

2. Puntuacions típiques i corba normal

Hem vist que les característiques importants d'una distribució de valors són la seva localització, representada normalment per la mitjana, i la seva dispersió, per a la qual se sol utilitzar la desviació típica. Amb aquesta informació podem deduir el valor relatiu d'una dada concreta respecte al grup al qual pertany. Però ens podem trobar en situacions com la següent.

Imaginem-nos que un estudiant ha fet quatre exàmens en quatre assignatures diferents i ha obtingut les puntuacions següents:

| | |
|------------------|-----|
| Matemàtiques | 6 |
| Llengua catalana | 8 |
| Anglès | 7,5 |
| Ciències socials | 6,5 |

Aparentment, sembla que la millor qualificació de l'estudiant és la de llengua. El que passa és que les puntuacions directes (s'anomenen així les puntuacions obtingudes directament a través de l'aplicació d'algun tipus de prova) difícilment són comparables entre si. Pot passar, per exemple, que l'examen de llengua catalana hagi estat molt més fàcil que els altres, de manera que tot el grup ha obtingut puntuacions elevades. O bé que l'examen d'anglès hagi puntuat sobre 20 punts i el de matemàtiques sobre 6. Aquestes qualificacions proporcionen informació sobre el valor directe obtingut, però no ens permeten saber fins a quin punt la puntuació és bona, ni si és bona en comparació de les puntuacions dels altres.

Ara suposem que, a més de la puntuació directa, coneixem la mitjana:

| | X | Mitjana |
|------------------|----------|----------------|
| Matemàtiques | 6 | 4 |
| Llengua catalana | 8 | 8,5 |
| Anglès | 7,5 | 6,5 |
| Ciències socials | 5,5 | 7 |

Mirant les mitjanes ja podem veure que hi ha unes notes que estan per sota de la mitjana i unes altres que estan per sobre. En fer les diferències de cada valor respecte a la mitjana ens surt que:

| | X | Mitjana | X - Mitjana |
|------------------|----------|----------------|--------------------|
| Matemàtiques | 6 | 4 | 2 |
| Llengua catalana | 8 | 8,5 | -0,5 |
| Anglès | 7,5 | 6,5 | 1 |
| Ciències socials | 5,5 | 7 | -1,5 |

Per tant, les notes de llengua catalana i les de ciències socials no seran gaire bones ja que se situen per sota de la mitjana obtinguda pel grup. El resultat més dolent de l'estudiant ha estat en ciències socials, ja que la puntuació de 5,5 es troba a una

distància d'1,5 punts per sota de la mitjana, i la puntuació de llengua catalana es troba mig punt per sota de la mitjana.

Tant les qualificacions de matemàtiques com les d'anglès estan per sobre de la mitjana. Podríem pensar que la millor qualificació és la de matemàtiques, que està dos punts per sobre de la mitjana, mentre que la d'anglès està només un punt per sobre. No obstant això, també hem de tenir en compte la dispersió de la distribució. Per tant, hem de mirar les desviacions típiques. Suposem que són les següents:

| | X | Mitjana | X - Mitjana | S |
|---------------------|------------|----------------|--------------------|------------|
| Matemàtiques | 6 | 4 | 2 | 0,5 |
| Llengua catalana | 8 | 8,5 | -0,5 | 1 |
| Anglès | 7,5 | 6,5 | 1 | 1 |
| Ciències socials | 5,5 | 7 | -1,5 | 1,5 |

Si ens fixem en les desviacions típiques, veiem que la dispersió és més elevada en anglès (1) que en matemàtiques. La desviació típica ens mostra que la dispersió en les qualificacions d'anglès és d'un punt respecte de la mitjana, la qual cosa vol dir que hi va haver puntuacions que van superar la mitjana en més d'un punt i altres puntuacions que estaven més d'un punt per sota de la mitjana. La puntuació de 7,5 es troba un punt exacte per sobre de la mitjana, la qual cosa indica que es troba a una distància d'una desviació típica respecte de la mitjana. La desviació típica de la prova de matemàtiques és 0,5. Igualment, això vol dir que la puntuació de 6 es troba a una distància de quatre desviacions típiques per sobre de la mitjana. Dit d'una altra manera, en la prova de matemàtiques, les puntuacions obtingudes estan molt més concentrades al voltant de la mitjana que en la prova d'anglès. Mirant els valors de les desviacions típiques respectives veiem que per a l'estudiant ha estat molt més difícil tenir una puntuació de 6 en matemàtiques que una puntuació de 7,5 en anglès.

En conclusió, malgrat que una primera mirada de les puntuacions ens feia pensar que la millor qualificació era la de llengua catalana, després d'aquesta anàlisi resulta que la millor puntuació en termes relatius ha estat la de matemàtiques, que era la més baixa en valor absolut.

El que s'ha fet, doncs, és transformar les puntuacions directes en unes altres puntuacions mesurades amb el mateix tipus d'escala, de manera que puguem comparar les puntuacions (puntuacions típiques).

2.1. Les puntuacions típiques o puntuacions Z

El procediment habitual per transformar puntuacions directes en típiques és convertir-les en unes puntuacions en una escala de valors que tingui de mitjana 0 i de desviació típica 1. D'aquesta manera, si la puntuació resultant és positiva, podem dir que està per sobre de la mitjana (ja que estarà per sobre de 0), i si la puntuació és negativa, podem afirmar que el valor està per sota de la mitjana. A més, el valor resultant ens indicarà a quantes desviacions típiques per sobre o per sota de la mitjana es troba el valor que estem transformant. La fórmula que s'utilitza és la següent:

$$Z = \frac{X - \bar{X}}{S} \quad (7)$$

Exercici

Imaginem-nos que tenim les notes del primer quadrimestre de l'assignatura de Bases de dos cursos diferents, les del curs 2003-2004 i les del curs 2004-2005. En el curs 2003-2004, la mitjana de les qualificacions va ser de 5,5 i la desviació típica, de 2,2. En el curs 2004-2005 la mitjana aritmètica va ser de 6 i la desviació típica, d'1,6.

Un estudiant que obtingués una nota de 7,5, a quin dels dos cursos representaria que ha aconseguit una millor qualificació?

$$Z_1 = \frac{X_1 - \bar{X}_1}{S_1} = \frac{7,5 - 5,5}{2,2} = 0,91 \quad (\text{Curs 2003-2004})$$

$$Z_2 = \frac{X_2 - \bar{X}_2}{S_2} = \frac{7,5 - 6}{1,6} = 0,93 \quad (\text{Curs 2004-2005})$$

Observem, doncs, que en termes relatius l'estudiant que hagi obtingut una qualificació de 7,5 l'any 2004-2005 ha tret una millor nota que el que ha aconseguit la mateixa puntuació l'any 2003-2004, tot i que la diferència és mínima.

2.2. Exercicis sobre puntuacions típiques

- Un mateix alumne obté en els exàmens d'Estadística i d'Avaluació de programes un 6,3 i un 7,1 respectivament. Si la mitjana d'Estadística de tota la classe és de 4,6 amb desviació típica d'1,2 i la mitjana d'Avaluació de programes és de 6,9 amb desviació típica de 0,8, calcula en quina assignatura aquest alumne ha tret comparativament una puntuació més alta.
- En una prova de domini de la lectura aplicada a 800 alumnes de 3r curs d'educació primària d'un centre escolar, la puntuació mitjana obtinguda és de 85 i la $S = 11$. En un altre centre, els alumnes de 3r obtenen una puntuació mitjana de domini de la lectura de 86 i una $S = 14$. Una alumna obté una puntuació de 88. En quin dels dos centres, comparativament, aquest valor representa una puntuació més alta de domini de la lectura?
- La marca de cotxes X té un consum mitjà de 5 litres de gasolina cada cent quilòmetres, amb una desviació típica d'1,3 litres. La marca de cotxes Y té un consum mitjà de 5,5 litres, amb una desviació típica d'1,2 litres. Un cotxe de cada marca que hagi consumit durant el mateix trajecte 8 litres de gasolina, a quina de les dues marques representa que ha tingut un consum més elevat?
- En un estudi sobre el procés d'inserció laboral dels diplomats en Educació Social de la UdG es va passar un qüestionari a tots els diplomats des de l'any 1996 fins a l'any 2005. En una de les preguntes del qüestionari se'ls demanava que valoressin de l'1 al 7 la feina que feien en el moment de respondre al qüestionari. Concretament, es demanava que valoressin els ítems següents:

| | Gens d'acord | | | | | Totalment d'acord | |
|---|--------------|---|---|---|---|-------------------|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1. El contingut de la feina és satisfactori. | | | | | | | |
| 2. Les perspectives de millora i promoció en la feina actual són bones. | | | | | | | |
| 3. El nivell de retribució et sembla satisfactori. | | | | | | | |
| 4. Els coneixements derivats de la formació universitària t'han estat útils per a la feina. | | | | | | | |
| 5. Les perspectives d'estabilitat són bones. | | | | | | | |
| 6. La jornada laboral (horaris, torns de treball, etc.) és satisfactòria. | | | | | | | |
| 7. Et sents realitzat com a educador social en aquesta feina. | | | | | | | |

A continuació tens una taula amb els estadístics descriptius una vegada analitzat cada un dels ítems.

| | N | Rang | Suma | Mitjana | Desv. típ. | Variància |
|--|-----|------|------|---------|------------|-----------|
| Valoració contingut (1-7) | 104 | 6 | 573 | 5,51 | 1,488 | 2,213 |
| Valoració perspectives millora (1-7) | 105 | 6 | 437 | 4,16 | 1,716 | 2,945 |
| Valoració retribució (1-7) | 103 | 6 | 485 | 4,71 | 1,752 | 3,071 |
| Valoració coneixements de formació (1-7) | 105 | 6 | 447 | 4,26 | 1,629 | 2,654 |
| Valoració perspectives estabilitat (1-7) | 105 | 6 | 543 | 5,17 | 1,816 | 3,297 |
| Valoració jornada laboral | 105 | 6 | 536 | 5,10 | 1,748 | 3,056 |
| Valoració realització personal (1-7) | 101 | 6 | 506 | 5,01 | 2,017 | 4,070 |

4.1. Una persona ha valorat amb un 6 tant les perspectives de millora com la retribució que rep. Aquesta puntuació, en quina de les dues variables representa una millor valoració tenint en compte les característiques de cada distribució?

4.2. Una persona ha valorat amb un 4 tant la utilitat dels coneixements derivats de la formació universitària per a la seva feina com el nivell de realització personal a la feina. En quina de les dues variables representa una puntuació pitjor?

4.3. Hi ha persones que han valorat amb un 3 la jornada laboral i persones que han valorat amb un 2 les perspectives de millora. En quina de les dues variables aquestes puntuacions representen una millor valoració?

4.4. Una puntuació de 3, en quina de les tres variables representa una millor puntuació, en la valoració de les perspectives de millora, en la valoració de les perspectives d'estabilitat o en la valoració del nivell de realització personal que suposa la feina?

4.5. Una persona ha valorat amb un 5 les perspectives de millora i amb un 6 el contingut de la feina. En quin dels dos ítems la seva valoració ha estat més positiva, tenint en compte les característiques de la distribució de valors de cada variable?

2.3. Respostes als exercicis sobre puntuacions típiques

Exercici 1

$$Z_1 = 1,42$$

$$Z_2 = 0,25$$

La puntuació més alta, comparativament, és la d'Estadística.

Exercici 2

Una alumna que aconsegueix 88 punts té una puntuació més alta de domini de la lectura al centre 1 perquè la puntuació típica del centre 1 (0,27) és més gran que la puntuació típica del centre 2 (0,14).

Exercici 3

$$Z_1 = \frac{X_1 - \bar{X}_1}{S_1} = \frac{8-5}{1,3} = 2,3$$

$$Z_2 = \frac{X_2 - \bar{X}_2}{S_2} = \frac{8-5,5}{1,2} = 2,1$$

Comparant les dues marques, veiem que el consum de 8 litres representa un consum més elevat en el cas de la marca X que en la marca Y.

Exercici 4

4.1.

$$Z_1 = \frac{X_1 - \bar{X}_1}{S_1} = \frac{6 - 4,16}{1,716} = 1,07$$

$$Z_2 = \frac{X_2 - \bar{X}_2}{S_2} = \frac{6 - 4,71}{1,752} = 0,74$$

Mirant les dues puntuacions tipificades veiem que $1,07 > 0,74$, la qual cosa vol dir que aquesta puntuació de 6 representa una valoració més bona de les perspectives de millora que no pas de la retribució que percep, tenint en compte les característiques de la distribució obtinguda a partir de les respostes a aquesta pregunta.

4.2.

$$Z_1 = -0,16$$

$$Z_2 = -0,5 \quad -0,5 < -0,16$$

Per tant, tenint en compte les característiques de les dues mostres, la puntuació de 4 atorgada a la variable “valoració del nivell de realització com a educador social” representa una pitjor valoració que la puntuació de 4 atorgada a la “valoració de la utilitat per a la feina dels coneixements universitaris”.

4.3.

$$Z_1 = -1,2 \text{ (puntuació típica per al valor 3 de “valoració de la jornada laboral”)}$$

$$Z_2 = -1,26 \text{ (puntuació típica per al valor 2 de “valoració de les perspectives de millora”)}$$

$$-1,26 < -1,2$$

El valor 3 representa una millor puntuació, si bé hi ha poca diferència. Vol dir que gairebé són equivalents les dues valoracions (el 2 i el 3) per a les dues variables diferents.

4.4.

$$Z_1 = -0,68 \text{ (puntuació típica per al valor 3 de “valoració de les perspectives de millora”)}$$

$$Z_2 = -1,20 \text{ (puntuació típica per al valor 3 de “valoració de les perspectives d'estabilitat”)}$$

$$Z_3 = -0,99 \text{ (puntuació típica per al valor 3 de “valoració del nivell de realització personal”)}$$

$$-1,20 < -0,99 < -0,68$$

Per tant, una puntuació de 3 és una millor puntuació en la valoració de les perspectives de millora.

4.5.

$$Z_1 = 0,49 \text{ (puntuació típica per al valor 5 de “valoració de les perspectives de millora”)}$$

$$Z_2 = 0,33 \text{ (puntuació típica per al valor 6 de “valoració del contingut de la feina”)}$$

$0,33 < 0,49$. Per tant, haver puntuat amb un 5 les perspectives de millora representa una millor valoració que haver puntuat amb un 6 el contingut de la feina.

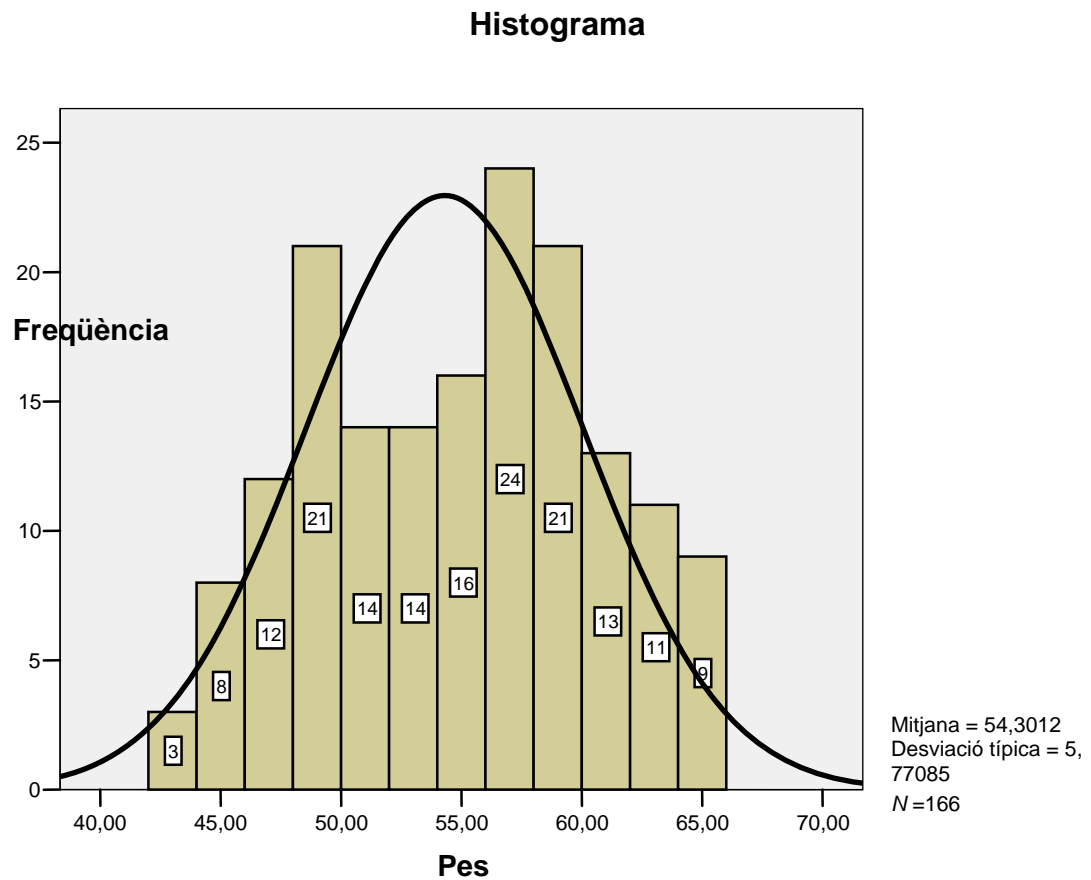
2.4. El model de la corba normal

En aquest document fem només una petita introducció al model de la corba normal, que ampliarem més endavant. De moment, el que ens cal saber és que el model de la corba normal és una distribució teòrica, és a dir, és una distribució que mai no coincidirà exactament amb la distribució d'unes dades empíriques. Les distribucions de valors es poden representar gràficament. Quan l'escala de mesura d'una variable és una escala d'interval, el tipus de gràfic que podem utilitzar és l'histograma. Així, imaginem-nos que tenim la següent distribució de valors de la variable “pes” de les dones d'entre 18 i 30 anys:

| Valors | f |
|--------|-----|
| 43,00 | 3 |
| 44,00 | 3 |
| 45,00 | 5 |
| 46,00 | 2 |
| 47,00 | 10 |
| 48,00 | 11 |
| 49,00 | 10 |
| 50,00 | 10 |
| 51,00 | 4 |
| 52,00 | 10 |
| 53,00 | 4 |
| 54,00 | 6 |
| 55,00 | 10 |
| 56,00 | 13 |
| 57,00 | 11 |
| 58,00 | 9 |
| 59,00 | 12 |
| 60,00 | 8 |
| 61,00 | 5 |
| 62,00 | 7 |
| 63,00 | 4 |
| 64,00 | 4 |
| 65,00 | 4 |
| 66,00 | 1 |
| Total | 166 |

Al gràfic 2.1 podem veure la distribució d'aquests valors. Cada columna representa un interval de 2 valors referits al pes: [42-44), [44-46); [46-48); [48-50), etc.

Gràfic 2.1: Distribució de la mostra segons el pes



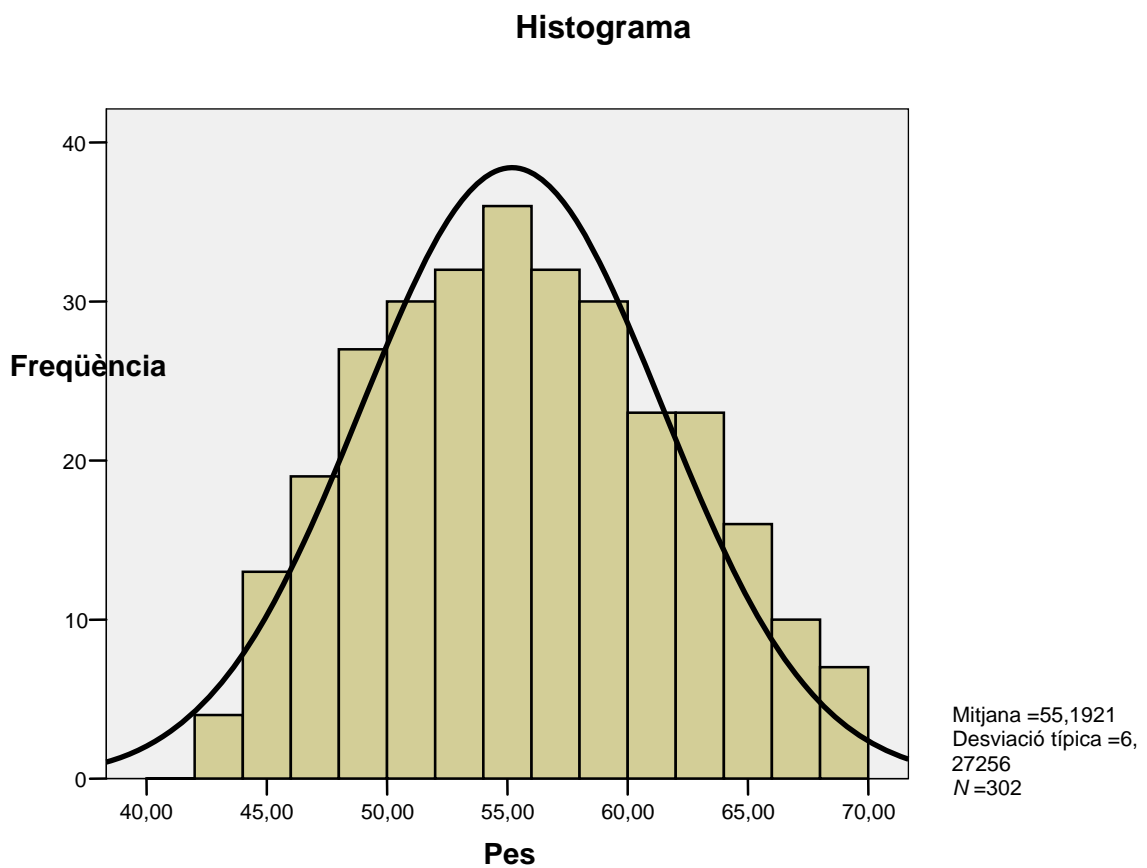
Ara imaginem-nos que hem augmentat la mida de la mostra. Els valors són els següents (i a continuació veiem l'histograma):

pes

| Valors | <i>f</i> | % |
|--------|----------|-----|
| 42,00 | 1 | ,3 |
| 43,00 | 3 | 1,0 |
| 44,00 | 6 | 2,0 |
| 45,00 | 7 | 2,3 |
| 46,00 | 6 | 2,0 |
| 47,00 | 13 | 4,3 |
| 48,00 | 15 | 5,0 |
| 49,00 | 12 | 4,0 |
| 50,00 | 18 | 6,0 |
| 51,00 | 12 | 4,0 |
| 52,00 | 16 | 5,3 |
| 53,00 | 16 | 5,3 |
| 54,00 | 16 | 5,3 |

| | | |
|-------|-----|-------|
| 55,00 | 20 | 6,6 |
| 56,00 | 15 | 5,0 |
| 57,00 | 17 | 5,6 |
| 58,00 | 13 | 4,3 |
| 59,00 | 17 | 5,6 |
| 60,00 | 13 | 4,3 |
| 61,00 | 10 | 3,3 |
| 62,00 | 15 | 5,0 |
| 63,00 | 8 | 2,6 |
| 64,00 | 8 | 2,6 |
| 65,00 | 8 | 2,6 |
| 66,00 | 6 | 2,0 |
| 67,00 | 4 | 1,3 |
| 68,00 | 4 | 1,3 |
| 69,00 | 1 | ,3 |
| 70,00 | 2 | ,7 |
| Total | 302 | 100,0 |

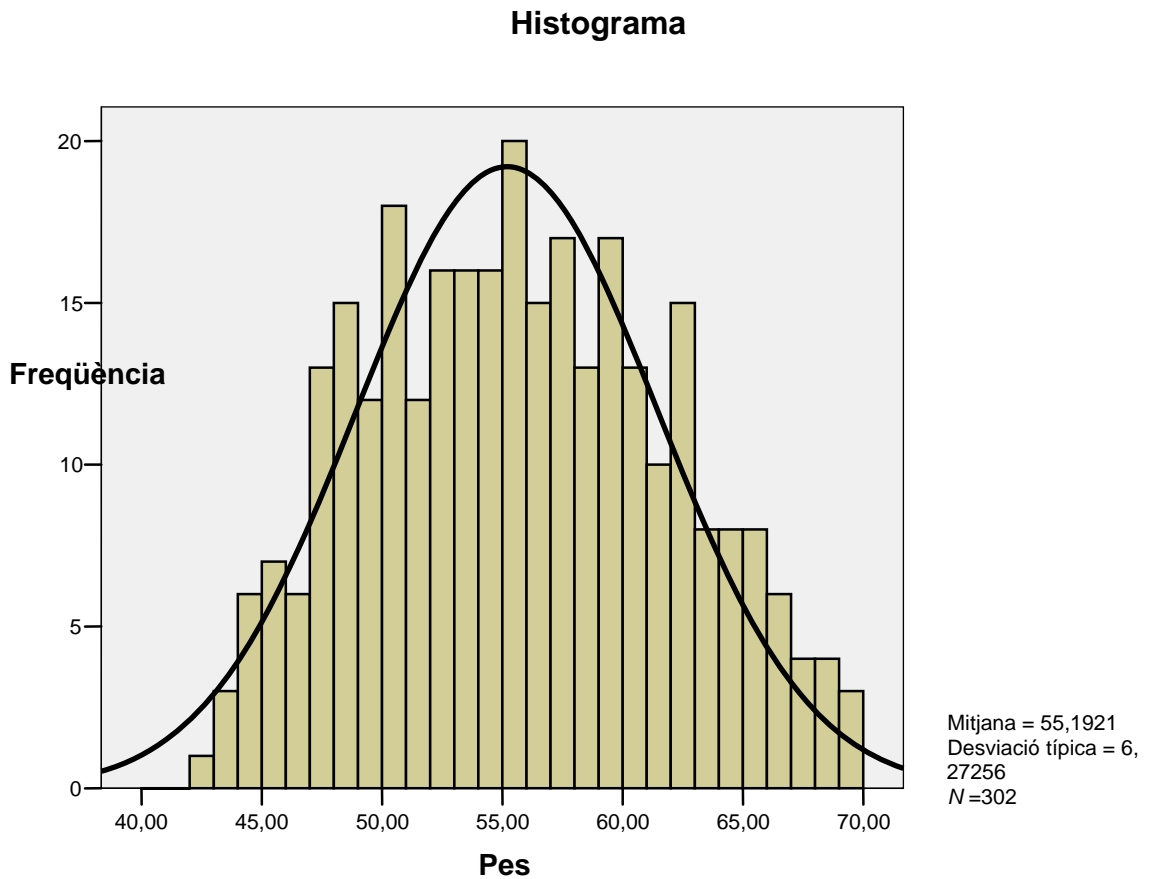
Gràfic 2.2: Distribució de la mostra segons el pes



En ambdós histogrames veiem que els intervals són de 2, és a dir, cada columna representa les freqüències sumades de dos valors. Si anéssim fent intervals més petits,

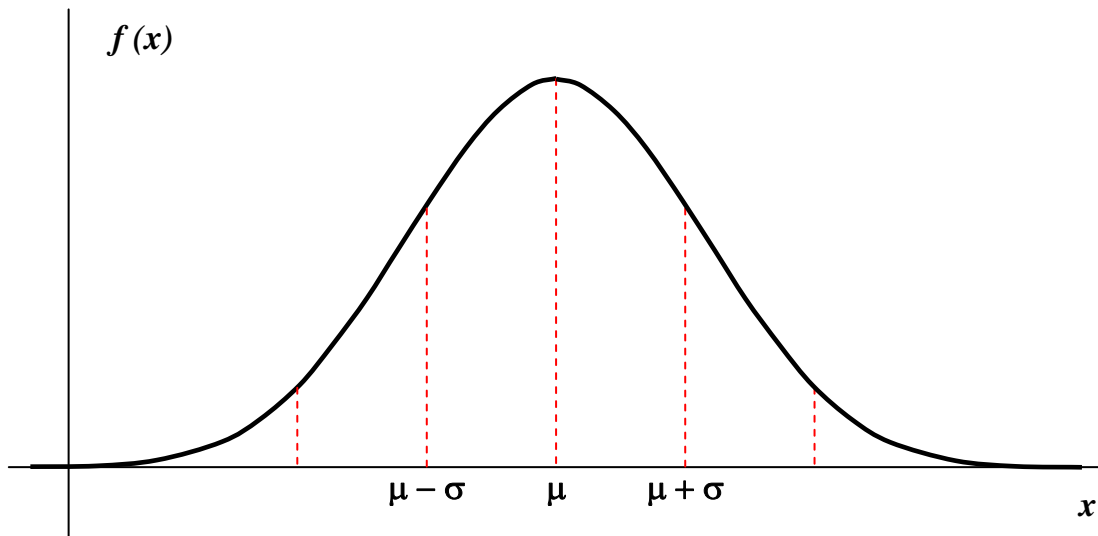
fins a tenir la distribució per a cada valor de pes, ens trobaríem que la forma de la gràfica cada vegada s'acosta més a la que hi ha dibuixada en el gràfic 2.3.

Gràfic 2.3: Distribució de la mostra segons el pes. Intervals d'1.



Si féssim la distribució de tots els valors de la població de dones d'entre 18 i 30 anys tindríem que, teòricament, la distribució seria molt similar a la que es representa en el gràfic 2.4. Al punt central trobaríem la mitjana de la població (μ). Entre la mitjana i una desviació típica per sobre i per sota seu, trobaríem la majoria de casos. A mesura que ens allunyem de la mitjana cada vegada hi ha menys casos, és a dir, menys freqüències. Així, als extrems és on trobem menys casos.

Gràfic 2.4: Corba normal



L'equació que descriu aquesta funció és la següent, que es coneix també com a *funció de densitat*:

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

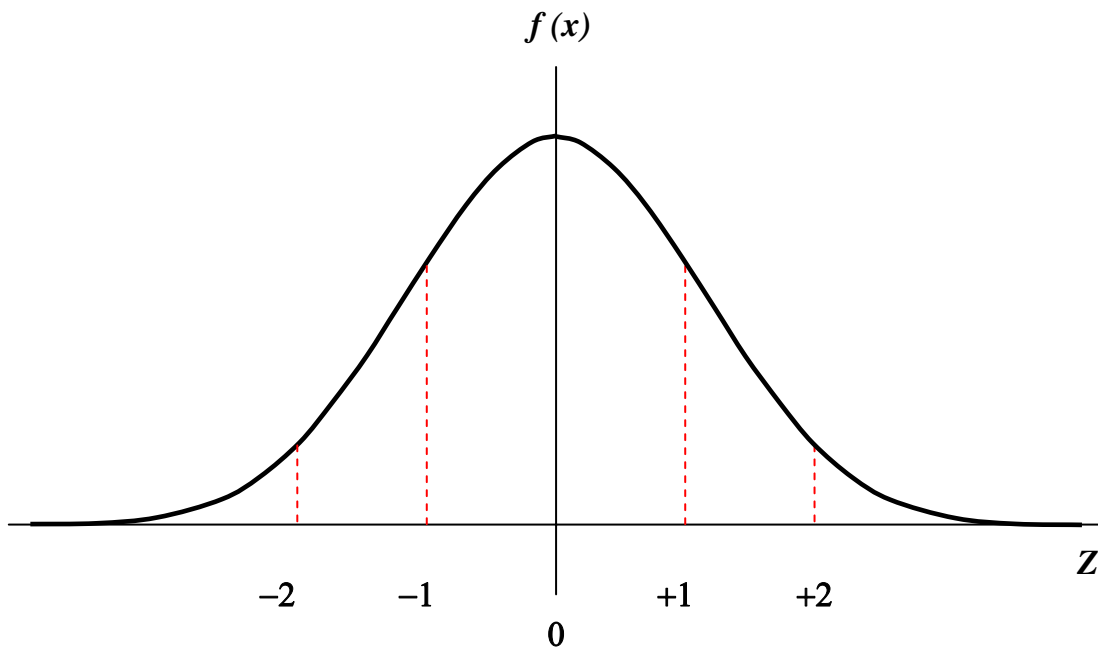
| | |
|---------------------------|----------------|
| μ mitjana | $\pi = 3,1416$ |
| σ desviació típica | $e = 2,7182$ |
| σ^2 variància | x abscissa |

Característiques de la corba normal:

1. La corba assoleix l'alçada màxima al punt central, que correspon al valor de la mitjana.
2. La corba és simètrica respecte a l'eix vertical, que passa per la mitjana.
3. En el model de la corba normal, la mitjana, la mediana i la moda coincideixen al mateix punt.

La distribució amb què es treballa normalment és la corba normal típica, és a dir, una distribució normal que té de mitjana 0 i de desviació típica 1.

Gràfic 2.5: Corba normal tipificada



Tal com expliquen Etxeberria i Tejedor (2005), la corba normal representa un histograma en el qual la mida de l'interval és molt i molt petita.

Quan la distribució és normal, la proporció de casos que es troben entre la mitjana menys una desviació típica i la mitjana més una desviació típica sempre és del 68 %. I la proporció de casos que estan entre la mitjana menys dues desviacions típiques i la mitjana més dues desviacions típiques sempre és del 95 %.

En el cas de la corba normal tipificada, l'àrea delimitada per l'eix d'abscisses (x) i la corba val 1. En general, l'àrea sota la corba normal representa el 100 % dels casos.

Gràfic 2.6: Àrees sota la corba normal

