Matemàtiques orientades a les ensenyances acadèmiques:

3r B d'ESO. Capítol 11: Estadística i probabilitat

www.apuntesmareaverde.org.es

Autor: Fernando Blasco

Revisor: David Hierro

Il·lustracions: Banc d'Imatges d'INTEF

Traducció: Pedro Podadera. Instituto Juan de Garay de Valencia

Estadística i probabilitat. 3r B ESO

Índex

Resum

L'Estadística és una Ciència que va sorgir per a portar la comptabilitat de l'Estat. D'ací ve el seu nom. Al segle XX es van desenrotllar les seues tècniques i es va separar de les Matemàtiques, passant a ser una ciència amb entitat pròpia. Als Mitjans de comunicació trobem freqüents estadístiques. En medicina es necessiten mètodes estadístics per a provar nous medicaments. En tot experiment científic, després de l'arreplegada de dades, es necessita utilitzar proves estadístiques que permeten traure informació d'aqueixes dades.

L'origen de la Probabilitat es troba en els jocs d'atzar. Cardano, Galileu, Pascal, Fermat són alguns dels matemàtics que es van ocupar als seus inicis.

1. LA PRESA DE DADES

1.1. Un exemple per a realitzar una anàlisi

Exemple:

- La Casa de la Moneda vol estudiar quantes monedes ha d'emetre, tenint en compte les que estan en circulació i les que es queden atresorades (bé en cases particulars, o en màquines de refrescos, o dipositades en un banc). S'ha fet una enquesta a peu de carrer a 60 persones i s'ha apuntat quantes monedes portava cada una d'elles a la butxaca. Hem obtingut aquestes dades:

12	7	11	8	8	9	6	12	7	7	13	0	10	9	13	18	7	6	11	12	16	0	10	10	8	8	9	11	10	8
16	8	5	2	12	8	14	14	16	6	2	0	18	10	10	12	14	6	7	3	12	11	10	18	9	7	12	1	15	8

El primer pas consisteix a fer un esquema per al recompte: usarem una taula i marcarem baquetes cada vegada que aparega aqueix nombre.

0	///	**7**	///// /	**14**	///
1	/	**8**	///// ///	**15**	/
2	//	**9**	////	**16**	///
3	/	**10**	///// //	**17**	
4		**11**	////	**18**	///
5	/	**12**	///// //	**19**	
6	////	**13**	//	**20**	

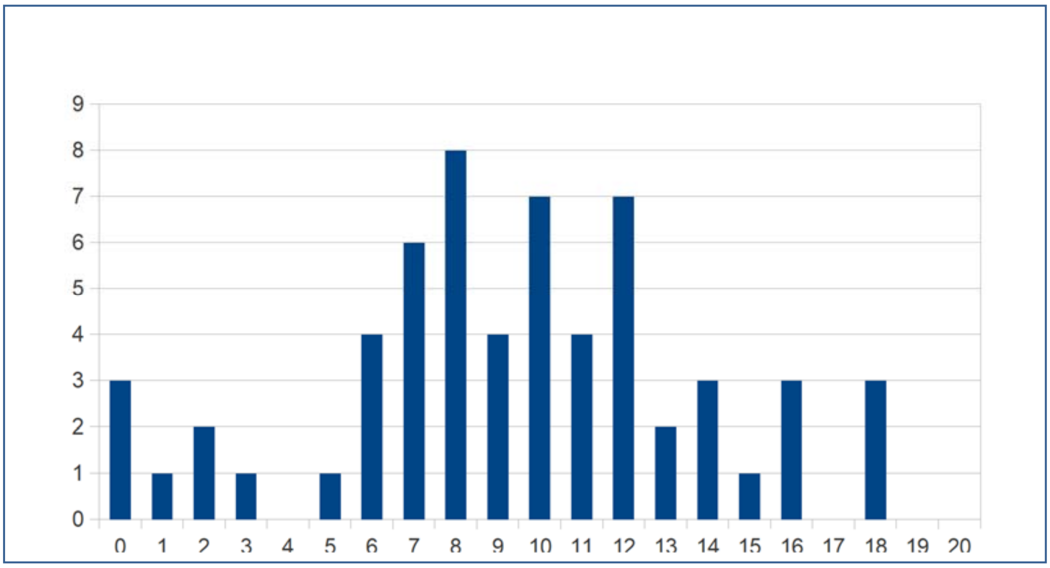
Passar d'aqueix recompte a una taula de **freqüències absolutes** és molt senzill: només cal substituir les baquetes pel nombre que representen.

0	3
1	1
2	2
3	1
4	0
5	1
6	4

7	6
8	8
9	4
10	7
11	4
12	7
13	2

14	3
15	1
16	3
17	0
18	3
19	0
20	0

És molt millor analitzar les dades de manera visual. Estem més acostumats a treballar d'aqueixa manera. Podem representar les dades de la taula de freqüències en un **diagrama de barres**, on l'altura de cada barra representa la freqüència d'aparició.



El processament de dades estadístiques s'utilitza molt. Òbviament no es fan les operacions a mà, sinó que s'utilitzen calculadors o fulls de càlcul. Disposar d'aqueixos mitjans tecnològics serà un bon complement per al capítol, encara que recordem que el més important és comprendre què es fa en cada moment.

Començarem introduint un poquet de **nomenclatura**. Quasi tots aquests noms els has escoltat ja que als Mitjans de comunicació els utilitzen moltíssim

Població és el col·lectiu sobre el qual es vol fer l'estudi.

Mostra és un subconjunt de la població de manera que a partir del seu estudi es poden obtindre característiques de la població completa.

Individu és cada un dels elements de la població o la mostra.

Exemple:

- Es vol fer un estudi sobre hàbits alimentaris dels estudiants de 3º d'ESO de tot Madrid. Però com és molt costós entrevistar tots els estudiants es decideix prendre un IES per cada districte i entrevistar als alumnes de 3º d'ESO d'aqueixos col·legis triats.

  La **població** objecte de l'estudi seran tots els estudiants madrilenys matriculats en 3º d'ESO.

  La **mostra** són els estudiants de 3º d'ESO matriculats als instituts triats.

  Cada un dels estudiants de 3º d'ESO és un **individu** per a aquest estudi estadístic.

Activitats proposades

a) Volem fer un estudi de la quantitat de monedes que porten a la butxaca els estudiants de la teua classe. Però per a no preguntar a tots tria 10 companys a l'atzar i anota en el teu quadern quantes monedes porta cada u.

a) Quina és la població objecte de l'estudi?

b) Quina és la mostra triada?

c) Especifica 5 individus que pertanguen a la població i no a la mostra.

1.2. Variables estadístiques

Exemple:

- En un estudi estadístic es pot preguntar coses tan molt variades com

b) Quines fruites menges al llarg d'una setmana?
1 .-Quantes peces de fruita menges al dia?
2 .-Quantes monedes portes a la butxaca?
3 .-Quina és la teua altura?
4 .-Quantes marques de xocolate recordes?
5 .-Quines són les marques de xocolate que recordes?
6 .-Quants germans tens?
7 .-Quin és el teu color favorit per a un cotxe?
8 .-Quant temps passes al dia veient la televisió?
9 .-Quants seguidors tens en twitter?

Aqueixes preguntes poden correspondre a estudis de salut, econòmics, publicitaris o socioeconòmics. Algunes es responen amb un nombre i altres es responen amb un nom o un adjectiu. Inclús hi ha diferències entre les que es responen amb nombres: el nombre de monedes que portes o el nombre de seguidors de *twitter* es contesten amb nombres enters, mentres que per a trobar la teua altura o les hores que panses davant del televisor necessitem utilitzar nombres reals (normalment amb representació decimal).

Una variable es diu **quantitativa** si els seus valors s'expressen amb nombres.

Les variables quantitatives poden ser

c) **discretes** si només admeten valors aïllats

contínues si entre dos valors poden donar-se també tots els intermedis

Una variable estadística és **qualitativa** quan els seus valors no s'expressen mitjançant un nombre, sinó amb una qualitat.

Activitats proposades

1. Classifica en variables qualitatives i quantitatives les que apareixen al primer exemple d'aquesta secció. Per a les quantitatives indica si són contínues o discretes.

1.3. Les fases d'un estudi estadístic.

En un estudi estadístic hi ha 6 fases fonamentals:

a) Determinació de l'objecte de l'estudi. Açò és, saber què volem estudiar.
b) Selecció de les variables que es van a estudiar.
c) Arreplega de les dades.
d) Organització de les dades.
e) Representació i tractament de les dades.
f) Interpretació i anàlisi.

En aquest llibre començarem els exemples a partir del punt 4, amb dades ja proporcionats en els enunciats, encara que a continuació reflexionarem un poc sobre la selecció d'una mostra,

1.4. Mètodes de selecció d'una mostra estadística. Representativitat d'una mostra

Per a arreplegar les dades i determinar els valors de la variable es pot utilitzar tota la població, tot l'univers sobre el qual es realitza l'estudi, o seleccionar una mostra. Moltes vegades no és convenient arreplegar valors de tota la població, perquè és complicat o massa costós, o inclús perquè és impossible com en el cas d'un control de qualitat en què es destruïsca l'objecte a analitzar. La part de l'Estadística que s'ocupa de com seleccionar adequadament les mostres es denomina *Teoria de Mostres*.

Exemples:

- Si estudiem el pes dels habitants d'una ciutat, la població serà el total de les persones de la dita ciutat.

I. Però el més normal serà no arreplegar informació sobre totes les persones de la ciutat (ja que seria una labor molt complexa i costosa), sinó que se sol seleccionar un subgrup (mostra) que s'entenga que és prou representatiu.

II. Per a conéixer la intenció de vot davant d'unes eleccions europees, municipals, autonòmiques… s'utilitzen mostres, perquè preguntar a tota la població seria molt costós (i això ja es fa en les eleccions).

III. Però si una fàbrica vol conéixer les hores de vida útil d'un tipus de pereta, no pot posar a funcionar a tota la població, totes les peretes, fins que s'espatllen perquè es queda sense producció. En aquest cas és imprescindible seleccionar una mostra.

IV. En *control de qualitat* es fan estudis estadístics i es prenen mostres.

Per a determinar la millor forma de seleccionar una mostra hi ha tota una part de l'Estadística, la Teoria *de Mostres*, que ens indica diversos detalls a tindre en compte:

- Com s'han de triar els elements de la mostra?
- Quin ha de ser la grandària de la mostra?
- Fins a quin punt la mostra és representativa de la població?

La forma de seleccionar la mostra, ***mostreig***, ha de reunir unes determinades característiques perquè puga caracteritzar la població, ser ***representativa*** de la població. Ha de ser un mostreig **aleatori**, és a dir, a l'atzar. Si la mostra està mal triada, no és representativa, es produeixen **biaixos**, errors als resultats de l'estudi.

Tots els individus de la població han de tindre les mateixes possibilitats de ser seleccionats per a la mostra.

Exemples:

- Es vol estudiar el nivell adquisitiu dels persones d'una ciutat, per al que passem una enquesta a la porta d'uns grans magatzems, et pareix un mostreig aleatori?

No ho és. Les persones que entren en un determinat establiment no representen a tota la població.

- Faràs un estudi sobre els gustos musicals dels jóvens, i per a això, preguntes a cinc d'entre les teues amistats, et pareix un mostreig aleatori?

No ho és. Les teues amistats poden tindre uns gustos diferents dels de la resta de la població.

Mètodes de selecció d'una mostra

Hi ha diversos mètodes per a seleccionar una mostra, que donarien per a analitzar en un llibre sobre "Mostreig". Però és convenient conéixer algun. Vegem tres d'ells:

Mostreig aleatori simple

Tots els individus de la població tenen la mateixa probabilitat de ser triats en la mostra.

Mostreig aleatori sistemàtic

S'ordenen els individus de la població. Es tria a l'atzar un individu, i se selecciona la mostra prenent individus mitjançant bots igualment espaiats.

Mostreig aleatori estratificat

Es divideix la població en grups homogenis d'una determinada característica, *estrats*, per exemple edat, i es pren una mostra aleatòria simple en cada estrat.

Exemple:

- S'estudia l'estat dels ossos de la població d'un país, i es divideix la població en "xiquets", "jóvens", "edat mitjana" i "tercera edat". En cada grup es fa un mostreig aleatori simple.

Representativitat d'una mostra

Quan es tria una mostra els dos Aspectes que cal tindre en compte són, la grandària i la representativitat de la mostra.

Si la mostra és massa xicoteta, encara que estiga ben triada, el resultat no serà fiable.

Exemple:

- Volem estudiar l'estatura de la població espanyola. Per a això triem a una persona a l'atzar i la mesurem.

Evidentment aquest resultat no és fiable. La mostra és massa xicoteta.

Si la mostra és massa gran els resultats seran molt fiables, però el gasto pot ser massa elevat. Inclús, de vegades, mostres massa grans no ens proporcionen millors resultats.

Quan una mostra tinga la grandària adequada, i haja sigut triada de forma aleatòria direm que és una **mostra representativa**.

Si la mostra no ha sigut triada de forma aleatòria direm que la mostra és **esbiaixada**.

Activitats proposades

2. Assenyalar en quin cas és més convenient estudiar la població o una mostra:

a) El diàmetre dels caragols que fabrica una màquina diàriament.

i) L'altura d'un grup de sis amics.

3. Es pot llegir el següent titular al periòdic que publica el teu institut: "*La nota mitjana dels alumnes de 3º ESO és de* 7'9". Com s'ha arribat a aquesta conclusió? S'ha estudiat a tota la població? Si hagueren seleccionat per al seu càlcul només a les alumnes, seria representatiu el seu valor?

4. En una sèrie de televisió tenen dubtes sobre què fer amb la protagonista, si que tinga un accident o si ha de casar-se. Faran una consulta. A tota la població o seleccionat una mostra representativa? Raona la resposta.

2. REPRESENTACIÓ DE LA INFORMACIÓ

2.1. Exemples per a treballar

La secció anterior la començàvem analitzant una variable discreta: el nombre de monedes que es porten a la butxaca. Pots repassar què fèiem allí: com recomptàvem les dades, com els portàvem després a una taula de freqüències i com representàvem la informació en un gràfic.

Farem ara el mateix procés amb una variable contínua.

Ja saps que:

Podem distingir entre **freqüències absolutes**, si, com en aquest exemple, fem un recompte del nombre de vegades que apareix cada dada. **Freqüències relatives**, que estudiarem amb més deteniment al final del capítol, i que consisteix a dividir cada freqüència absoluta pel nombre total d'observacions. **Freqüències acumulades**, tant freqüències absolutes acumulades com a freqüències relatives acumulades si es calculen tots els valors menors o iguals a ell.

Exemples:

- S'està realitzant un control del pes d'un grup de xiquets. Per a això, es comptabilitzen el nombre de vegades que mengen al dia una xocolatina 13 xiquets durant un mes, obtenint els nombres següents: 2, 5, 3, 2, 0, 4, 1, 7, 4, 2, 1, 0, 2.

La informació obtinguda es pot resumir en una taula de **freqüències absolutes** i **freqüències absolutes acumulades**:

Valors	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Freqüència absoluta	2	2	4	1	2	1	0	1
Freqüència absoluta acumulada	2	4	8	9	11	12	12	13

També es pot resumir en una taula de **freqüències relatives** i **freqüències relatives acumulades**:

Valors	0	1	2	3	4	5	6	7
Freqüència relativa	0'15	0'15	0'30	0'07	0'15	0'07	0	0'07
Freqüència relativa acumulada	0'15	0'30	0'61	0'69	0'84	0'92	0'92	1

- En una fàbrica es realitza un estudi sobre la grossària, en *mm*, d'un cert tipus de llandes de refresc. Amb aquest fi, selecciona una mostra de grandària $N = 25$, obtenint els valors següents: 7'8, 8'2, 7'6, 10'5, 7'4, 8'3, 9'2, 11'3, 7'1, 8'5, 10'2, 9'3, 9'9, 8'7, 8'6, 7'2, 9'9, 8'6, 10'9, 7'9, 11'1, 8'8, 9'2, 8'1, 10'5.

Aquesta informació es pot resumir fent cinc intervals i fent una taula de freqüències absolutes, freqüències absolutes acumulades, freqüències relatives i freqüències relatives acumulades

Intervals de classe	**(7,**	**(8,**	**(9,**	**(10,**	**(11,**
Marques de classe	7'5	8'5	9'5	10'5	11'5
Freqüència absoluta	6	8	5	4	2
Freqüència relativa	0'24	0'32	0'2	0'16	0'08
Freqüència relativa	0'24	0'56	0'76	0'92	1

Exemple:

- Les altures dels 12 jugadors de la Selecció Espanyola de Bàsquet (en metres) que van participar a l'Eurocopa 2013 s'arrepleguen a la taula següent:

2'03	1'96	1'91	2'11	1'91	1'93	2'08	1'99	1'90	2'16	2'06	2'03

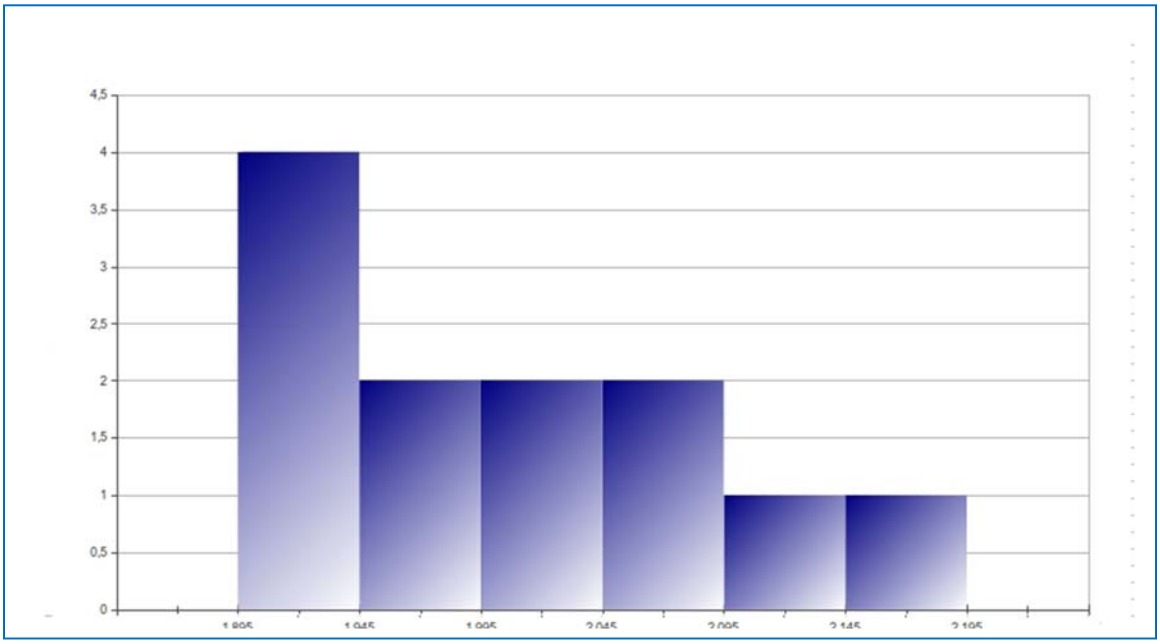
Com les dades són contínues, per a fer el recompte fixarem **intervals** d'altura:

1. entre 1'895 i 1'945 ////
2. entre 1'945 i 1'995 //
3. entre 1'995 i 2'045 //
4. entre 2'045 i 2'095 //
5. entre 2'095 i 2'145 /
6. entre 2'145 i 2'195 /

Ara portem les dades del recompte a un diagrama de freqüències:

entre 1'895 i 1'945	4
entre 1'945 i 1'995	2
entre 1'995 i 2'045	2
entre 2'045 i 2'095	2
entre 2'095 i 2'145	1
entre 2'145 i 2'195	1

En aquest cas la representació gràfica la fem amb un **histograma de freqüències**.

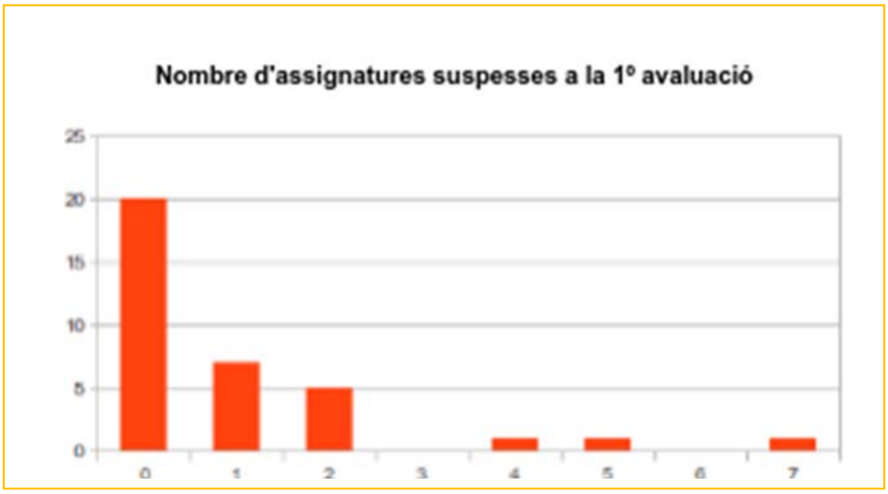


Observa la diferència entre aquest gràfic (corresponent a una variable contínua) i el que vam fer per al recompte de monedes (que representava una variable discreta). Aquest gràfic es denomina histograma de freqüències i és semblant a un diagrama de barres però ara representem unes barres pegades a altres, Per a recordar que es tracta d'intervals de classe i no de valors aïllats de les variables. Al nostre exemple tots els intervals tenen la mateixa longitud, 0'05 cm. Si les longituds dels intervals foren diferents les altures dels rectangles haurien de ser proporcionals a l'àrea.

2.2. Diagrama de barres

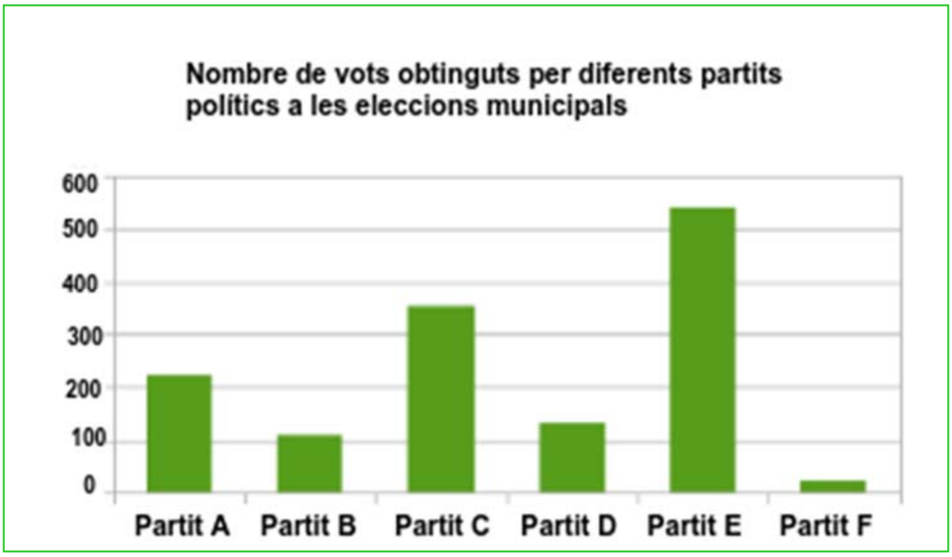
S'utilitza per a representar dades de variables estadístiques discretes o variables estadístiques qualitatives.

Al principi del capítol estudiant el nombre de monedes que es porten a la butxaca. Podem utilitzar aquest tipus de gràfic en altres situacions.



El gràfic anterior representa el nombre d'alumnes (d'una classe de 35) que han aprovat tot, el d'alumnes amb 1 assignatura suspensa, amb dues assignatures suspenses, etc. El millor de la representació gràfica és que d'una **sola ullada sabem que 20 alumnes han aprovat tot i que hi ha un alumne que té 7 assignatures suspeses.**

També podem utilitzar diagrames de barres per a representar variables qualitatives, com l'elecció de la modalitat de batxillerat que cursen els alumnes d'un IES o les preferències polítiques dels ciutadans d'un municipi.

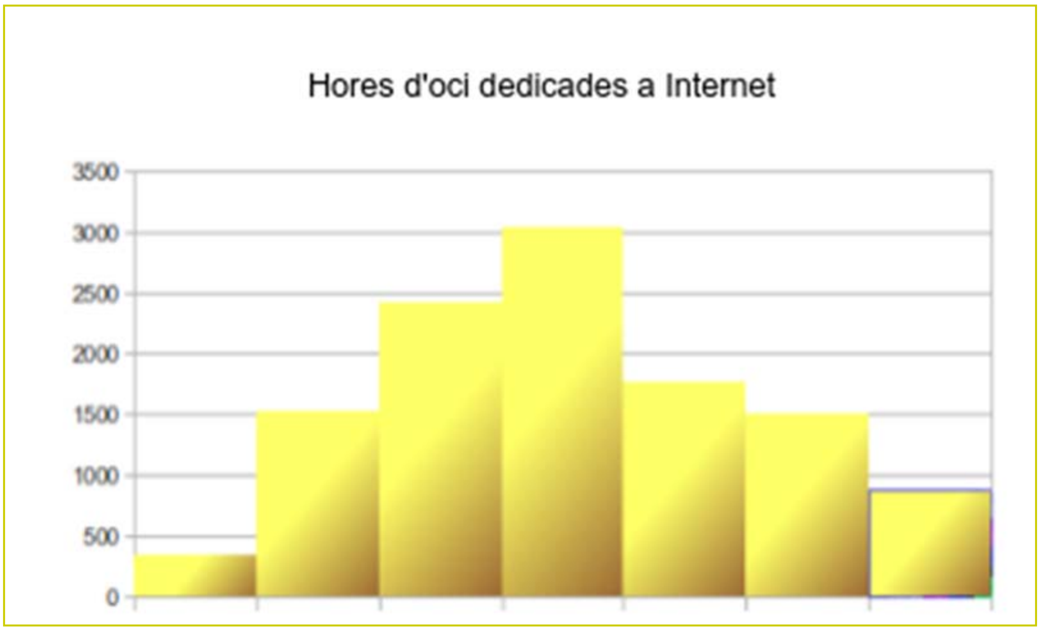


2.3. Histograma de freqüències

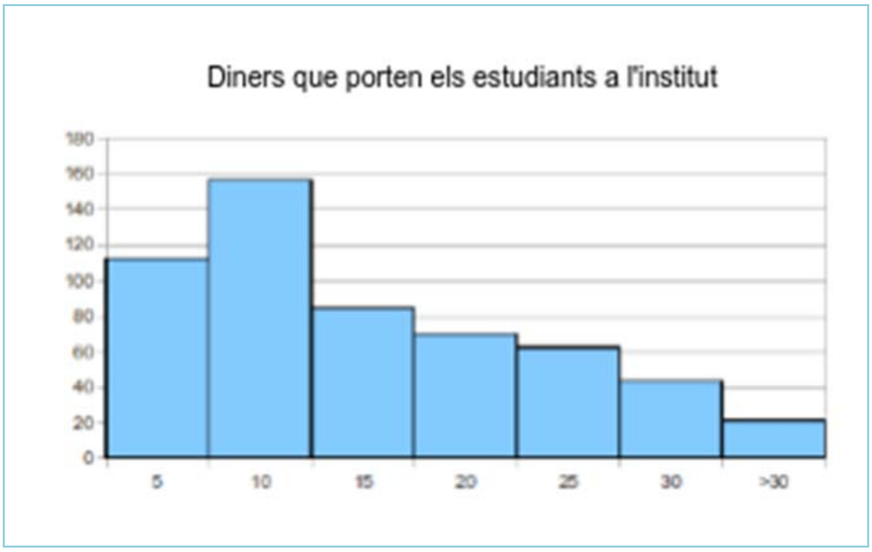
Aquest tipus de gràfic l'hem utilitzat abans per a representar les altures dels jugadors de la Selecció Espanyola de Bàsquet.

És semblant a un diagrama de barres però l'altura de cada barra ve donada pel nombre d'elements que hi ha a cada classe.

Altres variables que podem considerar com a variables contínues són el nombre d'hores que els jóvens d'una població dediquen a internet als seus moments d'oci o la quantitat de diners que es porta a la butxaca (ull, açò no és el nombre de monedes).



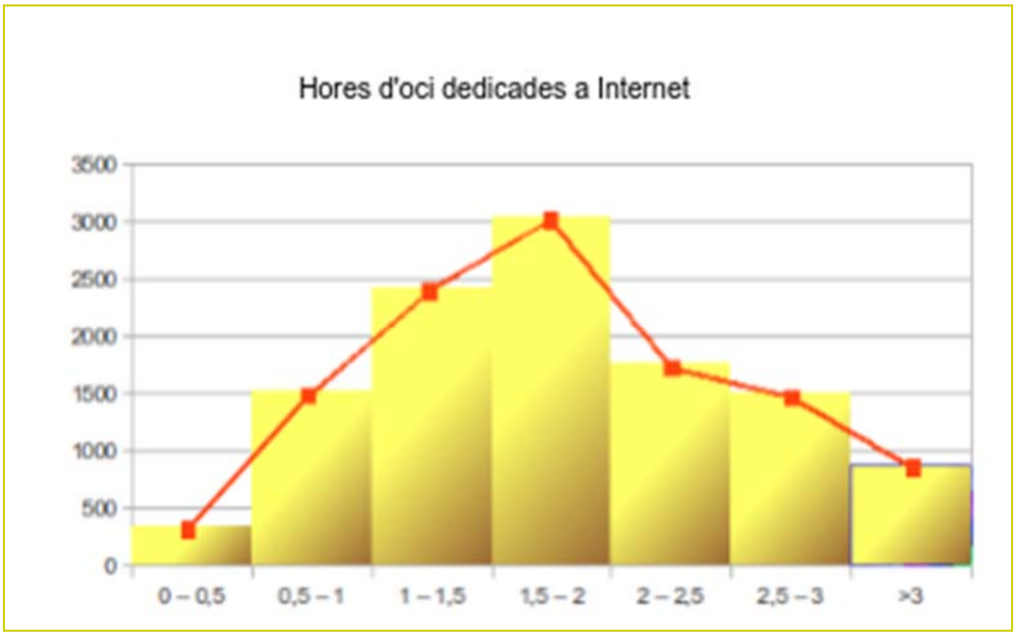
Al gràfic que incloem a continuació les marques de l'eix de les *x* es refereixen als trams de diners expressats de 5 en 5 euros. L'altura del gràfic es correspon amb la quantitat d'alumnes que porten aqueixa quantitat de diners. D'una simple ullada es veu que hi ha un poc més de 150 alumnes que porten entre 5 € i 10 € a l'institut i que poc més de 40 alumnes porten entre 25 € i 30 €.



Les barres són més amples i apareixen unes a continuació d'altres per a destacar que estem representant una variable contínua i que les altures es corresponen amb individus dins d'un interval de dades. Però recorda, si els intervals foren distints, les altures dels rectangles serien proporcionals a l'àrea.

2.4. Polígon de freqüències

S'utilitza als mateixos casos que l'histograma. Però dóna idea de la variació de la tendència. La línia poligonal es construeix unint els punts mitjans dels costats superiors dels rectangles.

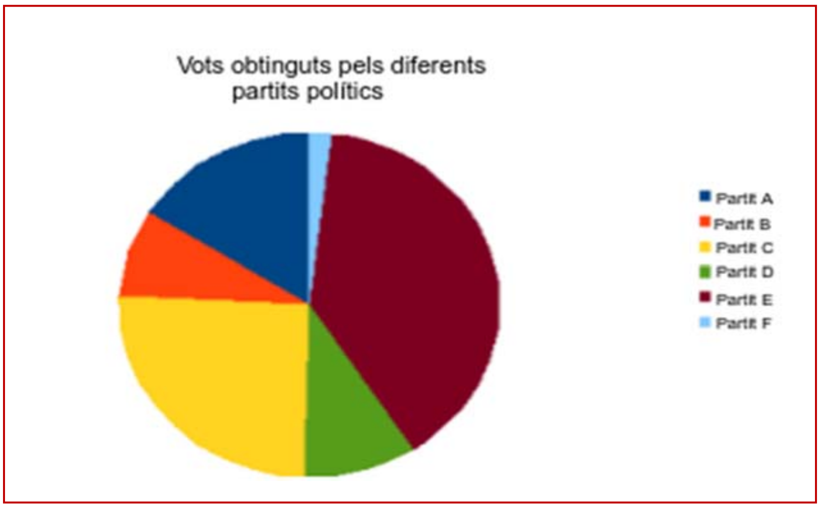


2.5. Diagrama de sectors

En algunes ocasions ens interessa fer-nos a la idea de la proporció que té cada resultat en relació amb els altres. S'utilitza molt amb variables qualitatives. Per exemple, aquesta representació s'utilitza per a mostrar els resultats d'unes les eleccions quan volem comparar els vots obtinguts pels diferents partits.

En un diagrama de sectors apareixen representats sectors circulars. L'angle d'aquests sectors és proporcional a la freqüència absoluta.

Reprenent l'exemple dels resultats obtinguts per diferents partits polítics representarem aqueixos mateixos resultats mitjançant un diagrama de sectors:



Activitats proposades

5. Reuneix a 10 amics. Reconta quantes monedes de cada valor (1cèntim, 2 cèntims, 5 cèntims, …) teniu entre tots. Representa mitjançant un gràfic adequat el nombre de monedes de cada classe que hi ha. Hi ha algun altre diagrama que et permeta veure quins tipus de monedes són més abundants en la mostra que has pres?

6. A la classe d'Educació Física el professor ha mesurat el temps que tarda cada alumne a recórrer 100 metres. Els resultats estan en aquesta taula:

14'92	13'01	12'22	16'72	12'06	10'11	10'58	18'58	20'07	13'15	20'10	12'43	17'51	11'59	11'79
16'94	16'45	10'94	16'56	14'87	17'59	13'74	19'71	18'63	19'87	11'12	12'09	14'20	18'30	17'64

Agrupa aquests resultats per classes començant en 10 segons i fent intervals de longitud 1 segon. Realitza una taula de freqüències i representa adequadament aquestes dades.

3. PARÀMETRES ESTADÍSTICS

3.1. Introducció

Segur que saps què és la mitjana de dos nombres i probablement saps calcular la mitjana d'una sèrie de dades. Però a més d'aqueixa mesura estadística hi ha altres mesures que poden ser interessants per a conéixer propietats de les dades que tenim.

Ara estudiarem les **mesures de centralització** (mitja, mitjana i moda) que ens proporcionen un valor de referència entorn del que es distribueixen les dades i les **mesures de dispersió** (recorregut, desviació mitjana, variància i desviació típica). Aquestes mesures ens indiquen com estan de separats les dades entorn de la mitja.

Exemple:

- Imagina que en dos exàmens de matemàtiques obtens un 6 i un 5. La mitja és 5.5. Suposa ara que les notes que has tingut són 10 i 1. La mitja també és 5.5 però hauràs d'estudiar-te la part en què has tret 1 per a recuperar. Les mesures de dispersió ens van a servir per a detectar quan tenim valors extrems, allunyats de la mitja.

3.2. Mesures de centralització

La **mitja** es calcula sumant tots els valors i dividint entre el nombre de dades.

Si $x_1, x_2, \ldots, x_n$ són els valors que pren la variable estadística que estem considerant, la mitja es representa per$\bar{x}$ i es calcula mitjançant la fórmula:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Aqueixa suma es pot escriure abreviadament com $\bar{x} = \frac{\sum x_i}{n}$. El símbol $\sum$ s'utilitza habitualment per a representar sumes de diversos sumands. L'utilitzaràs molt a partir d'ara.

Per a calcular la **mitjana** s'ordenen totes les dades de menor a major i ens quedem amb el que ocupa la posició central. Si tenim un nombre parell de dades, prenem com a mitjana la mitja dels dos nombres que ocupen les posicions centrals. La representarem per *Me*.

La ***mitjana*** *Me* és un valor tal que el 50 % de les observacions són inferiors a ell.

Els ***quartils*** Q_1, Q_2 i Q_3 són els valors tals que el 25 %, 50 % i 75 % (respectivament) dels valors de la variable són inferiors a ell. Per tant la mitjana coincideix amb el segon quartil.

Usem el terme **moda** per a referir-nos al valor que més es repeteix. La denotem per *Mo*.

Activitats resoltes

- Continuem utilitzant les dades d’estatura corresponents als 12 jugadors de la Selecció Espanyola de Bàsquet *(veure secció 2.1 d’aquest capítol).* ***L’estatura mitja*** es calcula sumant totes les altures i dividint entre el nombre de dades.

2’03 + 2’06 + 2’16 + 1’90 + 1’99 + 2’08 + 1’93 + 1’91 + 2’11 + 1’91 + 1’96 + 2’03 = 24’07

2’0058.

En aquest exemple no podem parlar de **moda**, ja que no hi ha un únic valor que siga el que més es repeteix.

La **mitjana** en aquest cas és 2’01. Per a calcular-la ordenem totes les dades de menor a major i ens quedem amb la que ocupa la posició central. Com en aquest cas tenim un nombre imparell de dades, prenem com a mitjana la mitja aritmètica dels 2 que ocupen les posicions centrals.

Les dades, després d’ordenar-les, quedarien així:

1’90	1’91	1’91	1’93	1’96	1’99	2’03	2’03	2’06	2’08	2’11	2’16

Mitja d’ambdós = 2’01

Per a calcular els **quartils** hem de dividir el total de dades, en aquest exemple 12, entre 4, (o multiplicar per 0’25 que és el mateix) i obtenim 3. Per tant el primer quartil observem que està entre 1’91 i 1’93, fem la mitja i obtenim que Q_1 = 1’92. Per a calcular el tercer quartil multipliquem per 3 i dividim per 4, (o multipliquem per 0’75) i en aquest cas s’obté el valor que està entre 9, 2’06, i 10, 2’08, per la qual cosa Q_3 = 2’07.

3.3. Mesures de dispersió

Recorregut és la diferència entre la dada major i la dada menor. També es denomina **rang**.

Desviació mitja és la mitja de les distàncies de les dades a la mitja dels dades de què disposem.

$$\text{DM} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \ldots + |x_n - \bar{x}|}{n} = \frac{\sum |x_i - \bar{x}|}{n}$$

Variança és la mitja dels quadrats de les distàncies de les dades a la mitja.

$$\text{Variança} = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots+(x_n-\bar{x})^2}{n} = \frac{\sum (x_i-\bar{x})^2}{n}$$

Equivalentment (desenrotllant els quadrats que apareixen en l’expressió) es pot calcular mitjançant aquesta altra expressió:

$$\text{Variança} = \frac{\sum x_{i2}}{n} - \bar{x}$$

Desviació típica és l'arrel quadrada de la variança.

Es representa per σ.

$$\sigma = \sqrt{\frac{\sum x_{i^2}}{n} - \bar{x}}$$

Recorregut interquartílic o interval **interquartílic** és la distància entre el tercer i el primer quartil:

R = Recorregut interquartílic = $Q_3 - Q_1$.

Aquestes fórmules provenen de diferents formes de mesurar les distàncies. Per al càlcul de la desviació mitja s'usen valors absoluts, que és com es mesura la distància entre nombres a la recta real. La desviació típica té a veure amb la forma de mesurar distàncies al pla (recordem que la hipotenusa d'un triangle és l'arrel quadrada de la suma dels quadrats dels catets). No cal que comprengues ara d'on ixen aquestes fórmules però sí que és convenient que sàpies que no és per capritx dels matemàtics que les van inventar. Cada cosa al seu temps...

Activitats resoltes

- Tornem a usar les dades de l'exemple de la Selecció Espanyola amb què treballem.

Recorregut: 2'16 – 1'90 = 0'26 (metres). Açò és la diferència d'altures entre el jugador més alt i el més baix.

Per a calcular la **desviació mitja** primer calcularem la suma que apareix al numerador. Després dividirem entre el nombre de dades.

|2'03 – 2'0058| + |2'06 – 2'0058| + |2'16 – 2'0058| + |1'90 – 2'0058| + |1'99 – 2'0058| +

|2'08 – 2'0058| + |1'93 – 2'0058| + |1'91 – 2'0058| + |2'11 – 2'0058| + |1'91 – 2'0058| +

|1'96 – 2'0058| + |2'03 – 2'0058| = 0'0242 + 0'0458 + 0'0958 + 0'1042 + 0'0958 + 0'0758 + 0'0742 + 0'0158 + 0'1058 + 0'1542 + 0'9458 + 0'0242 = 0'87

Així la **desviació mitja** és 0'87/12 = 0'0725

Per a calcular la **variança** primer calcularem la suma que apareix al numerador, de forma semblant a com acabem de fer. Després acabarem dividint entre el nombre de dades.

$(2'03 – 2'0058)^2 + (2'06 – 2'0058)^2 + (2'16 – 2'0058)^2 + (1'90 – 2'0058)^2 + (1'99 – 2'0058)^2 +$

$(2'08 – 2'0058)^2 + (1'93 – 2'0058)^2 + (1'91 – 2'0058)^2 + (2'11 – 2'0058)^2 + (1'91 – 2'0058)^2 +$

$(1'96 – 2'0058)^2 + (2'03 – 2'0058)^2 = 0'08934$

Així la **variança** és 0’08934/12 = 0’00744

La **desviació típica** és l’arrel quadrada de la variança: $\sigma = \sqrt{0'00744}$ = 0’08628.

Recorregut interquartílic o **interval interquartílic** es calcula restant $Q_3 - Q_1$ = 2’07 – 1’92 = 0’15.

Les mesures de posició ens permeten realitzar un altre tipus de gràfic estadístic que s’anomena el ***gràfic de caixa***.

3.4. Càlcul detingut dels paràmetres estadístics

El més còmode per a calcular paràmetres estadístics és utilitzar un full de càlcul. Les calculadores científiques també incorporen funcions per a obtindre els principals paràmetres estadístics. Per a saber com usar la teua calculadora pots llegir el manual que ve amb ella.

Ara veurem com es poden utilitzar les taules de freqüències per a calcular la mitja i la variança.

Quan hi ha valors repetits en compte de sumar aqueix valor diverses vegades podem multiplicar el valor per la seua freqüència absoluta. També, el nombre de dades és la suma de les freqüències.

D’aquesta manera obtenim la següent **fórmula per a la mitja**

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

Anàlogament, la **variança** es pot calcular mitjançant

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

o, alternativament, mitjançant l’expressió

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

(Aquestes dues fórmules són equivalents. La segona expressió s’obté desenrotllant els quadrats de la primera i simplificant).

Per tant la **desviació típica** es calcula:

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

Activitats resoltes

- Les notes de 15 alumnes en un examen de matemàtiques es reflecteixen a la següent taula

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

Volem calcular la seua mitja i la seua variança.

En primer lloc, elaborem una taula de freqüències amb aqueixes dades:

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

Afegim una columna en què escriurem el resultat de multiplicar la freqüència i el valor, açò és, $xi \cdot fi$.

x_i	f_i	$x_i \cdot f_i$
1	1	1
2	0	0
3	1	3
4	1	4
5	4	20
6	3	18
7	2	14
8	1	8
9	1	9
10	1	10
	$\sum f_i = n = 15$	$\sum x_i \cdot f_i = 87$

Sumant les freqüències (columna central) obtenim el nombre de dades.

Així la mitja és el quocient entre la suma de la columna de la dreta entre la suma de la columna central.

$$\bar{x} = \frac{87}{15} = 5'8$$

Per a calcular la variança afegirem una columna més a la taula anterior. En aqueixa columna escriurem el producte de la freqüència pel quadrat del valor.

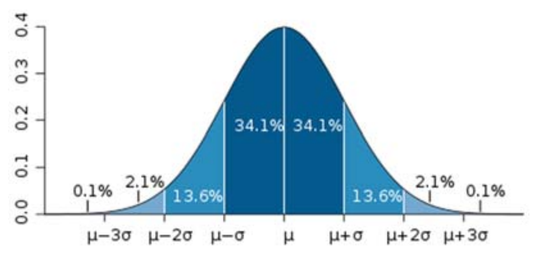
x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	1	1	1
2	0	0	0
3	1	3	9
4	1	4	16
5	4	20	100
6	3	18	108
7	2	14	98
8	1	8	64
9	1	9	81
10	1	10	100
	$\sum f_i = n = 15$	$\sum x_i \cdot f_i = 87$	$\sum x_i^2 \cdot f_i = 577$

Així la **variança** és $\sigma^2 = \frac{577}{12} - 5'8^2 = 14'4433$

I la **desviació típica** és $\sigma = \sqrt{14'4433} = 3'8004$.

3.5. Interpretació conjunta de la mitja i la desviació típica

Hem vist que la desviació típica ens mesura la distància de les dades respecte de la mitja. Ens dóna molta informació. Informa sobre com s'agrupen les dades al voltant de la mitja.



Mitja i desviació típica. Imatge de wikipedia

Si les dades que hem arreplegat tingueren una distribució normal (de moment no sabem el que açò significa exactament dins de l'Estadística, però pots suposar que significa això, que són normals, que no els passa gens estrany) resulta que a l'interval entre la mitja menys una desviació típica i la mitja més una desviació típica estan més del 68 % de les dades. A l'interval entre la mitja menys 2 desviacions típiques i la mitja més 2 desviacions típiques estan més del 95 % de les dades, i entre la mitja menys 3 desviacions típiques i la mitja més 3 desviacions típiques estan més del 99'7 % de les dades.

Es podria dir que quelcom, per exemple la intel·ligència d'una persona, l'altura d'una planta o el pes d'un animal... és normal si està dins d'aqueix interval $(\bar{x} - \sigma, \ \bar{x} + \sigma)$, que és intel·ligent, alt o pesat si està entre $(\bar{x} + \sigma, \ \ \bar{x} + 2\sigma)$, o que és un geni, gegant o molt pesat si està a l'interval $(\bar{x} + 2\sigma, \bar{x} + 3\sigma)$.

Observa que estem dient que pràcticament totes les dades disten de la mitja menys de 3 desviacions típiques i que més del 68 % disten menys d'una desviació típica. Açò serà de gran utilitat perquè connecta amb altres branques de l'Estadística. Fins ara hem estat descrivint el que ocorre. Ara podrem prendre decisions, inferir o predir amb una certa probabilitat el que ocorrerà. Per això estudiarem a continuació les probabilitats.

3.6. Diagrama de caixes o de bigots

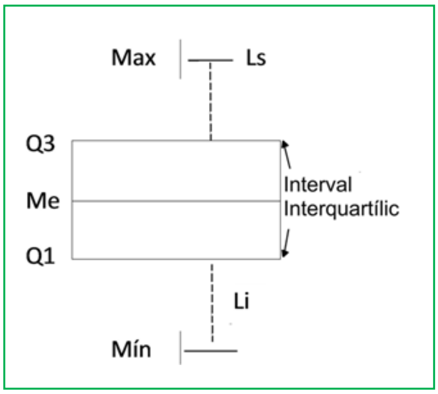
El **diagrama de caixes** és una representació gràfica en què s'utilitzen els quartils, la mitja, els valors màxims i mínims… intentant visualitzar tot el conjunt de dades.

Es forma un rectangle (o caixa) els costats del qual són els quartils (Q_1 i Q_3) i on s'assenyala al centre, la mitjana (*Me*). S'afigen dos braços (o *bigots*) on s'assenyalen els valors màxim (*Màx*) i mínim (*Mín*).

Es poden calcular, a més, uns límits superior i inferior. L'inferior, L_i; és Q_1 – 1'5 per l'interval interquartílic, i el superior Ls és Q_3 + 1'5 per l'interval interquartílic.

Exemple

- Neus ha tingut en Matemàtiques les notes següents: 8, 4, 6, 10 i 10. Calcula el seu recorregut, la variança, la desviació típica, els quartils i l'interval interquartílic.



Ordenem les dades: $4 \leq 6 \leq 8 \leq 10 \leq 10$, i calculem que:

Mitjana = *Em* = 8.

$Q_1 = 6$. $Q_3 = 10$.

Interval interquartílic = 10 – 6 = 4.

Els bigots ens indiquen:

Màx = 10. Mín = 4.

Ls = Q3 + 4*1'5 = 16. Li = Q1 – 4*1'5 = 0.

En aquest exemple el màxim és igual a 10, que és menor que el possible extrem superior, igual a 16. El mínim és 4, major que l'extrem inferior, per tant no hi ha *valors atípics* que siguen majors que el límit superior o menors que el límit inferior. Els extrems dels bigots, al nostre exemple són 10 i 4.

El diagrama de caixa és el de la figura del marge.

4. INTRODUCCIÓ AL CÀLCUL DE PROBABILITATS

4.1. Conceptes bàsics en probabilitat

Tots els dies apareixen a la nostra vida fets que tenen a veure amb la probabilitat. Si juguem al parxís, intuïm que *més o menys* una de cada 6 vegades eixirà un 5, amb la qual cosa podrem traure una fitxa a recórrer el tauler. Al 'Monopoly' traure un doble tres vegades seguides ens envia a la presó ("sense passar per la casella d'eixida"). Açò no ocorre moltes vegades; no obstant això, tots els que hem jugat a açò hem anat a la presó per aqueix motiu.

La **probabilitat** és una mesura de com és de factible que tinga lloc un determinat succés.

Per a estudiar la probabilitat, hem d'introduir alguns noms. Ho anem a fer amb ajuda d'un cas concret.

Exemple

- Imaginem que tenim una bossa amb 5 boles: 2 blanques, 2 roges i una negra. Fem el següent **experiment aleatori**: ficar la mà a la bossa i mirar el color de la bola que ha eixit.

Hi ha 3 *casos* possibles: "que la bola siga blanca", "que la bola siga roja" o "que la bola siga negra". Abreviadament els representarem per *blanca*, *roja* o *negra* (també podrem representar els colors o escriure B, R o N; recorda que en matemàtiques sempre s'ha de simplificar, inclús la manera d'escriure).

L'espai mostral és el conjunt de tots els casos possibles : {B, R, N}.

Els diferents **successos** són els subconjunts de l'espai mostral. Al nostre exemple els successos possibles són {B},{R}, {N}, {B,R}, {B,N}, {R,N}, {B,R,N}.

És segur que en el nostre experiment la bola que traiem és "blanca", "negra" o "roja". Per això a l'espai mostral se l'anomena també **succés segur**.

Recorda aquests noms:

Un **experiment aleatori** és una acció (experiment) el resultat de la qual depèn de l'atzar.

A cada un dels resultats possibles d'un experiment aleatori li cridarem **cas** o **succés individual**.

El conjunt de tots els casos possibles s'anomena **espai mostral** o **succés segur**.

Un **succés** és un subconjunt de l'espai mostral.

Exemples.

- Baralla espanyola de 40 cartes. Experiment: traiem una carta a l'atzar i mirem el seu pal.

  Espai mostral {ors, copes, espases, bastos}

1 Experiment: llancem simultàniament 1 moneda d'euro i una de 2 euros a l'aire.

  Espai mostral:{Cara-Cara, Cara-Creu Creu-Cara, Creu-Creu}

2 Experiment: llancem simultàniament 2 monedes d'1 euro (indistingibles)

Espai mostral: {Ixen 2 cares, Ixen 2 creus, Ix 1 cara i una creu}

3 Experiment: llancem una moneda d'1 euro i apuntem què ha eixit; la tornem a llançar i apuntem el resultat.

Espai mostral: {CC, CX, XC, XX}

4 Experiment: llancem simultàniament dos daus i sumem els nombres que es veuen a les cares superiors.

Espai mostral:{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

5 Experiment: llancem un dau usual i sumem els nombres que apareixen a la cara superior i la cara inferior (la que no es veu, que està sobre la taula).

Espai de successos: {7}

Als exemples anteriors, (2) i (4) són equivalents: els possibles resultats del llançament de 2 monedes que es distingeixen són els mateixos que els del llançament d'una mateixa moneda dues vegades (per exemple, equiparem el resultat del llançament de la moneda d'1 euro de l'exemple 3 amb el primer llançament de la moneda de l'exemple 4 i el resultat del llançament de la moneda de 2 euros amb el segon llançament).

A l'experiment 6 sempre ix el mateix resultat (per alguna raó els punts en els daus usuals es distribueixen sempre de manera que les cares oposades sumen 7). Tècnicament aquest no és un experiment aleatori, ja que el resultat no depèn de l'atzar.

Activitats proposades

7. Para cada un dels exemples 1 a 5 anteriors indica 3 successos diferents que no siguen successos individuals.

8. En una bossa tenim 10 boles roges numerades de l'1 al 10. Es fan els dos experiments següents:

EXPERIMENT A: Es trau una bola de la bossa i es mira el seu color.

EXPERIMENT B: Es trau una bola de la bossa i es mira el seu nombre.

Quin d'aquests experiments no és un experiment aleatori? Per què?

9. Per a l'experiment que sí que és un experiment aleatori indica el seu espai mostral.

10. Una baralla francesa té 52 cartes, distribuïdes en 13 cartes de piques, 13 de cors, 13 de trèvols i 13 de diamants. Les piques i els trèvols són cartes negres mentres que els cors i els diamants són cartes roges. Es mescla la baralla, es talla i es fa l'experiment següent: agafar les dues cartes que han quedat dalt del tot i observar de quin color són.

Descriu l'espai mostral.

4.2. Càlcul de probabilitats

Ja hem indicat que la probabilitat és una mesura que ens indica el grau de confiança que ocórrega un determinat succés.

La **probabilitat** s'expressa mitjançant un nombre comprés entre 0 i 1.

Si aqueix nombre està pròxim a 0 direm que és un succés improbable (ull, improbable no vol dir que siga impossible), mentres que si està pròxim a 1 direm que aqueix succés serà molt més probable.

Exemple

- En una bossa que conté 20 boles blanques introduïm una bola negra (indistingible al tacte). Mesclem bé les boles de la bossa, i realitzem l'experiment consistent a ficar la mà a la bossa i traure una bola.

Sense que hagem estudiat res formalment sobre probabilitat. Què penses que és més probable, que la bola treta és blanca o que és negra? Estarem d'acord en què és més probable traure una bola blanca.

Ara ja sí que podem plantejar-nos una pregunta: En quina mesura és més probable traure una bola blanca?

No és difícil de calcular. Les dades que tenim són els següents

b) la bossa té 21 boles

- 1 bola és negra
- 20 boles són blanques

La probabilitat de traure la bola negra és 1 d'entre 21. La probabilitat de traure una bola blanca és de 20 entre 21.

El que acabem d'utilitzar és conegut com a **Llei de Laplace.** Si tots els casos d'un espai mostral són **equiprobables** (açò és, tenen la mateixa probabilitat d'ocórrer), i S és un succés d'aqueix experiment aleatori es té que

$$P(S) = \frac{nombredecasosfavorablesalsuccésS}{nombredecasospossibles}$$

Exemple.

- Mesclem una baralla espanyola de 40 cartes (els pals són ors, copes, espases i bastos i en cada pal hi ha cartes numerades de l'1 al 7 a més d'una sota, un cavall i un rei).

Es realitza l'experiment consistent a tallar *la baralla i quedar-nos amb la carta superior*.

Considerarem els successos següents:

a) Obtindre una figura

- Obtindre una carta amb un nombre imparell

- Obtindre una carta d'espases
- Obtindre una carta d'espases o una figura
- Obtindre la sota d'ors

En principi les cartes no estaran marcades, amb la qual cosa la probabilitat que isca cada una d'elles és la mateixa. Açò és, estem davant d'un experiment aleatori amb tots els casos equiprobables.

a) A la baralla hi ha 12 figures (3 per cada pal). Així

   Casos favorables: 12

   Casos possibles: 40

   Probabilitat: 12/40= 3/10

b) Per cada pal hi ha 4 cartes amb nombres imparells: 1, 3, 5 i 7.

   Casos favorables: 16

   Casos possibles: 40

   Probabilitat: 16/40=2/5

c) Hi ha 10 cartes d'espases a la baralla

   Casos favorables: 10

   Casos possibles: 40

   Probabilitat: 10/40=1/4

d) Hi ha 10 cartes d'espases i a més altres 9 figures que no són d'espases (clar, les 3 figures d'espases ja les hem comptat).

   Casos favorables: 19

   Casos possibles: 40

   Probabilitat: 19/40

e) Només hi ha una sota d'ors

   Casos favorables: 1

   Casos possibles: 40

   Probabilitat: 1/40

El que és capaç de calcular probabilitats ràpidament té avantatge en alguns jocs en què es mescla atzar amb estratègia. Per exemple, jocs de cartes o de dòmino. Si sabem quines cartes o fitxes s'han jugat podem estimar la probabilitat que un altre jugador tinga una determinada jugada. Òbviament en aqueixos casos no *quantifiquem* (no fem els càlculs exactes) però sí que *estimem* si tenim la probabilitat al nostre favor o en contra nostre.

Per a aprendre més…

Girolamo Cardano (1501-1576) va ser un personatge inquiet i prolífic. A més de dedicar-se a les matemàtiques era metge, però també era un jugador. De fet ell va ser qui va escriure el primer treball que es coneix sobre jocs d'atzar. Un segle després el Caballer de Meré, un conegut jugador, va plantejar a Blas Pascal diversos problemes que li apareixien a les seues partides. Un dels problemes que li va plantejar és el del repartiment dels guanys quan una partida s'ha d'interrompre. Aquest problema ja havia sigut tractat amb anterioritat per Lucca Pacioli (el matemàtic que va inventar la taula de doble entrada per a ajudar als Medici a portar la comptabilitat de la seua Banca).

El problema enunciat i resolt per Pacioli és aquest:

- Dos equips juguen a la pilota de manera que gana el joc el primer equip que gana 6 partits. L'aposta és de 22 ducats, que se'ls portarà el guanyador. Per algun motiu cal interrompre el joc quan un equip ha guanyat 5 partits i l'altre 3. Es vol saber com repartir els 22 ducats de l'aposta, d'una manera justa.

Pensa-ho!

A pesar d'haver passat a la història de les matemàtiques, la solució que va donar Pacioli a aquest problema hui no es consideraria correcta per no tindre en compte la probabilitat. Què proposes tu? Aquest és un problema curiós, perquè no tenim totes les dades ni coneixem les probabilitats que intervenen a la seua resolució, però és un bonic exemple per a pensar en equip i discutir sobre el tema. Dir què és i què no és just és molt complicat.

Activitats resoltes

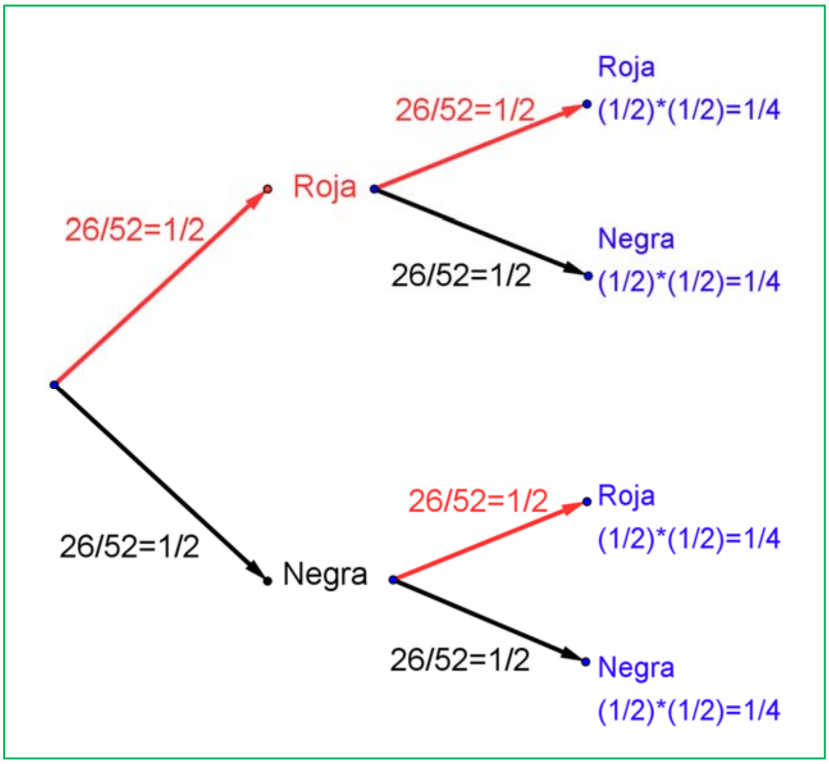
- Una bossa de boles conté 26 negres i 26 roges. Es mescla el contingut de la bossa, es fica la mà i es trau una bola, es mira el color i es torna a la bossa. A continuació es trau una altra bola i es mira el color. Quina és la probabilitat que hagen eixit una bola roja i una bola negra?

Abans de continuar llegint, pensa-ho. Si t'equivoques no passa res: el sentit de probabilitat no el tenim massa desenrotllat, però aquest és el moment de fer-ho.

Aquest problema l'hem plantejat moltes vegades a altres estudiants. Alguns diuen que la probabilitat és 1/3 perquè hi ha 3 casos possibles: Roja-Roja, Negra-Negra i Roja-Negra. Aqueixa resposta no és correcta.

En realitat el succés traure *una bola de cada color* consta de 2 casos Roja-Negra i Negra-Roja. Depenent de com haguérem escrit l'espai mostral o de com haguérem plantejat el problema aqueix detall es podria veure amb major o menor claredat.

Així, la probabilitat de traure una bola de cada color és, en realitat 1/2.



Si no t'ho creus pots fer un experiment: serà difícil que tingues 26 boles negres i 26 boles roges, però sí que és fàcil que tingues una baralla francesa. Mescla-la, talla i mira el color de la carta que ha quedat dalt al muntó. Apunta-ho. Torna a deixar les cartes en la baralla, torna a mesclar, talla de nou i mira el color de la carta que ha quedat dalt ara. Apunta els colors. Repeteix aquest experiment moltes vegades: 20, 50 o 100.

Si tens en compte els resultats veuràs que, aproximadament, la meitat de les vegades les dues cartes són del mateix color i l'altra meitat les cartes són de colors diferents. Amb això, hem pogut "comprovar" que la probabilitat d'aqueix succés era 1/2.

Una altra forma que et pot ajudar a raonar sobre aquest problema, i molts altres de probabilitat, és confeccionar un **diagrama en arbre**. La primera bola que traiem té una probabilitat de ser Roja igual a 26/52 = 1/2. Aqueix nombre l'escrivim a la branca de l'arbre. Si tornem a la bossa la bola i tornem a traure una altra bola de la bossa, la probabilitat que siga Roja torna a ser 26/52 = 1/2. Completem amb idèntic raonament la resta de les branques.

La probabilitat que les dues boles que hagem tret siguen roges és el producte de les seues branques: (1/2)·(1/2) = 1/4. La mateixa probabilitat obtenim per als successos Negra-Negra, Negra-Roja i Roja-Negra. La probabilitat de Roja-Negra és per tant 1/4, igual a la de Negra-Roja. Com són successos elementals la probabilitat que les dues boles siguen de distint color és la suma: 1/4 + 1/4 = 1/2.

4.3. Probabilitat i freqüència relativa

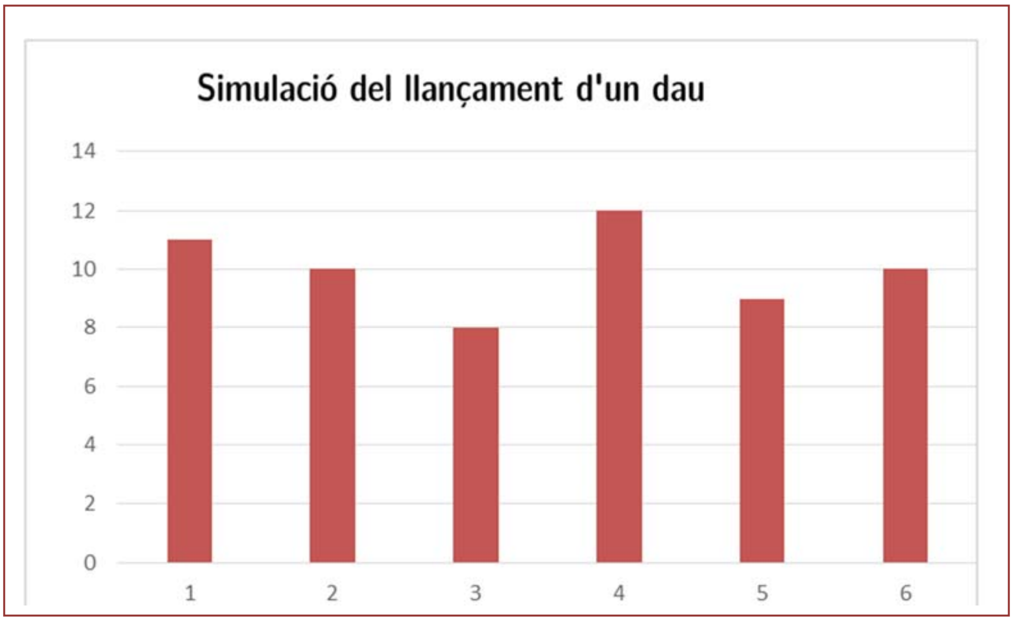
Al principi del capítol, quan introduíem els principals conceptes estadístics, parlàvem de la freqüència. A aqueixa freqüència se l'anomena **freqüència absoluta** per a distingir-la d'un altre concepte, que és molt més pròxim a la probabilitat.

Anomenarem **freqüència relativa** d'un resultat d'un experiment aleatori a la seua freqüència absoluta dividit entre el nombre de repeticions de l'experiment.

Exemple

- Llança un dau 60 vegades, copia aquesta taula al teu quadern i apunta el que ix:

Si dibuixes un diagrama de barres amb els resultats de l'experiment obtindràs un paregut a aquest:

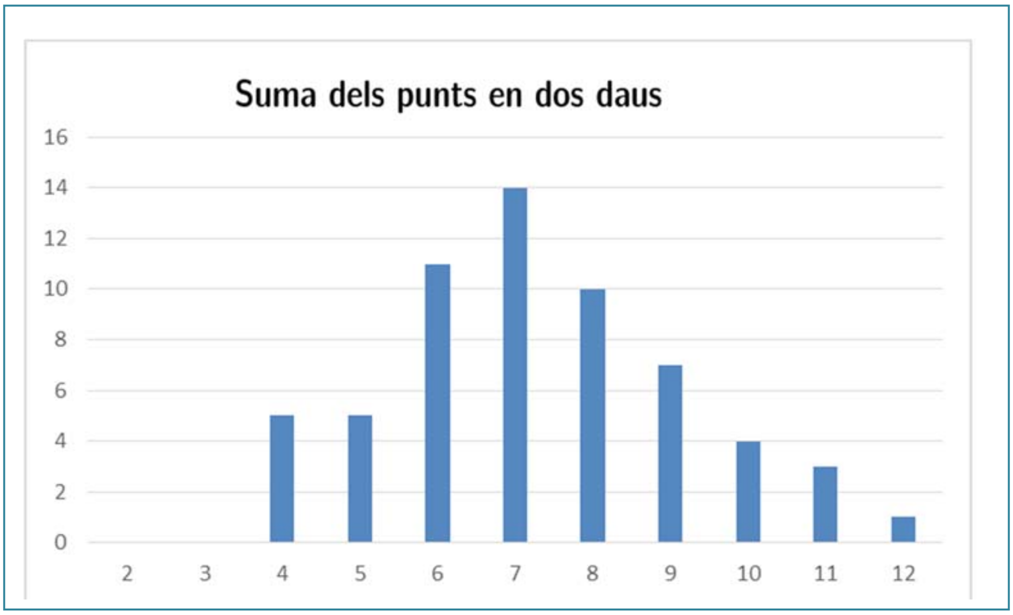


La freqüència relativa de cada un dels casos és prou pareguda a la probabilitat d'aqueix cas (que és 1/6).

Exemple.

- Feix ara un altre experiment: llança 2 daus 60 vegades i apunta la suma dels valors dels dos daus en aquesta taula.

Dibuixa ara un diagrama de barres. El que obtindràs serà un paregut a aquest:



Si la probabilitat "s'ha de parèixer" a les freqüències relatives, en aquest cas veiem que el succés *que la suma done 7* és més probable que qualsevol dels altres. I molt més probable que *que la suma done 2* o *que la suma done 12*.

La **llei dels grans nombres** ens diu que quan es repeteix moltes vegades un experiment aleatori la freqüència relativa de cada succés S s'aproxima a la seua probabilitat. Com més gran siga el nombre de repeticions, millor va sent l'aproximació.

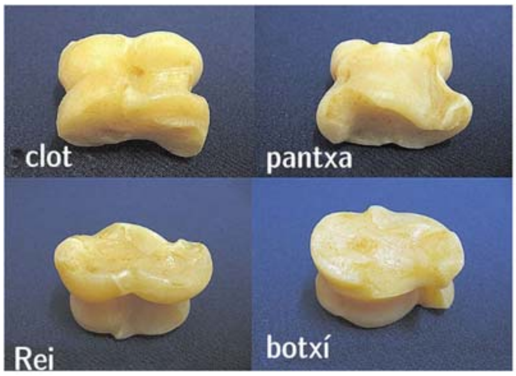
En aquest cas l'útil és utilitzar les freqüències relatives per a estimar probabilitats quan aquestes no són conegudes.

Activitats proposades

11. En alguns llocs d'Espanya es continua jugant a la taba. La taba és un os de corder que no és regular. Pot caure en quatre posicions distintes. Podem pensar en ella com si fóra un dau "rar".

Considera l'experiment "llançar la taba a l'aire i veure la que marca la seua cara superior: clot, panxa, rei i botxí".

Aproxima la probabilitat de cada un dels casos d'aquest experiment aleatori.



(Imatge: Wikimedia Commons)

12. La teua calculadora probablement tindrà una funció que serveix per a generar nombres aleatoris. Normalment dóna un nombre comprés entre 0 i 1.

Realitza l'experiment aleatori "genera un nombre aleatori i apunta el seu segon decimal". Fes 40 repeticions d'aquest experiment. Dibuixa un histograma de freqüències.

13. La probabilitat no és un concepte intuïtiu. Per a això farem una prova. Considerarem l'experiment aleatori *llançar una moneda*. Copia la taula al teu quadern

14. Escriu en la 1ª fila d'aquesta taula el que tu creus que eixiria en repetir l'experiment 30 vegades. Pensa-ho i ompli la taula. Com tu vulgues (inventa-t'ho, però "amb sentit").

1) A la 2ª fila de la taula escriu el resultat real de 30 llançaments de la moneda.

Què observes en ambdós casos? Alguna pauta? Presta atenció a aquestes qüestions per a cada una de les files de la taula.

Hi ha més o menys 15 cares i 15 creus?

Apareixen grups seguits de cares o creus?

Quin és el nombre més gran de cares que han eixit seguides? I el de creus?

Normalment quan "t'inventes" els resultats sí que sols posar la meitat de cares i la mitat de creus. En un experiment aleatori aquests nombres estan prop de la meitat però no solen ser la meitat exacta.

Quan t'ho inventes, en general poses pocs grups seguits de cares o creus.

El cervell ens enganya i en temes probabilístics hem d'educar-lo molt més. Per això aquest tema és molt important, encara que siga el que moltes vegades es queda sense donar. Ens ajuda a què, com a ciutadans, no ens enganyen. Ni amb loteries, ni amb cartes, ni amb estadístiques electorals.

CURIOSITATS. REVISTA

Un problema resolt: Les tres ruletes

Disposem de tres ruletes A, B i C cada una d'elles dividida en 32 sectors iguals amb distints punts:

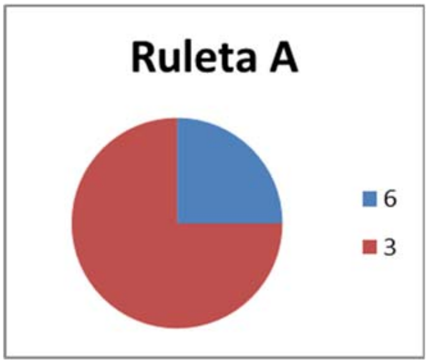
A: 8 sectors amb la xifra 6 i 24 sectors amb la xifra 3.

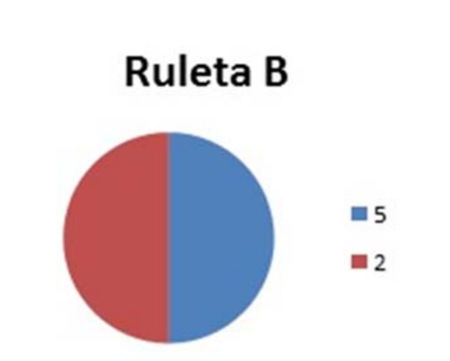
B: 16 sectors amb la xifra 5 i 16 sectors amb la xifra 2.

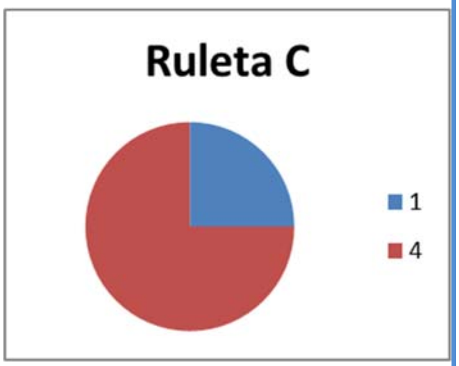
C: 8 sectors amb la xifra 1 i 24 sectors amb la xifra 4.

Dos jugadors seleccionen una ruleta cada u. Guanya qui obtinga major puntuació amb la ruleta.

Qui té avantatge en triar ruleta, la persona que tria primer o la que tria en segon lloc?







Solució: "Les tres ruletes"

Fes un **diagrama d'arbre** i comprova que:

Jugant amb la Ruleta A i la Ruleta B.

P(guanyar A) = $\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{2}=\frac{5}{8}$ P(guanyar B) = $\frac{3}{4}\cdot\frac{1}{2}=\frac{3}{8}$

Guanya el que juga amb la Ruleta A.

Jugant amb la Ruleta A i la Ruleta C.

P(guanyar A) = $\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{4}=\frac{7}{16}$ P(guanyar C) = $\frac{3}{4}\cdot\frac{3}{4}=\frac{9}{16}$

Guanya el que juga amb la Ruleta C.

Jugant amb la Ruleta B i la Ruleta C

P(ganar B) = $\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{4}=\frac{5}{8}$ P(ganar C) = $\frac{1}{2}\cdot\frac{3}{4}=\frac{3}{8}$

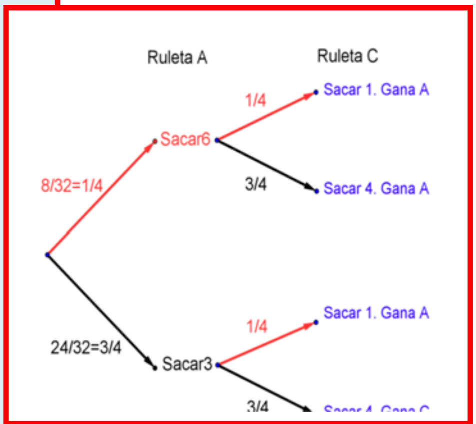
Guanya el que juega amb la Ruleta B.

Guanya el jugador que eligeix en segon lloc:

Si el primer eligeix la Ruleta A ⟶ El segon eligeix la Ruleta C i guanya.

Si el primer eligeix la Ruleta B ⟶ El segon eligeix la Ruleta A i guanya

Si el primer eligeix la Ruleta C ⟶ El segon eligeix la Ruleta B i guanya



Breu història de la Probabilitat

Girolamo Cardano (1501-1576)va ser un personatge inquiet i prolífic. A més de dedicar-se a les matemàtiques era metge, però també era un jugador. De fet ell va ser qui va escriure el primer treball que es coneix sobre jocs d'atzar.

Un segle després el Caballer de Méré li va plantejar a Blaise Pascal alguns problemes sobre **jocs** com el següent:

Un jugador intenta obtindre un 1 en 8 llançament successius d'un dau, però el joc s'interromp després de 3 llançaments fallits. En quina proporció ha de ser compensat el jugador?

Pascal va escriure a Fermat sobre aquest problema i la correspondència intercanviada es pot considerar com l'inici de la Teoria de Probabilitats, però no van publicar per escrit les seues conclusions. Aquest problema ja havia sigut tractat amb anterioritat per Lucca Pacioli (el matemàtic que va inventar la taula de doble entrada per a ajudar als Medici a portar la comptabilitat de la seua Banca).

Huygens en 1657 va publicar un breu escrit "*Els jocs d'atzar*" on narra la dita correspondència.

Però el primer llibre sobre Probabilitat és de 1713 de Jacques Bernoulli, *"L'art de la conjectura"* . En ell s'enuncia **la llei dels grans nombres** que ve a dir que la probabilitat d'un succés s'acosta a les freqüències relatives quan el nombre d'experiments és gran. Conéixer açò va portar a grans jugadors a guanyar al Casino de Montecarlo, com es narra més avall. L'Estadística i La Probabilitat es van usar en problemes socials com defendre la **vacunació de la pigota,** l'educació pública ...a la Il·lustració Francesa.

Fins ací, ja saps resoldre tots els problemes històrics. Però hi ha altres més difícils, que requereixen més coneixements de Matemàtiques, com el de **l'agulla de Buffon**, que s'ha utilitzat per a calcular xifres de π:

Tenim un feix de rectes paral·leles equidistants a una distància d. Es llança una agulla a l'atzar de grossor menyspreable i longitud L. Llavors la probabilitat que l'agulla talle alguna de les rectes és: 2L/πd.

La ruleta

William Jaggers va arribar a Montecarlo amb uns pocs francs a la butxaca i, durant un mes va anotar els nombres que eixien en cada ruleta, i en quatre dies va guanyar dos milions quatre-cents mil francs. *Jaggers* va aconseguir trencar la banca a *Montecarlo* analitzant les freqüències relatives de cada nombre de la ruleta i observant que s'havia desgastat alguna cosa del mecanisme d'una d'elles, amb la qual cosa tots els valors no tenien la mateixa probabilitat. Va apostar als nombres més probables i va

Luca Pacioli

Luca Pacioli (1445 – 1517), de nom complet **Frai Lucca Bartolomeo de Pacioli** o **Lucca Di Borgo San Sepolcro**, el cognom del qual també apareix escrit com ***Paccioli*** i ***Paciolo*** va ser un frare franciscà i matemàtic italià, precursor del càlcul de probabilitats. Ja hem parlat d'ell en aquestes revistes pels seus treballs sobre la proporció àuria o divina proporció com ell la va anomenar.

Va escriure un llibre amb 36 capítols sobre **comptabilitat** on utilitza la partida doble o taula de doble entrada per a ajudar als Medici a portar la comptabilitat de la seua Banca, defineix les seues regles, com ara no hi ha deutor sense creditor, o que la suma del que es deu ha de ser igual al que s'abona. No va ser el seu inventor, però sí el seu divulgador.

Ducat

El problema enunciat i resolt per Pacioli és aquest:

Dos equips juguen a la pilota de manera que guanya el joc el primer equip que guanya 6 partits. L'aposta és de 22 ducats, que se'ls portarà el guanyador. Per algun motiu cal interrompre el joc quan un equip ha guanyat 5 partits i l'altre 3. Es vol saber com repartir els 22 ducats de l'aposta, d'una manera justa.

Lucca sabia de proporcions, i la solució que va donar hui no es considera vàlida. No sabia probabilitats! Però tu, sí.

Partim de la hipòtesi que cada un dels jugadors té la mateixa probabilitat de guanyar: 1/2. Anomenem A al jugador que ja ha guanyat 5 partides i B al què porta guanyades 3.

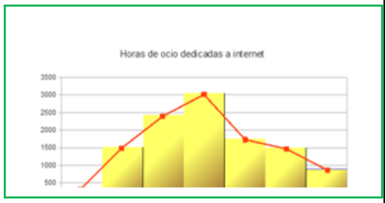
Si feren una nova partida podria guanyar A amb probabilitat 1/2 o B amb la mateixa probabilitat. Si guanya A ja es porta la bossa. Si guanya B llavors B portaria 4 jugades guanyades i A 5. Es continua el joc. Pot guanyar A o B. Observa el diagrama d'arbre.

La probabilitat que guanye B és (1/2)·(1/2)·(1/2)

Com repartiries els 22 ducats?

Estadística i probabilitat. 3r B ESO

RESUM

Població	Col·lectiu sobre el qual es fa l'estudi	Estudiants de tot Madrid
Mostra	Subconjunt de la població que permeta obtindre característiques de la població sencera.	Alumnes es 3º d'ESO seleccionats
Individu	Cada un dels elements de la població o mostra	Joan Pérez
Variables estadístiques	Quantitativa discreta Quantitativa contínua Qualitativa	Nombre de peu que falca Estatura Esport que practica
Gràfics estadístics	Diagrama de barres Histograma de freqüències Polígon de freqüències Diagrama de sectors 	
Mitja	$\bar{x} = \frac{\sum x_i}{n} = (x_1 + x_2 + \ldots + x_n)/n$	Amb les dades: 8, 2, 5, 10 i 10 *Mitja* = 35/5 = 7
Moda	És el valor més freqüent	*Mo* = 10
Mitjana	Deixa per davall la mitat	4 < 6 < **8** < 10 = 10. *Me* = 8.
Rang o recorregut	És la diferència entre la dada major i la dada menor.	10 – 2 = 8
Desviació mitjana	És la mitjana de les distàncies de les dades a la mitja de les dades de què disposem.	(|8–7|+|2–7|+|5–7|+|10–7|+|10–7|)/5 =(1+5+2+3+3)/5 = 14/5 = DM
Variança	És la mitja dels quadrats de les distàncies de les dades a la mitja: $\frac{\sum_{i=1}^{n}(x_i-m)^2}{n} = \sqrt{\frac{\sum_{i=1}^{n} x_i{}^2}{n} - m^2}$	V = (1 + 25 + 4 + 9 + 9)/5 = 47/5 = 9,4
Desviació típica	És l'arrel quadrada de la variança.	$\sigma = \sqrt{47/5} = 3{,}06$
Probabilitat	Valor entre 0 i 1 que ens dóna una mesura de com siga de factible que es verifique un determinat succés.	P(3) = 1/6 en tirar un dau
Espai mostral	El conjunt de tots els casos possibles	{1, 2, 3, 4, 5, 6}
Succés	Subconjunt de l'espai mostral	Traure parell: {2, 4, 6}
Llei de Laplace.	$P(S) = \frac{Nombre de casos favorables al succés}{Nombre de casos possibles}$	*P*(parell) = 3/6 = 1/2.

EXERCICIS I PROBLEMES

Estadística

1. S'han arreplegat les dades sobre el nombre de fills que tenen 20 matrimonis. Com és la variable utilitzada? Escriu una taula de freqüències de les dades arreplegades i representa les dades en un diagrama de sectors:

3, 1, 1, 2, 0, 2, 3, 1, 1, 1, 1, 0, 3, 2, 1, 2, 1, 2, 2, 3.

2. Amb les dades del problema anterior calcula la mitja, la mitjana, la moda i els quartils.

3. Amb les dades del problema anterior calcula el rang, la desviació mitja, la variança, la desviació típica i l'interval interquartílic.

4. Representa aqueixes dades en un diagrama de caixes.

5. La següent taula expressa les estatures, en metres, de 1000 soldats:

Estatura	1,50 - 156	1,56 – 1,62	1,62 - 168	1,68 - 1,74	1,74 - 1,80	1,80-1,92
Nr de soldats	10	140	210	340	210	90

a) Representa les dades en un histograma.

b) Calcula la mitja i la desviació típica.

c) Determina l'interval on es troben la mitjana.

6. Es pregunta a un grup de persones pel nombre de televisors que hi ha en el seu llar i els resultats són:

Nombre de televisors	**0**	**1**	**2**	**3**	**4**	**5**
Nombre de llars	2	27	15	4	2	1

Quin tipus de variables és? Representa les dades a la representació que et parega més adequada.

Calcula la mitja i la desviació típica.

7. Amb les dades del problema anterior calcula la mitjana i l'interval interquartílic.

8. En un centre escolar s'ha arreplegat informació sobre el nombre d'ordinadors a les cases de 100 famílies i s'han obtingut els resultats següents:

Nombre ordinadors	**0**	**1**	**2**	**3**	**4**
Nombre de famílies:	24	60	14	1	1

Representa les dades en un diagrama de barres i calcula la mitja, la mitjana i la moda.

9. Amb les dades del problema anterior calcula el rang, la desviació mitja, la variança i la desviació típica. Fes un diagrama de caixes.

10. Es pregunta a un grup de persones pel nombre de vegades que han visitat el dentista a l'últim any. Les respostes obtingudes s'arrepleguen en la taula següent:

Nombre de visites:	1	2	3	4	5
Nombre de persones:	13	18	7	5	7

Representa les dades en un diagrama de sectors i calcula la mitja, la mitjana i la moda.

11. Es pregunta a un grup de persones pel nombre de vegades que han visitat el dentista a l'últim any. Les respostes obtingudes s'arrepleguen a la taula següent:

Nombre de visites:	1	2	3	4	5
Nombre de persones:	13	18	7	5	7

Calcula el rang, la desviació mitja, la variança i la desviació típica.

12. En les eleccions de 2014 al Parlament Europeu es van obtindre els següents escons per grup parlamentari (DM: demòcrata – cristians; S: socialistes; L: Liberals; V: verds; C: conservadors; I: esquerra unitària; LD: Llibertat i democràcia; NI: No inscrits; Altres).

Partits	DM	S	L	V	C	I	LD	NI	Altres	Total
Escons	213	190	64	52	46	42	38	41	65	751

Què representació de les dades et pareix més adequada? Pots calcular la mitja o el rang? Quin tipus de variable és la de la taula?

13. En les eleccions de 2014 al Parlament Europeu es van obtindre els següents escons per algun dels estats membres:

Estat	Alemanya	Espanya	França	Itàlia	Polònia	Regne Unit	Portugal	Grècia	Altres	Total
Escons	96	54	74	73	51	73	21	21		751

Quina representació de les dades et pareix més adequada? Pots calcular la mitja o el rang? Quin tipus de variable és la de la taula? Determina el nombre d'escons dels altres països membres de la Unió Europea.

14. En les eleccions de 2004, 2009, 2014 al Parlament Europeu es van obtindre els següents percentatges de vots per alguns dels estats membres:

Estat	Alemanya	Espanya	França	Itàlia	Regne Unit	Portugal	Grècia	Bèlgica	% total
2004	43	45’14	42’76	71’72	38’52	38’6	63’22	90’81	45’47
2009	43’27	44’87	40’63	65’05	34’7	36’77	52’61	90’39	43
2014	47’6	45’9	43’5	60	36	34’5	58’2	90	43’09

Quina representació de les dades et pareix més adequada? Pots calcular la mitja o el rang? Quin tipus de variable és la de la taula? Ordena als països de major a menys percentatge de votants a les eleccions de 2014.

15. Amb les dades del problema anterior sobre les eleccions de 2004’ 2009’ 2014 al Parlament Europeu es van obtindre els següents percentatges de vots per alguns dels estats membres:

Estat	Alemanya	Espanya	França	Itàlia	Regne Unit	Portugal	Grècia	Bèlgica	% total
2004	43	45’14	42’76	71’72	38’52	38’6	63’22	90’81	45’47
2009	43’27	44’87	40’63	65’05	34’7	36’77	52’61	90’39	43
2014	47’6	45’9	43’5	60	36	34’5	58’2	90	43’09

Representa en un polígon de freqüències els percentatges de participació del total dels estats membres.

16. Amb les dades del problema anterior sobre les eleccions de 2004, 2009, 2014 al Parlament Europeu es van obtindre els següents percentatges de vots per alguns dels estats membres:

Estat	Alemanya	Espanya	França	Itàlia	Regne Unit	Portugal	Grècia	Bèlgica	% total
2004	43	45’14	42’76	71’72	38’52	38’6	63’22	90’81	45’47
2009	43’27	44’87	40’63	65’0	34’7	36’77	52’61	90’39	43

				5					
2014	47’6	45’9	43’5	60	36	34’5	58’2	90	43’09

Separa els Estats Membres en dos grups, els que van tindre un percentatge superior al percentatge mitjà i els que el van tindre tindre menor en 2004. Fes el mateix per a 2014. Són els mateixos? Analitza el resultat.

17. Amb les dades del problema anterior sobre les eleccions de 2004, 2009, 2014 al Parlament Europeu es van obtindre els següents percentatges de vots per alguns dels estats membres:

Estat	Alemanya	Espanya	França	Itàlia	Regne Unit	Portugal	Grècia	Bèlgica	% total
2004	43	45’14	42’76	71’72	38’52	38’6	63’22	90’81	45’47
2009	43’27	44’87	40’63	65’05	34’7	36’77	52’61	90’39	43
2014	47’6	45’9	43’5	60	36	34’5	58’2	90	43’09

Calcula el percentatge de participació mitjà per a Alemanya en aqueixes tres convocatòries i la desviació típica. El mateix per a Espanya, per a Bèlgica i per a Portugal.

18. En les eleccions de 2014 al Parlament Europeu els resultats d’Espanya han sigut:

Cens	Total de votants	Abstenció	Vots nuls	Vots en blanc
35.379.097	15.920.815	18.810,754	290.189	357.339

Representa en un diagrama de sectors aquestes dades. Fes una taula de percentatges: el cens és el 100 %. Determina els altres percentatges. Consideres que ha guanyat l’abstenció?

19. En les eleccions de 2014 al Parlament Europeu els resultats d’Espanya han sigut:

PP	PSOE	Esquerra plural	Podem	UPiD	Altres	Total de votants
4.074.363	8.001.754	1.562.567	1.245.948	1.015.994		15.920.815

Determina el nombre de vots dels altres partits. Representa en un diagrama de barres aquestes dades. Fes una taula de percentatges per a cada partit. Has de distribuir 54 escons, com els distribuiries per partits?

Probabilitat

20. Es considera l'experiment aleatori de tirar un dau dues vegades. Calcula les probabilitats següents:

a) Traure algun 1.

b) La suma dels dígits és 8.

c) No traure cap 2.

d) Traure algun 1 o bé no traure cap 2.

21. Es considera l'experiment aleatori traure dues cartes de la baralla espanyola. Calcula la probabilitat de:

a) Traure algun rei.

b) Obtindre almenys un basto.

c) No obtindre cap basto.

d) No obtindre el rei de bastos.

e) Traure alguna figura: sota, cavall, rei o as.

f) No traure cap figura.

22. Es considera l'experiment aleatori de tirar una moneda tres vegades. Calcula les probabilitats següents:

a) Traure cara a la primera tirada.

b) Traure cara a la segona tirada.

c) Traure cara a la tercera tirada.

b) Traure alguna cara.

c) No traure cap cara.

d) Traure tres cares.

23. Amb una baralla espanyola es fa l'experiment de traure tres cartes, amb reemplaçament, quina és la probabilitat de traure tres reis? I si l'experiment es fa sense reemplaçament, quina és ara la probabilitat de tindre 3 reis?

24. En una urna hi ha 6 boles blanques i 14 boles negres. Es trauen dues boles amb reemplaçament. Determina la probabilitat que:

a) Les dos siguen negres.

b) Hi haja almenys una negra.

c) Cap siga negra.

25. En una urna hi ha 6 boles blanques i 14 boles negres. Es trauen dues boles sense reemplaçament. Determina la probabilitat que:

a) Les dos siguen negres.
b) Hi haja almenys una negra.
c) Cap siga negra.
d) Compara els resultats amb els de l'activitat anterior.

26. En llançar quatre monedes a l'aire,

a) Quina és la probabilitat que les quatre siguen cares?
b) Quina és la probabilitat d'obtindre com a màxim tres cares?
c) Quina és la probabilitat de tindre exactament 3 cares?

27. Dos tiradors al plat tenen unes marques ja conegudes. El primer encerta amb una probabilitat de 0,7 i el segon de 0,5. Es llança un plat i ambdós disparen. Expressa mitjançant un diagrama d'arbre i les distintes possibilitats: a) Quina probabilitat hi ha de que un dels tiradors done al plat? b) Calcula la probabilitat que cap encerte. c) Calcula la probabilitat que els dos encerten.

28. Es llança una moneda fins que aparega cara dues vegades seguides. a) Calcula la probabilitat que l'experiència acabe al segon llançament. b) Calcula la probabilitat que acabe al tercer llançament.

29. En el llançament de naus espacials s'han instal·lat tres dispositius de seguretat A, B i C. Si falla A es posa automàticament en marxa el dispositiu B, i si falla aquest, es posa en marxa C. Se sap que la probabilitat que falle A és 0,1, la probabilitat que B funcione és 0,98 i la probabilitat que falle C és 0,05. Calcula la probabilitat que tot funcione bé.

30. Es fa un estudi sobre els incendis forestals d'una zona i es comprova que el 40 % són intencionats, el 50 % es deuen a negligències i el 10 % a causes naturals. S'han produït tres incendis, a) quina és la probabilitat que almenys un haja sigut intencionat? b) Probabilitat que els tres incendis es deguen a causes naturals. c) Probabilitat que cap incendi siga per negligències.

31. Es llança dues vegades un dau equilibrat amb sis cares. Trobar la probabilitat que la suma dels valors que apareixen a la cara superior siga múltiple de tres.

32. Se sap que s'han eliminat diverses cartes d'una baralla espanyola que té quaranta. La probabilitat d'extraure un as entre les que queden 0,12, la probabilitat que isca una copa és 0,08 i la probabilitat que no siga ni as ni copa és 0.84.

Calcular la probabilitat que la carta siga l'as de copes. Es pot afirmar que entre les cartes que no s'han eliminat està l'as de copes?

33. Una persona despistada té huit calcetins negres, sis blaus i quatre rojos, tots ells solts. Un dia amb molta pressa, tria dos calcetins a l'atzar. Trobar la probabilitat de:

a) que els calcetins siguen negres.

1) que els dos calcetins siguen del mateix color.
2) que almenys un d'ells siga roig.
3) que un siga negre i l'altre no.

34. Tres persones viatgen en un cotxe. Si se suposa que la probabilitat de nàixer en qualsevol dia de l'any és la mateixa i sabem que cap ha nascut en un any bixest,

a) trobar la probabilitat que només una d'elles celebre el seu aniversari aqueix dia.

➢ calcular la probabilitat que almenys dos complisquen anys aqueix dia.

AUTOAVALUACIÓ

1. Es fa un estudi sobre el color que prefereixen els habitants d'un país per a un cotxe. La variable utilitzada és:

a) quantitativa b) qualitativa c) quantitativa discreta d) quantitativa contínua

2. En un histograma de freqüències l'altura dels rectangles és:

a) proporcional a l'àrea b) igual a la freqüència absoluta

c) proporcional a la freqüència relativa d) proporcional a la freqüència acumulada

3. Anna ha obtingut en Matemàtiques les notes següents: 7, 8, 5, 10, 8, 10, 9 i 7. La seua nota mitja és de:

a) 7,6 b) 8,2 c) 8 d) 9

4. En les notes anteriors d'Anna la mitja és:

a) 9 b) 8 c) 7,5 d) 8,5

5. En les notes anteriors d'Anna la moda és:

a) 10 b) 8 c) 7 d) 7, 8 i 10

6. L'espai mostral de successos elementals equiprobables de l'experiment "llançar dues monedes i comptar el nombre de cares" és:

a) {2C, 1C, 0C} b) {CC, CX, XC, XX} c) {XX, XC, CC} d) {CC, CX, XC, CC}

7. Llancem dos daus i comptem els punts de les cares superiors. La probabilitat que la suma siga 7 és:

a) 1/6 b) 7/36 c) 5/36 d) 3/36

8. En traure una carta d'una baralla espanyola (de 40 cartes), la probabilitat que siga un or o bé un rei és:

a) 14/40 b) 13/40 c) 12/40 d) 15/40

9. En una bossa hi ha 7 boles roges, 2 negres i 1 bola blanca. Es trauen 2 boles. La probabilitat que les dos siguen roges és:

a) 49/100 b) 42/100 c) 49/90 d) 7/15

10. Llancem tres monedes a l'aire. La probabilitat que les tres en caure siguen cares és:

a) 1/5 b) 1/7 c) 1/8 d) 1/6