
Mètodes matemàtics

Tema 4: Estadística

NOTA: Les presentacions són la fusió de les presentacions de les seccions EPSEB i ESEIAAT del DMAT, mentre que els problemes han estat extrets del llibre [A. Ferrer et al., *Fonaments d'estadística aplicada*, Ed. Romargraf, 1995. ISBN 8460545857].

Grau en enginyeria geomàtica i geoinformació
Departament de Matemàtiques
EPSEB Universitat Politècnica de Catalunya

Índex

I	Estadística descriptiva	5
1	Estadística descriptiva: transparències	7
2	Estadística descriptiva: problemes	39
II	Probabilitat	43
3	Probabilitat: transparències	45
4	Probabilitat: problemes	59
III	Variables aleatòries discretes	63
5	Variables aleatòries discretes: transparències	65
6	Variables aleatòries discretes: problemes	105

Part I

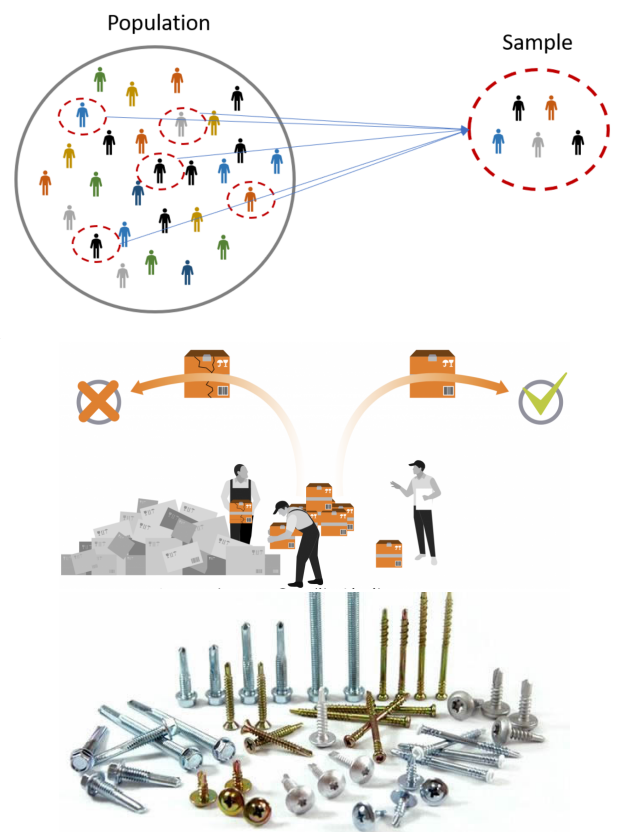
Estadística descriptiva

Capítol 1

Estadística descriptiva: transparències

Població i variables

- ▶ **Població:** conjunt d'elements objecte de l'estudi.
- ▶ **Mostra:** subconjunt de la població.
- ▶ **Variable:** propietat dels elements de la població.
 - V. qualitativa.
 - V. quantitativa (discreta o contínua).
- ▶ **Observació:** valor que pren la variable en un element de la població.



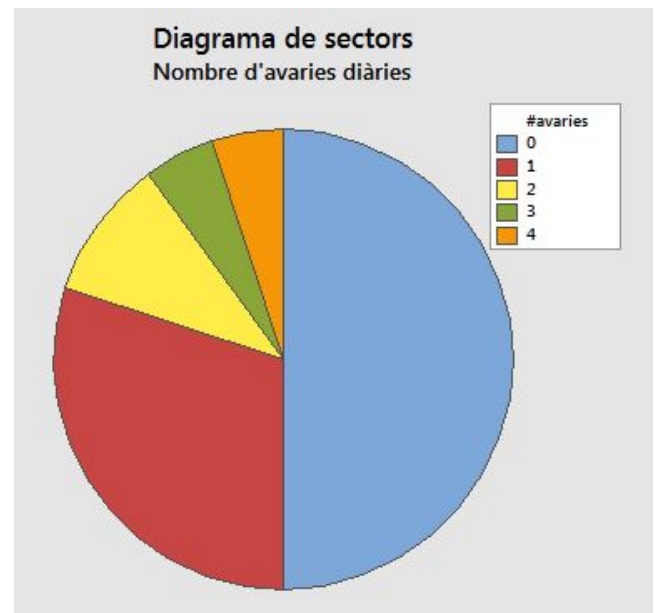
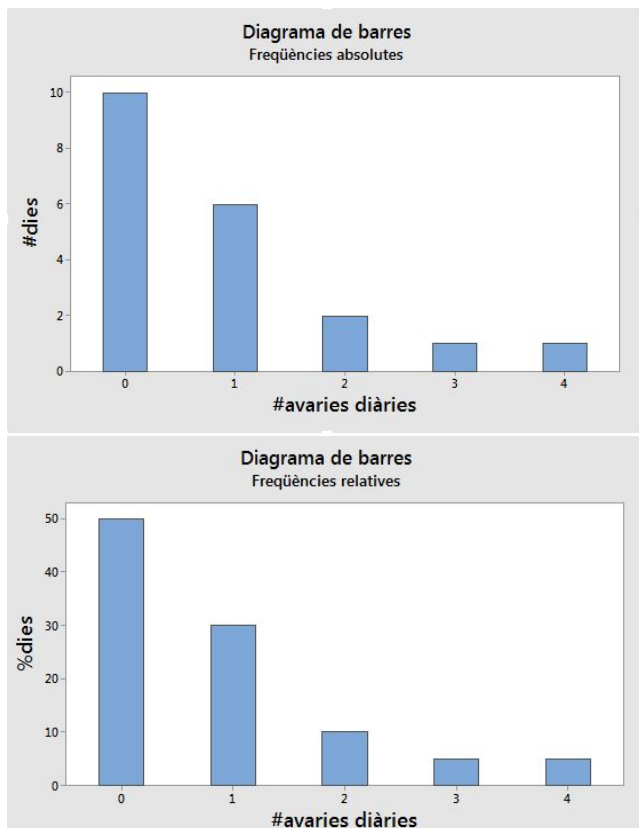
Variables discretes

- ▶ **Freqüència absoluta:** nombre de vegades (n_i) que apareix un valor determinat x_i en una mostra, és a dir, a les n observacions de la variable X .
- ▶ **Freqüència relativa:** $f_i := n_i/n$.
- ▶ **Freqüències acumulades:** $N_i := \sum_{j \leq i} n_j$ (**absoluta**); $F_i := \sum_{j \leq i} f_j$ (**relativa**).

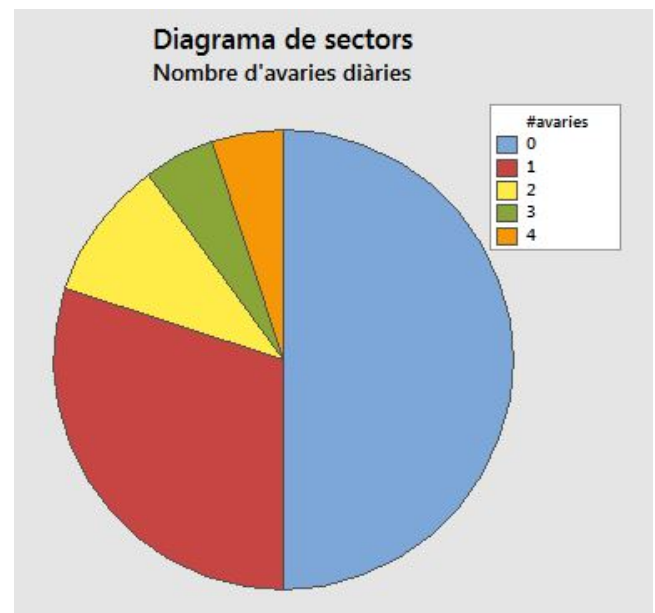
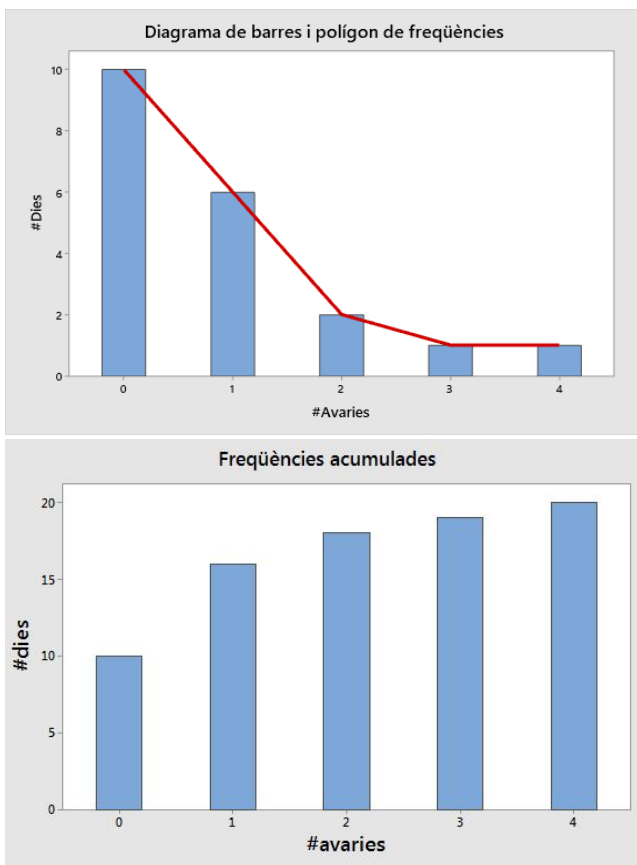
Exemple: Avaries diàries en una planta industrial al llarg de 20 dies.

valors	n_i	f_i	N_i	F_i
0	10	$10/20=0.5$	10	$10/20=0.5$
1	6	$6/20=0.3$	16	$16/20=0.8$
2	2	$2/20=0.1$	18	$18/20=0.9$
3	1	$1/20=0.05$	19	$19/20=0.95$
4	1	$1/20=0.05$	20	$20/20=1$
total	20	1		

Variables discretes: representacions gràfiques



Variables discretes: representacions gràfiques



Variables contínues

Exemple: Dades contínues: **tensió de ruptura** d'un tipus de fil (21 elements): 1.712, 1.735, 1.613, 1.76, 1.87, 1.815, 1.801, 1.765, 1.70, 1.73, 1.68, 1.683, 1.64, 1.62, 1.75, 1.69, 1.74, 1.715, 1.79, 1.77, 1.81 quilograms-força.

Variables contínues: arrodoniment, uniformització

Exemple: Dades contínues: **tensió de ruptura** d'un tipus de fil (21 elements): 1.71, 1.74, 1.61, 1.76, 1.87, 1.82, 1.80, 1.77, 1.70, 1.73, 1.68, 1.68, 1.64, 1.62, 1.75, 1.69, 1.74, 1.72, 1.79, 1.77, 1.81 quilograms-força.

Variables contínues: canvi d'escala

Exemple: Dades contínues: **tensió de ruptura** d'un tipus de fil (21 elements): 171, 174, 161, 176, 187, 182, 180, 177, 170, 173, 168, 168, 164, 162, 175, 169, 174, 172, 179, 177, 181 **deca**grams-força.

Variables contínues: intervals i marques de classe

Exemple: Dades contínues: **tensió de ruptura** d'un tipus de fil (21 elements): 171, 174, 161, 176, 187, 182, 180, 177, 170, 173, 168, 168, 164, 162, 175, 169, 174, 172, 179, 177, 181 decagrams-força.

- Agrupem les dades en **intervals de classe** (elecció no trivial, $O(\sqrt{n})$ intervals):

classes	n_i	f_i	N_i	F_i
(160, 165]	3	0.1429	3	0.1429
(165, 170]	4	0.1905	7	0.3333
(170, 175]	6	0.2857	13	0.6190
(175, 180]	5	0.2381	18	0.8571
(180, 185]	2	0.0952	20	0.9523
(185, 190]	1	0.0475	21	1
total	21	1		

- **Marca de classe:** valor mitjà de cada interval de classe.

Variables contínues: distribució de freqüències

Exemple: Dades contínues: **tensió de ruptura** d'un tipus de fil (21 elements): 171, 174, 161, 176, 187, 182, 180, 177, 170, 173, 168, 168, 164, 162, 175, 169, 174, 172, 179, 177, 181 decagrams-força.

- Agrupem les dades en **interval·ls de classe** (elecció no trivial, $O(\sqrt{n})$ interval·ls):

classes	marca	n_i	f_i	N_i	F_i
(160, 165]	162.5	3	0.1429	3	0.1429
(165, 170]	167.5	4	0.1905	7	0.3333
(170, 175]	172.5	6	0.2857	13	0.6190
(175, 180]	177.5	5	0.2381	18	0.8571
(180, 185]	182.5	2	0.0952	20	0.9523
(185, 190]	187.5	1	0.0475	21	1
total		21	1		

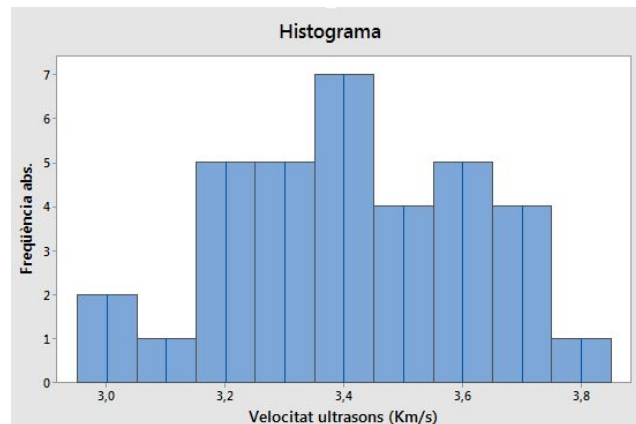
- **Marca de classe:** valor mitjà de cada interval de classe.

Variables contínues: representacions gràfiques

Exemple: Mesures de les **velocitats amb ultrasons** de 34 pilars d'una estructura (en Km/s): 3.60, 3.21, 3.41, 3.17, 3.40, 3.63, 3.02, 3.02, 3.27, 3.51, 3.60, 3.58, 3.65, 3.33, 3.34, 3.35, 3.35, 3.70, 3.58, 3.24, 3.52, 3.41, 3.14, 3.28, 3.18, 3.35, 3.33, 3.72, 3.53, 3.45, 3.43, 3.84, 3.23, 3.72.

► Diagrama de tija i fulles.

3.0	2	2				
3.1	4	7	8			
3.2	1	3	4	7	8	
3.3	3	3	4	5	5	5
3.4	0	1	1	3	5	
3.5	1	2	3	8	8	
3.6	0	0	3	5		
3.7	0	2	2			
3.8	4					

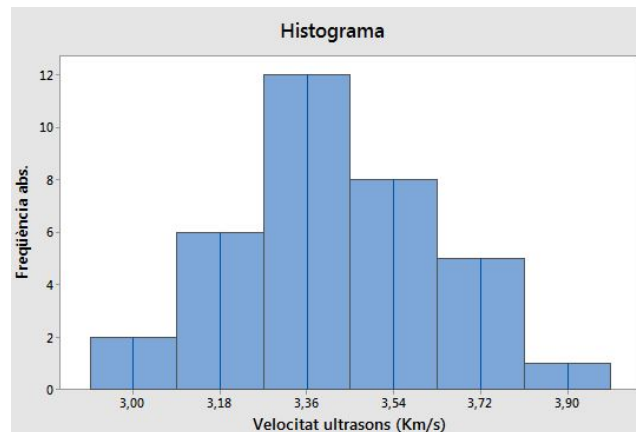


Variables contínues: representacions gràfiques

Exemple: Mesures de les **velocitats amb ultrasons** de 34 pilars d'una estructura (en Km/s): 3.60, 3.21, 3.41, 3.17, 3.40, 3.63, 3.02, 3.02, 3.27, 3.51, 3.60, 3.58, 3.65, 3.33, 3.34, 3.35, 3.35, 3.70, 3.58, 3.24, 3.52, 3.41, 3.14, 3.28, 3.18, 3.35, 3.33, 3.72, 3.53, 3.45, 3.43, 3.84, 3.23, 3.72.

► Diagrama de tija i fulles.

3.0	2	2				
3.1	4	7	8			
3.2	1	3	4	7	8	
3.3	3	3	4	5	5	5
3.4	0	1	1	3	5	
3.5	1	2	3	8	8	
3.6	0	0	3	5		
3.7	0	2	2			
3.8	4					



Descriptors d'un conjunt d'observacions

- ▶ Donada una mostra de n observacions, $x_1, x_2, \dots, x_n$, volem saber **com es distribueixen les dades**.
- ▶ Descriptors de **centralització** i **dispersió**.
- ▶ **Estadístics** d'una mostra: $\bar{x}, s_n, s_{n-1}, \dots$
- ▶ **Paràmetres** d'una població: $\mu, \sigma, \dots$

Mesures de tendència central

► **Mitjana** (*media, mean*):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

n dades

$$\bar{x} = \sum_{i=1}^m x_i \text{fr}_{rel}(x_i)$$

m valors diferents

- Velocitat ultrasons: $\bar{x} = \frac{3.02 + \dots + 3.84}{34} \approx 3.4144$
- Nombre avaries/dia:
 $\bar{x} = 0 \cdot 0.5 + 1 \cdot 0.3 + 2 \cdot 0.1 + 3 \cdot 0.05 + 4 \cdot 0.05 = 0.85$

► **Moda**: és el valor observat més freqüent.

Té més sentit en variables qualitatives o quantitatives discretes. En les contínues, s'aplica més als intervals de classe (**classe modal**).

- Nombre d'avaries diàries: 0 és la moda.
- Tensió de ruptura (161, 162, 164, **168, 168**, 169, 170, 171, 172, 173, **174, 174**, 175, 176, **177, 177**, 179, 180, 181, 182, 187): trimodal.

Mesures de tendència central → mesures de posició

Suposant les dades ordenades:

► **Mediana:**

$$M = Q_2 = P_{50} = \begin{cases} x_{\frac{n+1}{2}}, & \text{si } n \text{ és senar,} \\ \frac{1}{2} \left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right), & \text{si } n \text{ és parell.} \end{cases}$$

El **quantil** γ és **un** valor x tal que un $\gamma\%$ de les dades són menors o iguals que x . A part de la mediana, els més destacats són el **primer** i el **tercer quantil**, que denotarem per Q_1 i Q_3 , respectivament. Per a calcular-los:

- Ordenem les dades de la mostra.
- Les dividim en dos subconjunts S_1 i S_3 de la mateixa mida (si n és senar, ambdós contenen la mediana).
- Aleshores, $Q_1 = M(S_1)$ i $Q_3 = M(S_3)$, on M és la mediana.

Mesures de tendència central → mesures de posició

Notació per quan $\{x_i\}_i$ no estan ordenats: $x_{(i)}$ = element que ocupa la posició i -èsima un cop ordenats.

n	Q_1	$M = Q_2$	Q_3
$4k$	$\frac{x_{(k)} + x_{(k+1)}}{2}$	$\frac{x_{(2k)} + x_{(2k+1)}}{2}$	$\frac{x_{(3k)} + x_{(3k+1)}}{2}$
$4k + 1$	$x_{(k+1)}$	$x_{(2k+1)}$	$x_{(3k+1)}$
$4k + 2$	$x_{(k+1)}$	$\frac{x_{(2k+1)} + x_{(2k+2)}}{2}$	$x_{(3k+2)}$
$4k + 3$	$\frac{x_{(k+1)} + x_{(k+2)}}{2}$	$x_{(2k+2)}$	$\frac{x_{(3k+2)} + x_{(3k+3)}}{2}$

Mesures de tendència central → mesures de posició

Notació per quan $\{x_i\}_i$ no estan ordenats: $x_{(i)}$ = element que ocupa la posició i -èsima un cop ordenats.

20	Q_1	$M = Q_2$	Q_3
$4 \cdot 5$	$\frac{x_{(5)} + x_{(6)}}{2}$	$\frac{x_{(10)} + x_{(11)}}{2}$	$\frac{x_{(15)} + x_{(16)}}{2}$
$4k + 1$	$x_{(k+1)}$	$x_{(2k+1)}$	$x_{(3k+1)}$
$4k + 2$	$x_{(k+1)}$	$\frac{x_{(2k+1)} + x_{(2k+2)}}{2}$	$x_{(3k+2)}$
$4k + 3$	$\frac{x_{(k+1)} + x_{(k+2)}}{2}$	$x_{(2k+2)}$	$\frac{x_{(3k+2)} + x_{(3k+3)}}{2}$

Mesures de tendència central → mesures de posició

Notació per quan $\{x_i\}_i$ no estan ordenats: $x_{(i)}$ = element que ocupa la posició i -èsima un cop ordenats.

n	Q_1	$M = Q_2$	Q_3
$4k$	$\frac{x_{(k)} + x_{(k+1)}}{2}$	$\frac{x_{(2k)} + x_{(2k+1)}}{2}$	$\frac{x_{(3k)} + x_{(3k+1)}}{2}$
$4k + 1$	$x_{(k+1)}$	$x_{(2k+1)}$	$x_{(3k+1)}$
$4k + 2$	$x_{(k+1)}$	$\frac{x_{(2k+1)} + x_{(2k+2)}}{2}$	$x_{(3k+2)}$
$4k + 3$	$\frac{x_{(k+1)} + x_{(k+2)}}{2}$	$x_{(2k+2)}$	$\frac{x_{(3k+2)} + x_{(3k+3)}}{2}$

Mesures de tendència central → mesures de posició

Notació per quan $\{x_i\}_i$ no estan ordenats: $x_{(i)}$ = element que ocupa la posició i -èsima un cop ordenats.

21	Q_1	$M = Q_2$	Q_3
$4k$	$\frac{x_{(k)} + x_{(k+1)}}{2}$	$\frac{x_{(2k)} + x_{(2k+1)}}{2}$	$\frac{x_{(3k)} + x_{(3k+1)}}{2}$
$4 \cdot 5 + 1$	$x_{(6)}$	$x_{(11)}$	$x_{(16)}$
$4k + 2$	$x_{(k+1)}$	$\frac{x_{(2k+1)} + x_{(2k+2)}}{2}$	$x_{(3k+2)}$
$4k + 3$	$\frac{x_{(k+1)} + x_{(k+2)}}{2}$	$x_{(2k+2)}$	$\frac{x_{(3k+2)} + x_{(3k+3)}}{2}$

Mesures de tendència central → mesures de posició

Notació per quan $\{x_i\}_i$ no estan ordenats: $x_{(i)}$ = element que ocupa la posició i -èsima un cop ordenats.

n	Q_1	$M = Q_2$	Q_3
$4k$	$\frac{x_{(k)} + x_{(k+1)}}{2}$	$\frac{x_{(2k)} + x_{(2k+1)}}{2}$	$\frac{x_{(3k)} + x_{(3k+1)}}{2}$
$4k + 1$	$x_{(k+1)}$	$x_{(2k+1)}$	$x_{(3k+1)}$
$4k + 2$	$x_{(k+1)}$	$\frac{x_{(2k+1)} + x_{(2k+2)}}{2}$	$x_{(3k+2)}$
$4k + 3$	$\frac{x_{(k+1)} + x_{(k+2)}}{2}$	$x_{(2k+2)}$	$\frac{x_{(3k+2)} + x_{(3k+3)}}{2}$

Mesures de tendència central → mesures de posició

Notació per quan $\{x_i\}_i$ no estan ordenats: $x_{(i)}$ = element que ocupa la posició i-èsima un cop ordenats.

34	Q_1	$M = Q_2$	Q_3
$4k$	$\frac{x_{(k)} + x_{(k+1)}}{2}$	$\frac{x_{(2k)} + x_{(2k+1)}}{2}$	$\frac{x_{(3k)} + x_{(3k+1)}}{2}$
$4k + 1$	$x_{(k+1)}$	$x_{(2k+1)}$	$x_{(3k+1)}$
$4 \cdot 8 + 2$	$x_{(9)}$	$\frac{x_{(17)} + x_{(18)}}{2}$	$x_{(26)}$
$4k + 3$	$\frac{x_{(k+1)} + x_{(k+2)}}{2}$	$x_{(2k+2)}$	$\frac{x_{(3k+2)} + x_{(3k+3)}}{2}$

Mesures de posició (exemples)

► Tensió de ruptura, $n = 21$:

- $Q_2 = x_{(11)} = 174$.
- $S_1 = \{161, 162, 164, 168, 168, 169, 170, 171, 172, 173, 174\} \Rightarrow Q_1 = x_{(6)} = 169$.
- $S_3 = \{174, 174, 175, 176, 177, 177, 179, 180, 181, 182, 187\} \Rightarrow Q_3 = x_{(16)} = 177$.

► Nombre avaries/dia, $n = 20$: $Q_2 = (x_{(10)} + x_{(11)})/2 = 0.5$,
 $Q_1 = (x_{(5)} + x_{(6)})/2 = 0$, $Q_3 = (x_{(15)} + x_{(16)})/2 = 1$.

► Velocitat ultrasons, $n = 34$, recorrem al diagrama de tija i fulles:
 $Q_2 = (x_{(17)} + x_{(18)})/2 = 3.405$, $Q_1 = x_{(9)} = 3.27$, $Q_3 = x_{(26)} = 3.58$.

Exercici: En el cas de la velocitat per ultrasons, refeu el càlcul de les mesures de tendència central i de dispersió si una dada de valor 3.02 Km/s passa a ser de 2.65 Km/s, i compareu-los amb els obtinguts amb les dades originals.

Mesures de dispersió

► Rang o amplitud: $x_{(n)} - x_{(1)}$

► Rang interquartílic: $IQR = RIQ = Q_3 - Q_1$

► Variància (poblacional):

$$s_n^2 = \text{Var}(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

► Variància corregida (mostral, Minitab):

$$s_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{n-1} s_n^2$$

► Desviació típica i típica corregida: $s_n = +\sqrt{s_n^2}$, $s_{n-1} = +\sqrt{s_{n-1}^2}$

► Coeficient de variació: $CV = s_n / \bar{x}$.

Mesures de dispersió (exemples)

- ▶ Velocitat ultrasons:

$$s_n = \left(\frac{3.02^2 + \dots + 3.84^2}{34} - (3.4144)^2 \right)^{1/2} \approx 0.2021 \text{ km/s.}$$

- ▶ Nombre avaries/dia: $\bar{x} = (0^2 \cdot 0.5 + 1^2 \cdot 0.3 + 2^2 \cdot 0.1 + 3^2 \cdot 0.05 + 4^2 \cdot 0.05) - 0.85^2 \approx 1.23$, on hem usat la formula expressada en termes de freqüències.

$$s_n^2 = \sum_{i=1}^m x_i^2 fr_{rel}(x_i) - \bar{x}^2.$$

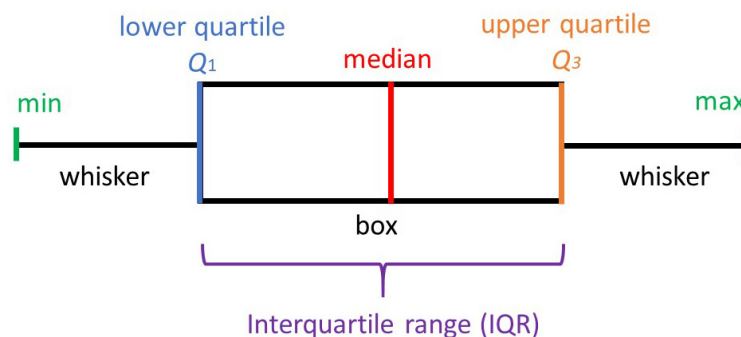
- ▶ Tensió de ruptura: $\bar{x} = 173.33$,

$$s_n^2 = \frac{1}{21} \sum_{i=1}^{21} x_i^2 - \bar{x}^2 = 30087.14 - (173.33)^2 = 43.85, \quad s_n = 6.53,$$

$$s_{n-1} = 6.70 \text{ decagrams-força.}$$

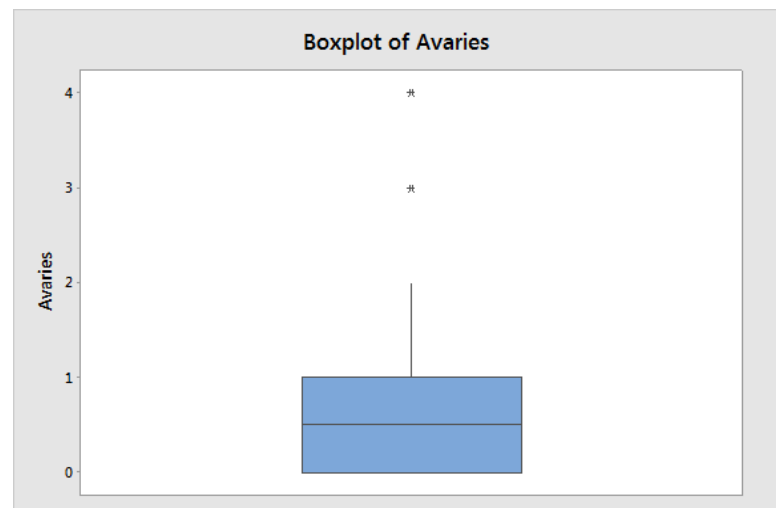
Box plot o diagrama de caixes

1. Càlcul del rang interquartílic: $RIQ = Q_3 - Q_1$.
2. Càlcul dels límits inferiors i superiors:
 - $f_1 = Q_1 - 1.5(Q_3 - Q_1)$, $F_1 = Q_1 - 3(Q_3 - Q_1)$.
 - $f_3 = Q_3 + 1.5(Q_3 - Q_1)$, $F_3 = Q_3 + 3(Q_3 - Q_1)$.
3. Es consideren les observacions adjuntes
 - adj_1 = menor valor observat a $(f_1, Q_1]$.
 - adj_3 = major valor observat a $[Q_3, f_3)$.
4. Si l'asimetria en les dades és gairebé inapreciable, els valors observats a $(F_1, f_1]$ o a $[f_3, F_3)$ es consideren anomalies (outliers) moderades i els observats $(-\infty, F_1]$ o a $[F_3, +\infty)$ es consideren anomalies extremes.



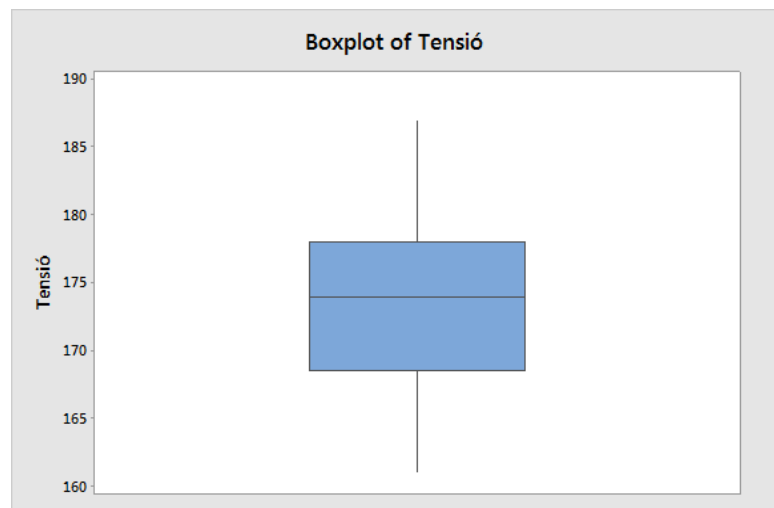
Box plot (examples)

- ▶ $R/IQ = 1$.
- ▶ $f_1 = -1.5$, $f_3 = 2.5$.
- ▶ $adj_1 = 0$, $adj_3 = 2$.
- ▶ Dades anòmales: 3, 4.



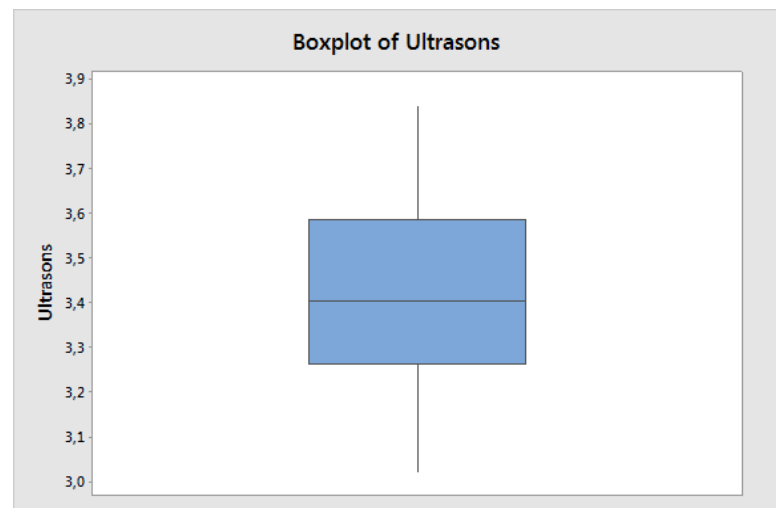
Box plot (examples)

- ▶ $R/IQ = 8$.
- ▶ $f_1 = 157$, $f_3 = 189$.
- ▶ $adj_1 = 161$, $adj_3 = 187$.
- ▶ Dades anòmales: no n'hi ha.



Box plot (examples)

- ▶ $R/IQ = 0.31$.
- ▶ $f_1 = 2.8$, $f_3 = 4.05$.
- ▶ $adj_1 = 3.02$, $adj_3 = 3.84$.
- ▶ Dades anòmales: no n'hi ha.



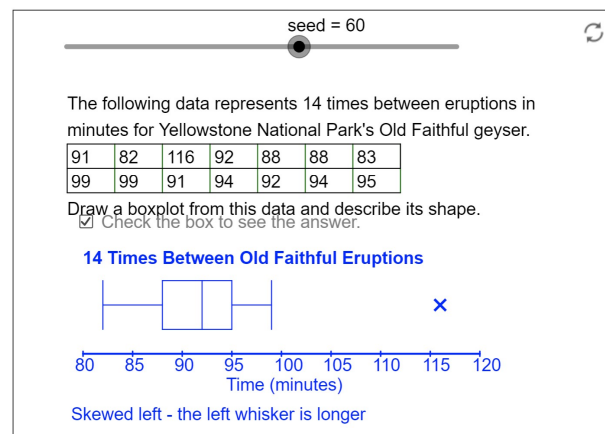
Box plot (applet de Geogebra)

Boxplot Construction

Autor: David Gurney

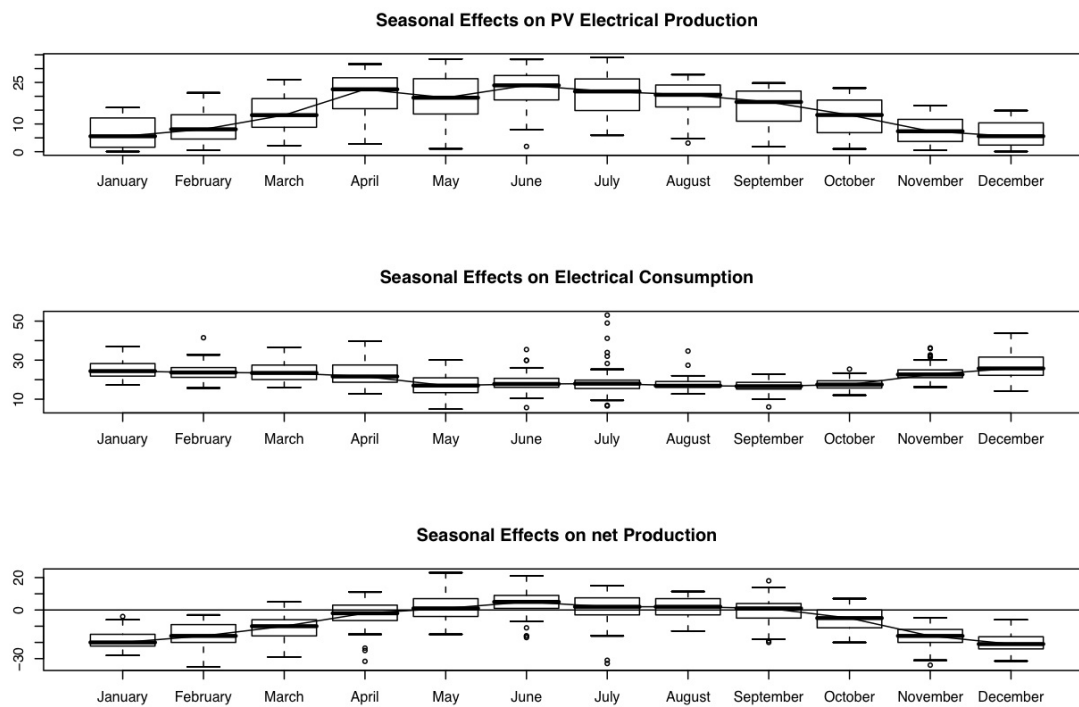
Tema/es: Diagrama de caixa

Move the slider to select a different problem.
Check the check box to see the answer.



<https://www.geogebra.org/m/pQ4Y6Jb4>

Evolució temporal de *box plots*: l'anomalia és relativa



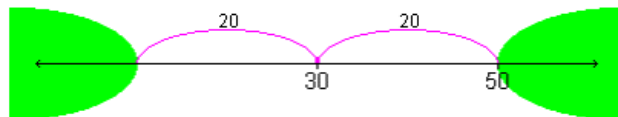
Font: <http://personality-project.org/r/r.plottingdates.html>

Desigualtat de Txeixov

Ens demostra que és molt poc freqüent trobar dades que estiguin “a moltes desviacions” de la mitjana.

$$\text{freq. rel.}\{x_i \text{ t.q. } |x_i - \bar{x}| < k s_n\} \geq 1 - \frac{1}{k^2}$$

Exemple 1: $\bar{x} = 30$, $s_n = 10$, $k = 2 \Rightarrow$
 $\text{fr. rel.}\{x_i \text{ t.q. } |x_i - 30| < 20\} \geq 0.75.$



Exemple 2: Velocitat d'ultrasons, on $\bar{x} = 3.4144$ i $s_n = 0.2021$:

k	Interval	No. dades	% aprox.	predicció Txeixov
1	(3.2123, 3.6165)	22	65%	≥ 0
2	(3.0102, 3.8186)	33	97%	≥ 0.75
3	(2.8081, 4.0207)	34	100%	≥ 0.89

Capítol 2

Estadística descriptiva: problemes

1.7 Problemes.

1. S'ha mesurat el diàmetre de 8 peces fabricades per una màquina automàtica. Els resultats obtinguts han estat:

6.30; 6.19; 6.21; 6.23; 6.21; 6.27; 6.33; 6.20

Calculeu la mitjana i la desviació tipus d'aquestes dades.

Sol: 6.2425; 0.04815

2. S'ha sotmès a una prova de combustió a 140 bles d'un metre de longitud cadascú. Els temps, mesurats en segons, són els següents:

x_i	82	83	84	85	86	87	88	89	90
n_i	3	10	15	24	28	28	17	7	8

- a) Calculeu els paràmetres de la tendència central i de la dispersió.

- b) Dibuixeu la gràfica de les freqüències relatives i de les freqüències acumulades.

*Sol: mitjana=86.135; des.tipus=1.90191; moda= 86 i 87; mediana = 86;
LQ = 85; UQ = 87*

3. En un estudi sobre l'alçada d'una certa població s'han obtingut, en una mostra de 120 individus escollits a l'atzar, les següents dades:

x_i	[1.50-1.54)	[1.54-1.58)	[1.58-1.62)	[1.62-1.66)	[1.66-1.70)	[1.70-1.74)	[1.74-1.78)	[1.78-1.82)	[1.82-1.86)	[1.86-1.90)	[1.90-1.94)
n_i	3	7	10	12	15	23	19	14	8	5	4

Representeu gràficament les dades i calculeu els paràmetres de la tendència central i de la dispersió.

*Sol: mitjana=1.7196; des.tipus=0.094303; moda=1.72; mediana = 1.72;
LQ=1.64; UQ=1.80*

4. Mesurats els diàmetres de 49 eixos s'ha obtingut:

D_i	[5.270-5.280)	[5.280-5.290)	[5.290-5.300)	[5.300-5.310)	[5.310-5.320)	[5.320-5.330)
n_i	5	14	10	7	10	3

Representeu gràficament les freqüències relatives i les acumulades i calculeu els paràmetres de la tendència central i de la dispersió.

Sol: mitjana=5.2974; des.tipus=0.0146427

5. Les mesures de les velocitats amb ultrasons de 34 pilars d'una estructura son:

3.60; 3.21; 3.41; 3.17; 3.40; 3.63; 3.02; 3.02; 3.27; 3.51; 3.60; 3.58; 3.65; 3.33; 3.34;
3.35; 3.35; 3.70; 3.58; 3.24; 3.52; 3.41; 3.14; 3.28; 3.18; 3.35; 3.33; 3.72; 3.53; 3.45;
3.43; 3.84; 3.23; 3.72Km/seg

- a) Calculeu el rang, la mediana, les quartils, el rang interquartil·lic, la mitjana, la desviació tipus, dibuixeu un diagrama de tija i fulles, i el box-plot.

b) Si una dada de valor 3.02 Km/seg passa a ser de 2.65 Km/seg, calculeu tots els paràmetres anteriors; compareu-los amb els obtinguts anteriorment.

Sol: a) Rang=0.82; LQ=3.25; UQ=3.59; mediana=3.405; mitjana=3.4144; IQR=0.33

b) Rang=1.19; LQ=3.25; UQ=3.59; mediana=3.405; mitjana=3.4035; IQR=0.33

6. A una empresa constructora les causes més freqüents de lesions laborals durant un determinat trimestre han estat les següents:

cops al cap	3
dits aixafats	8
peus torçats	5
talls en utilitzar màquines talladores	1
lesions als ulls	4

a) Construïu el corresponent diagrama de Pareto.

Després d'una campanya de seguretat, al següent trimestre els resultats de les lesions laborals han estat:

cops al cap	3
dits aixafats	5
peus torçats	2
talls en utilitzar màquines talladores	1
lesions als ulls	3

b) Construïu el nou diagrama de Pareto (utilitzeu la mateixa escala vertical que en el diagrama anterior).

c) Es pot dir que la campanya de seguretat ha resultat eficaç?

Part II

Probabilitat

Capítol 3

Probabilitat: transparències

Experiment aleatori, espai mostral i succés

- ▶ **Experiment aleatori** (resultat no previsible) vs. **experiment determinista** (resultat previsible).
- ▶ L'**espai mostral**, Ω , és el conjunt de tots els resultats possibles que es poden obtenir al realitzar un experiment aleatori. Cada un d'aquests resultats és un **esdeveniment elemental** ($a, b, c, \dots$).
- ▶ Un **succés** o **esdeveniment** ($A, B, C, \dots$) és qualsevol subconjunt de Ω . El conjunt d'esdeveniments, $\mathcal{A} = \text{Parts}(\Omega)$ inclou Ω (**esdeveniment segur**) i $\emptyset$ (**esdeveniment impossible**).
- ▶ **Exemple 1** (llançament d'un dau):
 - $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\Omega = \{\text{cara}, \text{creu}\}, \dots$ (segons experiment).
 - $\mathcal{A} = \{\Omega, \emptyset, \{1\}, \{2\}, \dots, \{1, 2\}, \{1, 3\}, \dots, \{1, 2, 3\}, \dots, \{1, 2, 3, 4\}, \dots\}$,
 $\mathcal{A} = \{\Omega, \emptyset, \text{cara}, \text{creu}\}, \dots$ (segons experiment).
 - Successos plausibles: $A = \{\text{resultat} \geq 3\}$, $A = \{\text{resultat parell}\}, \dots$

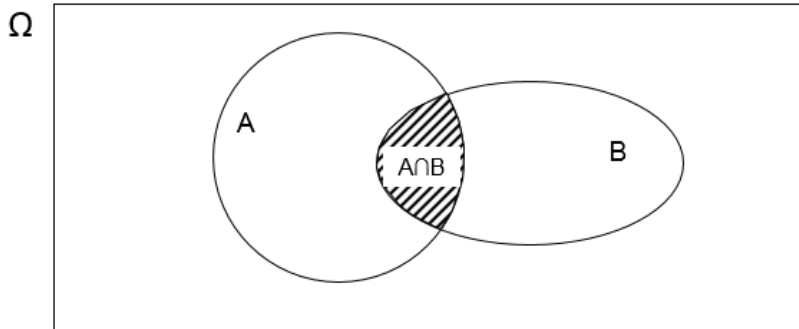
Experiment aleatori, espai mostral i succés: exemples

Ω **espai mostral**; $A, B, C, \dots$ **successos o esdeveniments** plausibles.

- ▶ **Exemple 2:** temps de vida τ d'un component electrònic.
 - $\Omega = [0, +\infty)$ unitats de temps.
 - $A = \{\tau : \tau \geq 600\} = [600, +\infty)$, $B = [0, 200)$,
 $C = [0, 200) \cup (400, \infty), \dots$
- ▶ **Exemple 3:** nombre de partícules emeses per minut per una font radioactiva.
 - $\Omega = \{0, 1, 2, \dots\} = \mathbb{N} \cup \{0\}$.
 - $A = \{\tau : \tau \geq 10\} = \{10, 11, \dots\} = \mathbb{N} \cap [10, +\infty)$,
 $B = \{10, 11, 12, 13, 14, 15\}, \dots$
- ▶ **Exemple 4:** percentatge d'encerts d'un treballador en la classificació de material tèxtil.
 - $\Omega = [0, 1]$.
 - $A = [0.8, 1]$, i qualsevol subconjunt de $[0, 1]$.

Operacions amb successos

- ▶ $A \cup B$ (**unió**) es compleix quan es compleix A o B (almenys un).
- ▶ $A \cap B$ (**intersecció**) es compleix quan es compleixen A i B .
- ▶ $\bar{A}$ (**complementari**) es compleix quan NO es compleix A .
- ▶ $A - B$ es compleix quan es compleix A , però no es compleix B .



A i B són **incompatibles** $\Leftrightarrow$
 $A \cap B = \emptyset$.

- ▶ Exemple 1 (llançament d'un dau): $A = \{\text{parell}\}$, $B = \{3, 6\}$, $C = \{\text{senar}\}$.
 A i C **incompatibles**, A i B **compatibles**,...

Probabilitat: axiomàtica i propietats

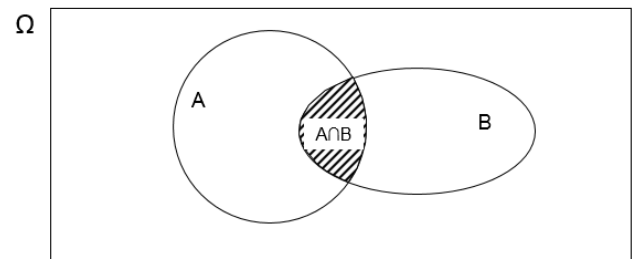
- ▶ Una **funció de probabilitat**, P , és una aplicació $P : \mathcal{A} \rightarrow [0, 1]$ tal que:
 - $P(\Omega) = 1$.
 - $P(A \cup B) = P(A) + P(B)$ sempre que $A \cap B = \emptyset$.
 - Si Ω no és finit, $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ sempre que tots els parells $\{A_i, A_j\}$ siguin disjunts.
- ▶ Direm que $(\Omega, \mathcal{A}, P)$ formen un **espai de probabilitat**.

Propietats:

- ▶ $P(\emptyset) = 0$; $P(\overline{A}) = 1 - P(A)$.
- ▶ Si $A \subseteq B$, aleshores $P(A) \leq P(B)$.
- ▶ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- ▶ $P(A - B) = P(A) - P(A \cap B)$.

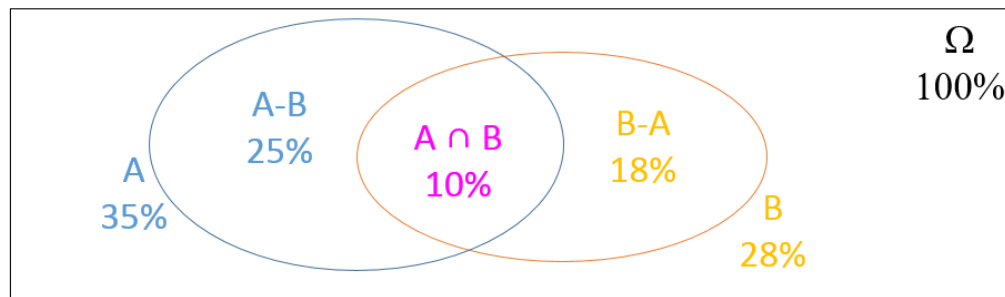
$$\begin{aligned} \text{▶ } P(A_1 \cup \dots \cup A_k) = & \sum_{i=1}^k P(A_i) - \sum_{1 \leq i < j \leq k} P(A_i \cap A_j) + \sum_{1 \leq i < j < s \leq k} P(A_i \cap A_j \cap A_s) + \\ & \dots + (-1)^{k-1} P(A_1 \cap \dots \cap A_k). \end{aligned}$$

Associem **probabilitats** amb **àrees** relatives ($\text{àrea}(\Omega) = 1$).



Probabilitat: exemple

Exemple 5: Suposem que $P(A) = 0.35$, $P(B) = 0.28$ i $P(A \cap B) = 0.1$.



► $P(A \cup B)$?

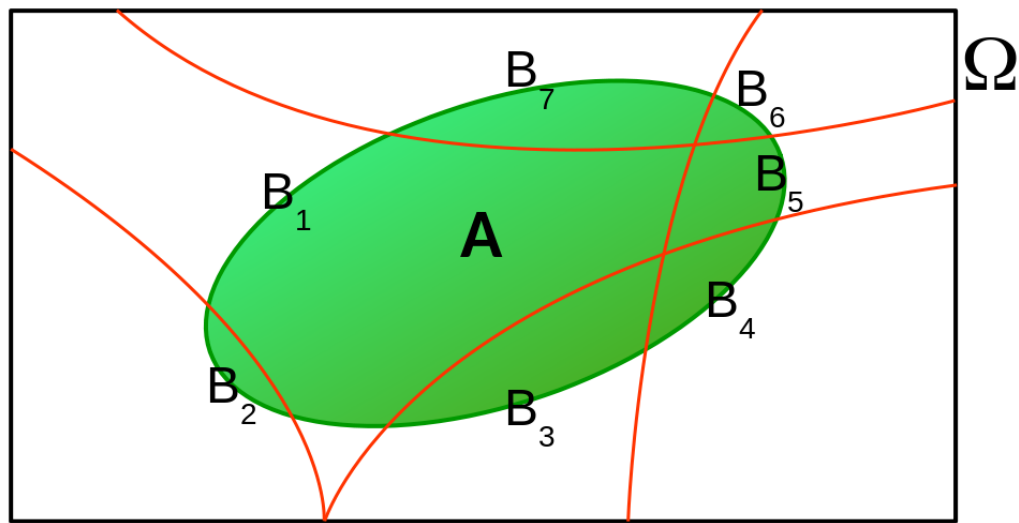
$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.35 + 0.28 - 0.10 = 0.53$$

► $P(A \text{ però no } B)$? $P(A - B) = 0.35 - 0.10 = 0.25$

► $P(\text{només un})$? $P(A - B) + P(B - A) = 0.25 + 0.18 = 0.43$

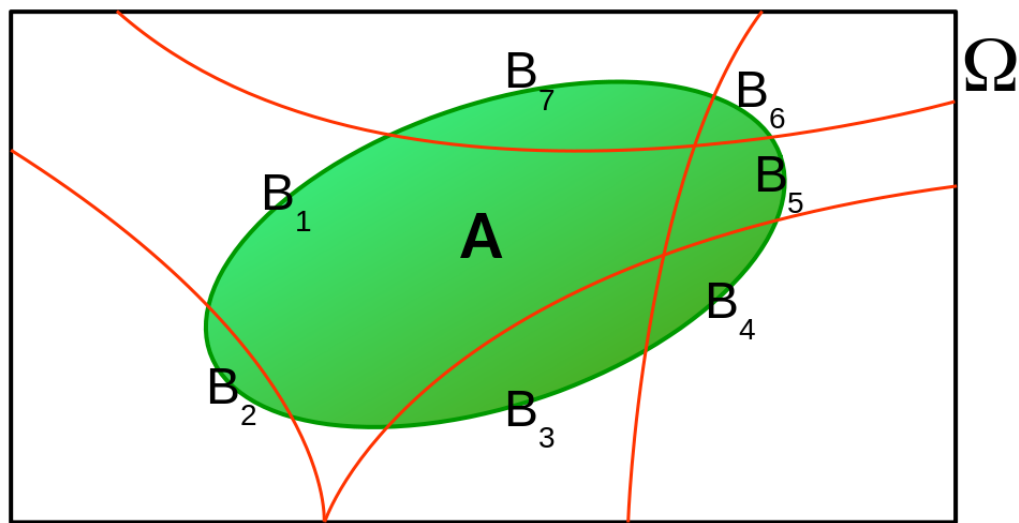
► $P(\text{cap})$? $P(\emptyset) = 1 - P(A \cup B) = 0.47$

Probabilitat condicionada



- Probabilitat **que passi** A ?
- Probabilitat **que passi** B_i ?
- Probabilitat **que passi** A sabent que passa B_i ?

Probabilitat condicionada



- Probabilitat **que passi** A ? $P(A)$
- Probabilitat **que passi** B_i ? $P(B_i)$
- Probabilitat **que passi** A **sabent que passa** B_i ? $P(A|B_i)$

Probabilitat condicionada

- **Probabilitat condicionada** d'A sabent que s'ha donat B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

- Observem que

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A).$$



$$P(A_1 \cap \dots \cap A_m) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots$$

- $P(\cdot|B) : \mathcal{A} \rightarrow [0, 1]$ és una nova **funció de probabilitat**.

Probabilitat condicionada: exemple

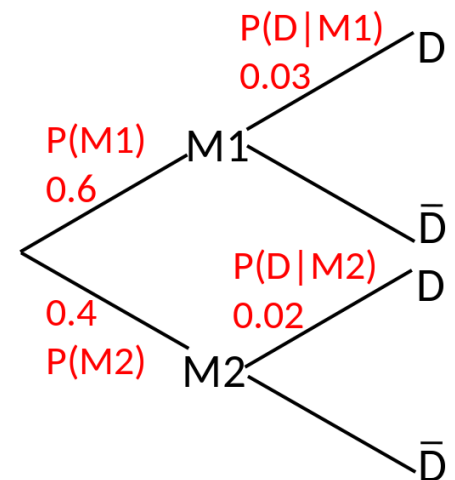
Exemple 6: $M1$ produeix un 60% de peces, $M2$ un 40%. De les que fabrica $M1$, un 3% són defectuoses. De les que fabrica $M2$, un 2% ho són.

- Probabilitat de trobar una peça defectuosa?

$$P(D) = P(M1 \cap D) + P(M2 \cap D) = 0.018 + 0.008 = 0.026.$$

- Probabilitat que una peça sigui de $M1$ sabent que és defectuosa? $P(M1|D) = P(M1 \cap D)/P(D) = 0.018/0.026 = 0.69$.

- D'una mostra de 1000 peces de la màquina $M1$, en prenem dues. Probabilitat que ambdues siguin defectuoses?



$$P(D_1 \cap D_2) = P(D_2|D_1)P(D_1) = \frac{29}{999} \frac{3}{100} = \frac{29}{33300} \approx 0.08787\% \dots$$

Successos independents i teorema de la probabilitat total

- ▶ A i B són **independents** $\Leftrightarrow P(A \cap B) = P(A) P(B)$.
- ▶ A i B són **independents** $\Leftrightarrow P(A|B) = P(A)$ i $P(B|A) = P(B)$.
- ▶ **Atenció!**: **independents** $\neq$ **incompatibles**
- ▶ $A_1, \dots, A_m$ són **conjuntament independents** si per a qualsevol subconjunt $\{A_{i_1}, \dots, A_{i_k}\}$, amb $i_j \in \{1, \dots, m\}$ diferents es compleix $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k})$.

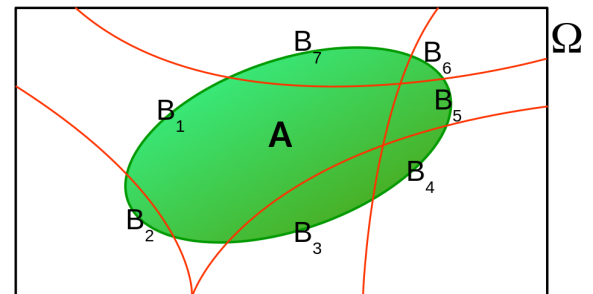
- ▶ **Teorema de la probabilitat total:**

Considerem una partició $\Omega = \bigcup_{i=1}^m B_i$,

on B_i són disjunts dos a dos.

Aleshores,

$$P(A) = \sum_{i=1}^m P(A|B_i) P(B_i).$$



Teorema de Bayes

► Teorema de Bayes:

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)} \quad P(B_i|A) = \frac{P(A|B_i) P(B_i)}{\sum_{j=1}^m P(A|B_j) P(B_j)}$$

Observació: Relacionem una probabilitat condicionada amb la seva contrària.

► Exemple 6 (continuació):

$$P(M1|D) = P(D|M1) P(M1)/P(D) = 0.03 \cdot 0.6 / 0.026 = 0.69.$$

► Exemple 7: Caixa de 100 peces amb 20 de defectuoses; A = primera defectuosa, B = segona defectuosa. Només comprovem la segona peça i veiem que és defectuosa. Quina és la probabilitat que la primera peça NO hagi estat defectuosa?

$$P(\bar{A}|B) = \frac{P(B|\bar{A}) P(\bar{A})}{P(B|A) P(A) + P(B|\bar{A}) P(\bar{A})} = \frac{\frac{20}{99} \frac{80}{100}}{\frac{19}{99} \frac{20}{100} + \frac{20}{99} \frac{80}{100}} = \frac{80}{99}.$$

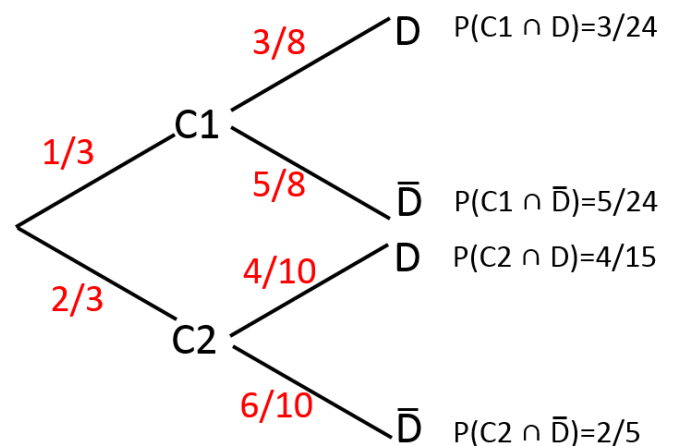
Diagrames d'arbre

Si tenim una sèrie d'experiments seqüencials, cada un d'ells amb m possibles resultats (sovint, $m = 2$), ho podem representar mitjançant **diagrames d'arbre**.

- **Exemple 8:** Tria d'una peça a partir d'un llançament d'un dau.
Si $x < 3$, triem una peça de $C1$; altrament, en triem una de $C2$.
En el subconjunt $C1$, hi ha 8 peces (3 d'elles, defectuoses); en $C2$, 10 peces (4 de defectuoses).

En llançar el dau, quina és la **probabilitat de trobar una peça defectuosa?**

$$P(D) = \frac{3}{24} + \frac{4}{30} = \frac{47}{120} \approx 0.392$$



Referències

- ▶ Presentació del tema “Repàs de probabilitat”. Professorat d’Estadística Aplicada del Grau d’Arquitectura Tècnica i Edificació (EPSEB, UPC).
- ▶ S. Forcada i J. Rubió. “Elements d’estadística”, Capítol 3, Edicions UPC, 2007.

Capítol 4

Probabilitat: problemes

3.5 Problemes.

1. Trobeu l'espai mostral Ω associat a les experiències aleatòries següents i digueu en cada cas si Ω és finit, infinit numerable o infinit no numerable.
 - a) De les famílies que tenen quatre fills comptar el nombre de nois.
 - b) El temps de desgast d'un pneumàtic.
 - c) Escollir tres boles d'una bossa que conté $\{5B, 4N, 2V\}$.
 - d) Nombre de vegades que cal llançar una moneda fins a obtenir tres cares seguides.
2. Considerem l'experiència aleatòria llançar un dau i una moneda. Si E_1 és l'esdeveniment "treure cara" i $E_2 = \{\text{treure 3 o 6}\}$, expresseu en paraules el significat de
 - a) $\overline{E_1}$
 - b) $\overline{E_2}$
 - c) $E_1 \cap E_2$
 - d) $E_1 \cap \overline{E_2} = E_1 - E_2$
3. Per a cadascuna de les experiències aleatòries de l'exercici 1, escriviu dos esdeveniments A, B i expresseu en paraules el significat de $A \cup B, A \cap B, \overline{A}$. Trobeu també esdeveniments impossibles associats a les experiències aleatòries de l'exercici esmentat.
4. Considerem l'experiència aleatòria: observar la puntuació obtinguda al llançar dos daus. I els esdeveniments:

$A = \{\text{Obtenir una suma de punts múltiple de 4}\}$
 $B = \{\text{Obtenir un producte de punts múltiple de 6}\}$
 $C = \{\text{Obtenir igual puntuació en tots dos daus}\}$

Escriviu els esdeveniments: $A, B, C, A \cup B, A \cap B, A \cup C, A \cap C, B \cup C, B \cap C, A \cup B \cup C, A \cap B \cap C$.
5. Demostreu les propietats 2,3 i 4 de la Proposició 1.
6. Es tria a l'atzar una bola d'una caixa que conté 6 boles vermelles, 4 de blanques i 5 de blaves. Trobeu la probabilitat que té la bola escollida de ser
 - a) de color vermell
 - b) de color blau
 - c) de color blanc
 - d) de color diferent de vermell
 - e) de color vermell o blanc.

Sol: a) 6/15; b) 5/15; c) 4/15; d) 9/15; e) 10/15.

7. Trobeu les probabilitats dels esdeveniments

$A, B, C, A \cup B, A \cap B, A \cup C, A \cap C, B \cup C, B \cap C, A \cup B \cup C, A \cap B \cap C$
de l'exercici 4 de l'apartat 1.

8. Tenim dues bosses U_1 i U_2 amb boles blanques i negres i respectivament amb les següents composicions: $U_1 = \{4B, 5N\}$ i $U_2 = \{6B, 4N\}$.

De la bossa U_1 traiem una bola i sense mirar-la es coloca en la bossa U_2 .

- a) Si traiem una bola d' U_2 , quina probabilitat té de ser blanca?
b) I si traiem dues boles d' U_2 , quina probabilitat tenen de ser blanques totes dues?

Sol: a) 0.586; b) 0.321

9. Una enquesta diu que les activitats en el lleure dels habitants d'una ciutat, es reparteix de la forma següent: al 56% els agrada anar a la platja, el 37% els agrada l'esquí i al 14% els agraden les dues coses. Trobeu la probabilitat que, en escollir un ciutadà a l'atzar, a aquest li agradi:

- a) Almenys una de les dues activitats.
b) Anar a la platja però no anar a esquiar.
c) Només una de les activitats.
d) Cap de les dues activitats.

Sol: a) 0.79; b) 0.42; c) 0.65; d) 0.21

10. Demostreu la Fórmula de Bayes a partir de la definició de la probabilitat condicionada.

11. Amb el mateix enunciat de l'exercici 9, trobeu la probabilitat que, en escollir un ciutadà a l'atzar, a aquest li agradi:

- a) Totes dues activitats si sabem que li agrada anar a la platja.
b) No anar a la platja si sabem que es esquiador.

Sol: a) 0.25; b) 0.62

12. S'ha repartit un grup de 200 persones en els conjunts A i B de 90 i 110 persones respectivament. S'observa que hi ha 48 persones del conjunt A i 47 del conjunt B que duen ulleres (U). Es tria una persona a l'atzar d'entre el grup inicial, trobeu

- a) Quina probabilitat té de ser del conjunt B.
b) Quina probabilitat té de portar ulleres i ser del conjunt A.
c) Quina probabilitat té de ser del conjunt B si sabem que porta ulleres. Són independents els esdeveniments {ser del conjunt B} i {portar ulleres}?
d) Quina probabilitat té de ser del conjunt A si sabem que no porta ulleres. Són independents els esdeveniments {ser del conjunt A} i {no portar ulleres}?

Sol: a) 0.55; b) 0.24; c) 0.4947, No; d) 0.4, No

13. La probabilitat que té un atleta de fer un salt de llargada de 6 metres en un intent, és de 0.3. Durant la competició pot fer tres intents. Si no supera aquesta marca queda eliminat.

- a) Calculeu la probabilitat que té de passar la prova durant la competició.
b) Si ens informen que ha superat la prova, quina és la probabilitat que té d'haver-la superat en un màxim de dos intents?

Sol: a) 0.657; b) 0.776

14. Les probabilitats que tenen les persones A, B i C de sortir un cap de setmana són $3/4$, $2/3$ i $3/7$, respectivament. Quina és la probabilitat que tenen de poder sortir junts un cap de setmana?

Sol: 0.214

15. La probabilitat que té una persona de viure d'aquí 25 anys és $\frac{3}{5}$ i la probabilitat que té el seu cònjuge de viure 25 anys més és $\frac{2}{3}$. Trobeu la probabilitat que d'aquí 25 anys:

- a) Visquin tots dos.
- b) En visqui un com a mínim.
- c) Visqui només el primer cònjuge.

Sol: a) $\frac{2}{5}$; b) $\frac{13}{15}$; c) $\frac{1}{5}$

16. Sabem que la producció d'un cert material és defectuosa en un 12%. Suposem que existeix una prova per a detectar-ne la qualitat que no és totalment fiable ja que dona que el material és acceptable en el 10% dels casos de material defectuós i també dona material no acceptable en un 5% de material no defectuós.

- a) Quina és la probabilitat que té el material de ser no defectuós si la prova ha donat que és acceptable?
- b) Si la prova ha donat que el material és acceptable, quina probabilitat té de ser defectuós?

Sol: a) 0.9858; b) 0.0142

17. Tenim tres joiers idèntics, cadascun dels quals té dos calaixos. El primer joier té un rellotge d'or a cada calaix. El segon, té un rellotge de plata a cada calaix. El tercer joier té un rellotge de plata en un calaix i un d'or a l'altre calaix. Escollim un joier a l'atzar, obrim un calaix i resulta contenir un rellotge de plata. Quina és la probabilitat de que en l'altre calaix hi hagi un rellotge d'or?

Sol: $\frac{1}{3}$

Part III

Variables aleatòries discretes

Capítol 5

Variables aleatòries discretes:
transparències

Variables aleatòries, discretes i contínues

- ▶ Determinarem i estudiarem **models de probabilitat** plausibles que representin la variabilitat de la població.
- ▶ Una **variable aleatòria (VA)** és tota aplicació $X : \Omega \longrightarrow \mathbb{R}$ que assigna un nombre real a cada esdeveniment elemental.
- ▶ Si $X(\Omega)$ és numerable es diu que X és **discreta (VAD)**. Si $X(\Omega)$ és no numerable es diu que X és **contínua (VAC)**.
- ▶ **Exemple VA discreta** (llançament de 3 monedes):
 - $X =$ “nombre de cares”.
 - $\Omega = \{ccc, cc+, c+c, +cc, c++ , +c+, ++c, +++\}$
 - $X(\Omega) = \{0, 1, 2, 3\}$
- ▶ **Exemple VA contínua** (llançament de dards a diana):
 - $X =$ “distància del dard al centre de la diana”
 - $X(\Omega) = [0, L)$ ($L =$ radi de la diana)

Funció/distribució de probabilitat d'una VAD

- ▶ La **funció o distribució de probabilitat** associada a una VAD X és l'aplicació $f : X(\Omega) \rightarrow [0, 1]$ tal que $f(x_i) = P(X = x_i)$.
- ▶ Exemple (llançament de 3 monedes):
 - $X =$ “nombre de cares”
 - $\Omega = \{ccc, cc+, c+c, +cc, c++, +c+, +++c, +++\}$
 - $X(\Omega) = \{0, 1, 2, 3\}$

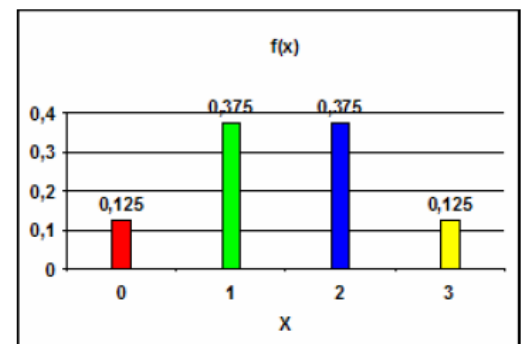
$$f(0) = P(X = 0) = 1/8$$

$$f(1) = P(X = 1) = 3/8$$

$$f(2) = P(X = 2) = 3/8$$

$$f(3) = P(X = 3) = 1/8$$

Hem establert un model teòric de comportament probabilístic.



- ▶ Observació: la suma de tots els valors que pren la distribució ha de ser 1, i.e.: $\sum f(x_i) = 1$.

Funció de distribució d'una VAD

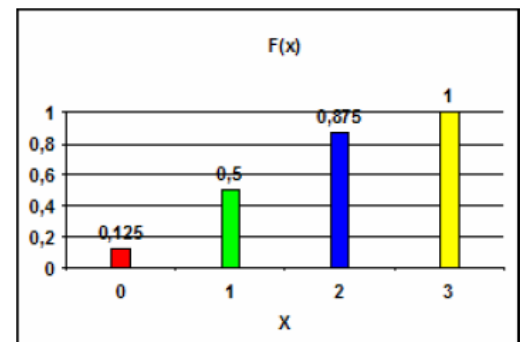
- ▶ La **funció de distribució** associada a una VAD X és l'aplicació $F : X(\Omega) \longrightarrow [0, 1]$ tal que $F(x_i) = P(X \leq x_i)$.
- ▶ Exemple (llançament de 3 monedes):
 - $X =$ “nombre de cares”
 - $\Omega = \{ccc, cc+, c + c, +cc, c + +, +c+, + + c, + + +\}$
 - $X(\Omega) = \{0, 1, 2, 3\}$

$$F(0) = P(X \leq 0) = 1/8$$

$$F(1) = P(X \leq 1) = 1/8 + 3/8 = 1/2$$

$$F(2) = P(X \leq 2) = 1/8 + 3/8 + 3/8 = 7/8$$

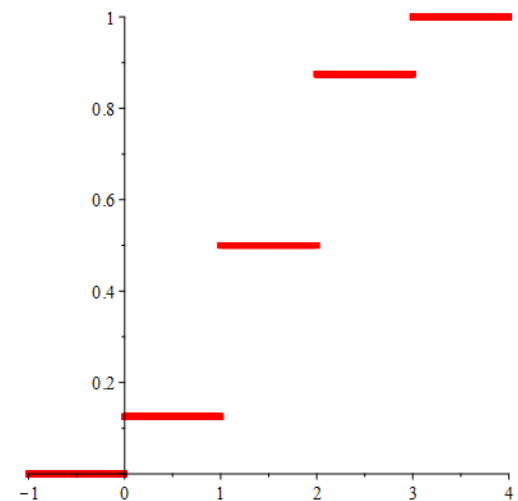
$$F(3) = P(X \leq 3) = 1$$



- ▶ Observació: la funció de distribució és **monòtona creixent** (de 0 a 1) i esglaonada.

Funció de distribució: propietats

- ▶ $0 \leq F(x) \leq 1$ per a tot $x \in \mathbb{R}$.
- ▶ $F(b) - F(a) = P(a < X \leq b)$ per a $a, b \in \mathbb{R}, a \leq b$.
- ▶ F és una funció **contínua per la dreta** en tot $x \in \mathbb{R}$.
- ▶ Tota funció de distribució és **monòtona creixent**.
- ▶ Exemple (llançament de 3 monedes):
 - $P(2 < X \leq 3) = F(3) - F(2) = 1 - 7/8 = 1/8$.
 - $P(2 \leq X \leq 3) = F(3) - F(1) = 1 - 1/2 = 1/2$.

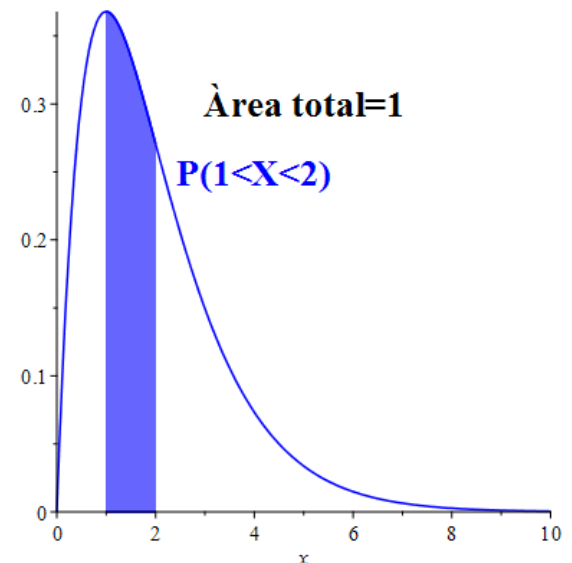


Funció de densitat de probabilitat d'una VAC

- ▶ La **funció de densitat de probabilitat** associada a una VAC X és una aplicació $f : \mathbb{R} \rightarrow [0, 1]$ tal que:
 - $f(x) \geq 0$ per a tota $x \in \mathbb{R}$.
 - $\int_{-\infty}^{+\infty} f(x) dx = 1$.

Observacions:

- ▶ $P(a \leq X \leq b) = \int_a^b f(x) dx$.
- ▶ $P(X = x) = 0$;
 $P(a \leq X \leq b) = P(a < X < b)$.
- ▶ **Exemples:** $f(x) = x e^{-x}$ si $x \geq 0$ i zero altrament.



Funció de distribució d'una VAC

- ▶ La **funció de distribució** associada a una VAC X es defineix com

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(\xi) d\xi.$$

Observacions:

- ▶ $F(x)$ és contínua i, si f és contínua, és derivable.
- ▶ Pel TFC: $F'(x) = f(x)$; $P(a \leq X \leq b) = F(b) - F(a)$.
- ▶ **Exemples:** $f(x) = x e^{-x}$ si $x \geq 0$ i zero altrament. Aleshores:
 $F(x) = 0$ si $x < 0$ i $F(x) = 1 - (x + 1) e^{-x}$ si $x \geq 0$.
 - $P(0 < X < 1) = \int_0^1 f(x) dx = F(1) = 1 - 2 e^{-1} \approx 0.2642$;
 $P(1 \leq X \leq 2) = \int_1^2 f(x) dx = F(2) - F(1) =$
 $1 - 3 e^{-2} - (1 - 2 e^{-1}) = 2 e^{-1} - 3 e^{-2} \approx 0.3298$
- ▶ Vegeu exemples “acadèmics” a pàgines 8–12 dels *Apunts del tema 4 (1a part)*.

Característiques (paràmetres) d'una VA X

	Discretes	Contínues
Esperança o mitjana $E[X] = \mu$	$\sum x_i f(x_i)$	$\int_{-\infty}^{+\infty} x f(x) dx$
Variància $Var[X] = \sigma^2 = E[X^2] - E[X]^2$	$\sum x_i^2 f(x_i) - \mu^2$	$\int_{-\infty}^{+\infty} x^2 f(x) dx - \mu^2$

La **desviació tipus** es defineix, en tots els casos, com a $\sigma = +\sqrt{\sigma^2}$.

Observem que quan més petita sigui σ , més concentrats al voltant de $E[X]$ estaran els valors.

► **Exemple** (VAD: llançament de 3 monedes):

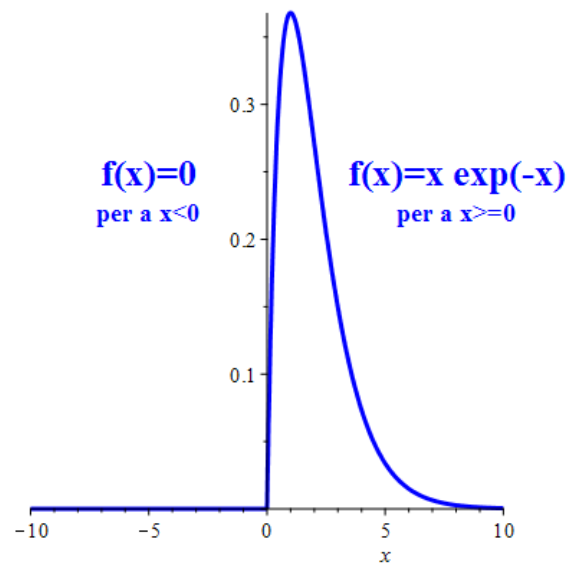
- $X = \text{"nombre de cares"}; X(\Omega) = \{0, 1, 2, 3\}$
- $f(0) = f(3) = 1/8; f(1) = f(2) = 3/8$.
- $E[X] = \sum x_i f(x_i) = 0 \cdot 1/8 + 1 \cdot 3/8 + 2 \cdot 3/8 + 3 \cdot 1/8 = 3/2 = 1.5$.
- $E[X^2] = \sum x_i^2 f(x_i) = 0^2 \cdot 1/8 + 1^2 \cdot 3/8 + 2^2 \cdot 3/8 + 3^2 \cdot 1/8 = 3$.
- $Var[X] = 3 - 1.5^2 = 0.75 \Rightarrow \sigma = 0.866$.

Característiques (paràmetres) d'una VA X

Exemple (VAC): $X =$ “nivell de sucre en sang després d'una ingesta”.

$$f(x) = \begin{cases} 0, & \text{si } x < 0, \\ x e^{-x}, & \text{si } x \geq 0, \end{cases}$$

- ▶ $E[X] = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x^2 e^{-x} dx = -(x^2 + 2x + 2) \exp(-x)]_0^{+\infty} = 2.$
- ▶ $E[X^2] = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^{+\infty} x^3 e^{-x} dx = -(x^3 + 3x^2 + 6x + 6) \exp(-x)]_0^{+\infty} = 6.$
- ▶ $Var[X] = 6 - 2^2 = 2 \Rightarrow \sigma = \sqrt{2} \approx 1.4142.$



Propietats de l'esperança i la variància

Considerem dues variables aleatòries X i Y , i dos nombres reals qualssevol a, b . Aleshores,

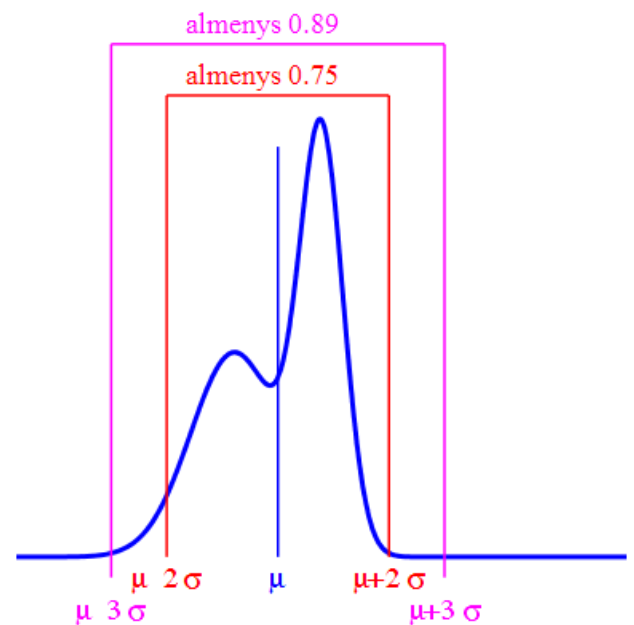
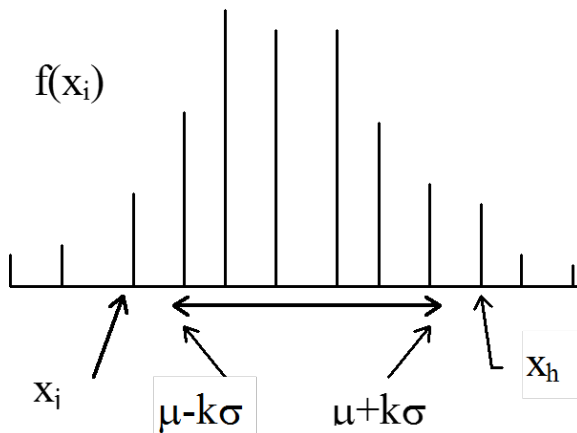
- ▶ $E[aX + b] = aE[X] + b$.
- ▶ $E[X + Y] = E[X] + E[Y]$.
- ▶ $Var[X] = E[(X - E[X])^2]$ (de fet, és la definició original).
- ▶ $Var[aX + b] = a^2 Var[X]$.

Teorema de Txeboxov

► Teorema de Txeboxov:

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}, \quad \forall k \in (0, +\infty).$$

- Interpretació: *és molt poc probable trobar dades que estiguin molt allunyades de la mitjana.*



Teorema de Txeboxov: exemple

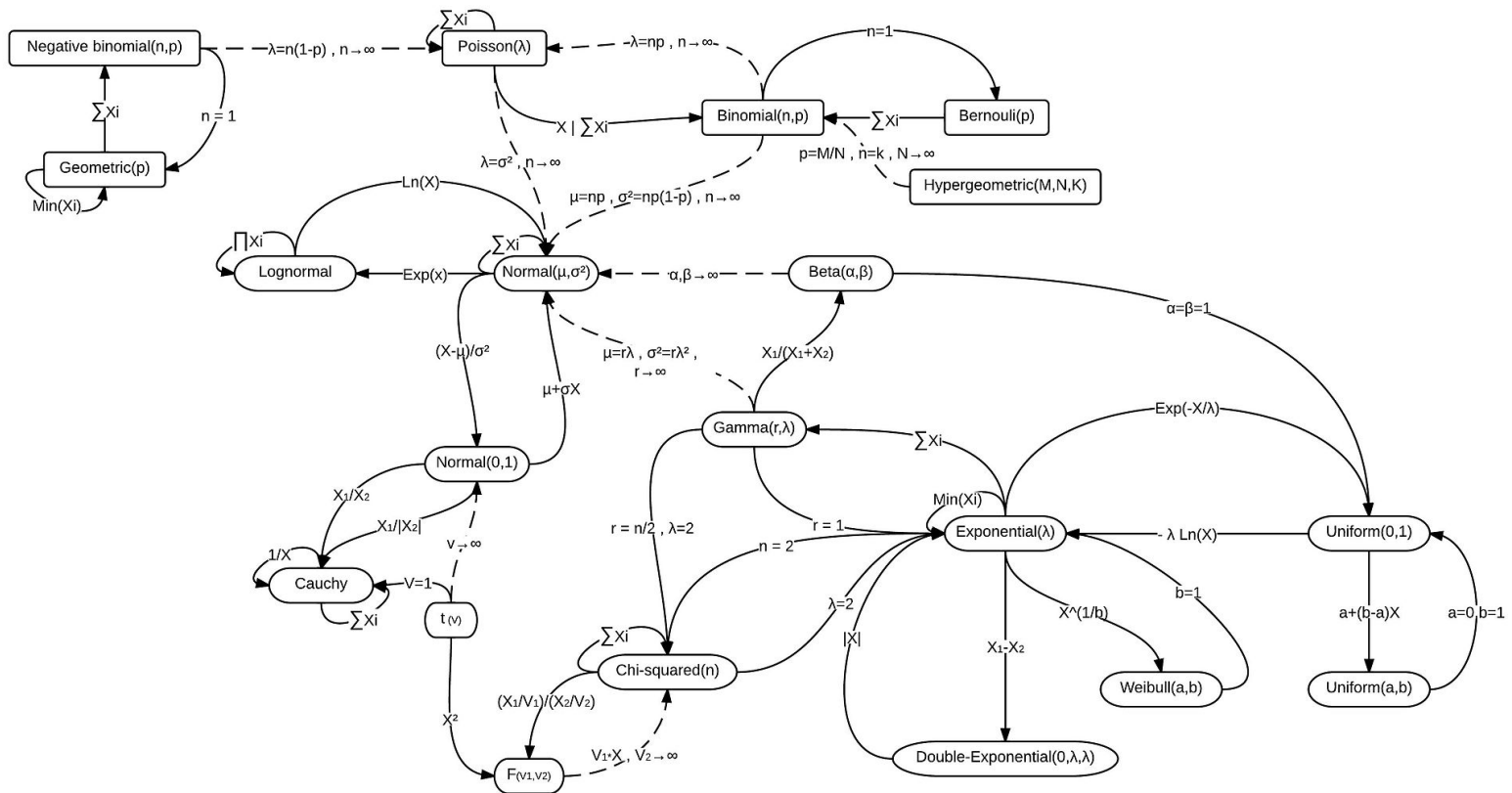
El gruix d'una capa aïllant és una variable aleatòria d'esperança $\mu = 10$ i desviació típica $\sigma = 1$. Es vol **acotar la probabilitat que el gruix sigui inferior a 7.5 o superior a 12.5**.

$$1 - P(7.5 < X < 12.5) < ???$$

Com que $7.5 = \mu - 2.5\sigma$ i $12.5 = \mu + 2.5\sigma$, podem aplicar la **desigualtat de Txeboxov** per a $k = 2.5$. Així:

$$1 - P(7.5 < X < 12.5) \leq 1 - \left(1 - \frac{1}{6.25}\right) = \frac{1}{6.25} \approx 0.16.$$

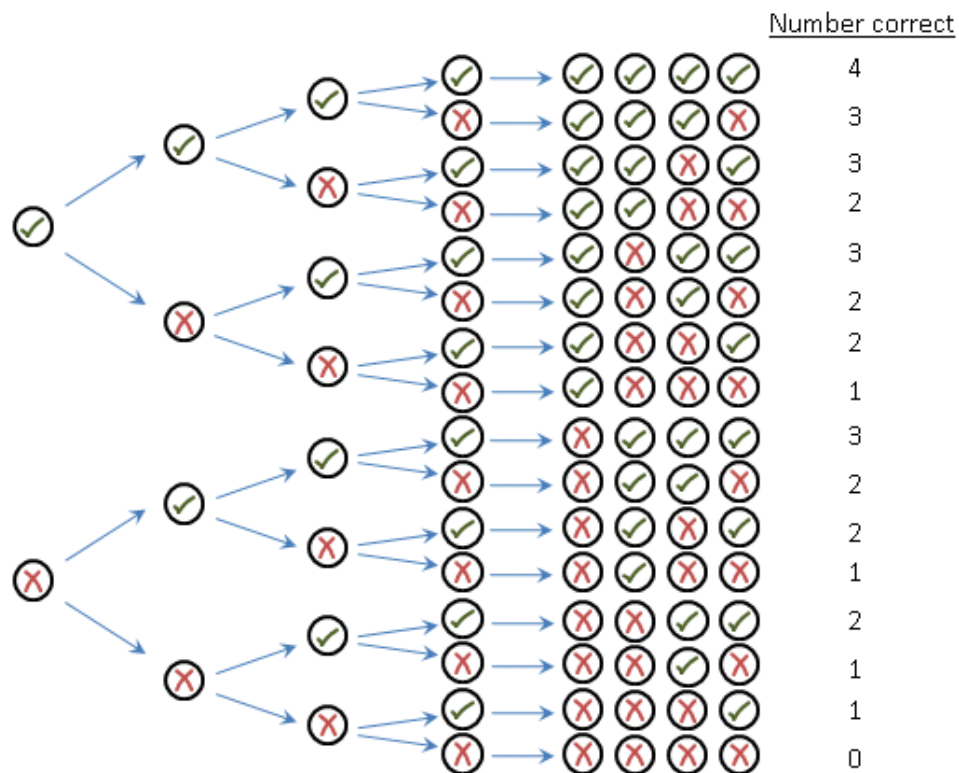
Variables aleatòries amb nom: preàmbul dels propers temes



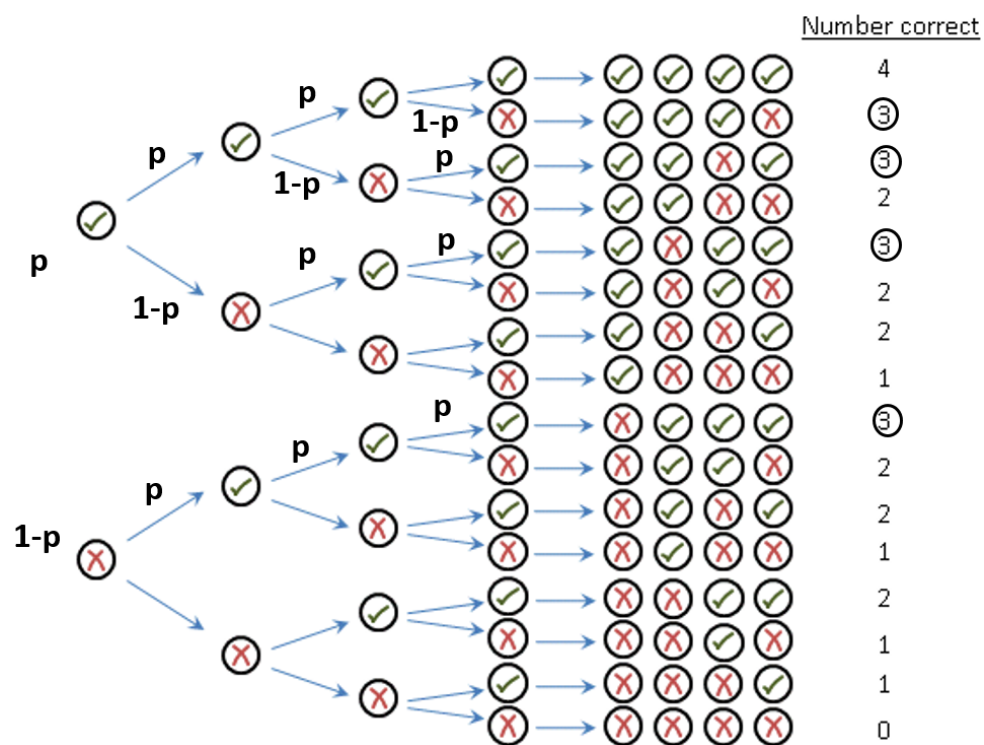
Models discrets de probabilitat

- ▶ Distribució binomial.
- ▶ Distribució geomètrica.
- ▶ Distribució binomial negativa.
- ▶ Distribució hipergeomètrica.
- ▶ Distribució de Poisson.
- ▶ Distribució multinomial.
- ▶ Distribució uniforme discreta.
- ▶ ...

Distribució binomial: exemple 0



Distribució binomial: exemple 0



$X = \# \text{ encerts}; P(X = 3) = 4 p^3 (1 - p)$. I si $n = 3$? I si $n = 5$?...

Distribució binomial

- ▶ **Experiència/assaig de Bernoulli**: experiment aleatori amb 2 possibles resultats, èxit o fracàs (1/0, blanc/negre, on/off...), i amb probabilitat d'èxit constant p . Denotarem la corresponent VAD com $X \sim B(p)$.
- ▶ Si es realitzen n experiències de Bernoulli independents i X és la VA que compta el nombre d'èxits en les n proves, aleshores direm que X segueix una **distribució binomial**, $X \sim B(n, p)$.
- ▶ Si $X \sim B(n, p)$ i $k =$ nombre d'èxits en les n proves, llavors:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

- ▶ $X \sim B(n, p) \longrightarrow E[X] = np, \text{Var}[X] = np(1 - p).$

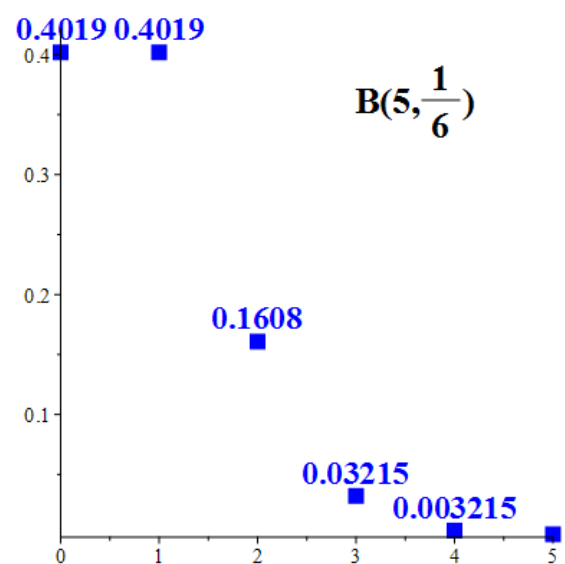
Distribució binomial: exemple 1

Volem estudiar la probabilitat de **treure exactament tres vegades 1 punt** al llançar 5 vegades un dau.

- ▶ $A =$ “treure un 1 en un llançament”.
Experiència de Bernoulli amb $p = 1/6$.
- ▶ Com que el llancem 5 vegades, $n = 5$.
- ▶ $X =$ “treure n vegades 1 en 5 llançaments” és una VAD binomial, $B(5, 1/6)$.

$$P(X = k) = \binom{5}{k} (1/6)^k (5/6)^{5-k}$$

$$P(X = 3) = \binom{5}{3} (1/6)^3 (5/6)^2 = \frac{125}{3888} \approx 0.03215.$$



Distribució binomial: exemple 2

Un sistema conté 40 circuits integrats independents i la probabilitat que un circuit sigui defectuós és 0.01; perquè funcioni el sistema han de funcionar, com a mínim, 35 circuits. Calculem la probabilitat que el sistema funcioni. X = “nombre de circuits que fallen d’un total de 40”: $X \sim B(40, 0.01)$.

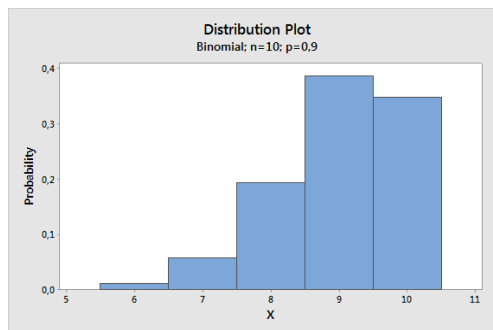
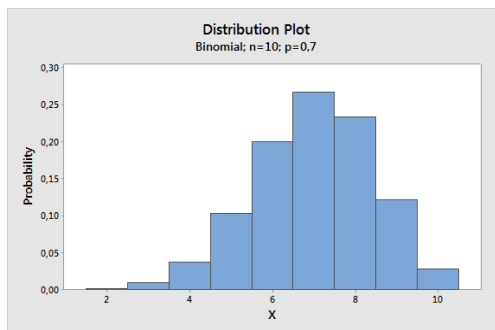
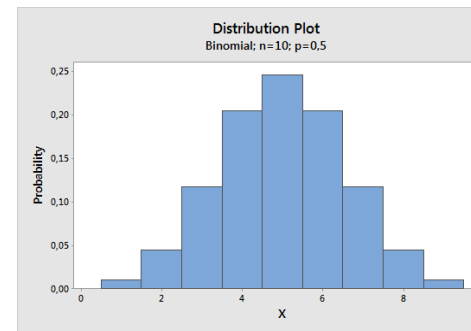
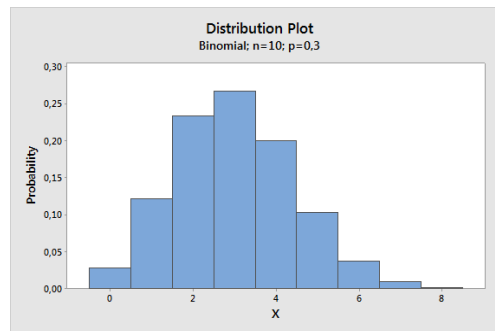
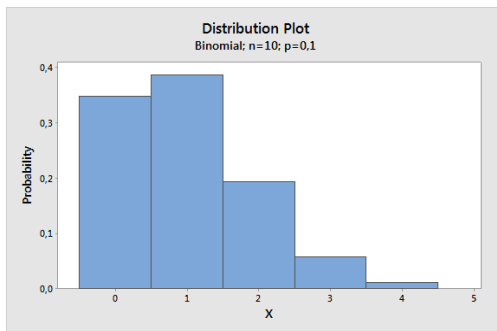


$$P(X \leq 5) = \sum_{k=0}^5 \binom{40}{k} 0.01^k 0.99^{40-k} \approx 0.99997.$$

- ▶ $E[X] = n p = 40 \cdot 0.01 = 0.4$. En mitjana, fallaran 0.4 circuits.
- ▶ $Var[X] = n p (1 - p) = 40 \cdot 0.01 \cdot 0.99 = 0.396$;
 $\sigma = \sqrt{Var[X]} \approx 0.6293$.
- ▶ **Observació:** Si apliquéssim Txeboxov amb $k = 2$, per exemple, tindríem que $P(-0.859 < X < 1.658) \geq 0.75$, predicció que se es millora si emprem el coneixement precís del model de probabilitat. En aquest cas
 $P(-0.859 < X < 1.658) = P(X = 0) + P(X = 1) \approx 0.6690 + 0.2703 = 0.9393$.

$B(n, p)$: variacions del paràmetre p

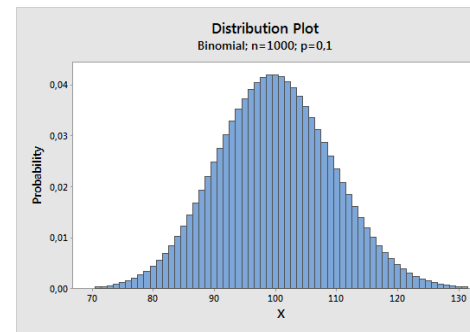
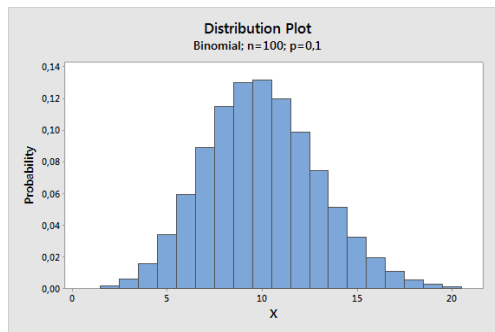
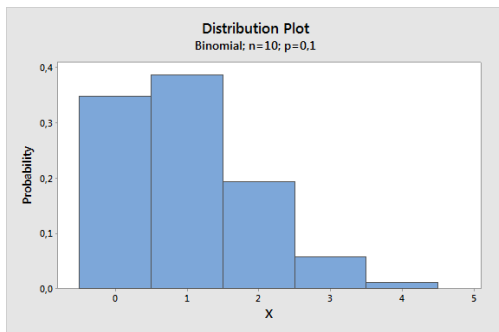
$n = 10$ $p = 0.1, 0.3, 0.5, 0.7, 0.9$



- Observem com la **moda** passa de valors baixos a valors alts
- Observem la **simetria**: $p = 0.5$ simètrica, p i $1 - p$ simètriques entre sí.

$B(n, p)$: variacions del paràmetre n

$p = 0.1$



- Observem com **augmenta la simetria** a l'augmentar n , i va tendint a la forma de *campana*. Hi tornarem en el tema de variables aleatòries contínues.

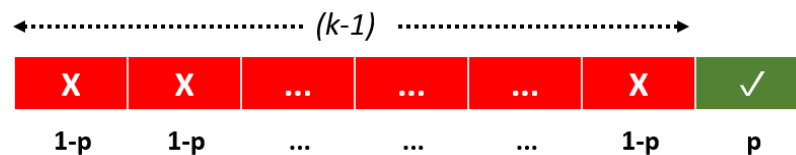
Altres models discrets de probabilitat

- ▶ Distribució geomètrica.
- ▶ Distribució binomial negativa.
- ▶ Distribució hipergeomètrica.
- ▶ Distribució de Poisson.

Distribució geomètrica

Secció 4.3 dels apunts de l'assignatura

- ▶ En el context dels **experiments de Bernoulli independents** interessa també poder valorar probabilísticament quan s'obtindrà el **primer èxit**.



- ▶ Per exemple, al tirar una moneda repetidament quina és la probabilitat que la primera cara no surti abans de la quarta tirada?
- ▶ **Variable aleatòria geomètrica:** $Y \sim G(p)$ = nombre d'experiments de Bernoulli independents amb probabilitat d'èxit p necessaris per a obtenir el primer èxit:

$$P(Y = k) = (1 - p)^{k-1} p, \quad k = 1, 2, \dots$$

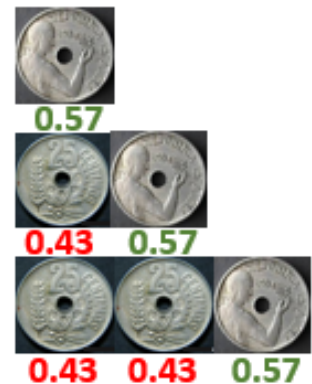
- ▶ $E(Y) = 1/p$; $\text{Var } Y = \frac{1-p}{p^2}$.

Distribució geomètrica

Exemple

- Prenem una moneda tal que la probabilitat de treure cara al tirar-la és 0.57.
- Sigui Y = nombre de tirades independents necessàries per a obtenir la primera cara.
- Si la moneda es tira repetidament, la probabilitat que la primera cara no surti abans de la quarta tirada és:

$$\begin{aligned}P(Y \geq 4) &= 1 - P(Y < 4) = 1 - P(Y \leq 3) \\&= 1 - P(Y = 1) - P(Y = 2) - P(Y = 3) \\&= 0.079507,\end{aligned}$$

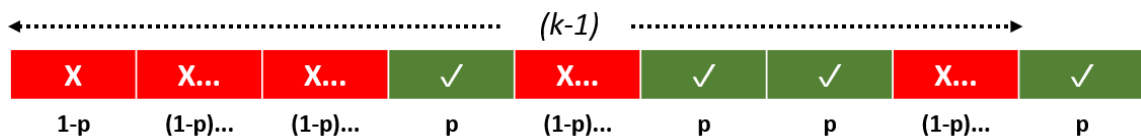


$$\begin{aligned}P(Y = 1) &= p = 0.57, \\ \text{ja que } P(Y = 2) &= (1 - p)p = 0.43 \cdot 0.57 = 0.2451, \\ P(Y = 3) &= (1 - p)^2 p = 0.1849 \cdot 0.57 = 0.105393.\end{aligned}$$

Distribució binomial negativa

Secció 4.4 dels apunts de l'assignatura

- ▶ En el context dels **experiments de Bernoulli independents** interessa també poder valorar probabilísticament quan s'obtindran els **primers r èxits**.



- ▶ Per exemple, al tirar una moneda repetidament, quina és la probabilitat que la 3^a cara no surti abans de la 5^a tirada?
- ▶ **Variable aleatòria binomial negativa:** $N \sim BN(r, p)$ = nombre de proves de Bernoulli independents amb probabilitat d'èxit p , necessàries per a obtenir exactament r èxits:

$$P(N = k) = \binom{k-1}{r-1} (1-p)^{k-r} p^r, \quad k = r, r+1, \dots$$

- ▶ $E(N) = r/p$; $\text{Var}(N) = r(1-p)/p^2$.

Distribució binomial negativa

Exemple

- ▶ Prenem una moneda tal que la probabilitat de treure cara al tirar-la és 0.57.
- ▶ Sigui N = nombre de tirades independents necessàries per a obtenir les 3 primeres cares, és a dir $N \sim BN(3, 0.57)$.
- ▶ Vegem que la probabilitat que les tres primeres cares no surtin abans de la cinquena tirada és aproximadament 0.57591:

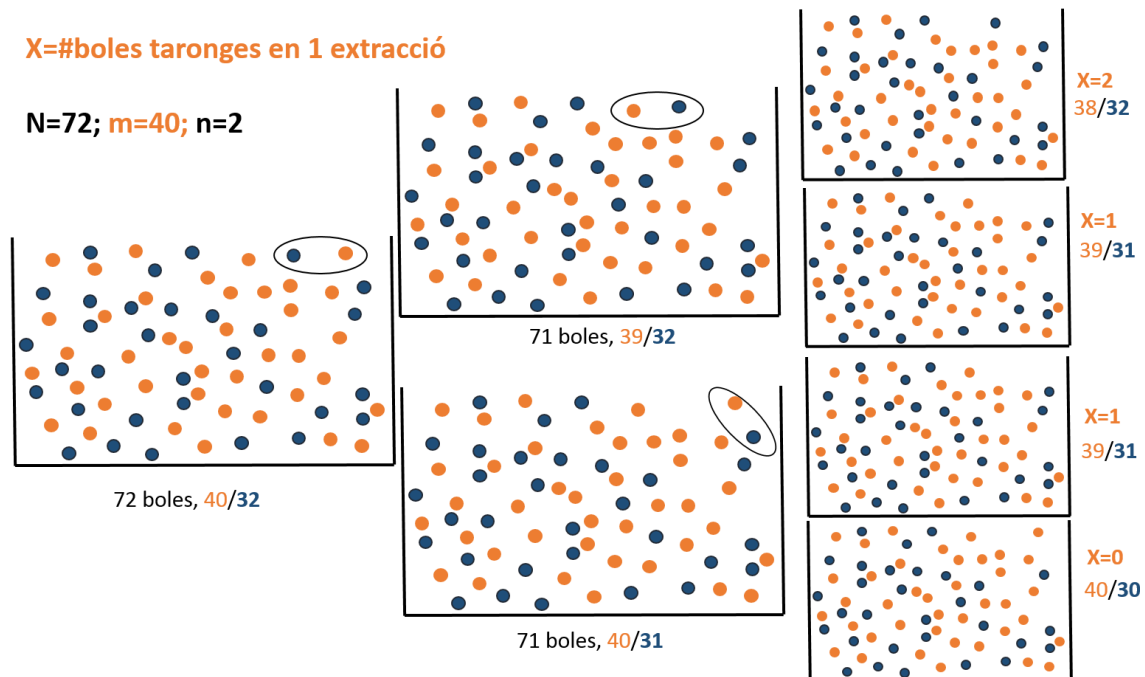
$$P(N \geq 5) = 1 - P(N < 5) = 1 - P(N = 4) - P(N = 3), \quad \text{on}$$

$$P(N = 3) = \binom{3-1}{3-1} (1-p)^{3-3} p^3 = 0.57^3 \approx 0.18519,$$

$$P(N = 4) = \binom{4-1}{3-1} (1-p)^{4-3} p^3 = 3 \cdot 0.43 \cdot 0.57^3 \approx 0.23890.$$

Distribució hipergeomètrica

Explicació gràfica



Distribució hipergeomètrica

- ▶ Una variable aleatòria **hipergeomètrica** $X \sim H(N, m, n)$ descriu el nombre d'èxits en una seqüència de n extraccions **sense reposició**¹ d'una població finita de mida N amb probabilitat d'èxit m/N .
- ▶ La seva probabilitat s'expressa per

$$P(X = k) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}}, \quad \text{amb } k = 0, \dots, \min(m, n).$$

Observem que depèn de 3 paràmetres (N , m i n) tals que $N \in \mathbb{N}$, $m = 0, 1, \dots, N$ i $n = 1, \dots, N$.

¹Es contraposa a la binomial, que considera extraccions amb reposició.

Distribució hipergeomètrica

$$P(X = k) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}}, \quad \text{amb } k = 0, \dots, \min(m, n).$$

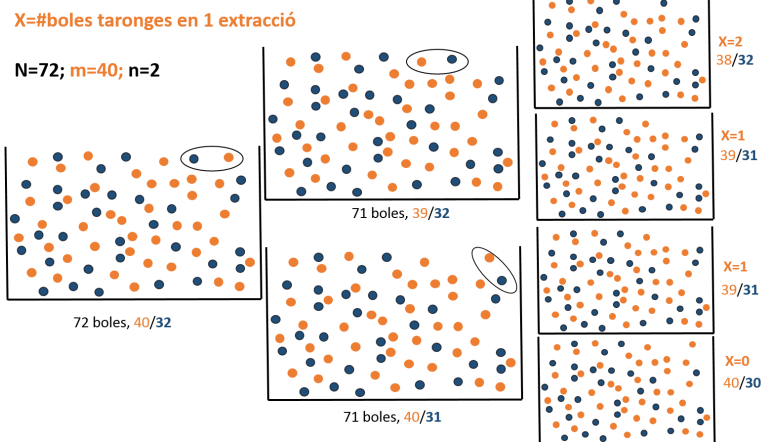
- **Explicació de la fórmula:** tenim $\binom{N}{n}$ extraccions possibles (sense reposició), hi ha $\binom{m}{k}$ formes d'obtenir k èxits i existeixen $\binom{N-m}{n-k}$ formes de combinar els fracassos.
- $E(X) = \frac{m n}{N}$;
 $\text{Var}(X) = \frac{m n (N-m) (N-n)}{N^2 (N-1)} = \frac{m n}{N} \left(1 - \frac{m}{N}\right) \frac{N-n}{N-1}.$
- Un exemple clàssic és el cas d'**extreure n boles** (sense reposició) **d'una urna que conté N boles en total, de les quals m són del color C_1 (=èxit).**

Distribució hipergeomètrica

Exemple

- Una urna conté 72 boles, 40 de taronges, 32 de blaves. Si es treuen 2 boles (de cop o una per una sense reposició), quina és la probabilitat que k d'elles siguin taronges i $n - k$ siguin blaves?

$$P(k \text{ taronges i } 2 - k \text{ blaves}) = \frac{(\# \text{mostres amb } k \text{ taronges}) (\# \text{mostres amb } 2-k \text{ blaves})}{\# \text{mostres de 2 boles}}$$



$$= \frac{\binom{40}{k} \binom{32}{2-k}}{\binom{72}{2}}$$

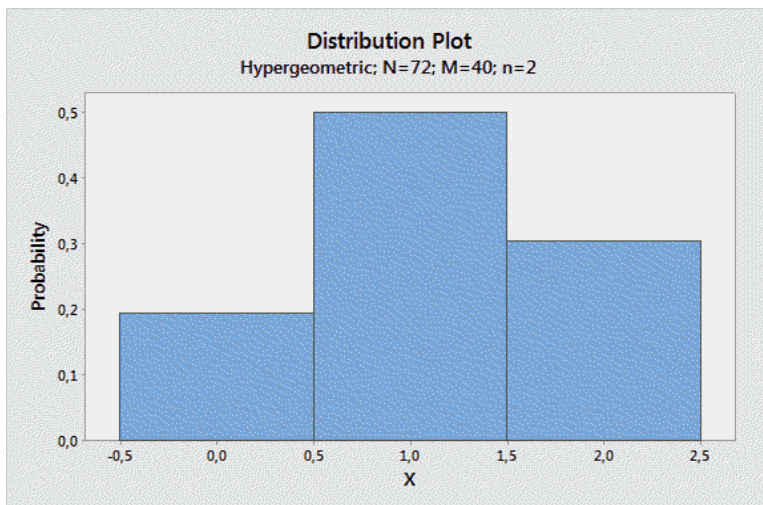
$$= \begin{cases} 0.194 & \text{si } k = 0, \\ 0.501 & \text{si } k = 1, \\ 0.305 & \text{si } k = 2. \end{cases}$$

Distribució hipergeomètrica

Exemple

- Una urna conté 72 boles, 40 de taronges, 32 de blaves. Si es treuen 2 boles (de cop o una per una sense reposició), quina és la probabilitat que k d'elles siguin taronges i $n - k$ siguin blaves?

$$P(k \text{ taronges i } 2 - k \text{ blaves}) = \frac{(\# \text{mostres amb } k \text{ taronges}) (\# \text{mostres amb } 2-k \text{ blaves})}{\# \text{mostres de 2 boles}}$$

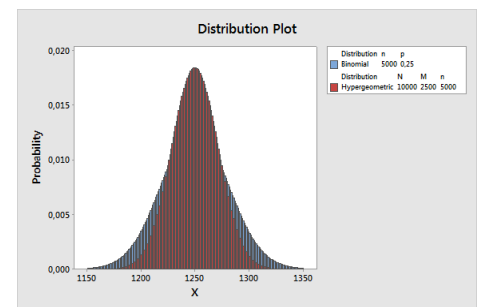
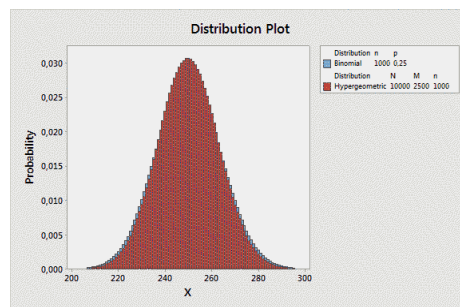
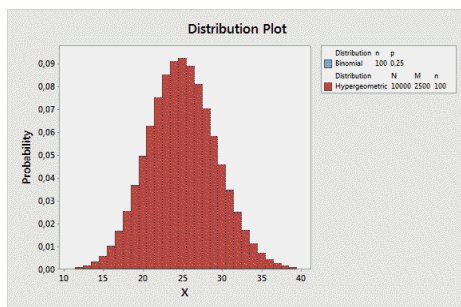


$$= \frac{\binom{40}{k} \binom{32}{2-k}}{\binom{72}{2}}$$
$$= \begin{cases} 0.194 & \text{si } k = 0, \\ 0.501 & \text{si } k = 1, \\ 0.305 & \text{si } k = 2. \end{cases}$$

Distribució hipergeomètrica

Aproximació per una binomial

- La distribució Hipergeomètrica(N, m, n) es pot aproximar per una distribució binomial $B(n, m/N)$ si N és prou gran i $n \ll N$. Empíricament, es considera una aproximació vàlida per a $N \geq 50$ i $n/N \leq 0.1$.



Hipergeomètrica vs binomial per a $N = 10^4$, $p = 0.25$ ($m = 2500$) i $n \in \{100, 1000, 5000\}$.
Veiem que en créixer n , l'aproximació perd validesa.

Distribució de Poisson

Secció 4.6 dels apunts de l'assignatura

- ▶ **Experiència de Poisson:** experiment aleatori consistent en enregistrar el nombre d'incidències d'un esdeveniment puntual i independent en un interval continu, generalment de temps o d'espai (p.e.: arribades de cotxes a un peatge, arribada de peticions a un servidor web, aparició de fallades en una estructura, ...)
- ▶ La variable aleatòria $X \sim P(\lambda)$ = “nombre d'ocurrències en un interval determinat” s'anomena **distribució de Poisson de paràmetre λ** , on λ és la mitjana d'observacions per unitat de temps.
- ▶ Per exemple, nombre de cotxes que arriben al llarg de 3 hores, nombre de clients que arriben al llarg d'un matí, ...
- ▶ Si $X \sim \text{Poisson}(\lambda)$ i $k \in \mathbb{N} \cup \{0\}$, llavors:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}; \quad E[X] = \text{Var}[X] = \lambda.$$

Distribució de Poisson

La distribució de Poisson és un model adequat sota les següents hipòtesis:

- ▶ Els esdeveniments es produeixen de manera **independent**.
- ▶ λ és la **mitjana** en què els esdeveniments es produeixen en un cert interval, i assumim que és **constant**.
- ▶ No es poden produir dos esdeveniments al mateix instant.
- ▶ Sempre podrem trobar un interval prou petit en què es produeixi exactament un esdeveniment o no es produeixi.

Distribució de Poisson: exemple

En una línia de control de qualitat es cometien, de mitjana, 4 errors diaris en la classificació d'un producte.

Observem que la variable X = "errors diaris en la classificació d'un producte" segueix una **distribució de Poisson de paràmetre $\lambda = 4$** .

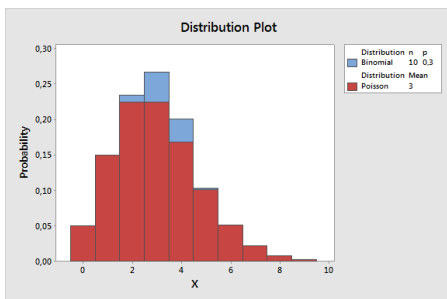
- Quina és la probabilitat que en un dia es cometin 4 o més errors?

$$\begin{aligned}P(X \geq 4) &= P(X > 3) = 1 - P(X < 4) \\&= 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) \\&= 1 - e^{-4} \left(\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!} + \frac{4^3}{3!} \right) \\&= 1 - e^{-4} \left(1 + 4 + 8 + \frac{32}{3} \right) \approx 0.5665.\end{aligned}$$

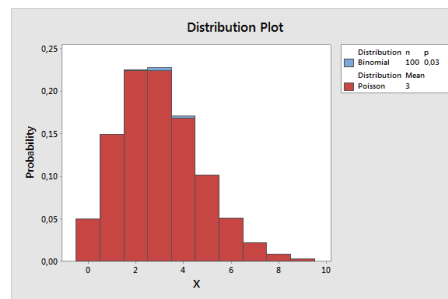
- Quina és la probabilitat que en un dia no es cometin cap error?

$$P(X = 0) = e^{-4} 4^0 / 0! = e^{-4} \approx 0.01832.$$

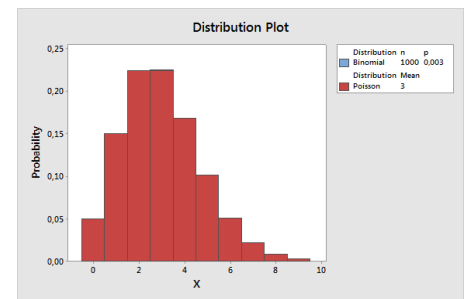
Aproximació d'una distribució binomial per una Poisson



$$n = 10, p = 0.3$$
$$\lambda = 3$$



$$n = 100, p = 0.03$$
$$\lambda = 3$$



$$n = 1000, p = 0.003$$
$$\lambda = 3$$

- Observem que a mesura que augmentem n (i baixem p es van assemblant més.
- Es pot demostrar que acaben coincidint quan $n \rightarrow +\infty$, $p \rightarrow 0$ i $np \sim 1$, és a dir, $B(n, p) \sim P(\lambda)$, amb $\lambda = np$.
- Regles empíriques:
 - $n \geq 30$, $p \leq 0.1$.
 - $n \geq 20$, $p \leq 0.05$.
 - $n \geq 100$, $np \leq 10$.

Aproximació d'una distribució binomial per una Poisson

Exemple

Una companyia d'assegurances sap que hi ha un 0.005% d'incendis en una determinada ciutat cada any. Quina és la probabilitat que té la companyia d'haver de pagar a més de 2 clients dels 100 000 que té assegurats?

- ▶ Si $X =$ “nombre d'assegurats que pateix un incendi”, tenim que $X \sim B(10^5, 5 \cdot 10^{-5})$ ja que $p = 0.005\% = 0.00005$.
- ▶ $P(X > 2) = 1 - P(X \leq 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$, però observem que cal fer operacions *escabroses* com ara $P(X = 2) = \binom{10^5}{2} (0.00005)^2 (0.99995)^{99998}$.
- ▶ Aproximem per una Poisson ja que es compleix alguna (totes de fet) de les regles empíriques. Determinem $\lambda = np = 10^5 \cdot 5 \cdot 10^{-5} = 5$.
- ▶ Llavors:

$$\begin{aligned} P(X > 2) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - e^{-5} \left(\frac{5^0}{0!} + \frac{5^1}{1!} + \frac{5^2}{2!} \right) \approx 0.8753. \end{aligned}$$

Treball autònom

De la llista d'exercicis:

- ▶ **1–3**: Variables aleatòries discretes, en general.
- ▶ **4, 5a-b, 6, 14**: Binomial.
- ▶ **5c, 7, 9a**: Geomètrica.
- ▶ **9b, 10, 15**: Binomial negativa.
- ▶ **8**: Hipergeomètrica.
- ▶ **11–13**: Poisson.

Referències

- ▶ Presentació del tema “Repàs de probabilitat”. Professorat d’Estadística Aplicada del Grau d’Arquitectura Tècnica i Edificació (EPSEB, UPC).
- ▶ S. Forcada i J. Rubió. “Elements d’estadística”, Capítol 4, Edicions UPC, 2007.

Capítol 6

Variables aleatòries discretes: problemes

4.7 Problemes.

- Un joc consisteix en llançar dos daus i calcular la suma dels punts obtinguts. Es cobra el que aquest número indiqui en pessetes. Participar en aquest joc costa 10 pts.
 - Descriu aquest joc utilitzant una variable aleatòria.
 - Trobeu la distribució de probabilitat d'aquesta variable aleatòria.
 - Trobeu l'esperança i la desviació tipus del joc. Interpreteu els resultats.

Sol: $E(X)=7$, $\sigma_X=2.42$

- Donada la variable aleatòria amb distribució de probabilitat:

X	-2	1	3	5
P	1/7	1/8	1/9	z

Trobeu z, $F(2)$, $E(X)$ i σ_X

Sol: $z = 0.621$, $F(2) = 0.268$, $E(X)=3.28$, $\sigma_X=2.54$

- Quin és el recorregut o conjunt de valors que poden prendre les variables aleatòries següents?
 - X ="nombre de cares obtingudes al llançar 5 monedes"
 - X ="nombre de vegades que s'ha de llançar un dau fins a obtenir un resultat parell"
 - X ="nombre de persones amb ulleres d'un grup de 30"

En els apartats a) i b) trobeu la corresponent funció de distribució de probabilitat.

- Una variable aleatòria pren valors -3,-2,-1,0,1,2,3 i la funció de distribució de probabilitat és $f(x)=k(2.5)^x$ per a aquests valors.
 - Quin és el valor de k?
 - Trobeu $E(X)$ i σ_X .
 - Calculeu $P(0 \leq X \leq 2)$.

Sol: a) $k=0.03846$; b) $E(X)=2.34$, $\sigma_X=1.01$; c) 0.375

- Tres participants en una competició de tir amb arc tenen probabilitats 0.1, 0.2 i 0.3 respectivament de fer diana. Trobeu la funció de distribució de probabilitat de la variable aleatòria "número de dianes aconseguides" quan tots tres fan una tirada.

- Una variable aleatòria segueix una llei binomial del tipus $B(5,0.2)$.

Determineu la funció de distribució de probabilitat, l'esperança i la desviació tipus.

Sol: $E(X) = 1$; $\sigma = \sqrt{0.8}$

- Suposem que els motors d'un avió operen independentment i que la probabilitat que un d'aquests motors té de fallar és 0.3. Un avió pot volar amb seguretat si la meitat dels motors funciona. Determineu si és més probable que acabi feliçment el vol un avió bimotor, o un quadrimotor.

Sol: $P(\text{acabar vol en bimotor})=0.91 < P(\text{acabar en quadrimotor})=0.9163$

8. En una residència d'estudiants hi ha una cambra de bany per cada 10 persones. A les 7 del matí la probabilitat que una persona vulgui dutxar-se és 0.6. Quantes dutxes hi hauria d'haver a la cambra de bany perquè un estudiant en trobi una de lliure el 90% de les vegades que vulgui dutxar-se a les 7 del matí?

Sol: 8 dutxes per cambra de bany

9. El 30% de les caixes de rajoles de bany d'una determinada marca, en tenen alguna de trencada. Si es trien 10 caixes, trobeu:

- a) La probabilitat que exactament 5 caixes tinguin alguna rajola trencada.
b) Quantes d'aquestes 10 caixes és d'esperar que tinguin alguna rajola trencada.

Sol: $P(X=5)=0.103$, $E(X)=3$

10. Una empresa té 10 màquines excavadores iguals. S'ha comprovat que aquest tipus de màquina està avariada un dia de cada cinc.

- a) Quina és la probabilitat que en un mateix dia hi hagi més de dues màquines avariades?
b) Què és més probable que totes funcionin correctament o que tres estiguin avariades.
c) Si les pèrdues diàries per excavadora avariada són de 10000 pts. Calculeu la pèrdua mitjana diària de l'empresa per aquest concepte.

Sol: a) 0.32; b) $P(\text{funcionen totes})=0.1$; $P(3 \text{ avariades})=0.2$; c) 20000 ptes/dia

11. Una companyia d'assegurances sap que hi ha un 0.005% d'incendis domèstics en una determinada ciutat cada any.

- a) Quina és la probabilitat que té la companyia d'haver de pagar a més de 2 clients dels 100000 que té assegurats.
b) Quin és el nombre d'incendis domèstics esperat en aquesta ciutat durant un any?

Sol: a) 0.87; b) 5

12. La mitjana de persones que passen la prova de patró de vaixell esportiu és 0.6. Trobeu la probabilitat que té una persona de passar la prova:

- a) en el tercer intent.
b) abans del quart intent.

Sol: a) 0.096; b) 0.936

13. Certa regió volcànica té una mitjana de 15 erupcions en un període de 50 anys. Determineu la probabilitat que en els propers 50 anys hi hagi de 15 a 20 erupcions.

Sol: 0.45

14. Es va fer un test de qualitat de 400 peces de roba produïdes per una empresa tèxtil. Els resultats obtinguts són:

Nº defectes per peça	0	1	2	3	4	5	6	Total
Nº de peces	92	142	96	46	18	6	0	400

Mostreu que la variable aleatòria X = nombre de defectes per peça segueix, aproximadament, una distribució de Poisson.

Quantes peces, d'una mostra de 1000, podem esperar que tinguin 4 o més defectes?

Sol: 57 peces amb 4 o més defectes.

15. Una empresa d'assegurances té una mitjana de 3 accidents per setmana.
 a) Quina és la probabilitat que, en una setmana, no hi hagi cap accident?
 b) Quina és la probabilitat que, en una setmana, hi hagin més de 3 accidents?
Sol: a) 0.05; b) 0.35
16. Si la probabilitat que un individu pateixi reacció per una determinada vacuna és 0.001, quina és la probabilitat que, d'un total de 2000 individus,
 a) 3 tinguin reacció.
 b) més de 2 tinguin reacció.
Sol: a) 0.18; b) 0.32
17. Una central telefònica rep una mitjana de 240 trucades per hora i pot establir un màxim de 8 connexions per minut. Calculeu la probabilitat de que, en un minut donat,
 a) rebi una sola trucada.
 b) quedi col·lapsada.
Sol: a) 0.07; b) 0.02
18. Una màquina que fabrica un cert tipus de peça, produeix una freqüència relativa de 0.01 peces defectuoses. En una mostra de 100 peces, quina és la probabilitat de tenir-ne 99 de bones?
Sol: 0.37
19. En una empresa, s'ha calculat que la probabilitat que un operari s'accidenti en una jornada laboral és 0.001.
 a) Quina és la probabilitat que, en un any amb 280 dies laborables, un operari tingui un accident?
 b) I que tingui algun accident?
Sol: a) 0.21; b) 0.24
20. Els accidents de treball per any que es produeixen en una fàbrica segueixen una llei de Poisson tal que la probabilitat que n'hi hagi 5 és igual a 16/15 vegades la que n'hi hagi 2. Calculeu
 a) el paràmetre de la llei de Poisson.
 b) el nombre màxim d'accidents anyals amb una probabilitat del 90%
 c) la probabilitat que no hi hagi cap accident en 4 anys.
Sol: a) $\lambda = 4$; b) 6; c) e^{-16}
21. Per un determinat lloc d'una autopista passa una mitjana de 6 vehicles per minut. Si un gos comença a travessar la carretera immediatament després que passi un vehicle, i tarda 10 segons en fer-ho, quina és la probabilitat que passi algun vehicle durant aquest temps?
Sol: $1 - e^{-1} = 0.63$