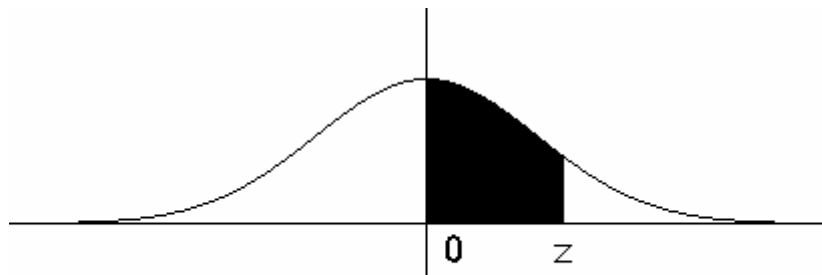


IES Arquitecte Manuel Raspall

Matemàtiques

Estadística

UNA APROXIMACIÓ A LA LLEI NORMAL



Batxillerat

A Recollida de dades

En termes generals podem dir que, amb l'Estadística Descriptiva, hom persegueix dos objectius fonamentals:

- Donar mitjançant un nombre limitat de característiques (nombres) una descripció simple i tan completa com sigui possible d'un conjunt (població) d'elements (individus).
- Interpretar les característiques a fi de treure conclusions pel que fa a les propietats del conjunt estudiat.

Començarem l'estadística fent una petita simulació d'estudi estadístic recollint dades dels alumnes de 1r de Batxillerat de l'institut. Aprofitarem aquestes dades en diversos temes per estudiar i comparar algunes característiques vostres.

Abans de començar és precís que tots plegats utilitzem un vocabulari adient en que el significat de cada paraula estigui molt ben delimitada:

Anomenarem **població** al conjunt dels *individus* sobre el que es vol fer un estudi estadístic (La paraula *individu* no està necessàriament lligada a persones, per exemple si es vol fer un estudi estadístic sobre l'alçada dels castanyers del Montseny, la població seran els castanyers del Montseny)

Moltes vegades és impossible fer un mostreig sobre el total de la població, aleshores s'estudia un subconjunt prou representatiu de la població i s'extreuen conclusions generals de tota la població (amb cert marge d'error) Aquest subconjunt s'anomena **mostra**.

L'aspecte de la població que s'estudia s'anomena **caràcter**.

No tots els caràcters que es poden estudiar requereixen el mateix tractament matemàtic, per exemple l'alçada d'un individu és un caràcter numèric mentre que el color dels ulls no. Tot i així el ventall de possibilitats d'un caràcter sempre l'anomenarem **valor del caràcter** (en l'exemple anterior dels ulls el caràcter pot prendre els valors: blau, verd, marró...). Els diferents tipus de caràcters ens obliga a classificar-los.

Caràcter quantitatiu : És de tipus numèric. És pot classificar en dos tipus:

Discret: Només pot prendre valors concrets (per exemple el nombre de fills)

Continu: Pot prendre qualsevol valor en un interval (per exemple l'alçada d'un castanyer)

Caràcter qualitatiu: És un aspecte de la població al que no es pot assignar un valor numèric (per exemple el color dels ulls) També es classifica en dos tipus:

Ordinal: Es pot ordenar (Per exemple la valoració subjectiva del *sex-apeal* es pot classificar per poc, una mica, bastant o molt)

Nominal: No es pot ordenar (per exemple el color dels ulls)

Una successió de dades, que ens donen els valors del caràcter estudiat, amb les corresponents freqüències absolutes o relatives, l'anomenarem una **distribució estadística**.

Ara, que ja disposem d'un vocabulari mínim, podem començar el nostre estudi sobre alguns caràcters dels alumnes de 1r de Batxillerat de L'institut. Començarem classificant els caràcters:

A.1

Classifica els següents caràcters i digues quins valors poden prendre cada un d'ells:

Sexe

Alçada

Pes

Nº de calçat

Nombre de Germans

Color dels ulls

Quin canal de televisió veus més

¿T'agrada la música màquina?

¿Quina part del cos dels membres del sexe oposat et crida més l'atenció en un primer cop d'ull?

A.2

Analitzeu el següent model d'enquesta per passar a l'alumnat. Comenteu amb els companys i companyes si es fàcil d'entendre i contestar, i si serà fàcil de fer el buidatge de les dades. Modifiqueu el que creieu convenient.

Els alumnes d'estadística 1r de Batxillerat fem aquesta enquesta amb l'únic objectiu d'estudiar les tècniques matemàtiques dels tractaments estadístics de les dades. T'agraïm que la contestis amb tota sinceritat.

1. Sexe: H ☐ D ☐
 2. Alçada:
 3. Pes:
 4. N° de calçat:
 5. Dimensió de la teva llar (¿quanta gent viu a casa teva?):
 6. ¿T'agrada, en general, el teu institut?
Gens ☐ Poc ☐ Regular ☐ bastant ☐ Molt ☐
 7. ¿Quin color d'ulls tens?:
Marró ☐ Blau ☐ Verd ☐ Castany (Marró verdos) ☐
-

BUIDAT DE LES DADES

Sexe	n_i
Home	
Dona	

Alçada	Nois	Noies
menys de 155		
de 155 a 159		
de 160 a 164		
de 165 a 169		
de 170 a 174		
de 175 a 179		
de 180 a 184		
de 185 a 190		
més de 190		

Pes	Nois	Noies
menys de 38		
de 38 a 42		
de 43 a 47		
de 48 a 52		
de 53 a 57		
de 58 a 62		
de 63 a 67		
de 68 a 72		
de 73 a 77		
de 78 a 83		
més de 82		

Institut	n_i
Gens	
Poc	
Regular	
Bastant	
Molt	

Nº de Calçat	Nois	Noies
menys de 33		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
més de 45		

Dim. Llar	n_i
1	
2	
3	
4	
5	
6 o més	

Color Ulls	Fr. Abs
Marrons	
Blaus	
Verds	
Castanys	

B Organització de les dades. Gràfics

D'antuï veurem de quina manera hom pot presentar i disposar les dades que porten la informació sobre la població que volem estudiar, considerada sota un determinat caràcter, en forma de taules i gràfics, per tenir-ne una visualització tan clara i breu com sigui possible.

B.1

El **Cens** és un recull de dades estadístiques de tota mena que es fa sobre el total de la població cada 10 anys i permet conèixer dades exactes de la situació socio-econòmica dels països. L'últim cens correspon a l'any 1991 i d'aquest últim cens s'han recollit les següents dades sobre les dimensions de les llars (nombre de persones en un habitatge)

Nombre de persones	Espanya	Catalunya	Alumnes de 1r de l'institut
1	1585477	262719	
2	2621944	488970	
3	2424360	428596	
4	2801837	451154	
5	1433357	201742	
6 o més	969342	99863	

Font web de l'Institut d'Estadística de Catalunya: <http://idescat.es>

- Utilitzant les dades de la nostra enquesta omple les dades que falten
- ¿Quina és la població d'aquest problema?
- ¿Quin és el caràcter estudiat en aquesta enquesta? ¿quin tipus de caràcter és? ¿Quins valors pot prendre?
- Farem uns gràfics per tal de tenir una descripció més clara de les dades i poder comparar les tres poblacions. Agafa per exemple la primera columna i representa sobre l'eix d'abscisses els diferents valors del caràcter i per a cadascun d'aquests valors dibuixa un segment perpendicular a l'eix de longitud proporcional a la freqüència absoluta. Fes el mateix amb les altres dues columnes.
- ¿Podem comparar aquests gràfics?
- Calcula la **freqüència relativa** de les tres columnes i torna a fer els gràfics anteriors utilitzant freqüències relatives. Aquests tipus de representacions solen anomenar-se

diagrames de barres. ¿Quines conseqüències podem treure de la comparació de les tres poblacions?

- g) En general els resultats de les enquestes venen expressats en percentatges. ¿Veus alguna relació amb el que hem anomenat freqüència relativa i que ja vam estudiar amb les probabilitats?
- h) ¿Quin percentatge de llars de Catalunya hi viuen menys de tres persones? Aquest nombre l'anomenarem **freqüència absoluta acumulada** del valor dos del caràcter i es denota amb la lletra **N**
- i) ¿Quina és la freqüència acumulada del valor dos per a la columna d'Espanya? Compara'l amb el que has obtingut abans.
- j) ¿Creus que aquest caràcter de la població descriu, el nivell socioeconòmic del país? ¿De quina manera?
- k) Imagina que volem comparar la situació de les dimensions de les llars a l'any 1991 amb la situació actual. ¿Seria l'última columna una **mostra** representativa? ¿Per què? ¿com hauriem de triar la mostra perquè sigui representativa?

B.2

Utilitzant totes les dades de la nostra enquesta sobre els/les alumnes de l'institut.

- a) Calcula la freqüència relativa de cada caràcter
- b) Introdueix les dades en un full de càlcul i tria el tipus de gràfic que creguis més adient al tipus de caràcter. Fes una impressió i justifica raonadament el perquè de la teva elecció.

B.3

Ompla la taula següent amb les dades de la nostra enquesta sobre els alumnes de 1r de Batxillerat de l'institut:

Alçada	Nois	Noies
menys de 155		
de 155 a 159		
de 160 a 164		
de 165 a 169		
de 170 a 174		
de 175 a 179		
de 180 a 184		
de 185 a 190		
més de 190		

- a) ¿Quina és la població que s'estudia en aquest problema? ¿Sota quin caràcter? ¿Es quantitatiu o qualitatiu?
- b) ¿Per què et sembla que les alçades s'han agrupat en intervals de 5 en 5 centímetres? Cada interval l'anomenarem classe, i el valor central de la classe marca de classe.

c) Quines són en aquest problema les marques de classe?

Fixa't que en el problema B 1 el caràcter podia prendre els valors 0, 1, 2, 3, 4 ... però no valors entre nombres 1 i 2 per exemple. Es tracta, per tant d'un caràcter quantitatiu **discret**. En canvi en aquest problema l'alçada d'un individu pot prendre qualsevol valor, per això en aquest cas és un caràcter quantitatiu **continu**. No obstant, si un noi té una alçada compresa entre 150 i 151 cm., el que podem suposar és que s'ha agafat la millor aproximació.

d) Si dos individus mesuren 150 cm., podem dir que són exactament de la mateixa alçada?

e) Quins són els extrems reals de les classes de la distribució?

f) Fes la distribució de freqüències acumulades dels membres del teu sexe

g) Representa sobre l'eix d'abscisses les classes i fes correspondre a cada extrem superior d'una classe la freqüència acumulada d'aquesta classe; obtindràs així un conjunt de punts. Uneix-los amb segments rectilinis. Té sentit fer-ho? La figura així obtinguda és el polígon acumulatiu.

h) Utilitza el gràfic anterior per a calcular quants nois/noies tenen una alçada inferior a 168 cm.

B.4

En una fàbrica de cotxes una de les peces del motor d'un determinat model es fabrica automàticament amb una màquina. Per perfecta que sigui aquesta màquina és impossible que totes les peces siguin iguals. Volem saber si és tolerable o no el nombre de peces defectuoses. Les longituds en mil·límetres de cent de les peces fabricades en una setmana són donades per la taula següent:

20.52	.50	.48	.51	.49	.49	.52	.45	.49	.48	.45
.43	.53	.53	.51	.48	.50	.52	.47	.53	.55	.45
.46	.53	.53	.57	.46	.52	.56	.51	.48	.48	.54
.55	.47	.47	.53	.50	.50	.46	.52	.49	.47	.50
.51	.49	.54	.49	.53	.53	.47	.50	.51	.49	.54
.46	.46	.50	.47	.49	.50	.52	.59	.46	.50	.57
.48	.52	.44	.50	.56	.51	.53	.57	.47	.47	.53
.54	.49	.45	.49	.46	.45	.51	.42	.49	.45	.47
.52	.48	.51	.47	.51	.50	.55	.44	.54	.50	.50
.46										

La peça hauria de tenir una longitud de 20.50 mm. Es consideren defectuoses les peces de longitud més petita que 20.45 mm. i més gran que 20.54 mm.

a) ¿Quina és la població que s'estudia i sota quin caràcter? ¿Es discret o continu?

b) Agrupa les dades en classes. Fes-ho de dues maneres diferents, però tenint en compte que una mateixa classe no hi hagi peces bones i defectuoses.

c) Per a cada diferent divisió en classes es un gràfic de la manera següent: representa sobre l'eix d'abscisses les classes (amb segments) i sobre cada classe dibuixa un

rectangle d'alçada igual a la freqüència absoluta de la classe. Fes el mateix per a la freqüència relativa. Aquest gràfic s'anomena **histograma d'alçades**.

- d) Afegeix, en el extrems de l'histograma d'alçades, dues classes de freqüència nul·la i uneix els punts mitjans de les bases superiors dels rectangles amb segments rectilinis. Així s'obté el **polígon de freqüències**.
- e) A partir dels gràfics raona si creus que la màquina te una marge d'error tolerable o no.

B.5

A la taula següent editada per l'I.N.E., veiem la distribució d'analfabetisme a l'Estat espanyol per edat i sexe en els censos dels anys 1981 i 1991. Ens proposem de construir un gràfic que ens doni una visió de la distribució de la població analfabeta per als anys 1981 i 1991.

<u>CENS 1981</u>	Homes		Dones		Total	
Edats	Analfabets	Taxa d'analfabets	Analfabets	Taxa d'analfabets	Analfabets	Taxa d'analfabets
de 10 a 14	10.337	0,6	9.549	0,6	19.886	0,6
de 15 a 19	14.280	0,9	13.097	0,8	27.337	0,8
de 20 a 24	14.652	1	17.417	1,2	32.069	1,1
de 25 a 34	32.884	1,3	53.769	2,2	86.653	1,7
de 35 a 44	54.891	2,6	132.482	6,1	187.373	4,4
de 45 a 54	110.240	4,8	253.645	10,8	363.885	7,9
de 55 a 64	104.050	6,1	264.410	13,7	368.460	10,1
Més de 65	210.483	12,2	695.395	27,7	905.878	21,4
<u>CENS 1991</u>						
de 10 a 14	5.078	0,3	4.397	0,3	9.474	0,3
de 15 a 19	6.988	0,4	6.064	0,4	13.052	0,4
de 20 a 24	8.082	0,5	7.469	0,5	15.551	0,5
de 25 a 34	17.990	0,6	20.102	0,7	38.092	0,6
de 35 a 44	21.746	0,9	35.695	1,5	57.441	1,2
de 45 a 54	36.778	1,8	93.113	4,4	129.891	3,1
de 55 a 64	77.723	3,7	182.919	8,1	260.642	6
Més de 65	124.036	5,6	443.037	14,1	567.073	10,6

- a) Fes un histograma per a la població total d'analfabets de l'any 1991. ¿Per què és enganyós aquest gràfic?
- b) Construeix un histograma en què l'àrea dels rectangles sigui la freqüència relativa. ¿Quina serà l'alçada de cada rectangle? Aquest nou histograma l'anomenarem **histograma d'àrees**. ¿Creus que aquest exemple és justificat que les classes tinguin diferents longituds? ¿Quins avantatges suposa l'histograma d'àrees en front del d'alçades?

C Característiques de posició i de dispersió

En aquest apartat volem calcular uns nombres que ens puguin servir com a característiques descriptives de les distribucions estadístiques. Així en veurem que ens donaran informació sobre el centre de la distribució (**característiques de posició**) i d'altres que en donaran sobre com s'agrupen les dades entorn del centre (**característiques de dispersió**)

C.1

Pretenem fer l'estudi de la distribució de sous en un taller d'arts gràfiques. Els sous mensuals dels treballadors són els següents:

Sous €	Nombre de treballadors
1.070	2
1.220	2
1.700	2
1.910	3
1.940	2
2.060	2
2.180	6
2.300	2
2.390	1
2.900	1
5.000	1

- Si ordenem els treballadors per l'ordre dels diners que guanyen ¿Quants diners guanyaria el treballador que està al mig? Aquest valor hom l'anomena **mediana** de la distribució estadística.
- ¿Quina és la mediana de l'exercici B.3? Observa que pots utilitzar el polígon acumulatiu.
- Si el total dels sous es repartís equitativament entre tot el personal, ¿quan cobraria cadascun? Aquest valor hom l'anomena **mitjana** de la distribució.

- d) Quin és el sou més freqüent en aquesta empresa? Aquest valor hom l'anomena **mode** de la distribució. ¿Quin dels sous obtinguts als apartats a) c) o d) et sembla més representatiu? ¿Per què?

C.2

Una companyia d'assegurances de cotxes vol establir una pòlissa d'accidents a tot risc per a la categoria de cotxes utilitaris. Per això passa una enquesta a 1.000 propietaris d'aquests cotxes, escollits a l'atzar (per sorteig entre matrícules), en la qual es pregunta quants diners han esmerçat durant l'any en reparacions per accident. Les dades obtingudes es resumeixen a la taula següent:

Diners esmentats (€)	Nombre de Propietaris
de 0 a 200	167
de 200 a 400	150
de 400 a 600	145
de 600 a 800	131
de 800 a 1.000	106
de 1.000 a 1.200	93
de 1.200 a 1.600	85
de 1.600 a 2.000	57
de 2.000 a 2.400	20
de 2.400 a 2.800	24
de 2.800 a 3.200	12
de 3.200 a 3.600	6
de 3.600 a 4.400	4

- a) ¿Quina és la *classe modal* d'aquesta distribució?
- b) ¿Quina és la *mediana*? ¿Quin significat té?
- c) ¿Quin preu ha de tenir la pòlissa d'accidents a tot risc si no es vol ni guanyar ni perdre diners? ¿Et sembla que aquest valor té alguna relació amb allò que abans hem anomenat *mitjana*?
- d) Fes l'*histograma* de la distribució.
- e) Si el preu de la pòlissa és l'obtingut a c), ¿pots preveure algun benefici per a l'empresa si s'aplica una **franquícia** de 400 €? Calcula els beneficis mitjans per pòlissa, que obtindrà la companyia si apliqués aquest sistema de franquícia. (S'anomena **franquícia** a una quantitat de diners de la factura que no paga la companyia sinó el client)
- f) Si no s'apliqués cap franquícia, ¿quin haurà d'ésser el preu de la pòlissa si la companyia vol obtenir els mateixos beneficis que aplicant-la?

C.3

Fins ara hem anat calculant les mesures de dispersió d'una manera pràctica a partir d'exemples. Ja és hora de que donem unes definicions el més correctes possibles d'aquests conceptes. Considera la distribució:

Dades (x_i)	n_i
x_1	n_1
x_2	n_2
.	.
.	.
.	.
x_k	n_k

a) ¿Sabries escriure'n la **mitjana** $\bar{x}$?

Si les dades són agrupades en classes, representarem aquestes a efectes del càlcul de la mitjana, per la marca de classe. Es a dir, considerarem tota la freqüència concentrada sobre la marca de la classes.

b) ¿Sabries explicar què entens per **mode**, o **classe modal**?

c) ¿Què entens per **mediana**? ¿Com la calcularies en general? (et cal explicar, per una banda com es calcula la mediana en una distribució discreta amb nombre parell o imparell de dades i per altra banda com es calcula en una distribució contínua a partir del polígon acumulatiu)

C.4

Les notes de Matemàtiques de dos grups de 40 alumnes són donades per la taula següent:

Notes	Grup I	Grup II
0	0	0
1	0	0
2	2	0
3	3	0
4	5	0
5	7	12
6	8	24
7	7	4
8	5	0
9	3	0
10	0	0
Total:	40	40

a) Troba la *mitjana* per cada grup i fes els diagrames de barres corresponents.

Veuràs que tot i que tenen la mateixa mitjana, les dues distribucions són ben diferents. Les notes del grup II estan molt més agrupades entorn de la mitjana que no les del grup I. Mentre al grup II no hi ha suspensos, al grup I n'hi ha 10, és dir, un 25 per cent.

En molts casos no n'hi ha prou de donar un valor mitjà d'una distribució i cal estudiar com es distribueixen les dades al seu entorn.

Anomenem **Rang** d'una distribució estadística a la diferència entre el valor més gran i el valor més petit que agafa el caràcter estudiat.

- b) Quin és el *rang* de les dues distribucions?
- c) ¿Et sembla que dona bon compte de la dispersió de les dues distribucions? En general, ¿et sembla que el rang és un valor representatiu de la dispersió d'una distribució?
- d) Intentarem construir una mesura de la dispersió que tingui en compte totes les dades. Agafa el grup I, calcula quant es desvia cada dada respecte a la mitjana i troba el valor mitjà d'aquestes desviacions. Fes el mateix per al grup II.
- e) Et sembla útil aquest valor? Per què?
- f) Agafa el grup I, calcula el valor absolut de la desviació de cada dada respecte a la mitjana i troba el valor mitjà d'aquestes desviacions. Fes el mateix per el grup II. Aquest valor hom l'anomena **desviació mitjana D.M.**

L'anorèxia i la Bulímia són malalties psicològiques molt greus fruit d'una incapacitat de reconèixer el propi pes com a normal (agreujat per una exaltació social d'una bellesa prima)

Determinar si un pes és o no normal des d'un punt de vista mèdic és un problema força complex, doncs hi intervenen molts factors: L'edat, l'alçada, la constitució física, psíquica i endocrinòloga,... Les matemàtiques contribueixen donant eines als especialistes amb les que puguin disposar de prou paràmetres per a fer diagnòstics fiables. En aquest tema i durant el curs estudiarem algunes d'aquestes eines.

Per començar els matemàtics redefinim la paraula **normal** donant-li un significat purament matemàtic que anirem descobrint en els exercicis següents. Respecte al pes, que un individu sigui o no *normal* des d'un punt de vista matemàtic no vol dir que ho sigui des d'un punt de vista mèdic, això tan sols ho podrà determinar un endocrinòleg.

Un primer pas en la recerca d'aquests conceptes és analitzar si un individu s'allunya poc o molt de la mitjana, per tant cal quantificar aquesta desviació.

C.5 Distribució del pes dels alumnes de 1r de Batxillerat.

Utilitzant les dades dels membres del teu mateix sexe de l'enquesta dels alumnes de 1r de Batxillerat omple la taula següent seguint les indicacions dels diferents apartats.

Distribució del pes dels alumnes de 1r de Batxillerat

Pes (Kg)	marca x_i	n_i	$x_i \cdot n_i$	f_i	l_i (Kg)	h_i (f_i / l_i)	δ_i	$\delta_i \cdot n_i$	$ \delta_i $	$ \delta_i \cdot n_i$	δ_i^2	$\delta_i^2 \cdot n_i$
[32.5, 37.5)												
[37.5, 42.5)												
[42.5, 47.5)												
[47.5, 52.5)												
[52.5, 57.5)												
[57.5, 62.5)												
[62.5, 67.5)												
[67.5, 72.5)												
[72.5, 77.5)												
[77.5, 82.5)												
[82.5, 87.5)												
Total:												

- Calcula la mitjana $\bar{x}$ de la distribució. Utilitzeu la columna $x_i \cdot n_i$
- ¿Quant es desvia cada dada respecte a la mitjana? Escriu-ho a la columna δ_i
($\delta_i = x_i - \bar{x}_i$)
- ¿Quina és la mitjana d'aquestes desviacions? Pots fer servir la columna $\delta_i \cdot n_i$.
- Utilitza les columnes $|\delta_i|$ i $|\delta_i| \cdot n_i$ per a trobar la desviació mitjana **D.M.** de la distribució.
- Una altra manera d'evitar la diferència de signes en les desviacions i poder així obtenir un altre valor útil, seria calcular els quadrats de les desviacions per cada dada. Escriu-ho a la taula en la columna δ_i^2 .
- Calcula la mitjana dels quadrats de les desviacions que has calculat a d) utilitzant la columna $\delta_i^2 \cdot n_i$. Aquest valor hom l'anomena **variància**. ¿Et sembla útil aquest valor? ¿Presenta alguna dificultat per utilitzar-lo com a mesura de la separació normal? Per respondre a aquesta pregunta tingues en compte quines són les unitats que té (has elevat els pesos al quadrat).
- L'arrel quadrada de la variància l'anomenarem **desviació tipus** i la representarem per σ . (Per tal de facilitar la notació a la es denota la variància amb el símbol σ^2). Calcula la desviació tipus. ¿Et sembla ara útil aquest valor? ¿Coincideix amb la desviació mitjana?

- h) Fes el diagrama d'àrees. Utilitza la columna f_i per a la freqüència relativa, la columna l_i per a la longitud o amplada de l'interval i la h per l'altura dels rectangles ($h_i = \frac{f_i}{l_i}$)
- i) Calcula, utilitzant el diagrama d'àrees, quantes noies/nois de la distribució tenen pesos compresos dins l'interval $(\bar{x} - \sigma, \bar{x} + \sigma)$. ¿Quin percentatge representa?
- j) Fes el mateix per a l'interval $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$.

C.6

Els mateixos grups d'alumnes del problema B.4 han tret, a Ciències Naturals, les notes següents:

Nota	Grup I	Grup II
0	0	0
1	2	0
2	1	0
3	2	0
4	3	12
5	5	7
6	10	0
7	8	11
8	5	10
9	3	0
10	1	0

- a) Troba la *mitjana* i la *desviació mitjana* de cadascuna de les distribucions.
- b) ¿Et sembla que la desviació mitjana és una bona mesura de la dispersió d'una distribució?
- c) Troba i compara les *desviacions tipus*.
- d) ¿Et sembla que la desviació tipus expressa millor la dispersió d'una distribució? Raona la resposta.

C.7

A la distribució estadística donada a la taula següent:

Dades: x_i	Fr. Ab: n_i	δ_i	$\delta_i \cdot n_i$	$ \delta_i $	$ \delta_i \cdot n_i$	δ_i^2	$\delta_i^2 \cdot n_i$
x_1	n_1						
x_2	n_2						
.	.						
.	.						
.	.						
x_k	n_k						
Suma:	N						

a) Completa la columna δ_i de les desviacions. ¿ Són tots positius els nombres d'aquesta columna?

b) Comprova per aquesta distribució que la mitjana de les desviacions respecte a la mitjana val zero:

$$\bar{\delta} = \overline{(x_i - \bar{x})} = 0$$

c) Completa la columna dels valors absoluts de les desviacions respecte a la mitjana $|\delta_i|$ i calcula la fórmula de la desviació mitjana:

$$\text{D.M.} =$$

d) Completa la columna dels quadrats de les desviacions respecte a la mitjana δ^2 i calcula la fórmula de la variància:

$$\sigma^2 =$$

e) Quins inconvenients presenta la variància pel que fa a les seves unitats? Es pot comparar directament amb la mitjana?

f) Escriu la fórmula de la desviació tipus:

$$\sigma =$$

g) Creus que la desviació tipus té algun avantatge enfront de la desviació mitjana?

C.8

L'únic inconvenient que te, fins ara la desviació mitjana és que és força complicada de calcular. Si aconseguim simplificar la fórmula eliminarem aquest únic inconvenient.

Demosta que la desviació tipus d'una distribució és l'arrel quadrada de la mitjana dels quadrats de les dades menys el quadrat de la mitjana, és a dir:

$$\sigma = \sqrt{\frac{\sum_{i=1}^k n_i \cdot x_i^2}{N} - \bar{x}^2} = \sqrt{\overline{x^2} - \bar{x}^2}$$

C.9

Observa que a partir d'ara totes les taules que farem es reduiran a la taula següent:

Dades: x_i	Fr. Abs: n_i	$x_i \cdot n_i$	x_i^2	$x_i^2 \cdot n_i$
x_1	n_1			
x_2	n_2			
.	.			
.	.			
.	.			
x_k	n_k			
Suma:	N			

A partir d'aquesta taula es calcula ràpidament la mitjana i la desviació tipus.

Fes en un full de càlcul amb ordinador una taula com aquesta que et calculi automàticament la mitjana i la desviació tipus.

C.10

Una eina fonamental és la **calculadora científica** que disposa d'una funció *MODE SD* o *STAT...* que adapta a la calculadora a fer càlculs estadístics introduint les dades amb la tecla *data* o *dt* o *D*,... i calcula:

Paràmetre estadístic	Tecla
Quantitat de dades: N	n
Suma de les dades $x_i \cdot n_i$	Σx
Suma dels quadrats de les dades $x_i^2 \cdot n_i$	Σx^2
Mitjana	$\bar{x}$
Desviació tipus	$x\sigma_n$

Cada model de calculadora te una manera diferent de treballar els paràmetres estadístics. Investiga com ho fa la teva calculadora.

D La distribució normal

D.1

Una empresa fabrica malles tubulars de plàstic de 10 fils. Al laboratori de control de qualitat es mesura, fil per fil, la resistència a la tracció d'aquesta malla. La prova es fa amb un dinamòmetre fins que el fil es trenca. Cada setmana s'agafa una mostra de la producció i es controla. El resultat al cap d'un any és el següent:

Resistència (Kg.)	Freqüència
[7, 8)	1
[8, 9)	7
[9, 10)	49
[10, 11)	230
[11, 12)	581
[12, 13)	645
[13, 14)	381
[14, 15)	120
[15, 16)	9

- Representa aquesta distribució gràficament i troba la resistència mitjana i la desviació tipus.
- ¿Quin percentatge de fils tenen una resistència a la tracció que es desvia de la mitjana menys d'una desviació tipus? ¿i menys de dues?

D.2 Un model experimental

Observa que al fer dos problemes radicalment diferents el C- 5 sobre pesos i el D-1 sobre resistències de fils, els resultats han estat molt semblants (el percentatge d'individus que es troben tancats entre una desviació tipus es troba al voltant del 64% i entre dues desviacions tipus al voltant del 95 %) Això ens indueix a pensar que és possible que hi hagi un determinat tipus de problemes que produeixin aquests percentatges pràcticament idèntics.

Si això és així seria molt interessant calcular quins són exactament aquests percentatges i així els podríem aplicar a aquests problemes directament sense fer càlculs.

Ens hem d'inventar, per tant, un **model experimental** fiable que ens proporcioni aquests percentatges. La millor manera de que el model sigui fiable és generar un experiment totalment aleatori que no pugui estar influenciat per cap factor extern: Llançarem 50 monedes 10000 vegades i contarem quantes cares surten cada vegada.

Aquest model és totalment aleatori i és fruit de moltes repeticions, per tant serà un model molt fiable però té un greu inconvenient: la variable no és contínua. Com que estem d'acord que si la variable fos contínua el model seria molt bo farem *trampa*: **suposarem que la variable és contínua** (es a dir per exemple el resultat 32 cares serà en realitat resultat entre 31,5 i 32,5 cares)

Per estar totalment segurs que els percentatges que surten són fruit d'un model experimental fiable i no de l'atzar caldria repetir el model experimental varies vegades. Per tant cadascú de vosaltres farà un **experiment individual** (o simularà el joc amb un ordinador). Com a exemple, el resultat de l'experiment haurà de tenir un aspecte, més o menys com el següent.

JOC ALEATORI: TIRAR 50 MONEDES 10000 VEGADES

<u>CARES</u>	<u>VEGADES</u>
11	0
12	1
13	3
14	9
15	23
16	51
17	95
18	152
19	262
20	411
21	591
22	786
23	933
24	1101
25	1195
26	1062
27	941
28	793
29	574
30	416
31	283
32	159
33	91
34	39
35	19
36	4
37	5
38	0
39	1

10000 TIRADES EN TOTAL

$N = 10000$; MITJANA = 24,992; DESVIACIÓ TIPUS = 3,5288434

- Representa gràficament la distribució. Fes-ho amb un **diagrama d'àrees** (recorda que em acordat imaginar que es tractava d'un **caràcter continu**)
- ¿Quin percentatge dels resultats es separa de la mitjana menys d'una desviació tipus? ¿i de dues? Compara aquests resultats amb els dels problemes anteriors i amb els dels teus companys.

D.3

Després de fer els problemes anteriors ja hem pogut observar que hi ha uns tipus de distribucions estadístiques en les que la majoria d'individus es troben al voltant de la mitjana doncs els percentatges d'individus entre una i dues desviacions tipus són molt semblants als del nostre model experimental. Aquestes distribucions les anomenaren **normals**

Per saber si una distribució és o no *normal* caldrà comprovar si els percentatges entre una i dues desviacions s'aproxima al del nostre model experimental.

- a) Comprova si el nombre de calçat dels/les alumnes de 1r de Batxillerat de l'institut es distribueixen *normalment* (ompli la taula amb les dades del vostre sexe de la nostra enquesta)

Nº de Calçat	n _i
menys de 33	
33,0000	
34,0000	
35,0000	
36,0000	
37,0000	
38,0000	
39,0000	
40,0000	
41,0000	
42,0000	
43,0000	
44,0000	
45,0000	
més de 45	

- b) ¿Creus que els alumnes de 1r de Batxillerat de l'institut representen una bona mostra de la població juvenil pel que fa al nombre de calçat?
- c) Un venedor de sabates de Cardedeu considera que representen una bona mostra pel que fa al nombre de calçat i vol encarregar 100 sabates d'un determinat model, quants n'han d'agafar de cada nombre?
- d) Un fabricant de mitjons que també us considera una bona mostra vol saber quin percentatge de mitjons ha de fabricar de cada una de les talles:

Mitjons de dona	
talla 1	de 33 a 35
talla 2	de 36 a 38
talla 3	de 39 a 41

Mitjons d'home	
talla 1	de 37 a 49
talla 2	de 40 a 42
talla 3	de 43 a 45

(cada u fa els que corresponen al seu sexe)

D.4

S'ha fet un estudi de la llargària de la producció de rajoles d'una fàbrica que pressuposem es distribueixen d'una manera *normal* (per tant pots utilitzar resultats obtinguts en el **teu model**

experimental) S'ha obtingut que la mitjana és de 30,12 cm. i la desviació tipus és de 0,15 cm. Considerem defectuoses les peces que es desvien més de dues desviacions tipus de la mitjana, ¿entre quines llargàries estan les rajoles acceptables? ¿Quin percentatge aproximat representen aquestes rajoles sobre el total?

D.5

En un test fet a uns estudiants, la mitjana de puntuació és de 500 i la desviació típica 100 punts. Si la distribució estadística obtinguda segueix la **lleï normal** calcula, utilitzant els resultats i el gràfic del teu model experimental, el percentatge aproximat d'estudiants que han tret les puntuacions següents:

- a) entre 500 i 600
- b) entre 400 i 600
- c) entre 600 i 700
- d) entre 300 i 400
- e) més de 700
- f) entre 500 i 550

D.6 : Model teòric de la distribució normal

Ja hem comentat que el concepte matemàtic de **normal** es correspon amb una distribució estadística en la que la freqüència relativa de les dades que es desvien una i dues desviacions típiques és aproximadament igual.

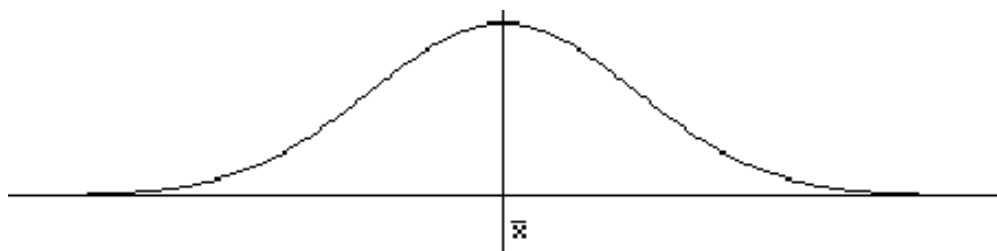
- a) Anota els valors obtinguts en els diferents problemes:

valors entre	C5	D1	D2	D3
$\bar{x} - \sigma$ i $\bar{x} + \sigma$				
$\bar{x} - 2\sigma$ i $\bar{x} + 2\sigma$				

Hem trobat un bon model experimental que ens donava aquests percentatges d'una manera força correcta, però cada u tenia un valor lleugerament diferent. Ens interessaria aconseguir un **model teòric** precís i vàlid per a tots.

Pensa que els percentatges sempre els hem tret del histograma d'àrees, per tant ens

interessa buscar una funció matemàtica que tingui un gràfic que s'ajusti al màxim a l'histograma propi d'una distribució normal:



- b) Enumera totes les propietats que creus que ha de tenir una funció matemàtica d'aquest tipus

La funció $y = e^{-x^2}$ té aquesta forma de *campana* però no compleix totes les propietats que ha de complir la funció que busquem. Cal modificar-la. La funció que surt s'anomena funció de Laplace-Gauss (la forma d'aquesta funció s'anomena **campana de Gauss**)

$$y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

Una vegada coneguda la mitjana i la desviació tipus d'una distribució normal es pot substituir a la fórmula, dibuixar el gràfic corresponent i calcular el percentatge d'individus que hi ha entre dos valors qualsevol calculant l'àrea que tanca el gràfic.

La complexitat d'aquest procés és força elevada. Per això hi ha hagut algú que s'ha entretingut en calcular aquests percentatges per a una distribució concreta i els ha escrit en una taula. Aquesta distribució concreta, anomenada **distribució normal típica**, és una en que la mitjana és $\bar{z} = 0$ i la desviació tipus $\sigma = 1$.

La notació que utilitzarem per a fer càlculs amb la normal es la següent:

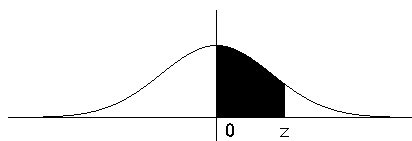
- En general una distribució normal de mitjana $\bar{x}$ i desviació tipus s es denota amb $N(\bar{x}, \sigma)$
- Per escriure la probabilitat o percentatge d'individus x que es troben entre dos valors, per exemple 2 i 4 escriurem: $P(2 \leq x \leq 4)$.

Si es tracta de la distribució típica utilitzarem la lletra **Z** en comptes de la lletra **X**.

D.7

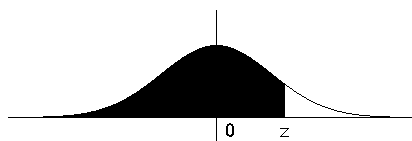
- a) En una distribució normal típica, calcula, utilitzant la taula, els percentatges entre els valors següents: $P(0 \leq z \leq 1)$; $P(-1.3 \leq z \leq 0)$; $P(0 \leq z \leq 2.12)$
- b) En una distribució normal típica, calcula, utilitzant la taula, la probabilitat que un valor z estigui entre els valors següents: $P(-1.52 \leq z \leq 1.52)$; $P(1.94 \leq z)$
- c) Troba a la taula de la llei normal típica la freqüència relativa de les dades que es desvien una i dues desviacions típiques. Compara el resultat amb els valors obtinguts als problemes anteriors (sobre tot al teu model experimental D.2)

Taula normal tipificada 1:



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0039	0,0079	0,0119	0,0159	0,0199	0,0239	0,0279	0,0318	0,0358
0,1	0,0398	0,0438	0,0477	0,0517	0,0556	0,0596	0,0635	0,0674	0,0714	0,0753
0,2	0,0792	0,0831	0,0870	0,0909	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998

Taula normal tipificada 2:



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998

Tipificació de la variable

Evidentment aquesta taula és, en principi, del tot insuficient doncs és molt excepcional trobar-nos amb una distribució de mitjana 0 i desviació tipus 1. Però amb un senzill **canvi de variable** es pot transformar qualsevol distribució en una distribució típica. Aquest procés s'anomena **tipificació de la variable** i la fórmula de tipificació és:

$$Z = \frac{X - \bar{X}}{\sigma}$$

Al restar la mitjana centrem la nova variable que tindrà mitjana 0. Al dividir per σ (desviació tipus) fem que la nova variable tingui desviació tipus 1.

D.8

Considera la distribució de qualificacions de l'exercici D.5. La mitjana de les puntuacions era de 500 punts i la desviació tipus 100 punts. Omple la taula següent on podràs veure el significat de tipificar la variable:

Variable original	Variable centrada	Variable tipificada
X	$X - \bar{X}$	$Z = \frac{X - \bar{X}}{\sigma}$
500		
600		
650		
400		
300		
530		
		-1,5
		3
		0,7
		2,15

D.9

Escriu la fórmula per passar de la variable tipificada z a la variable original x .

D.10

Tipifica la variable per tal de calcular en una distribució normal de mitjana $\bar{x} = 25$ i desviació tipus $\sigma = 2$ la probabilitat que un valor estigui entre 26 i 28, es a dir $P(26 \leq x \leq 28)$

D.11

Torna a fer el problema D.5 a partir de les taules de la distribució normal. A més, contesta les següents qüestions:

- a) ¿Quin percentatge d'estudiants ha tret més de 315 punts?
- b) ¿Quin percentatge n'ha tret entre 415 i 640?

La taula de la normal també es pot llegir a l'inrevés, és a dir a partir d'un percentatge es pot trobar el valor z , després aplicant la fórmula del canvi de variable a la inversa trobem el valor x .

D.12

Resol la qüestió següent referent també al problema D.5: ¿Quina és la mínima puntuació que s'ha de treure per estar entre el 10% dels millors tests?

D.13

Per a tenir un ús àgil de la taula resol, ajudant-te de dibuixos un exemple de les 9 combinacions diferents amb les que et pots trobar:

$P(0 \leq z \leq 0,32);$	$P(-0,32 \leq z \leq 0);$	$P(-1,21 \leq z \leq 0,32);$
$P(0,32 \leq z \leq 1,21);$	$P(-1,21 \leq z \leq -0,32);$	$P(1,21 \leq z);$
$P(z \leq 1,21);$	$P(z \leq -1,21);$	$P(-1,21 \leq z)$

D.14

Els camions que passen per certa autopista tenen una alçada, en metres, que es distribueix $N(3.5; 0.6)$. Es construeix un pont sobre l'autopista amb una alçada de 4.40 m. ¿Quin percentatge de camions no podran passar per l'autopista?

D.15

Una ciutat instal·la per a la seva il·luminació 20 000 bombetes. La durada d'una bombeta segueix una distribució normal de mitjana 392 dies i desviació tipus 40 dies. Si l'empresa que les ha instal·lat es compromet a canviar el cap d'un any totes les bombetes foses.

- a) Quantes bombetes necessitarà aproximadament?
- b) Quantes bombetes podem esperar que s'hagin fos just el dia abans de canviar-les, o sigui just un dia abans de complir-se un any de la instal·lació?

D.16

Una fàbrica fa unes peces de precisió que passen per un control de qualitat. La mitjana de les mides és de 105,24 mm. i la desviació tipus és de 0,21 mm. ¿Entre quines mides al voltant de la mitjana hi haurà el 50% de les peces?. ¿I el 80%?. Si es consideren inservibles les peces que facin més de 105,65 o menys de 104,90; ¿quin tan per cent de peces haurem de refusar?

D.17

Amb un milió de llumins s'han omplert 2.000 caixes. La màquina que ho fa no és massa precisa, de manera que després s'ha fet un recompte en una mostra representativa i s'ha comprovat que hi ha un 20% de caixes que tenen menys de 490 llumins. Suposant que la distribució sigui normal troba la mitjana i la desviació tipus del nombre de llumins que hi ha a les caixes.

D.18

El nivell de glucosa a la sang dels homes adults es distribueix normalment amb mitjana 86 mg/100ml. amb una desviació tipus de 9 mg/100ml. Quan la concentració de glucosa a la sang d'un individu supera els 110 mg/100ml. es considera que és diabètic.

- a) ¿Quin percentatge de entre els homes adults és diabètic?
- b) Si triem un home adult al atzar. ¿Quina probabilitat hi ha de que el seu nivell de glucosa estigui entre 77 i 88 mg/100ml?
- c) Una baixa concentració de glucosa a la sang produeix anèmia. Sabem que el 10 % dels membres de la població són anèmics. ¿Quin nivell de glucosa a la sang és el que determina quines concentracions inferiors produeixen anèmia?

D.19

Tornem a la qüestió plantejada al problema C.5. Per fer el seguiment del creixement normal dels seus pacients, els pediatres disposen d'una gràfica on s'indica l'alçada i pes mitjans dels joves així com la seva distribució en percentatges. Les noies contesteu a les preguntes que fan referència a les noies utilitzant la taula de les noies, i els nois amb la dels nois

- a) Interpreta els gràfics i dóna una explicació de com es poden utilitzar.

Si ordenem els valors del caràcter d'una distribució estadística contínua, al valor que te un N% per sota es diu que te **centil·la N**.

- b) Fes una anàlisi clínic respecte el pes i l'alçada de cada una de les 4 noies/nois següents: