

DISTRIBUCIONS UNIDIMENSIONALS

Per Roger Mauricio Grañó

Introducció

En qualsevol estudi estadístic cal tenir en compte els següents conceptes: població, caràcter estadístic i mostra.

- **Població:** Conjunt d'elements sobre el que fem l'estudi.
- **Caràcter estadístic:** Valor o propietat a estudiar.
- **Mostra:** Subconjunt de la població que es pren com a referent de la població.

Tipus de caràcters estadístics

1. Caràcter qualitatiu: Propietat no mesurable.
2. Caràcter quantitatiu: Mesurable.
 - a) Variables discretes. Només prenen valors discrets.
 - b) Variables contínues. Només prenen valors reals.

Per Roger Mauricio Grañó

Freqüència absoluta i relativa.

Un cop es recullen les dades d'un estudi estadístic (suposem que són dades numèriques discretes) s'ha de mirar el nombre de vegades que es repeteixen certs resultats en el conjunt. Aleshores definim:

- **Freqüència absoluta d'un valor x_i :** És el nombre de vegades que es repeteix el valor x_i . S'indica per f_i .
- **Freqüència relativa d'un valor x_i :** És la freqüència absoluta del valor x_i entre el total de resultats obtinguts (N). Es representa per h_i .

D'aquí es deueix que: $N = \sum f_i$ i $h_i = \frac{f_i}{N}$; $h_i \leq 1$

Per Roger Mauricio Grañó

Freqüència absoluta i relativa acumulada.

- **Freqüència absoluta acumulada d'un valor x_i :** És el nombre de valors (repeticions incloses) menors que el valor x_i . S'indica per F_i .
- **Freqüència relativa acumulada d'un valor x_i :** És la freqüència absoluta acumulada entre el total de resultats obtinguts (N). Es representa per H_i .

D'aquí es deueix que: $F_i = \sum_{j=1}^i f_j$ i $H_i = \frac{F_i}{N}$; $H_i \leq 1$

Per Roger Mauricio Grañó

EXEMPLE

La taula adjunta ens mostra els resultats obtinguts en llançar un dau 30 vegades.

xi	fi	Fi	hi	Hi
1	3	3	0,10	0,10
2	4	7	0,13	0,23
3	2	9	0,07	0,30
4	6	15	0,20	0,50
5	8	23	0,27	0,77
6	7	30	0,23	1,00
	30		1	

Quantes vegades ha sortit el valor 4?

Quina és la freqüència absoluta acumulada de 5?

Per Roger Mauricio Grañó

EXEMPLE a resoldre

Les altures en cm de 25 alumnes d'una classe són:

171, 177, 176, 160, 186, 170, 169, 178, 163, 166, 173, 162, 167, 179, 157, 173, 168, 167, 172, 168, 173, 182, 176, 183, 188

Agrupa els resultats en classes d'amplada 5 cm.

Classes	fi	Fi	hi	Hi
[155,160)				
[160,165)				
[165,170)				
[170,175)				
[175,180)				
[180,185)				
[185,190)	Per Roger Mauricio Grañó			

Classes. Marca d'una classe.

Quan treballem amb moltes dades numèriques, s'acostuma a agrupar els resultats per classes. Una classe és un interval definit per dues cotes, una inferior i una superior (a_i, a_{i+1}].

Aleshores, definim:

-**Marca de la classe** i : És el representant de la classe i . El valor que li correspon és:

$$\bar{x}_i = \frac{a_i + a_{i+1}}{2}$$

- **Amplada de la classe**: És la diferència entre les cotes de la classe

$$a = a_{i+1} - a_i$$

Per Roger Mauricio Grañó

EXEMPLE a resoldre

Les altures en cm de 25 alumnes d'una classe són:

171, 177, 176, 160, 186, 170, 169, 178, 163, 166, 173, 162, 167, 179, 157, 173, 168, 167, 172, 168, 173, 182, 176, 183, 188

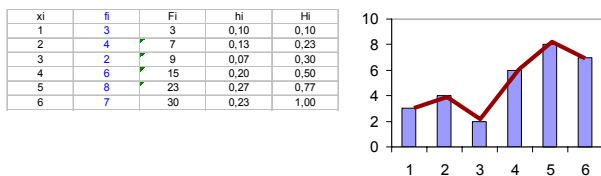
Agrupa els resultats en classes d'amplada 5 cm.

Classes	fi	Fi	hi	Hi
[155,160)	1	1	1/25	1/25
[160,165)	3	4	3/25	4/25
[165,170)	6	10	6/25	10/25
[170,175)	6	16	6/25	16/25
[175,180)	5	21	1/5	21/25
[180,185)	2	23	2/25	23/25
[185,190)	2	25	2/25	1

Representació gràfica de les dades estadístiques.

Diagrames de barres: S'usen en distribucions de freqüències de tipus discret per a dades no agrupades. Tan es poden considerar els diagrames de freqüències absolutes o acumulades, relatives o no.

Polígon de freqüències: Es formen unint els valors de l'ordenada a partir dels diferents valors de les freqüències a considerar.



Per Roger Mauricio Grañó

Mesures de posició

Mitjana aritmètica: És el quocient de la suma dels productes de tots els valors observats pel nombre de repeticions i el nombre total de dades.

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{N}$$

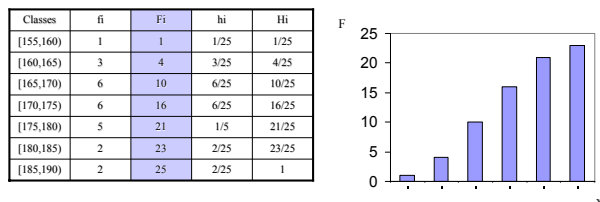
En cas de treballar amb les dades agrupades x_i és la marca de cada classe.

Mediana: És el valor de la variable estadística tal que la freqüència absoluta acumulada és $N/2$.

Obtenció de la mediana:

Per Roger Mauricio Grañó

Histogrames: S'usen en distribucions de freqüències amb les dades agrupades per classes. Tan es poden considerar els diagrames de freqüències absolutes o acumulades, relatives o no.



Per Roger Mauricio Grañó

1. Distribució amb dades discretes no agrupades.

- Si N és senar, la mediana és el valor x_i que ocupa la posició central.
- Si N és parell, la mediana és el promig dels dos valors centrals.

Si la mediana pren un valor impossible dins la distribució, diem que la mediana és doble i és igual als dos valors centrals.

2. Distribucions amb dades agrupades per classes.

- Primer cal trobar la classe mediana. *Idem que 1.*

- Després apliquem l'equació:

$$M = L_i + c \cdot \frac{\frac{N}{2} - F_{i-1}}{f_i}$$

Per Roger Mauricio Grañó

$$M = L_i + c \cdot \frac{\frac{N}{2} - F_{i-1}}{f_i}$$

L_i és el límit inferior de la classe mediana.

f_i és la freqüència absoluta de la classe mediana.

F_{i-1} és la freqüència absoluta acumulada de la classe anterior a la classe mediana

c és l'amplada de les classes.

Sempre $M \in [a_i, a_{i+1})$.

Per Roger Mauricio Grañó

Moda (m) : És el valor o valors de la distribució amb la freqüència absoluta més altes. Podem parlar de distribucions bimodals, trimodals...

Quantils: Són els valors de la distribució que divideixen el conjunt de dades en parts que tenen igual nombre de dades.

En general parlarem de:

Quartils: Quan el conjunt de dades es divideix en quatre parts. En aquest cas tenim tres quartils: Q_1 , Q_2 i Q_3 que es troben amb de la següent forma:

$$Q_1 = x_{\frac{N}{4}}; \quad Q_2 = x_{\frac{N}{2}}; \quad Q_3 = x_{\frac{3N}{4}}$$

Nota: $Q_2 = M$

Per Roger Mauricio Grañó

Quan les dades estan agrupades per classes haurem d'aplicar l'equació:

$$Q_i = L_i + c \cdot \frac{\frac{i \cdot N}{4} - F_{i-1}}{f_i}$$

- L_i és el límit inferior de la classe quartil.

- c és l'amplada de les classes

- f_i és la freqüència de la classe quartil

- F_{i-1} és la freqüència acumulada de la de la classe anterior.

Percentils: Quan el conjunt de dades queda agrupat en 100 parts. En aquest cas hi haurà 99 percentils. La seva posició es determinarà per:

$$P_k = x_{\frac{kN}{100}}$$

Nota: $P_{50} = M$; $P_{25} = Q_1$; $P_{75} = Q_3$

Per Roger Mauricio Grañó

Quan les dades estan agrupades per classes haurem d'aplicar l'equació:

$$P_i = L_i + c \cdot \frac{\frac{i \cdot N}{100} - F_{i-1}}{f_i}$$

- L_i és el límit inferior de la classe quartil.

- c és l'amplada de les classes

- f_i és la freqüència de la classe quartil

- F_{i-1} és la freqüència acumulada de la de la classe anterior.

Per Roger Mauricio Grañó

EXEMPLE a resoldre

Les altures en cm de 25 alumnes estan agrupades per classes tal i com es veuen a la taula de la figura. Determineu:

- La mitjana aritmètica de les altures.
- La mediana
- La classe modal

Classes	fi	Fi	hi	Hi
[155,160)	1	1	1/25	1/25
[160,165)	3	4	3/25	4/25
[165,170)	6	10	6/25	10/25
[170,175)	6	16	6/25	16/25
[175,180)	5	21	1/5	21/25
[180,185)	2	23	2/25	23/25
[185,190)	2	25	2/25	1

Desviació mitjana. És la mitjana aritmètica de les desviacions en valor absolut.

$$\bar{D} = \frac{\sum |D_i| \cdot f_i}{N}$$

Variància: És la mitjana aritmètica de les desviacions al quadrat.

$$S^2 = \frac{\sum D_i^2 \cdot f_i}{N}$$

Desviació tipus: És l'arrel quadrada de la variància.

$$S = \sqrt{\frac{\sum D_i^2 \cdot f_i}{N}}$$

Coefficient de variació. És el quocient entre la desviació tipus i la mitjana aritmètica.

$$d = \frac{S}{\bar{x}}$$

Per Roger Mauricio Grañó

Mesures de dispersió

Les mesures de posició serveixen per situar les dades de l'estudi.

Amb les mesures de dispersió ens serviran per conèixer el grau d'agrupament de les dades.

Rang: És la diferència entre $x_{\max}$ i $x_{\min}$.

$$R = x_{\max} - x_{\min}$$

Rang interquartilic: És la diferència entre els quartils 3 i 1.

$$Q = Q_3 - Q_1$$

Rang entre percentils: És la diferència entre els percentils 99 i 1.

$$P = P_{99} - P_1$$

Desviació de xi: És la distància entre el valor x_i i la mitjana aritmètica.

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

Per Roger Mauricio Grañó

EXEMPLE a resoldre

Volem obtenir el temps de caiguda d'una pedra des de la teulada de l'escola. Per això usem un cronòmetre digital que mesura fins a les dècimes de segon. Per tal que el càlcul sigui el més exacte possible decidim repetir l'experiència 500 vegades. A la taula adjunta obtenim els resultats indicats.

Calculeu:

t(s)	f
	500
12,3	36
13,1	34
13,5	9
13,8	75
13,9	65
14	70
14,8	98
15	12
15,1	3
15,7	25
15,9	65
16,6	8

Per Roger Mauricio Grañó

a) La Mediana de la distribució. Posició central : $N = 500 \Rightarrow \frac{N}{2} = 250$

$$x_{250} \text{ i } x_{251} \Rightarrow \begin{cases} x_{250} = 14,0s \\ x_{251} = 14,0s \end{cases} \Rightarrow M = 14,0s$$

$$Q_1 = x_{N/4} = x_{125} = 13,8s$$

b) Els quartils. $Q_2 = M = 14,0s$

$$Q_3 = x_{3N/4} = x_{375} = 14,8s$$

c) El percentil 22 $P_{22} = x_{22N/100} = x_{110} = 13,8s$

d) El rang, el rang interquartilic i entre percentils

$$R = x_{\max} - x_{\min} = 16,6 - 12,3 = 4,3s$$

$$Q = Q_3 - Q_1 = 14,8 - 13,8 = 1,0s$$

$$P = P_{99} - P_1 = x_{99N/100} - x_{N/100} =$$

$$x_{495} - x_5 = 16,6 - 12,3 = 4,3s$$

t(s)	f	F
	500	
12,3	36	36
13,1	34	70
13,5	9	79
13,8	75	154
13,9	65	219
14	70	289
14,8	98	387
15	12	399
15,1	3	402
15,7	25	427
15,9	65	492
16,6	8	500

Per Roger Mauricio Grañó

Problemes del llibre.

Matemàtiques cc. ss. Pg 174. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Matemàtiques

Pg 222. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Per Roger Mauricio Grañó

a) La mitjana aritmètica.

$$\bar{x} = \frac{\sum x_i \cdot f_i}{N} = \frac{7162,7}{500} = 14,33s$$

b) La desviació mitjana

$$\bar{D} = \frac{\sum |D_i| \cdot f_i}{N} = \frac{423,6812}{500} = 0,8476s$$

c) La variancia.

$$S^2 = \frac{\sum D^2 \cdot f}{N} = \frac{523,86742}{500} = 1,04773s^2$$

d) La desviació tipus. $S = \sqrt{S^2} = \sqrt{1,04773s^2} = 1,02359s$

e) El coeficient de variació. $d = \frac{S}{\bar{x}} = \frac{1,02359s}{14,33s} = 0,07142 \Rightarrow 7,142\%$

Per Roger Mauricio Grañó

t	f	xf	x-<x>	D ·f	D ² ·f
	500	7162,7		423,6812	523,86742
12,3	36	442,8	-2,0	72,9	147,7
13,1	34	445,4	-1,2	41,7	51,1
13,5	9	121,5	-0,8	7,4	6,1
13,8	75	1035	-0,5	39,4	20,7
13,9	65	903,5	-0,4	27,7	11,8
14	70	980	-0,3	22,8	7,4
14,8	98	1450,4	0,5	46,5	22,1
15	12	180	0,7	8,1	5,5
15,1	3	45,3	0,8	2,3	1,8
15,7	25	392,5	1,4	34,4	47,2
15,9	65	1033,5	1,6	102,3	161,2
16,6	8	132,8	2,3	18,2	41,4