente.

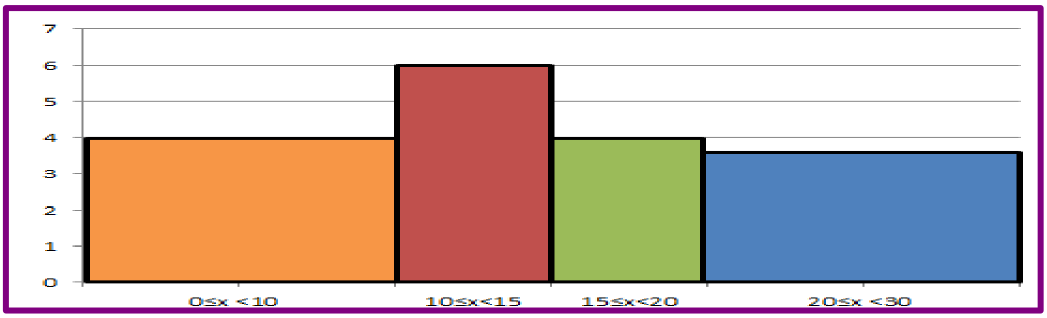
Como as clases teñen diferente anchura, debemos ver a altura que ten que ter cada rectángulo, xa que a frecuencia de cada intervalo vén dada pola área do mesmo e non pola súa altura.
Como os intervalos 1º e 4º teñen o dobre de anchura cós outros, para que a área sexa 40, teñen que ter como altura a metade da frecuencia.

A táboa é:

Intervalos	f_i
0 ≤ x < 10	40
10 ≤ x < 15	30
15 ≤ x < 20	20
20 ≤ x < 30	36
TOTAL	N = 126

Para calcular a altura de cada intervalo dividimos a frecuencia absoluta entre a anchura do intervalo e obtemos:

Intervalos	f_i	Altura do histograma neste intervalo f_i / ancho intervalo
0 ≤ x < 10	40	40 / 10 = 4
10 ≤ x < 15	30	30 / 5 = 6
15 ≤ x < 20	20	20 / 5 = 4
20 ≤ x < 30	36	36 / 10 = 3,6
TOTAL	N = 126	



E o que indica a cantidade de datos que hai en cada un dos intervalos é a área do rectángulo e non a súa altura.

Actividades propostas

S16. A frecuencia con que acode por semana á biblioteca o alumnado dun centro escolar pódese observar na táboa seguinte. Realice un diagrama de barras, un de sectores e un polígono de frecuencias.

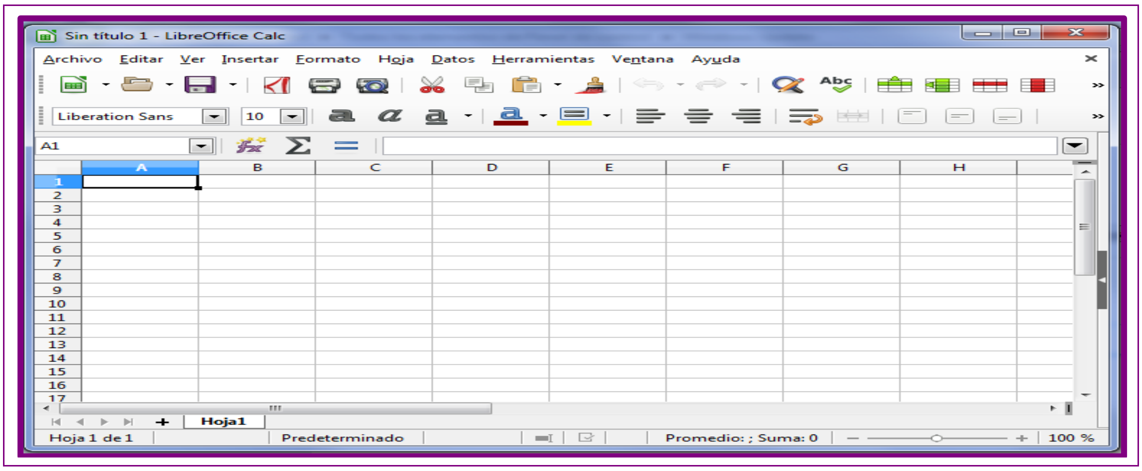
Frecuencia	Nunca	Unha vez	Dúas veces	Máis de dúas veces
Nº persoas	480	540	120	60

S17. Nun concurso de crebacabezas participan 54 persoas: 8 persoas tardan menos de 5 minutos en facelo, 12 tardan entre 5 e 10 minutos, 15 tardan entre 10 e 15 minutos, 10 tardan entre 15 e 20 minutos e 9 tardan entre 20 e 25 minutos. Debuxe o histograma que representa esta distribución.

2.4 Aplicacións informáticas para a representación gráfica de datos estatísticos

2.4.1 Diagramas de barras con LibreOffice Calc

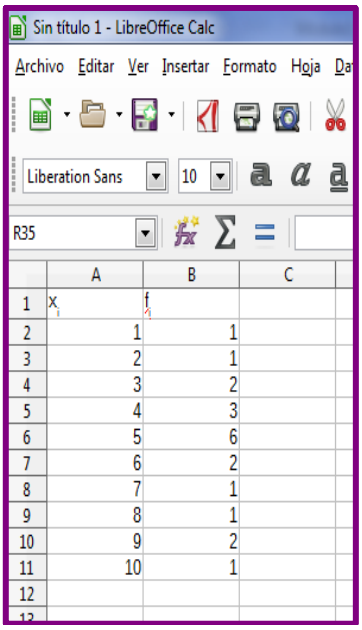
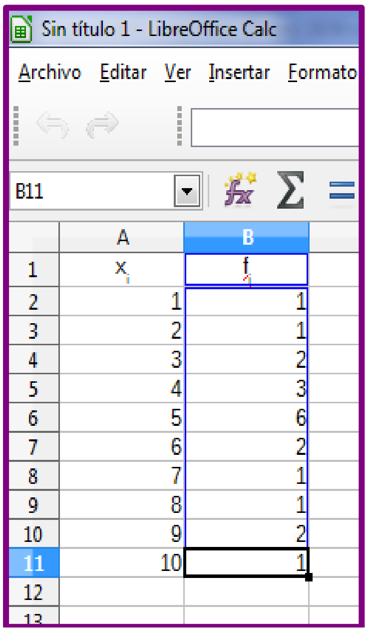
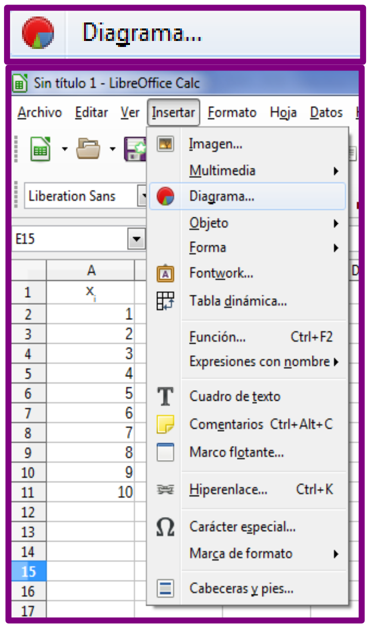
Para realizar estas representacións gráficas utilizaremos unha folla de cálculo. Unha folla de cálculo é un cadro formado por celas no que se poden colocar números, textos ou fórmulas. Cada cela identifícase cunha letra, que indica a columna, e un número, que indica a fila.



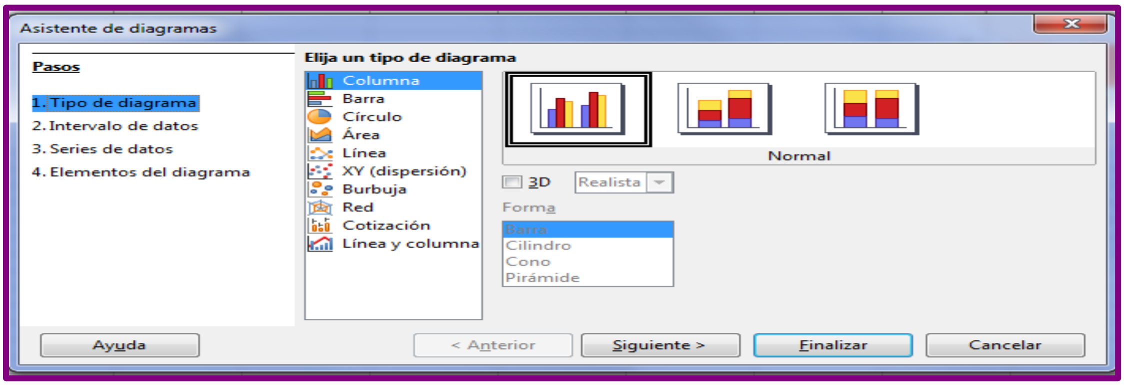
Debuxaremos o diagrama de barras para as frecuencias absolutas correspondentes á variable estatística “nota obtida no exame de matemáticas” nunha clase de 20 estudantes, xa feita no apartado anterior, onde as notas acadadas son as seguintes:

10	1	3	9	7	5	4	5	6	2	5	5	3	4	5	9	5	4	6	8

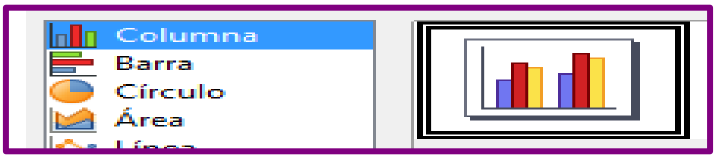
Para facelo, introducimos nas celas da cuadrícula o valor da variable na primeira columna e a súa frecuencia absoluta na segunda.

	Seleccionamos os datos da columna B	A continuación, seleccionamos no menú a opción “Inserir”
		

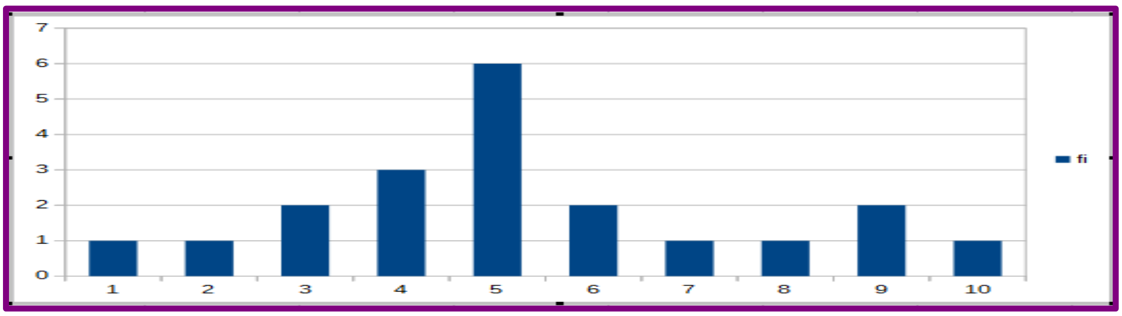
Aparece, entón, o seguinte cadro de diálogo:



Como queremos facer un diagrama de barras, escollemos a opción columna.

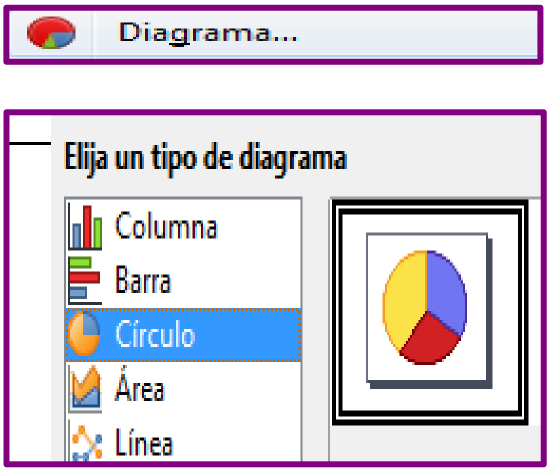


Prememos en “Finalizar” e aparécenos o gráfico seguinte:

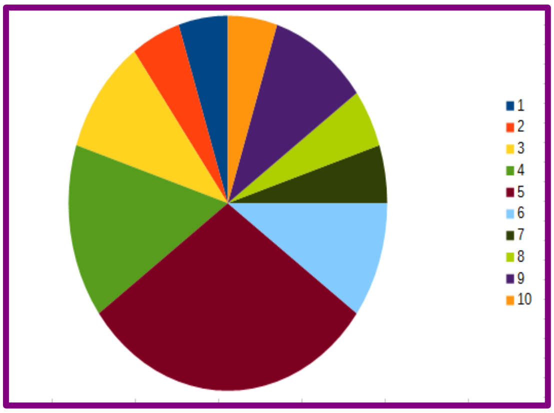


2.4.2 Diagramas de sectores con LibreOffice Calc

O proceso é o mesmo que para facer os diagramas de barras, pero cando seleccionamos “Inserir” escollemos:



Prememos en “Finalizar” e aparécenos o gráfico seguinte:



2.5 Parámetros estatísticos

Despois de obter os datos dunha distribución, necesitamos sintetizar a información para a súa posterior análise. Para iso, obteremos os parámetros estatísticos que serán de tres tipos: de centralización, de dispersión e de posición.

- Os **parámetros de centralización** indícannos arredor de que valor se distribúen os datos.
- Os **parámetros de dispersión** infórmannos sobre canto se afastan do centro os valores da distribución.
- Os **parámetros de posición** dividen un conxunto de datos en grupos co mesmo número de individuos. Para calculalos, cómpre que os **datos** estean ordenados de **menor a maior**.

2.5.1 Parámetros de centralización

Dentro dos parámetros de centralización temos a media, a mediana e a moda.

Media

A **media** é o valor central da distribución. As diferenzas que lles faltan aos valores que non alcanzan a media compénsanse co que lles sobra aos valores que son maiores cá media.

A media denótase por $\bar{x}$.

Para obtela, sumamos todos os valores que toma a variable (o número de veces que aparece cada un deles) e dividimos entre o número total de datos.

Outra forma máis fácil de obtela é multiplicar cada valor da variable (x_i) pola súa frecuencia absoluta (f_i), sumar eses produtos e o resultado dividilo entre o número total de datos (N).

É dicir:

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

Pero a suma de varios sumandos coa mesma estrutura, como pasa neste caso, represéntase utilizando o signo Σ, que indica desde que valor ata que valor vai a suma. Polo tanto, temos que:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \cdots + x_n \cdot f_n}{N} = \frac{\sum_{i=1}^{n} x_i \cdot f_i}{N}$$

Actividade resolta

Calcule a media da variable estatística "nota obtida no exame de matemáticas" nunha clase de 20 estudantes onde as notas acadadas son as seguintes:

10	1	3	9	7	5	4	5	6	2	5	5	3	4	5	9	5	4	6	8

Unha vez que temos as columnas dos valores da variable (x_i) e das frecuencias absolutas (f_i), debemos calcular unha nova columna que vai conter os produtos dos números das dúas columnas anteriores ($x_i \cdot f_i$).

x_i	f_i	$x_i \cdot f_i$
1	1	1
2	1	2
3	2	6
4	3	12
5	6	30
6	2	12
7	1	7
8	1	8
9	2	18
10	1	10
TOTAL	N = 20	$\sum_{i=1}^{10} x_i \cdot f_i = 106$

Así, temos que a media é:

$$\bar{x} = \frac{106}{20} = 5{,}3$$

Mediana

Se ordenamos os datos da distribución de menor a maior, a **mediana** é o valor que se atopa no medio; é dicir, é o valor que deixa tantos individuos antes como despois. Se o número de datos fose par, á mediana asignáselle o valor medio dos dous termos centrais.

A mediana denótase por Me.

Para obtela deberiamos ordenar todos os valores da variable e ver cal ocupa o lugar central. Pero isto pode levar moito tempo se o número de datos é moi grande, polo que debemos atopar unha maneira máis eficaz para calculala.

Calcularemos a mediana na actividade resolta anterior. Para iso vannos facer falta as frecuencias absolutas acumuladas.

Actividade resolta

Calcule a mediana da variable estatística “nota obtida no exame de matemáticas” nunha clase de 20 estudantes da actividade resolta anterior.

Unha vez que temos as columnas dos valores da variable (x_i) e das frecuencias absolutas (f_i), debemos ter tamén a columna da frecuencia absoluta acumulada (F_i).

x_i	f_i	F_i
1	1	1
2	1	2
3	2	4
4	3	7
5	6	13
6	2	15
7	1	16
8	1	17
9	2	19
10	1	20
TOTAL	N = 20	

Se interpretamos a columna das frecuencias absolutas acumuladas, temos que o valor máis pequeno da variable é 1, o segundo é 2, o terceiro e o cuarto valen 3, os tres seguintes (é dicir, o 5º, o 6º e o 7º) valen 4, os seis seguintes valen 5, e así sucesivamente.

É dicir, a columna das frecuencias absolutas acumuladas ordenounos os datos.

Só falta saber cal é o lugar do medio:

- Se o número de datos é impar, hai un único lugar no medio.
- Se o número de datos é par, hai dous lugares no medio.

Neste caso, o número de datos é 20, que é par, polo que hai dous valores nos lugares do medio.

Os lugares do medio son a metade de N e o seguinte, é dicir, 10 e 11.

Como a frecuencia absoluta acumulada do 4 é 7 e a do 5 é 13, temos que os lugares 8º, 9º, 10º, 11º, 12º e 13º valen 5. Polo tanto, os lugares 10º e 11º están ocupados por dous cincos.

Así, temos que a mediana é:

$$Me = \frac{5+5}{2} = 5$$

Moda

A **moda** é o valor que ten a maior frecuencia absoluta.

A moda denótase por Mo.

A moda non ten por que ter un único valor. No caso de que haxa varios valores que teñan a mesma frecuencia absoluta e que esta sexa a maior que aparece, teriamos unha variable con varias modas.

Actividade resolta

Calcule a moda da variable estatística “nota obtida no exame de matemáticas” nunha clase de 20 estudantes da actividade resolta anterior.

x_i	f_i
1	1
2	1
3	2
4	3
5	6
6	2
7	1
8	1
9	2
10	1
TOTAL	N = 20

Unha vez que temos as columnas dos valores da variable (x_i) e das frecuencias absolutas (f_i), non necesitamos nada máis para calcular a moda, xa que só temos que buscar cal é o valor que ten a frecuencia absoluta maior.

Como o valor que ten a frecuencia absoluta maior é $x_i = 5$, xa que a súa $f_i = 6$, temos que a moda é:

$$Mo = 5$$

A

Actividade resolta

As idades dos asistentes á festa de aniversario da súa filla son:

12	10	11	13	12	11	13	12	13	13	12	13	11	12	13
13	11	12	10	12	11	14	12	14	12	10	14	13	11	13

Calcule a media, a mediana e a moda.

- Para calcular a media necesitamos tres columnas da táboa: x_i, f_i e $x_i \cdot f_i$.
- Para calcular a mediana necesitamos tres columnas: x_i, f_i e F_i (en realidade f_i só é necesaria para calcular F_i, non para calcular a mediana).
- Para calcular a moda necesitamos dúas columnas: x_i e f_i.

Polo tanto, para calcular estes tres parámetros de centralización necesitamos catro columnas da táboa, que son:
x_i, f_i, F_i e $x_i \cdot f_i$.

x_i	f_i	F_i	$x_i \cdot f_i$
10	3	3	30
11	6	9	66
12	9	18	108
13	9	27	117
14	3	30	42
TOTAL	N = 30		363

Así temos:

- Media:

$$\bar{x} = \frac{363}{30} = 12{,}1$$

- Mediana: os lugares do medio (15º e 16º) están ocupados polo valor $x_i = 12$

$$Me = \frac{12+12}{2} = 12$$

- Moda: hai dúas modas. Dise que é unha distribución bimodal. As modas son:

$$Mo = 12 \text{ e } Mo = 13$$

Esta información que ofrecen os parámetros de centralización dinos que estes datos están todos arredor do valor 12. Xorde, entón, a seguinte pregunta: se todos están arredor do valor 12, son todos próximos a 12?

Esta pregunta ten sentido se pensamos que para obtermos 12 de media, podemos partir de 2 e 22 ou ben de 10 e 14. En ambos os casos a media é 12, pero os datos de partida son ben diferentes. Isto fai necesario coñecer máis sobre os datos da distribución, e para iso temos os parámetros de dispersión, que nos informarán de como están de aproximados os datos da táboa.

Actividades propostas

S18. As notas obtidas nun exame de matemáticas da ESA coas súas frecuencias absolutas son as seguintes:

x_i	0	1	2	3	4	5	6	7	8	9	10
f_i	0	5	4	2	2	1	1	2	3	4	8

Calcule a media, a moda e a mediana.

S19. As estaturas de varias persoas (xa agrupadas en clases coas marcas de clase correspondentes) e as súas frecuencias absolutas son as seguintes:

x_i	151	156	161	166	171	176
f_i	2	5	11	14	5	3

Calcule a media, a moda e a mediana.

2.5.2 Parámetros de dispersión

Dentro dos parámetros de dispersión temos a desviación media, a varianza e a desviación típica.

O que calculan os parámetros de dispersión é a media das distancias dos datos con respecto á media. Pero non se pode ter en conta o signo destas distancias (negativo se o dato é menor ca media e positivo cando o dato é maior ca media). Lembre que cando definimos a media dixemos que "*as diferenzas que lles faltan aos valores que non alcanzan a media compénsanse co que lles sobra aos valores que son maiores cá media*." Polo que, se consideramos o signo das diferenzas, a súa suma sempre será cero.

Polo tanto, para calcularmos a media das distancias dos datos respecto á media da distribución debemos transformar esas distancias en positivas. Iso podémolo facer de dúas formas.

- Mediante o valor absoluto, e así calculamos a desviación media.
- Elevando as distancias ao cadrado, e así calculamos a varianza.

Pero antes de ver como se calculan a desviación media e a varianza, veremos que é o rango ou percorrido dunha variable estatística.

Rango ou percorrido

O **rango** ou **percorrido** é a diferenza entre os datos maior e menor da variable. Vén sendo a lonxitude do tramo dentro do cal están os datos.

Actividade resolta

As idades dos asistentes á festa de aniversario da súa filla son:

12	10	11	13	12	11	13	12	13	13	12	13	11	12	13
13	11	12	10	12	11	14	12	14	12	10	14	13	11	13

Calcule o percorrido desta variable.

Para calcular o percorrido, necesitamos localizar os valores máximo e mínimo que toma esta variable.

- O valor mínimo é 10.
- O valor máximo é 14.
- Polo tanto, o rango é:

$$\text{rango} = 14 - 10 = 4$$

Desviación media

A **desviación media** é a media das distancias dos datos á media. Calcúlase mediante a media das diferenzas en valor absoluto. A desviación media denótase por DM.

$$DM = \frac{|x_1 - \bar{x}| \cdot f_1 + |x_2 - \bar{x}| \cdot f_2 + \cdots + |x_n - \bar{x}| \cdot f_n}{N} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}| \cdot f_i}{N}$$

Actividade resolta

Na mesma actividade resolta anterior, as idades dos asistentes á festa de aniversario da súa filla eran:

12	10	11	13	12	11	13	12	13	13	12	13	11	12	13
13	11	12	10	12	11	14	12	14	12	10	14	13	11	13

Calcule a desviación media

Para calcular a desviación media necesitamos:

- Calcular a media, polo que necesitamos tres columnas da táboa: x_i, f_i e $x_i \cdot f_i$.
- Calcular a media de $|x_i - \bar{x}|$, polo que necesitamos tres columnas: $|x_i - \bar{x}|$, f_i e $|x_i - \bar{x}| \cdot f_i$.

Polo tanto, para calcular a desviación media necesitamos cinco columnas da táboa que son:

x_i, f_i, $x_i \cdot f_i$, $|x_i - \bar{x}|$, e $|x_i - \bar{x}| \cdot f_i$.

x_i	f_i	$x_i \cdot f_i$
10	3	30
11	6	66
12	9	108
13	9	117
14	3	42
TOTAL	N = 30	363

Así temos.

- Media:

$$\bar{x} = \frac{363}{30} = 12{,}1$$

Como xa calcularamos anteriormente (ver o apartado de medidas de centralización).

Unha vez calculada a media, xa podemos calcular a desviación media.

x_i	f_i	$x_i \cdot f_i$	$|x_i - \bar{x}|$	$\boldsymbol{|x_i - \bar{x}| \cdot f_i}$
10	3	30	2,1	6,3
11	6	66	1,1	6,6
12	9	108	0,1	0,9
13	9	117	0,9	8,1
14	3	42	1,9	5,7
TOTAL	N = 30	363		27,6

Así, temos:

- Desviación media:

$$DM = \frac{27{,}6}{30} = 0{,}92$$

Varianza e desviación típica

A **varianza** é a media dos cadrados das distancias dos datos á media. A varianza denótase por σ^2.

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 \cdot f_1 + (x_2 - \bar{x})^2 \cdot f_2 + \cdots + (x_n - \bar{x})^2 \cdot f_n}{N} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot f_i}{N}$$

A varianza ten o problema de que as unidades en que se expresa, ao estaren elevadas ao cadrado, desvirtúan as medidas. Así, por exemplo, se estudamos a estatura en cm, ao elevarmos ao cadrado, as unidades serían cm^2 e isto non representa unha lonxitude, senón unha superficie. Por iso extraemos a súa raíz cadrada, que é o que chamamos desviación típica.

A **desviación típica** é a raíz cadrada da varianza. A desviación típica denótase por σ.

$$\sigma = \sqrt{\sigma^2}$$

Actividade resolta

Na mesma actividade resolta anterior, as idades dos asistentes á festa de aniversario da súa filla eran:

12	10	11	13	12	11	13	12	13	13	12	13	11	12	13
13	11	12	10	12	11	14	12	14	12	10	14	13	11	13

Calcule a desviación típica.

Para calcular a desviación típica necesitamos calcular a varianza e para calcular a varianza necesitamos:

- Calcular a media, polo que necesitamos tres columnas da táboa: x_i, f_i e $x_i \cdot f_i$.
- Calcular a media de $(x_i - \bar{x})^2$, polo que necesitamos catro columnas: $(x_i - \bar{x})$, $(x_i - \bar{x})^2$, f_i e $(x_i - \bar{x})^2 \cdot f_i$.

Polo tanto, para calcular a desviación típica necesitamos seis columnas da táboa, que son:

x_i, f_i, $x_i \cdot f_i$, $(x_i - \bar{x})$, $(x_i - \bar{x})^2$, e $(x_i - \bar{x})^2 \cdot f_i$.

x_i	f_i	$x_i \cdot f_i$
10	3	30
11	6	66
12	9	108
13	9	117
14	3	42
TOTAL	N = 30	363

Así, temos:

- Media:

$$\bar{x} = \frac{363}{30} = 12{,}1$$

Como xa calcularamos anteriormente (ver o apartado de medidas de centralización).

Unha vez calculada a media, xa podemos calcular a varianza e a desviación típica.

x_i	f_i	$x_i \cdot f_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
10	3	30	-2,1	4,41	13,23
11	6	66	-1,1	1,21	7,26
12	9	108	-0,1	0,01	0,09
13	9	117	0,9	0,81	7,23
14	3	42	1,9	3,61	10,83
TOTAL	N = 30	363			38,64

Así, temos:

- Varianza:

$$\sigma^2 = \frac{38{,}64}{30} = 1{,}288$$

E tamén:

- Desviación típica:

$$\sigma = \sqrt{\sigma^2} = \sqrt{1{,}288} = 1{,}1349$$

Actividade resolta

Calcule os parámetros de dispersión da variable estatística "nota obtida no exame de matemáticas" nunha clase de 20 estudantes onde as notas acadadas son as seguintes:

10	1	3	9	7	5	4	5	6	2	5	5	3	4	5	9	5	4	6	8

Para calcular os parámetros de dispersión necesitamos:

- Para o rango, determinar os valores máximo e mínimo da variable.
- Para a desviación media, x_i, f_i e $x_i \cdot f_i$ para calcular a media e $|x_i - \bar{x}|$, f_i e $|x_i - \bar{x}| \cdot f_i$ para a desviación media.
- Para a desviación típica, x_i, f_i e $x_i \cdot f_i$ para calcular a media e $(x_i - \bar{x})$, $(x_i - \bar{x})^2$, f_i e $(x_i - \bar{x})^2 \cdot f_i$ para a varianza.

Polo tanto, para calcular os parámetros de dispersión necesitamos oito columnas da táboa que son:

x_i, f_i, $x_i \cdot f_i$, $|x_i - \bar{x}|$, $|x_i - \bar{x}| \cdot f_i$, $(x_i - \bar{x})$, $(x_i - \bar{x})^2$, e $(x_i - \bar{x})^2 \cdot f_i$.

As columnas $|x_i - \bar{x}|$ e $(x_i - \bar{x})$ son a mesma agás o signo, polo que só calcularemos a columna de $|x_i - \bar{x}|$. En definitiva, son sete columnas.

x_i	f_i	$x_i \cdot f_i$
1	1	1
2	1	2
3	2	6
4	3	12
5	6	30
6	2	12
7	1	7
8	1	8
9	2	18
10	1	10
TOTAL	N = 20	106

Así, temos:

- Media:

$$\bar{x} = \frac{106}{20} = 5{,}3$$

Agora xa podemos calcular os parámetros de dispersión:

- Rango:

$$\text{rango} = 10 - 1 = 9$$

x_i	f_i	$x_i \cdot f_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
1	1	1	4,3	4,3	18,49	18,49
2	1	2	3,3	3,3	10,89	10,89
3	2	6	2,3	4,6	5,29	10,58
4	3	12	1,3	3,9	1,69	5,07
5	6	30	0,3	1,8	0,09	0,54
6	2	12	0,7	1,4	0,49	0,98
7	1	7	1,7	1,7	2,89	2,89
8	1	8	2,7	2,7	7.29	7,29
9	2	18	3,7	7,4	13,69	27,38
10	1	10	4,7	4,7	22.09	22,09
TOTAL	N = 20	106		35,8		106,2

- Desviación media:

$$DM = \frac{35{,}8}{20} = 1{,}79$$

- Varianza:

$$\sigma^2 = \frac{106{,}2}{20} = 5{,}31$$

- Desviación típica:

$$\sigma = \sqrt{\sigma^2} = \sqrt{5{,}31} = 2{,}3043$$

Outra forma de calcular a varianza é utilizando a expresión:

$$\sigma^2 = \frac{{x_1}^2 \cdot f_1 + {x_2}^2 \cdot f_2 + \cdots + {x_n}^2 \cdot f_n}{N} - \bar{x}^2 = \frac{\sum_{i=1}^{n} {x_i}^2 \cdot f_i}{N} - \bar{x}^2$$

Actividade resolta

Calcule a varianza e a desviación típica da actividade resolta anterior utilizando esta última expresión.

Para calcular a varianza deste xeito, en vez de utilizar as columnas $(x_i - \bar{x})$, $(x_i - \bar{x})^2$, f_i e $(x_i - \bar{x})^2 \cdot f_i$ debemos utilizar ${x_i}^2$ e ${x_i}^2 \cdot f_i$.

Xa tiñamos calculada a media que era:

$$\bar{x} = \frac{106}{20} = 5{,}3$$

x_i	f_i	$x_i \cdot f_i$	${x_i}^2$	${x_i}^2 \cdot f_i$
1	1	1	1	1
2	1	2	4	4
3	2	6	9	18
4	3	12	16	48
5	6	30	25	150
6	2	12	36	72
7	1	7	49	49
8	1	8	64	64
9	2	18	81	162
10	1	10	100	100
TOTAL	N = 20	106		668

- Varianza:

$$\sigma^2 = \frac{668}{20} - 5{,}3^2 = 33{,}4 - 28{,}09 = 5{,}31$$

- Desviación típica:

$$\sigma = \sqrt{\sigma^2} = \sqrt{5{,}31} = 2{,}3043$$

Que, naturalmente, coincide co calculado anteriormente.

Actividades propostas

S20. Calcule o rango, a desviación media, a varianza e a desviación típica da variable da actividade proposta número 18.

S21. Calcule o rango, a desviación media e a desviación típica da variable da actividade proposta número 19.

2.5.3 Parámetros de posición. Os cuartís

Dentro dos parámetros de posición só veremos os cuartís.

Os **cuartís** dunha serie estatística son Q_1, Q_2, e Q_3, de tal xeito que dividen a distribución dos datos xa ordenados en catro partes iguais:

- Q_1 deixa á súa esquerda o 25 % dos datos.

- Q_2 deixa á súa esquerda o 50 % dos datos e, polo tanto, coincide coa mediana.
- Q_3 deixa á súa esquerda o 75 % dos datos.

A diferenza $Q_3 - Q_1$ chámase **percorrido intercuartílico**.

Actividade resolta

As idades dos asistentes á festa de aniversario da súa filla son:

12	10	11	13	12	11	13	12	13	13	12	13	11	12	13
13	11	12	10	12	11	14	12	14	12	10	14	13	11	13

Calcule os cuartís

Para calcular os cuartís necesitamos a frecuencia absoluta acumulada.

x_i	f_i	F_i
10	3	3
11	6	9
12	9	18
13	9	27
14	3	30
TOTAL	N = 30	

Así, temos:

- Q_1 deixa tras el a cuarta parte da distribución. Como N = 30, temos que $\frac{N}{4} = 7,5$. Así, Q_1 debe deixar 7,5 elementos á súa esquerda, polo que debe ser o oitavo elemento. Polo tanto, $Q_1 = 11$.
- Q_2 é a mediana que, como N é par, será a media dos elementos que ocupan as posicións centrais, que son as posicións 15ª e 16ª. Polo tanto, $Q_2 = \frac{12+12}{2} = 12$.
- Q_3 deixa tras el as tres cuartas partes da distribución. Como N = 30, temos que $\frac{3N}{4} = 22,5$. Así, Q_3 debe deixar 22,5 elementos á súa esquerda, polo que debe ser o 23º elemento. Polo tanto, $Q_3 = 13$.

Actividades propostas

S22. Calcule os cuartís da variable da actividade proposta número 18.

S23. Calcule os cuartís da variable da actividade proposta número 19.

2.5.4 Diagrama de caixa e bigotes

O **diagrama de caixa e bigotes** é unha forma moi clara de representar os cuartís e o percorrido da variable. Vexámolo nas seguintes actividades resoltas:

Actividade resolta

Calcule o percorrido e os cuartís e faga o diagrama de caixa e bigotes correspondente á variable estatística "nota obtida no exame de matemáticas" nunha clase de 20 estudantes onde as notas acadadas son as seguintes:

10	1	3	9	7	5	4	5	6	2	5	5	3	4	5	9	5	4	6	8

Para calcular os cuartís, necesitamos ordenar os valores e calcular a frecuencia absoluta acumulada.

x_i	f_i	F_i
1	1	1
2	1	2
3	2	4
4	3	7
5	6	13
6	2	15
7	1	16
8	1	17
9	2	19
10	1	20
TOTAL	N = 20	

Así, temos que:

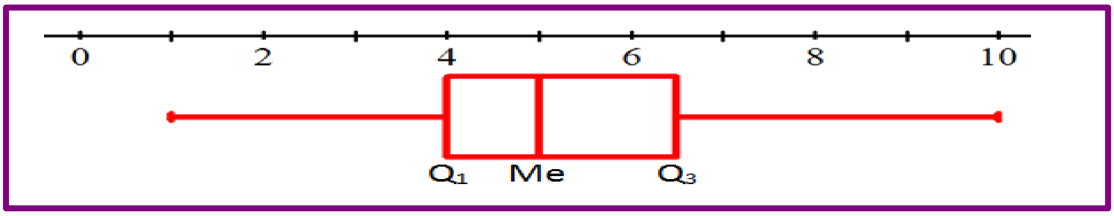
Percorrido: 10 - 1= 9.

Q_1 deixa tras el a cuarta parte da distribución. Como N = 20, temos que $\frac{N}{4}=5$. Así, Q_1 debe deixar 5 elementos á súa esquerda, polo que debe ser a media do 5º e 6º elemento. Polo tanto, $Q_1=\frac{4+4}{2}=4$.

Q_2 é a mediana que, como N é par, será a media dos elementos que ocupan as posicións centrais, que son as posicións 10ª e 11ª. Polo tanto, $Q_2=\frac{5+5}{2}=5$.

Q_3 deixa tras el as tres cuartas partes da distribución. Como N = 20, temos que $\frac{3N}{4}=15$. Así, Q_3 debe deixar 15 elementos á súa esquerda, polo que debe ser a media entre os elementos 15º e o 16º. Polo tanto, $Q_3=\frac{6+7}{2}=6{,}5$.

O diagrama de caixa e bigotes é o seguinte:

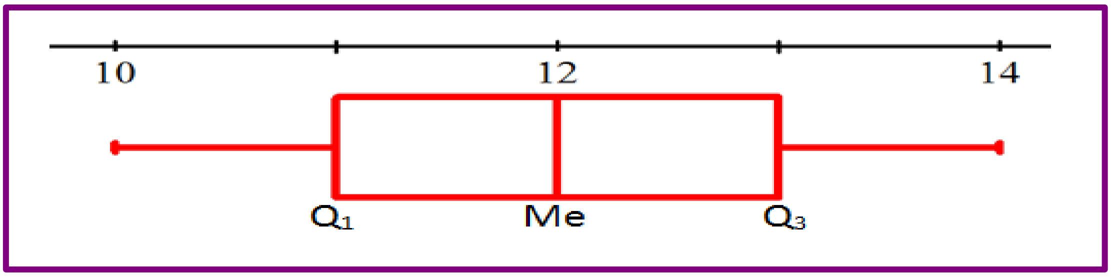


A caixa deste diagrama comprende o percorrido intercuartílico no que se marca a mediana e os bigotes alcanzan os valores máximos e mínimos da distribución.

Actividade resolta

Debuxe o diagrama de caixa e bigotes da actividade resolta do apartado anterior (2.5.3)

Tiñamos calculado que $Q_1 = 11$, $Q_2 = 12$ e $Q_3 = 13$. O valor mínimo era 10 e o máximo era 14.



Actividades propostas

S24. Debuxe o diagrama de caixa e bigotes correspondente á variable da actividade proposta número 18.

S25. Debuxe o diagrama de caixa e bigotes correspondente á variable da actividade proposta número 19.

2.5.5 Interpretación conxunta da media e da desviación típica

O **coeficiente de variación** é o cociente entre a desviación típica e a media.

$$CV = \frac{\sigma}{\bar{x}}$$

Exprésase en porcentaxe e serve para medir a dispersión relativa. En xeral, é menos dispersa a variable que presenta menor coeficiente de variación, non a de menor desviación típica, pois a dispersión depende tamén da media. Non é o mesmo unha desviación típica de 5 cando a media vale 10 que unha desviación típica de 5 cando a media vale 1 000. As desviacións típicas son iguais, pero no caso de que a media sexa 1 000, hai menos dispersión relativa que cando a media é 10.

Actividade resolta

As anotacións dos xogadores de dous equipos na liga ACB de baloncesto na última tempada foron os seguintes:

	Xog1	Xog2	Xog3	Xog4	Xog5	Xog6	Xog7	Xog8	Xog9	Xog10	Xog11	Xog12
Equipo 1	570	433	110	266	269	316	204	162	73	36	31	24
Equipo 2	352	388	304	194	197	293	267	182	186	188	148	145

Que equipo ten menor dispersión relativa.

Equipo 1:

A súa media é:

$$\overline{x_1} = \frac{570+433+110+266+269+316+204+162+73+36+31+24}{12} = \frac{2494}{12} = 206{,}83$$

A súa desviación típica é (antes calcularemos a varianza):

$$\sigma_1{}^2 = \frac{570^2+433^2+110^2+266^2+269^2+316^2+204^2+162^2+73^2+36^2+31^2+24^2}{12} - 206{,}83^2 =$$

$$= \frac{843484}{12} - 206{,}83^2 = 70290{,}33 - 42778{,}65 = 27511{,}68 \Rightarrow \sigma_1 = 165{,}87$$

Equipo 2:

A súa media é:

$$\overline{x_2} = \frac{352+388+304+194+197+293+267+182+186+188+148+145}{12} = \frac{2844}{12} = 237$$

A súa desviación típica é (antes calcularemos a varianza):

$$\sigma_2{}^2 = \frac{352^2+388^2+304^2+194^2+197^2+293^2+267^2+182^2+186^2+188^2+148^2+145^2}{12} - 237^2 =$$

$$= \frac{746440}{12} - 237^2 = 62228{,}33 - 56169 = 6059{,}33 \Rightarrow \sigma_2 = 77{,}84$$

Os coeficientes de variación son:

$CV_1 = \frac{\sigma_1}{\overline{x_1}} = \frac{165{,}87}{206{,}83} = 0{,}80$ para o equipo 1 e $CV_2 = \frac{\sigma_2}{\overline{x_2}} = \frac{77{,}84}{237} = 0{,}33$ para o equipo 2.

Polo que, podemos concluír que a dispersión relativa é moitísimo menor no equipo 2.

Actividades propostas

S26. Calcule o coeficiente de variación da variable da actividade proposta 18.

S27. Calcule o coeficiente de variación da variable da actividade proposta 19.

3. Actividades finais

S28. Para cada unha das seguintes variables aleatorias, indique se é cualitativa, cuantitativa discreta ou cuantitativa continua.

Lonxitude dunha uña.	Tempo que dura unha película.
Puntos obtidos en total nun partido de baloncesto.	Idade dunha árbore.
Cor dunha flor.	Olor dun perfume.
Sabor dunha froita.	Número de trens diarios entre A Coruña e Vigo.
Temperatura dun forno.	Peso dun paquete de azucre.

S29. As notas de 20 persoas nunha proba de inglés son as seguintes: 5, 3, 4, 8, 5, 2, 1, 9, 6, 5, 4, 3, 8, 4, 2, 6, 9, 7, 8 e 5. Calcule a media, mediana, moda, rango, desviación media, desviación típica e coeficiente de variación; represente as frecuencias absolutas nun diagramas de sectores e as absolutas acumuladas nun diagrama de barras. Constrúa tamén o correspondente diagrama de caixa e bigotes.

S30. As notas dun grupo de persoas nunha proba de matemáticas son as seguintes:

Puntos	3	4	5	6	7	8	9
Nº persoas	20	48	77	45	24	17	19

Calcule a media, mediana, moda, rango e desviación típica.

S31. As notas dun grupo de persoas nunha proba de lingua galega son as seguintes:

Puntos	3	4	5	6	7	8	9
Nº persoas	5	12	116	53	29	17	18

Calcule a media, mediana, moda, rango e desviación típica.

S32. O prezo do gasóleo nos diferentes países da Unión Europea o 3 de xuño de 2017 eran:

País	Prezo	País	Prezo
Alemaña	1.15 €/l	Grecia	1.329 €/l
Austria	1.018 €/l	Hungría	1.117 €/l
Bélxica	1.313 €/l	Reino Unido	1.359 €/l
Bulgaria	1.045 €/l	Irlanda	1.229 €/l
Chipre	1.166 €/l	Italia	1.404 €/l
Croacia	1.165 €/l	Letonia	1.01 €/l
Dinamarca	1.314 €/l	Lituania	0.942 €/l
Eslovaquia	1.035 €/l	Luxemburgo	0.972 €/l

Eslovenia	1.146 €/l	Malta	1.18 €/l
España	1.101 €/l	Polonia	0.987 €/l
Estonia	1.159 €/l	Portugal	1.289 €/l
Finlandia	1.26 €/l	República Checa	1.115 €/l
Francia	1.219 €/l	Romanía	0.996 €/l
Holanda	1.285 €/l	Suecia	1.33 €/l

Agrupe os datos en clases de amplitude 0,05 sendo a primeira delas a que vai de 0,925 €/l ata 0,975 €/l. Despois debuxe o histograma correspondente.

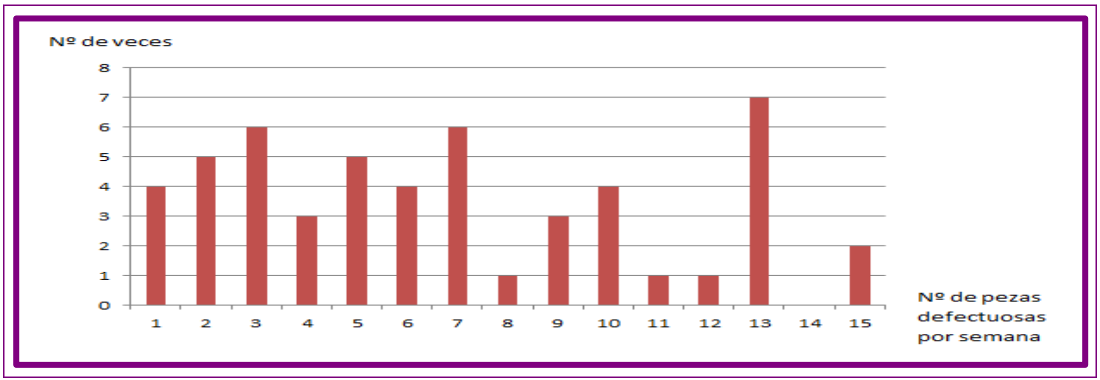
S33. Cos datos do exercicio anterior agrupados nas clases que se indicaron, calcule o coeficiente de variación e debuxe o diagrama de caixa e bigotes.

S34. Nos doce meses dun ano, nun centro meteorolóxico obtivéronse os seguintes valores de temperatura máxima e mínima:

	Xan	Feb	Mar	Abr	Mai	Xuñ	Xull	Ago	Set	Out	Nov	Dec
Temp. min.	4,6	7,8	5,8	6,2	10,8	12,2	13,6	14,0	11,8	10,4	5,0	1,8
Temp. máx.	15,8	23	25,6	24	24,8	23	31,6	28,0	28,2	26,2	20,4	16,2

Indique cal destas dúas variables ten menor dispersión relativa.

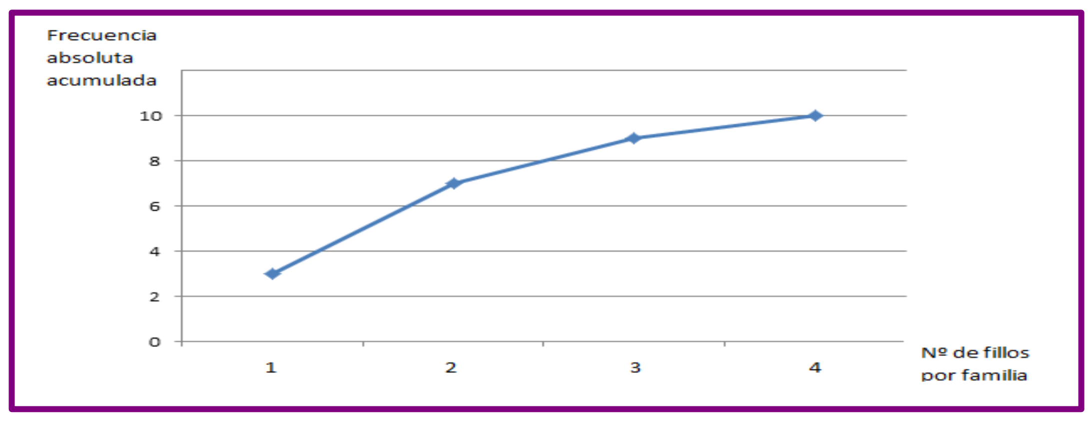
S35. Nunha fábrica de pezas para motores de coches controlouse durante un ano o número de pezas defectuosas que se fabricaban cada semana. Os datos obtidos aparecen recollidos no seguinte diagrama de barras:



Calcule os parámetros de centralización, os de dispersión desta variable e constrúa o seu diagrama de caixa e bigotes.

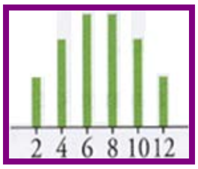
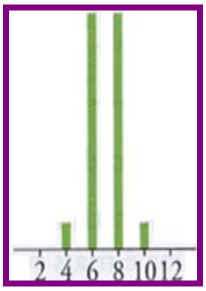
S36. Contabilizáronse o número de materias suspensas na segunda avaliación polo alumnado dun curso de 2º de ESO. Os resultados foron: 1, 3, 1, 0, 4, 0, 1, 1, 2, 3, 1, 1, 2, 1, 2, 4, 1, 0, 2, 3, 2, 3, 1, 1, 6, 0, 0, 2, 1, 4. Calcule o rango, a moda, a mediana, a media e a desviación media.

S37. Preguntouse a dez familias polo número de fillos e fillas que tiñan e o resultado móstrase no seguinte polígono de frecuencias no que se representan as frecuencias absolutas acumuladas:



Calcule coeficiente de variación desta variable.

S38. As tres distribucións seguintes teñen a mesma media, cal é?

A	B	C
		

As súas desviacións típicas son 3'8, 1'3 e 2'9. Observando as gráficas, a cal corresponde cada unha?

4. Solucionario

4.1 Solucións das actividades propostas

S1. As preguntas 1ª e 3ª non son apropiadas para a recollida de datos e a posterior elaboración dun estudo estatístico debido á gran variedade de respostas que podemos obter. As preguntas 2ª e 4ª, ao estar máis dirixidas, van facilitar o tratamento dos datos obtidos coas enquisas.

S2. É unha mostraxe por conglomerados. A poboación son as 30 persoas da clase que ten 30 individuos. A mostra está formada por 3 individuos.

S3. É unha mostraxe de tamaño 5 sobre a poboación "clientes que teñen a nómina domiciliada nalgunha oficina do banco". O banco debe elaborar unha mostraxe aleatoria simple.

S4. É unha mostraxe estratificada.

S5. Trátase dunha mostraxe oportunista e non fiable; a opinión dos viaxeiros entrevistados será, case seguro, mellor cá dos que tomaron voos que sufriron demora.

S6. É unha mostraxe aleatoria simple.

S7. É unha mostraxe aleatoria con reposición de tamaño 7 sobre unha poboación de tamaño 5.

S8. Persoas menores de 40 anos: 1 000; mulleres: 1 090; homes de 40 anos ou máis: 460.

S9.

Horas diarias de sono dos habitantes dunha provincia. A POBOACIÓN SON OS HABITANTES DA PROVINCIA E É NECESARIO OBTER UNHA MOSTRA PARA FACER O ESTUDO DEBIDO AO GRAN TAMAÑO DA POBOACIÓN.	Preferencias literarias das persoas maiores de idade que viven nun edificio. A POBOACIÓN SON AS PERSOAS MAIORES DE IDADE QUE VIVEN NUN EDIFICIO E NON É NECESARIO OBTER UNHA MOSTRA PARA FACER O ESTUDO, XA QUE NADA IMPIDE PREGUNTAR A TODOS OS INDIVIDUOS.

S10. Cualitativas: a música preferida, a cor do pelo. Cuantitativas: o peso, a idade, as cualificacións dun exame.

S11. Discreta, xa que só pode tomar valores naturais.

S12. Continua, xa que a mazá pode tomar calquera valor dentro dun intervalo.

S13.

x_i	f_i	F_i	h_i	H_i
0	9	9	0,18	0,18
1	10	19	0,20	0,38
2	12	31	0,24	0,62
3	9	40	0,18	0,80
4	8	48	0,16	0,96
5	2	50	0,04	1
TOTAL	50		1	

S14.

x_i	f_i	h_i
MB	6	0,12
B	9	0,18
R	27	0,54
M	5	0,10
MM	3	0,06
TOTAL	50	1

Non calculamos as frecuencias acumuladas por non ser datos numéricos.

a) 8 persoas.

b) 30 %.

S15.

x_i	f_i	F_i	h_i	H_i
0	0	0	0	0
1	2	2	0,02	0,02
2	6	8	0,06	0,08
3	9	17	0,09	0,17
4	18	35	0,18	0,35
5	22	57	0,22	0,57
6	24	81	0,24	0,81
7	12	93	0,12	0,93
8	7	100	0,07	1
TOTAL	100		1	

a) 22 persoas que representan o 22 %.

b) 43 persoas que representan o 43 %.

S16.

Diagrama de barras

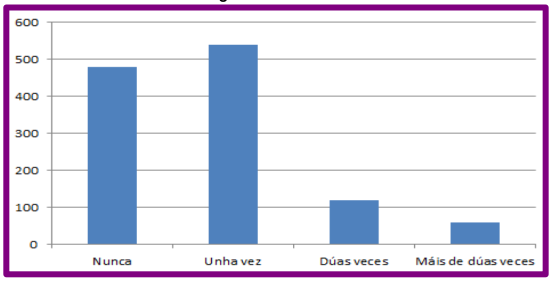
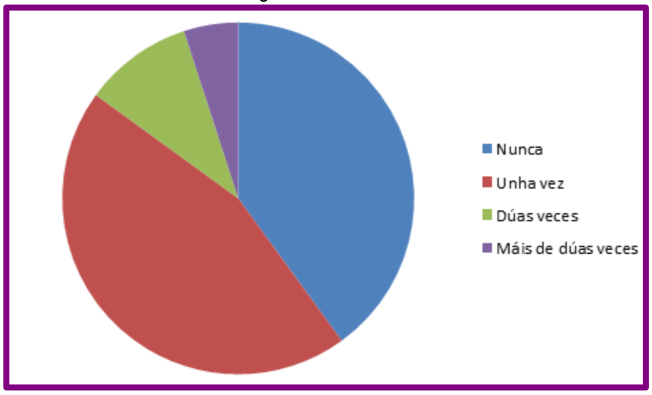
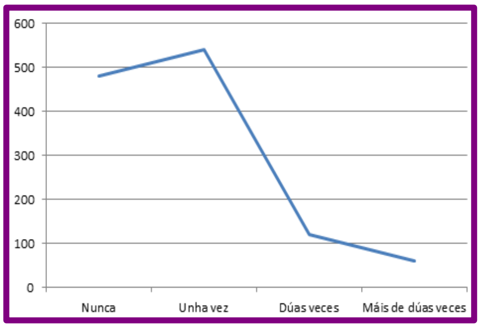


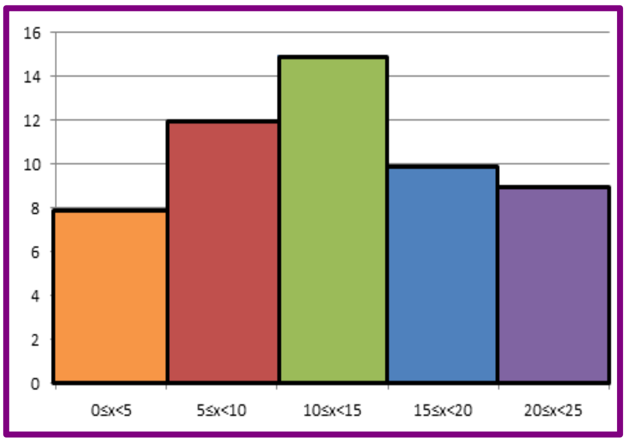
Diagrama de sectores



Polígono de frecuencias



S17. Como todos os intervalos teñen a mesma lonxitude, é igual que facer un diagrama de barras.



S18.

- Para calcular a media necesitamos tres columnas da táboa: x_i, f_i e $x_i \cdot f_i$.
- Para calcular a mediana necesitamos tres columnas: x_i, f_i e F_i (en realidade f_i só é necesaria para calcular F_i, non para calcular a mediana).
- Para calcular a moda necesitamos dúas columnas: x_i e f_i.

Polo tanto, para calcular estes tres parámetros de centralización necesitamos catro columnas da táboa que son: x_i, f_i, F_i e $x_i \cdot f_i$.

x_i	f_i	F_i	$x_i \cdot f_i$
0	0	0	0
1	5	5	5
2	4	9	8
3	2	11	6
4	2	13	8
5	1	14	5
6	1	15	6
7	2	17	14
8	3	20	24
9	4	24	36
10	8	32	80
TOTAL	32		192

Así, temos:

- **Media:**
$$\bar{x} = \frac{192}{32} = 6$$
- **Mediana**: os lugares do medio (16º e 17º) están ocupados polo valor $x_i = 7$.
$$Me = \frac{7+7}{2} = 7$$
- **Moda**: hai unha única moda que é:
$$Mo = 10$$

S19.

- Para calcular a media necesitamos tres columnas da táboa: xi, fi e xi · fi.
- Para calcular a mediana necesitamos tres columnas: xi, fi e Fi (en realidade fi so é necesaria para calcular Fi, non para calcular a mediana).
- Para calcular a moda necesitamos dúas columnas: xi e fi.

Polo tanto, para calcular estes tres parámetros de centralización necesitamos catro columnas da táboa que son: xi, fi, Fi e xi · fi.

x_i	f_i	F_i	$x_i \cdot f_i$
151	2	2	302
156	5	7	780
161	11	18	1771
166	14	32	2324
171	5	37	855
176	3	40	528
TOTAL	40		6560

Así, temos:

- **Media:**
$$\bar{x} = \frac{6560}{40} = 164$$
- **Mediana**: os lugares do medio (20º e 21º) están ocupados polo valor $x_i = 166$.
$$Me = \frac{166+166}{2} = 166$$
- **Moda**: hai unha única moda que é:
$$Mo = 166$$

S20.

Para calcular os parámetros de dispersión necesitamos:

- Para o rango, determinar os valores máximo e mínimo da variable.
- Para a desviación media, x_i, f_i e $x_i \cdot f_i$ para calcular a media e $|x_i - \bar{x}|$, f_i e $|x_i - \bar{x}| \cdot f_i$ para DM.
- Para a desviación típica, x_i, f_i e $x_i \cdot f_i$ para a media e $(x_i - \bar{x})$, $(x_i - \bar{x})^2$, f_i e $(x_i - \bar{x})^2 \cdot f_i$. para a varianza.

Polo tanto, para calcular os parámetros de dispersión necesitamos oito columnas da táboa que son:

x_i, f_i, $x_i \cdot f_i$, $|x_i - \bar{x}|$, $|x_i - \bar{x}| \cdot f_i$, $(x_i - \bar{x})$, $(x_i - \bar{x})^2$, e $(x_i - \bar{x})^2 \cdot f_i$.

As columnas $|x_i - \bar{x}|$ e $(x_i - \bar{x})$ son a mesma agás o signo, polo que só calcularemos a columna de $|x_i - \bar{x}|$.

En definitiva son sete columnas. Sabemos, pola actividade 18, que $\bar{x} = 6$.

x_i	f_i	F_i	$x_i \cdot f_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
0	0	0	0	6	0	36	0
1	5	5	5	5	25	25	125
2	4	9	8	4	16	16	64
3	2	11	6	3	6	9	18
4	2	13	8	2	4	4	8
5	1	14	5	1	1	1	1
6	1	15	6	0	0	0	0
7	2	17	14	1	2	1	2
8	3	20	24	2	6	4	12
9	4	24	36	3	12	9	36
10	8	32	80	4	32	16	128
TOTAL	32		192		104		394

- Rango: $\text{rango} = 10 - 0 = 10$
- Desviación media: $DM = \frac{104}{32} = 3{,}25$
- Varianza: $\sigma^2 = \frac{394}{32} = 12{,}31$
- Desviación típica: $\sigma = \sqrt{12{,}31} = 3{,}51$

S21.

Necesitamos as mesmas columnas da táboa que na actividade anterior e sabemos, pola actividade 19, que $\bar{x} = 164$.

x_i	f_i	F_i	$x_i \cdot f_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
151	2	2	302	13	26	169	338
156	5	7	780	8	40	64	320
161	11	18	1771	3	33	9	99
166	14	32	2324	2	28	4	56
171	5	37	855	7	35	49	245
176	3	40	528	12	36	144	432
TOTAL	40		6560		198		1490

▪ Rango: $\text{rango} = 176 - 151 = 25$	▪ Desviación media: $DM = \frac{198}{40} = 4{,}95$	▪ Varianza: $\sigma^2 = \frac{1490}{40} = 37{,}25$	▪ Desviación típica: $\sigma = \sqrt{37{,}25} = 6{,}10$

S22.

Para calcular os cuartís necesitamos a frecuencia absoluta acumulada.

x_i	f_i	F_i
0	0	0
1	5	5
2	4	9
3	2	11
4	2	13
5	1	14
6	1	15
7	2	17
8	3	20
9	4	24
10	8	32
TOTAL	32	

Así, temos:

- Q_1 deixa tras el a cuarta parte da distribución. Como N = 32, temos que $\frac{N}{4} = 8$. Así, Q_1 debe deixar 8 elementos á súa esquerda, polo que debe estar entre o 8º e o 9º elemento. Polo tanto, $Q_1 = \frac{2+2}{2} = 2$.
- Q_2 é a mediana e, como N é par, será a media dos elementos que ocupan as posicións centrais, que son as posicións 16ª e 17ª. Polo tanto, $Q_2 = \frac{7+7}{2} = 7$.
- Q_3 deixa tras el as tres cuartas partes da distribución. Como N = 32, temos que $\frac{3N}{4} = 24$. Así, Q_3 debe deixar 24 elementos á súa esquerda, polo que debe estar entre o 24º e o 25º elemento. Polo tanto, $Q_3 = \frac{9+10}{2} = 9{,}5$.

S23.

Para calcular os cuartís necesitamos a frecuencia absoluta acumulada.

x_i	f_i	F_i
151	2	2
156	5	7
161	11	18
166	14	32
171	5	37
176	3	40
TOTAL	40	

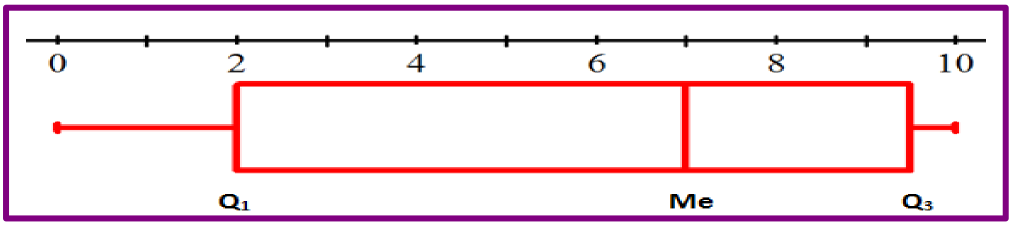
Así, temos:

- Q1 deixa tras el a cuarta parte da distribución. Como N = 40, temos que $\frac{N}{4} = 10$. Así, Q1 debe deixar 10 elementos á súa esquerda, polo que debe estar entre o 10º e o 11º elemento. Polo tanto, Q1 = $\frac{161+161}{2} = 161$.
- Q2 é a mediana e, como N é par, será a media dos elementos que ocupan as posicións centrais, que son as posicións 20ª e 21ª. Polo tanto, Q2 = $\frac{166+166}{2} = 166$.
- Q3 deixa tras el as tres cuartas partes da distribución. Como N = 40, temos que $\frac{3N}{4} = 30$. Así, Q3 debe deixar 30 elementos á súa esquerda, polo que debe estar entre o 30º e o 31º elemento. Polo tanto, Q3 = $\frac{166+166}{2} = 166$.

S24.

Pola actividade 22 sabemos que $Q_1 = 2$, $Q_2 = 7$ e $Q_3 = 9{,}5$. Tamén sabemos que o valor mínimo da variable é 0 e o maior é 10. Así:

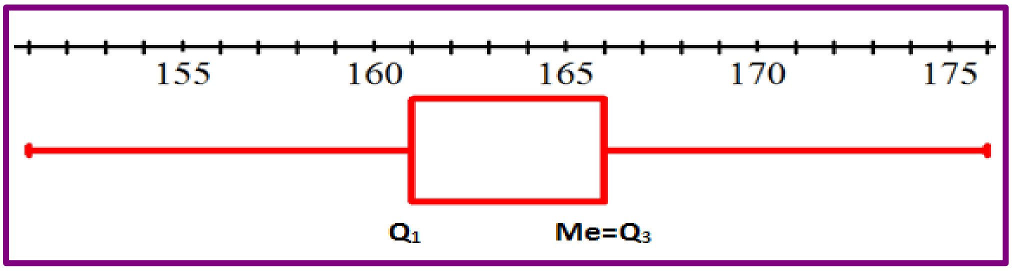
O diagrama de caixa e bigotes é o seguinte:



S25.

Pola actividade 23 sabemos que $Q_1 = 161$, $Q_2 = 166$ e $Q_3 = 166$. Tamén sabemos que o valor mínimo da variable é 151 e o maior é 176. Así:

O diagrama de caixa e bigotes é o seguinte:



S26. Pola actividade 18 sabemos que $\bar{x} = 6$ e pola actividade 20 sabemos que $\sigma = 3{,}51$. Polo tanto, o coeficiente de variación é $CV = \frac{3{,}51}{6} = 0{,}59$.

S27. Pola actividade 19 sabemos que $\bar{x} = 164$ e pola actividade 21 sabemos que $\sigma = 6{,}10$. Polo tanto, o coeficiente de variación é $CV = \frac{6{,}10}{164} = 0{,}04$.

4.2 Solucións das actividades finais

S28.

Lonxitude dunha uña. *Cuantitativa continua.*	Tempo que dura unha película. *Cuantitativa continua.*
Puntos obtidos en total nun partido de baloncesto. *Cuantitativa discreta.*	Idade dunha árbore. *Cuantitativa continua*, aínda que o normal é tratala como discreta.
Cor dunha flor. *Cualitativa.*	Olor dun perfume. *Cualitativa.*
Sabor dunha froita. *Cualitativa*	Número de trens diarios entre A Coruña e Vigo. *Cuantitativa discreta.*
Temperatura dun forno. *Cuantitativa continua,* aínda que o normal é tratala como discreta.	Peso dun paquete de azucre. *Cuantitativa continua.*

S29.

x_i	f_i	F_i	$x_i \cdot f_i$	$\lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert \cdot f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
1	*1*	*1*	*1*	*4,2*	*4,2*	*17,64*	*17,64*
2	*2*	*3*	*4*	*3,2*	*6,4*	*10,24*	*20,48*
3	*2*	*5*	*6*	*2,2*	*4,4*	*4,84*	*9,68*
4	*3*	*8*	*12*	*1,2*	*3,6*	*1,44*	*4,32*
5	*4*	*12*	*20*	*0,2*	*0,8*	*0,04*	*0,16*
6	*2*	*14*	*12*	*0,8*	*1,6*	*0,64*	*1,28*
7	*1*	*15*	*7*	*1,8*	*1,8*	*3,24*	*3,24*
8	*3*	*18*	*24*	*2,8*	*8,4*	*7,84*	*23,52*
9	*2*	*20*	*18*	*3,8*	*7,6*	*14,44*	*28,88*
TOTAL	*20*		*104*		*38,8*		*109,2*

- Media: $\bar{x} = \frac{104}{20} = 5,2$
- Mediana: $Me = \frac{5+5}{2} = 5$
- Moda: $Mo = 5$
- Desviación media: $DM = \frac{38,8}{20} = 1,94$
- Varianza: $\sigma^2 = \frac{109,2}{20} = 5,46$
- Desviación típica: $\sigma = \sqrt{5,46} = 2,34$
- Rango: $rango = 9 - 1 = 8$
- Coeficiente de variación: $CV = \frac{2,34}{5,2} = 0,45$
- 1º cuartil: $Q_1 = 3,5$
- 3º cuartil: $Q_3 = 7,5$
- Percorrido intercualtílico: $7,5 - 3,5 = 4$

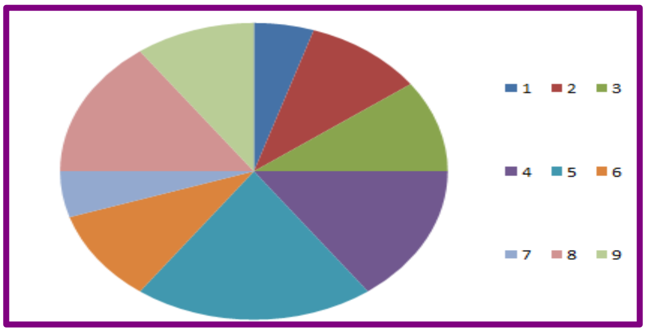


Diagrama de sectores das frecuencias absolutas

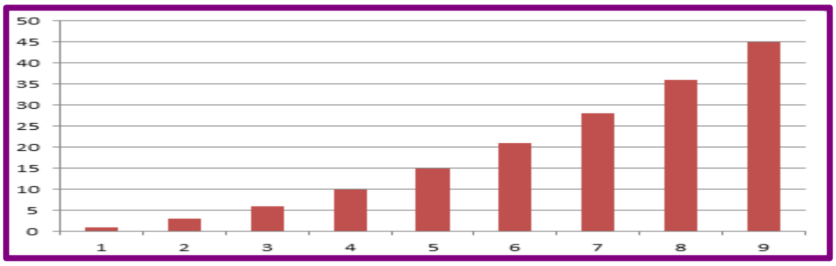
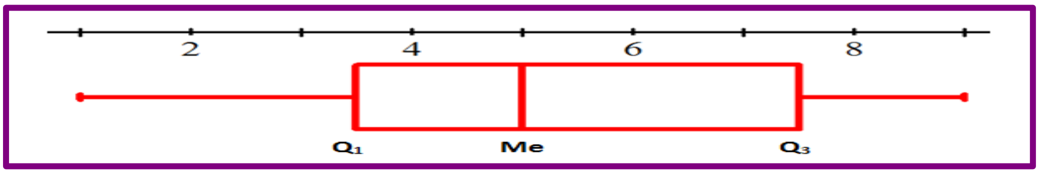


Diagrama de barras das frecuencias absolutas acumuladas

Diagrama de caixa e bigotes



S30.

x_i	f_i	F_i	$x_i \cdot f_i$	x_i^2	$x_i^2 \cdot f_i$
3	20	20	60	9	180
4	48	68	192	16	768
5	77	145	385	25	1925
6	45	190	270	36	1620
7	24	214	168	49	1176
8	17	231	136	64	1088
9	19	250	171	81	1539
TOTAL	250		1382		8296

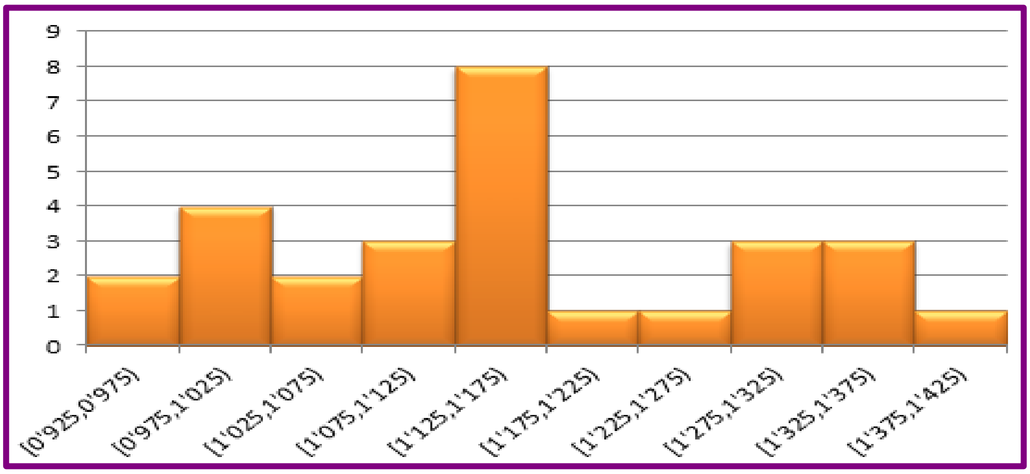
- Media: $\bar{x} = \frac{1382}{250} = 5{,}53$
- Mediana: $Me = \frac{5+5}{2} = 5$
- Moda: $Mo = 6$
- Varianza: $\sigma^2 = \frac{8296}{250} - 5{,}53^2 = 2{,}63$
- Desviación típica: $\sigma = \sqrt{5{,}46} = 1{,}62$
- Rango: $rango = 9 - 3 = 6$

S31.

x_i	f_i	F_i	$x_i \cdot f_i$	$\boldsymbol{x_i^2}$	$\boldsymbol{x_i^2 \cdot f_i}$
3	5	5	15	9	45
4	12	17	48	16	192
5	116	133	580	25	2900
6	53	186	318	36	1908
7	29	215	203	49	1421
8	17	232	136	64	1088
9	18	250	162	81	1458
TOTAL	250		1462		9012

- Media: $\bar{x} = \frac{1462}{250} = 5{,}85$
- Mediana: $Me = \frac{5+5}{2} = 5$
- Moda: $Mo = 5$
- Varianza: $\sigma^2 = \frac{9012}{250} - 5{,}85^2 = 1{,}85$
- Desviación típica: $\sigma = \sqrt{5{,}46} = 1{,}36$
- Rango: $rango = 9 - 3 = 6$

S32.



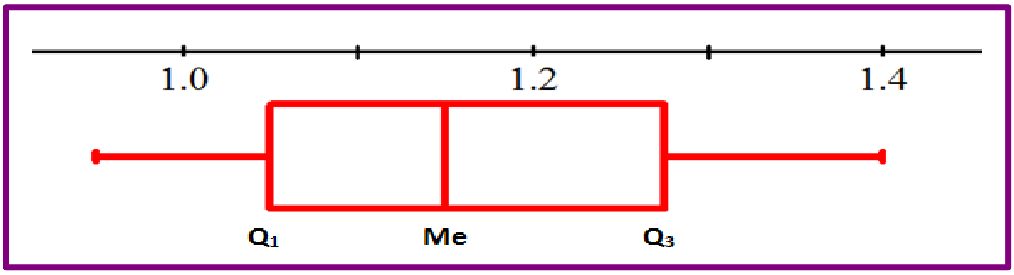
S33.

Clase	x_i	f_i	F_i	$x_i \cdot f_i$	x_i^2	$x_i^2 \cdot f_i$
[0'925,0'975)	0,95	2	2	1,9	0,9025	1,805
[0'975,1'025)	1	4	6	4	1	4
[1'025,1'075)	1,05	2	8	2,1	1,1025	2,205
[1'075,1'125)	1,1	3	11	3,3	1,21	3,63
[1'125,1'175)	1,15	8	19	9,2	1,3225	10,58
[1'175,1'225)	1,2	1	20	1,2	1,44	1,44
[1'225,1'275)	1,25	1	21	1,25	1,5625	1,5625
[1'275,1'325)	1,3	3	24	3,9	1,69	5,07
[1'325,1'375)	1,35	3	27	4,05	1,8225	5,4675
[1'375,1'425)	1,4	1	28	1,4	1,96	1,96
	TOTAL	28		32,3		35,915

- Media:
  $\bar{x} = \frac{32,3}{28} = 1,15$
- Mediana:
  $Me = \frac{1,15 + 1,15}{2} = 1,15$
- Rango:
  $rango = 1,4 - 0,95 = 0,45$

- Varianza:
  $\sigma^2 = \frac{35,915}{28} - 1,15^2 = 0,0164$
- Desviación típica:
  $\sigma = \sqrt{5,46} = 0,13$

- Coeficiente de variación:
  $CV = \frac{0,13}{1,15} = 0,11$
- 1º cuartil:
  $Q_1 = \frac{1,05 + 1,05}{2} = 1,05$
- 3º cuartil:
  $\bar{x} = \frac{1,25 + 1,3}{2} = 1,275$



S34.

x_i	f_i	$x_i \cdot f_i$	x_i^2	$x_i^2 \cdot f_i$
4,6	1	4,6	21,16	21,16
7,8	1	7,8	60,84	60,84
5,8	1	5,8	33,64	33,64
6,2	1	6,2	38,44	38,44
10,8	1	10,8	116,64	116,64
12,2	1	12,2	148,84	148,84
13,6	1	13,6	184,96	184,96
14	1	14	196	196,00
11,8	1	11,8	139,24	139,24
10,4	1	10,4	108,16	108,16
5	1	5	25	25,00
1,8	1	1,8	3,24	3,24
TOTAL	12	104		1076,16

TEMPERATURAS MÍNIMAS

- Media:
$$\bar{x} = \frac{104}{12} = 8,67$$
- Varianza:
$$\sigma^2 = \frac{1076,16}{12} - 8,67^2 = = 14,57$$
- Desviación típica:
$$\sigma = \sqrt{14,57} = 3,82$$
- Coeficiente de variación:
$$CV = \frac{3,82}{8,67} = 0,44$$

x_i	f_i	$x_i \cdot f_i$	x_i^2	$x_i^2 \cdot f_i$
15,8	1	15,8	249,64	249,64
23	1	23	529	529,00
25,6	1	25,6	655,36	655,36
24	1	24	576	576,00
24,8	1	24,8	615,04	615,04
23	1	23	529	529,00
31,6	1	31,6	998,56	998,56
28	1	28	784	784,00
28,2	1	28,2	795,24	795,24
26,2	1	26,2	686,44	686,44
20,4	1	20,4	416,16	416,16
16,2	1	16,2	262,44	262,44
TOTAL	12	286,8		7096,88

TEMPERATURAS MÁXIMAS

- Media:
$$\bar{x} = \frac{286,8}{12} = 23,9$$
- Varianza:
$$\sigma^2 = \frac{7096,88}{12} - 23,9^2 = = 20,20$$
- Desviación típica:
$$\sigma = \sqrt{20,20} = 4,49$$
- Coeficiente de variación:
$$CV = \frac{4,49}{23,9} = 0,19$$

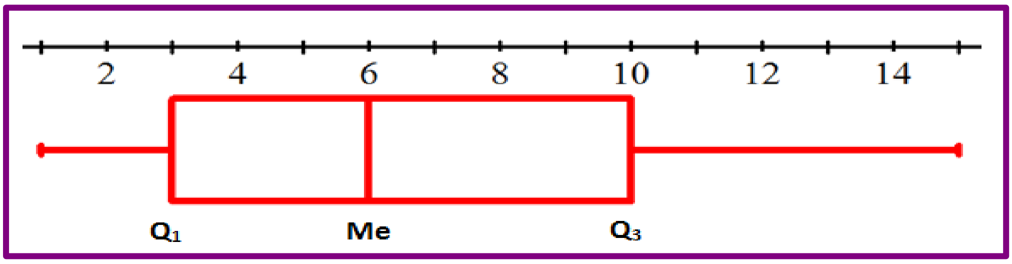
Ten menos dispersión relativa a das temperaturas mínimas por ter o coeficiente de variación máis pequeno.

S35.

x_i	f_i	F_i	$x_i \cdot f_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
1	4	4	4	5,81	23,24	33,7561	135,02
2	5	9	10	4,81	24,05	23,1361	115,68
3	6	15	18	3,81	22,86	14,5161	87,10
4	3	18	12	2,81	8,43	7,8961	23,69
5	5	23	25	1,81	9,05	3,2761	16,38
6	4	27	24	0,81	3,24	0,6561	2,62
7	6	33	42	0,19	1,14	0,0361	0,22
8	1	34	8	1,19	1,19	1,4161	1,42
9	3	37	27	2,19	6,57	4,7961	14,39
10	4	41	40	3,19	12,76	10,1761	40,70
11	1	42	11	4,19	4,19	17,5561	17,56
12	1	43	12	5,19	5,19	26,9361	26,94
13	7	50	91	6,19	43,33	38,3161	268,21
14	0	50	0	7,19	0	51,6961	0
15	2	52	30	8,19	16,38	67,0761	134,15
TOTAL	52		354		181,62		884,08

- Media:
  $\bar{x} = \frac{354}{52} = 6,81$
- Mediana:
  $Me = \frac{6+6}{2} = 6$
- Moda:
  $Mo = 13$
- Desviación media:
  $DM = \frac{181,62}{52} = 3,49$
- Varianza:
  $\sigma^2 = \frac{884,08}{52} = 17,00$
- Desviación típica:
  $\sigma = \sqrt{17,00} = 4,12$
- 1º cuartil:
  $Q_1 = 3$
- 3º cuartil:
  $\mathbf{Q_3 = 10}$
- Rango:
  $rango = 15 - 1 = 14$

Diagrama de caixa e bigotes



S36.

x_i	f_i	F_i	$x_i \cdot f_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot f_i$
0	5	5	0	6,81	34,05
1	11	16	11	5,81	63,91
2	6	22	12	4,81	28,86
3	4	26	12	3,81	15,24
4	3	29	12	2,81	8,43
5	0	29	0	1,81	0
6	1	30	6	0,81	0,81
TOTAL	30		53		151,3

- Media:
  $\bar{x} = \frac{53}{30} = 1,77$
- Mediana:
  $\mathrm{Me} = \frac{1+1}{2} = 1$
- Moda:
  $\mathrm{Mo} = 1$
- Desviación media:
  $\mathrm{DM} = \frac{151,3}{30} = 5,04$
- Rango:
  $\mathrm{rango} = 6 - 0 = 6$

S37.

x_i	f_i	$x_i \cdot f_i$	x_i^2	$x_i^2 \cdot f_i$
0	3	0	0	0
1	4	4	1	4
2	2	4	4	8
3	1	3	9	9
TOTAL	10	11		21

- Media:
  $\bar{x} = \frac{11}{10} = 1{,}1$
- Varianza:
  $\sigma^2 = \frac{21}{10} - 1{,}1^2 = 0{,}89$
- Desviación típica:
  $\sigma = \sqrt{0{,}89} = 0{,}94$
- Coeficiente de variación:
  $CV = \frac{0{,}94}{1{,}1} = 85{,}45$

S38. $\bar{x} = 7$. A desviación típica de 3,8 corresponde á gráfica C, a de 2,9 é a da gráfica A e a de 1,3 é a da gráfica B.

5. Glosario

C	▪ Coeficiente de variación	O cociente entre a desviación típica e a media.
	▪ Cuartís	Números que dividen os datos da distribución xa ordenados en catro partes iguais.
D	▪ Desviación media	É a media aritmética das distancias dos datos á media.
	▪ Desviación típica	É a raíz cadrada da varianza.
	▪ Diagrama de caixa e bigotes	Forma moi clara de representar os cuartís e o percorrido da variable.
E	▪ Estatística descritiva ou dedutiva	Parte da estatística que trata do reconto, a ordenación e a clasificación dos datos obtidos nas observacións.
	▪ Estatística inferencial ou indutiva	Parte da estatística que formula e resolve o problema de establecer previsións e deducións xerais sobre unha poboación a partir resultados obtidos dunha mostra.
F	▪ Frecuencia absoluta	Número de veces que aparece un valor.
	▪ Frecuencia absoluta acumulada dun valor x_i	Número de veces que aparece o valor x_i ou calquera dos anteriores a el.
	▪ Frecuencia relativa	Frecuencia absoluta dun valor determinado dividida polo número total de datos.
	▪ Frecuencia relativa acumulada	Frecuencia absoluta acumulada dividida entre o número total de datos.
I	▪ Individuo	Cada un dos membros dunha poboación.
	▪ Intervalos de clase	Cada un dos intervalos nos que se agrupan os datos, normalmente dunha variable continua.
M	▪ Marca de clase	Punto medio do intervalo de clase.
	▪ Media	É o valor central da distribución.
	▪ Mediana	É o valor que se atopa no medio, é dicir, é o valor que deixa tantos individuos antes, como despois.
	▪ Moda	É o valor que ten a maior frecuencia absoluta.
	▪ Mostra	Parte dunha poboación.
P	▪ Parámetros de centralización	Son os que indican arredor de que valor se distribúen os datos.
	▪ Parámetros de dispersión	Son os que informan sobre canto se afastan do centro os valores da distribución.
	▪ Parámetros de posición	Son os que dividen un conxunto de datos en grupos co mesmo número de individuos.
	▪ Percorrido ou rango	É a diferenza entre os datos maior e menor da variable.
	▪ Percorrido intercuartílico	Distancia entre os cuartís 3º e 1º.
	▪ Poboación	Conxunto de persoas, obxectos ou individuos sobre os que se quere estudar unha determinada característica.

R	▪ Rango ou percorrido	É a diferenza entre os datos maior e menor da variable.
V	▪ Variable estatística	Cada unha das propiedades ás que se pretenda aplicar un estudo estatístico.
	▪ Variables estatísticas cualitativas	Variables estatísticas que non se poden expresar como valores numéricos.
	▪ Variables estatísticas cuantitativas	Variables estatísticas que poden tomar valores numéricos.
	▪ Varianza	É a media aritmética dos cadrados das distancias dos datos á media.

6. Bibliografía e recursos

Bibliografía

- *Libros para a educación secundaria a distancia de adultos. Ámbito tecnolóxico-matemático.* Consellería de Educación e Ordenación Universitaria.
- *Matemáticas ESO 1.* Ed. Anaya, 2016.
- *Matemáticas ESO 2.* Ed. Anaya, 2016.
- *Matemáticas orientadas a las enseñanzas académicas ESO 3.* Ed. Anaya, 2016.
- *Matemáticas orientadas a las enseñanzas aplicadas ESO 3.* Ed. Anaya, 2016.
- *Matemáticas enseñanzas académicas ESO 3.* Ed. Santillana.
- *Matemáticas enseñanzas aplicadas ESO 3.* Ed. Santillana.

Ligazóns de Internet

Nestas ligazóns pode atopar trucos e información que pode consultar para mellorar a súa práctica.

- https://matematicasiesoja.wordpress.com/estadistica/
- http://www.vitutor.com/estadistica/descriptiva/estadistica.html
- http://recursostic.educacion.es/descartes/web/materiales_didacticos/estadistica_1_ciclo/indice.htm
- https://campusvirtual.ull.es/ocw/pluginfile.php/6104/mod_resource/content/1/tema6/PR6-estdescriptiva.pdf
- http://www3.uji.es/~mateu/t1-alumnos.pdf
- https://matematicasiesoja.files.wordpress.com/2013/10/ejerciciosresueltos1.pdf
- http://www.est.uc3m.es/esp/nueva_docencia/getafe/ciencias_empresariales/metodestad_empresa/doc_generica/archivos/Ejercicios%20resueltos%20Tema%201.pdf
- http://www.est.uc3m.es/esp/nueva_docencia/getafe/ciencias_empresariales/metodestad_empresa/doc_generica/archivos/Ejercicios%20resueltos%20Tema%202.pdf