INTRODUCCIÓ A L'ESTADÍSTICA DESCRIPTIVA I A LA TEORIA DE LES PROBABILITATS

Arnau Mir Jaume Suñer

8 de juliol de 2002

Prefaci

El primer dia de classe, els professors solem donar, a més del contingut de l'assignatura, una bibliografia bastant extensa en la qual els alumnes trobaran tot el que exposam durant el curs, i molt més. Sovint, aquests llibres recomanats són cars i difícils de trobar, i quan són a la biblioteca, normalment els té en préstec qualcú desconegut, moltes de vegades el professor.

D'altra banda, durant tot el curs inundam els alumnes amb llistes i més llistes de problemes dels quals, per motius estrictament de temps, només resolem una petita part. És per tot això que els autors vàrem decidir escriure un llibre que resolgués aquests dos problemes. D'una banda, tenir un llibre que estàs a l'abast dels alumnes i que cobrís el que s'explica a classe, i de l'altra proveir l'alumne d'una gran quantitat d'exercicis resolts que li permetin agafar confiança amb l'assignatura. Aquest ha estat el nostre ànim a l'hora d'escriure aquest llibre.

L'estructura del llibre està dividida en tres parts:

a) un resum teòric, on potser el nom no és del tot adequat, ja que cada capítol es desenvolupa en tota l'extensió que creiem adequada per a la gent a qui es dirigeix, l'únic que es troba a faltar són les demostracions dels resultats que s'hi presenten. De totes formes, si hi apareguessin, el llibre perdria el seu esperit inicial de llibre de problemes;

b) una àmplia col·lecció de problemes resolts, seleccionats pel seu interès didàctic de les llistes de problemes que any rere any s'han entregat als alumnes;

c) una bona col·lecció de problemes proposats que, si s'han treballat els problemes resolts, s'haurien de resoldre sense gaire inconvenient.

El material presentat en aquest llibre és adequat per a un primer curs de Probabilitats en qualsevol carrera universitària, tenint sempre en compte que en l'àmbit teòric no és ben bé un llibre de text, sinó un llibre de consulta, i que la seva utilitat en l'àmbit pràctic sí que és prou extensa.

L'obra es complementa amb el llibre "Introducció a la inferència estadística" dels mateixos autors. Per aquest motiu, hem inclòs un capítol dedicat a l'estadística descriptiva, de manera que els dos llibres formen un material prou adequat per a qualsevol assignatura d'estadística.

Índex

Índex de figures

Part I

Estadística descriptiva

Capítol 1

Variables unidimensionals

1.1 Resum teòric

L'estadística descriptiva és la part de l'estadística que s'encarrega de donar una descripció numèrica, una ordenació i una simplificació de la informació recollida.

S'anomena **població** el conjunt de referència sobre el qual recauran les observacions. Direm **unitat estadística** o **individu** a cada element de la població. Una **mostra** és un subconjunt de la població. Quan l'observació es pot quantificar, els valors numèrics de l'observació s'anomenen **variables estadístiques**. Aquestes poden ser **discretes**, si prenen un conjunt finit o numerable de valors, i **contínues**, si prenen un conjunt no numerable de valors.

Per tal d'ordenar les dades obtengudes en l'observació d'una mostra o població, necessitam conèixer els conceptes següents:

1. **Freqüència absoluta** n_i d'un valor x_i: el nombre de vegades que apareix en l'observació.

2. **Freqüència relativa** f_i d'un valor x_i: la freqüència absoluta dividida pel nombre d'observacions fetes.

3. **Freqüència absoluta acumulada** N_i en el valor x_i: la suma de les freqüències absolutes dels valors menors o iguals al x_i.

4. **Freqüència relativa acumulada** F_i en el valor x_i: la freqüència absoluta acumulada dividida pel nombre d'observacions fetes.

Propietats de les freqüències

Sigui N el nombre total d'observacions fetes i k, el nombre d'observacions diferents.

1. $\displaystyle\sum_{i=1}^{k} n_i = N$.
2. $\displaystyle\sum_{i=1}^{k} f_i = 1$.
3. $N_k = N$.
4. $F_k = 1$.
5. $0 \leq n_i \leq N,\ i = 1, \ldots, k$.
6. $0 \leq f_i \leq 1,\ i = 1, \ldots, k$.
7. $n_i = N_i - N_{i-1},\ i = 2, \ldots, k$.
8. El percentatge corresponent a un valor x_i s'obté com $f_i \times 100$.

Una **taula de freqüències** s'obté ordenant els valors de menor a major i anotant les diferents freqüències al costat de cada valor x_i:

x_i	n_i	N_i	f_i	F_i
x_1	n_1	N_1	f_1	F_1
x_2	n_2	N_2	f_2	F_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_k	N_k	f_k	F_k
Suma $\sum$	N		1	

Quan s'han fet moltes observacions i la variable estadística pren molts de valors diferents, l'estudi es fa tractant d'agrupar els valors de la variable estadística en intervals anomenats **intervals de classe**, que poden ser d'amplitud constant o variable. Aquests intervals es poden elegir de dos tipus: a) o bé intervals semioberts de la forma $[a, b)$ que no s'encavalquin, de manera que l'extrem obert d'un interval coincideixi amb l'extrem tancat del següent i que recobreixin tot el recorregut de la variable; o bé b) quan ens donin intervals que no s'encavalquin però de manera que l'extrem obert d'un interval no coincideixi amb l'extrem tancat del següent, prendrem el punt mitjà entre aquests dos extrems dels intervals consecutius com a nou extrem per als dos intervals.

Direm **marca de classe** al punt mitjà de cada interval. Sempre que vulguem resumir la informació continguda dins cada interval, agafarem la seva marca de classe.

Les diferents representacions gràfiques que permeten observar la informació recollida són:

1. **Diagrama de barres:** posam en l'eix de les abscisses els valors de la variable i damunt cada un una línia perpendicular d'alçada igual a la freqüència absoluta (o relativa) del valor en qüestió.

2. **Histograma:** quan les variables estadístiques estan agrupades en intervals de classe, aixecam damunt cada interval un rectangle d'àrea igual a la freqüència absoluta (o relativa) de l'interval.

3. **Polígon de freqüències:** s'obté unint els extrems superiors de les barres en el diagrama de barres. Si la variable està agrupada en intervals, el polígon de freqüències s'obté unint els punts mitjans de les bases superiors de cada interval.

4. **Diagrama de freqüències acumulades:** posam a les abscisses els valors de la variable i sobre cada un una línia perpendicular d'alçada igual a la freqüència (absoluta o relativa) acumulada. Finalment s'uneixen amb segments horitzontals l'extrem superior de cada línia amb la línia següent.

5. **Polígon de freqüències acumulades:** si la variable està agrupada en intervals, col·locam en les abscisses els intervals de classe, després posam damunt l'extrem dret de cada interval una línia perpendicular d'alçada equivalent a la seva freqüència (absoluta o relativa) acumulada. Finalment, unim els extrems superiors de totes aquestes línies.

De vegades interessa resumir la informació recollida a un o uns quants valors per tal de poder comparar diferents mostres o poblacions. Veurem ara quins són els valors més utilitzats per a això. En tot el que segueix, suposarem que la variable estadística X pren els valors $x_1, \ldots, x_k$ amb freqüències absolutes respectives $n_1, \ldots, n_k$. Sigui $N = n_1 + \cdots + n_k$.

- **Mitjana aritmètica:** $\bar{x} = \dfrac{\sum\limits_{i=1}^{k} x_i n_i}{N} = \sum\limits_{i=1}^{k} x_i f_i.$

- **Mitjana geomètrica:** $\bar{x}_G = \sqrt[N]{x_1^{n_1} \cdots x_k^{n_k}}.$

- **Mitjana quadràtica:** $\bar{x}_Q = \dfrac{\sum\limits_{i=1}^{k} x_i^2 n_i}{N}.$

- **Mitjana harmònica:** $\bar{x}_A = \dfrac{N}{\sum\limits_{i=1}^{k} \frac{n_i}{x_i}}.$

Resulta que $\bar{x}_A \leq \bar{x}_G \leq \bar{x} \leq \bar{x}_Q$.

- **Mediana:** és el valor M_e que, una vegada ordenats els valors de la variable, deixa igual nombre d'observacions a la seva esquerra que a la seva dreta. Si hi ha un nombre parell de valors, la mediana és la mitjana aritmètica dels dos valors centrals.

- **Moda:** és el valor M_d amb freqüència més alta. No té perquè ser única. Si la variable ve agrupada en intervals de classe, parlam de l'interval modal, que en l'histograma correspon al rectangle d'àrea més gran per unitat de base.

- **Quartils:** són tres valors que divideixen les observacions ordenades en quatre parts iguals. El primer quartil $P_{1/4}$ és el que deixa la quarta part de les observacions menors o iguals que ell i les tres quartes parts majors que ell. El segon quartil $P_{2/4}$ és la mediana. El tercer quartil $P_{3/4}$ és el recíproc del primer.

- **Decils:** el decil i-èssim és el valor de la variable que deixa $i/10$ parts de les observacions menors o iguals que ell. Anàlogament es defineix el **percentil** i-èssim com el valor de la variable que deixa $i/100$ parts de les observacions menors o iguals que ell.

- **Variància:** la dóna $S_x^2 = \dfrac{\sum\limits_{i=1}^{k} (x_i - \bar{x})^2 n_i}{N}$. La seva arrel quadrada positiva s'anomena la **desviació típica** o **estàndard**: $S_x = +\sqrt{S_x^2}$.

- **Desviació mitjana:**

  – respecte d'una mitjana p: $D_{M_p} = \dfrac{\sum\limits_{i=1}^{k} |x_i - p| n_i}{N}$.

  – respecte de la mitjana aritmètica: $D_{M_{\bar{x}}} = \dfrac{\sum\limits_{i=1}^{k} |x_i - \bar{x}| n_i}{N}$.

  – respecte de la mediana: $D_{M_{M_e}} = \dfrac{\sum\limits_{i=1}^{k} |x_i - M_e| n_i}{N}$.

- **Coeficient de variació de Pearson:** $CV = \dfrac{S_x}{\bar{x}}$.

- **Coeficient de variació mitjana:**

  – respecte d'una mitjana p: $CVM_p = \dfrac{D_{M_p}}{|p|}$.

  – respecte de la mitjana aritmètica: $CVM_{\bar{x}} = \dfrac{D_{M_{\bar{x}}}}{|\bar{x}|}$.

  – respecte de la mediana: $CVM_{M_e} = \dfrac{D_{M_{M_e}}}{|M_e|}$.

- **Recorregut:** és la diferència R entre el valor màxim que pren la variable i el valor mínim.

- **Recorregut interquartílic:** $R_I = P_{3/4} - P_{1/4}$.

- **Recorregut semiinterquartílic:** $R_{SI} = \dfrac{P_{3/4} - P_{1/4}}{2}$.

- **Moment d'ordre r respecte del paràmetre c:** $M_r(c) = \dfrac{\sum\limits_{i=1}^{k} (x_i - c)^r n_i}{N}$.

- **Moments centrals d'ordre r:** quan $c = 0$: $a_r = \dfrac{\sum\limits_{i=1}^{k} x_i^r n_i}{N}$. Per exemple, $a_1 = \bar{x}$.

- **Moments d'ordre r respecte de la mitjana:** $m_r = \dfrac{\sum_{i=1}^{k}(x_i - \bar{x})^r n_i}{N}$. Per exemple, $m_2 = S_x^2$.

Finalment parlarem de mesures d'asimetria i apuntament. Direm que una distribució de freqüències és **simètrica** si els valors equidistants d'un valor central tenen les mateixes freqüències. En aquest cas, resulta: $\bar{x} = M_e$. En el cas de distribucions unimodals, tenim que $\bar{x} = M_d = M_e$. Una distribució asimètrica pot ser:

- **asimètrica per la dreta** o **positiva**, quan la gràfica de les freqüències presenta cua a la dreta, és a dir, les freqüències descendeixen més lentament per la dreta que per l'esquerra. En aquest cas, resulta: $\bar{x} \geq M_e$. Si la distribució és unimodal, resulta $\bar{x} \geq M_e \geq M_d$.
- **asimètrica per l'esquerra** o **negativa**, en cas recíproc. En aquest cas, resulta: $\bar{x} \leq M_e$. Si la distribució és unimodal, resulta $\bar{x} \leq M_e \leq M_d$.

Els següents coeficients permeten conèixer l'asimetria d'una distribució sense haver-la de representar.

- **Coeficient d'asimetria de Pearson:** (només per a distribucions unimodals) $A_P = \frac{\bar{x} - M_d}{S_x}$.
  - Si $A_P > 0$, la distribució és asimètrica per la dreta.
  - Si $A_P = 0$, la distribució és simètrica.
  - Si $A_P < 0$, la distribució és asimètrica per l'esquerra.
- **Coeficient d'asimetria de Fisher:** $A_F = \frac{m_3}{S_x^3}$.
  - Si $A_F > 0$, la distribució és asimètrica per la dreta.
  - Si $A_F = 0$, la distribució és simètrica.
  - Si $A_F < 0$, la distribució és asimètrica per l'esquerra.

Per acabar, donarem un coeficient d'apuntament o **curtosi**, que ens indica si la gràfica d'una distribució, comparada amb la de la normal (campana de Gauss), és poc apuntada (**platicúrtica** o apuntament baix), punxeguda (**leptocúrtica** o apuntament alt) o mitjanament apuntada (**mesocúrtica**). El coeficient es defineix com:

$$g_2 = \frac{m_4}{S_x^4}.$$

Aleshores resulta:

- $g_2 > 3 \rightarrow$ leptocúrtica.
- $g_2 = 3 \rightarrow$ mesocúrtica.
- $g_2 < 3 \rightarrow$ platicúrtica.

1.2 Problemes resolts

1.1.- Considerem les puntuacions de 50 aspirants a un lloc de treball:

8	11	11	8	9	10	16	6	12	19
13	6	9	13	15	9	12	16	8	7
14	11	15	6	14	14	17	11	6	9
10	19	12	11	12	6	15	16	16	12
13	12	12	8	17	13	7	12	14	12

Trobau la taula de freqüències prenent intervals d'amplada 3. Trobau també l'histograma de freqüències absolutes amb el corresponent polígon de freqüències.

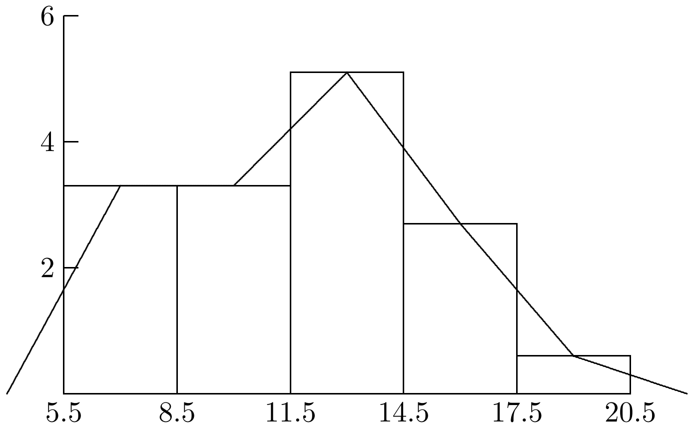

Figura 1.1: Histograma de freqüències absolutes per al problema 1.1

Resolució. El valor màxim és 19 i el mínim, 6. Tenint en compte que consideram límits reals per als intervals de classe, la taula serà:

intervals	X_j	n_j	N_j	f_j	F_j
$[5.5, 8.5)$	7	11	11	0.22	0.22
$[8.5, 11.5)$	10	11	22	0.22	0.44
$[11.5, 14.5)$	13	17	39	0.34	0.78
$[14.5, 17.5)$	16	9	48	0.18	0.96
$[17.5, 20.5)$	19	2	50	0.04	1.00

L'histograma de freqüències absolutes amb el corresponent polígon de freqüències serà com s'indica en el gràfic 1.1.

Fixau-vos que les alçades dels rectangles depenen del fet que l'amplada dels intervals és 3:

$$h_1 = \frac{n_1}{3} = \frac{11}{3} = 3.6666, \qquad h_2 = \frac{n_2}{3} = \frac{11}{3} = 3.666,$$

$$h_3 = \frac{n_3}{3} = \frac{17}{3} = 5.6666, \qquad h_4 = \frac{n_4}{3} = \frac{9}{3} = 3,$$

$$h_5 = \frac{n_5}{3} = \frac{2}{3} = 0.666.$$

1.2.- Considerem les dades següents:

10	5	2	7	9	5	7	6	5	9
12	2	6	6	9	12	6	6	6	4
9	7	12	11						

Trobau la mitjana aritmètica de les dades anteriors sense agrupar en intervals i agrupant les dades en intervals d'amplada 3.
Trobau també la mediana i els percentils $P_{1/4}$ i $P_{3/4}$ de les dades agrupades.

Resolució. La mitjana aritmètica de les dades anteriors sense agrupar en intervals serà:

$$\overline{X} = \frac{10 + 5 + 2 + \cdots + 12 + 11}{24} = \frac{173}{24} = 7.20833$$

Si les agrupam en intervals d'amplada 3, la mitjana serà (fem primer la taula de freqüències):

intervals	X_j	n_j	n_jX_j
$[1.5, 4.5)$	3	3	9
$[4.5, 7.5)$	6	12	72
$[7.5, 10.5)$	9	5	45
$[10.5, 13.5)$	12	4	48
Suma	24		174

La mitjana serà, doncs:

$$\overline{X} = \frac{174}{24} = 7.25$$

Notem que els valors difereixen ja que l'agrupament provoca una pèrdua d'informació.

A continuació trobam la mediana i els percentils $P_{1/4}$ i $P_{3/4}$:

Tenim que $N = 24$. Per tant, $\frac{N}{2} = 12$. L'interval crític, serà, en aquest cas: $[4.5, 7.5)$. La mediana valdrà:

$$Md = 4.5 + 3\frac{12 - 3}{12} = 6.75.$$

Percentil 25: $P_{1/4} \Rightarrow \frac{NP}{100} = 6$. Interval crític:

$[4.5, 7.5)$.

$$P_{1/4} = 4.5 + 3\frac{(6-3)}{12} = 5.25$$

Percentil 75: $P_{3/4} \Rightarrow \frac{NP}{100} = 18$. Interval crític:

$[7.5, 10.5)$.

$$P_{3/4} = 7.5 + 3\frac{(18-15)}{5} = 9.3$$

1.3.- Considerem la distribució de freqüències següent:

intervals	n_j
$[9.5, 29.5)$	38
$[29.5, 49.5)$	18
$[49.5, 69.5)$	31
$[69.5, 89.5)$	20

Trobau la variància i el coeficient de variació respecte de la mitjana.

Resolució. Primer hem de trobar la mitjana.

Per fer-ho, construïm la taula següent:

intervals	X_j	n_j	n_jX_j
$[9.5, 29.5)$	19.5	38	741
$[29.5, 49.5)$	39.5	18	711
$[49.5, 69.5)$	59.5	31	1844.5
$[69.5, 89.5)$	79.5	20	1590
Sumes		107	4886.50

La mitjana valdrà:

$$\overline{X} = \frac{4886.5}{107} = 45.6682.$$

Per trobar la variància hem d'afegir dues columnes més a la taula anterior:

X_j	X_j^2	$n_j X_j^2$
19.5	380.25	14449.50
39.5	1560.25	28084.50
59.5	3540.25	109747.75
79.5	6320.25	126405.00
Suma		278686.75

La variància i la desviació típica valen:

$$
\begin{aligned}
S_X^2 &= \frac{278686.75}{107} - 45.6682^2 = 518.962, \\
S_X &= \sqrt{518.962} = 22.7807.
\end{aligned}
$$

El coeficient de variació val:

$$CV = \frac{S_X}{\overline{X}} = \frac{22.6807}{45.6682} = 0.4988.$$

1.4.- Trobau els coeficients de simetria i apuntament de la distribució de freqüències següent:

intervals	n_j
$[14.5, 19.5)$	4
$[19.5, 24.5)$	6
$[24.5, 29.5)$	8
$[29.5, 34.5)$	11
$[34.5, 39.5)$	35
$[39.5, 44.5)$	100
$[44.5, 49.5)$	218

Resolució. Trobem primer la mitjana i la variància. Per fer-ho construïm la taula següent:

intervals	X_j	n_j	n_jX_j	$n_jX_j^2$
$[14.5, 19.5)$	17	4	68	1156
$[19.5, 24.5)$	22	6	132	2904
$[24.5, 29.5)$	27	8	216	5832
$[29.5, 34.5)$	32	11	352	11264
$[34.5, 39.5)$	37	35	1295	47915
$[39.5, 44.5)$	42	100	4200	176400
$[44.5, 49.5)$	47	218	10246	481562
Sumes		382	16509	727033

La mitjana i la variància valen:

$$\overline{X} = \frac{16509}{382} = 43.2173, \quad S_X^2 = \frac{727033}{382} - \left(\frac{16509}{382}\right)^2 = 35.49.$$

Trobem ara el coeficient d'asimetria A_F. Per fer-ho, hem d'afegir una columna més a la taula anterior:

X_j	n_j	$n_j(X_j - \overline{X})^3$
17	4	-72081.33
22	6	-57308.66
27	8	-34121.16
32	11	-15525.84
37	35	-8411.41
42	100	-180.37
47	218	11799.67
Sumes	382	-175829.09

El moment de tercer ordre val:

$$m_3 = \frac{-175829.09}{382} = -460.285.$$

A continuació calculam el coeficient d'asimetria de Fisher:

$$A_F = \frac{m_3}{S_X^3} = \frac{-460.285}{\left(\sqrt{35.49}\right)^3} = -2.18.$$

Per tant, podem dir que es tracta d'una distribució asimètrica per l'esquerra o asimètrica negativa.

A continuació trobem el coeficient d'apuntament g_2. Per fer-ho, hem d'afegir una columna més a la taula anterior:

intervals	X_j	n_j	$n_j(X_j - X)^4$
$[14.5, 19.5)$	17	4	1889776.11
$[19.5, 24.5)$	22	6	1215933.65
$[24.5, 29.5)$	27	8	553352.38
$[29.5, 34.5)$	32	11	174157.64
$[34.5, 39.5)$	37	35	52296.07
$[39.5, 44.5)$	42	100	219.56
$[44.5, 49.5)$	47	218	44634.88
Sumes		382	3930370.29

El moment de quart ordre val:

$$m_4 = \frac{3930370.29}{382} = 10288.93.$$

A continuació, calculam el coeficient d'apuntament

$$g_2 = \frac{m_4}{S_X^4} - 3 = \frac{10288.93}{35.49^2} - 3 = 5.17.$$

Per tant, es tracta d'una distribució punxeguda o leptocúrtica.

1.3 Problemes proposats

1.1.- A 50 aspirants a un determinat lloc de treball se'ls va sotmetre a una prova. Les puntuacions obtengudes varen ser:

$$\begin{array}{cccccccccc}
4, & 4, & 2, & 10, & 1, & 9, & 5, & 3, & 4, & 5, \\
6, & 6, & 7, & 6, & 8, & 7, & 6, & 8, & 7, & 6, \\
5, & 4, & 4, & 4, & 5, & 6, & 6, & 7, & 5, & 6, \\
6, & 7, & 5, & 6, & 6, & 7, & 5, & 6, & 4, & 3, \\
2, & 6, & 6, & 7, & 7, & 8, & 8, & 9, & 8, & 7.
\end{array}$$

a) Construïu la taula de freqüències i la representació gràfica corresponent.

b) Trobau la puntuació que seleccioni el 20% dels millors candidats.

1.2.- En la població d'estudiants de la facultat es va seleccionar una mostra de 20 alumnes i es varen obtenir les següents talles en centímetres:

162,	168,	174,	168,	166,	170,	168,	166,	170,	172,
188,	182,	178,	180,	176,	168,	164,	166,	164,	172.

Es demana:

a) Descripció numèrica i representació gràfica.

b) Mitjana aritmètica, mediana i moda.

1.3.- Agrupant les dades de l'exercici 1.2 en intervals d'amplada 10 cm., es demana:

a) Descripció numèrica i representació gràfica.

b) Mitjana aritmètica, mediana i moda.

c) Analitzau els càlculs fets i els errors d'agregació i comparau-los amb els de l'exercici anterior.

1.4.- Les tres factories d'una indústria han produït en l'últim any el següent nombre de motocicletes per trimestre:

	factoria 1	factoria 2	factoria 3
1r. trimestre	600	650	550
2n. trimestre	750	1200	900
3r. trimestre	850	1250	1050
4t. trimestre	400	800	650

Obteniu:

a) Producció mitjana trimestral de cada factoria i de tota la indústria.

b) Producció mitjana diària de cada factoria i de tota la indústria tenint en compte que durant el primer trimestre hi va haver 68 dies laborables, durant el segon, 78, durant el tercer, 54 i durant el quart, 74.

1.5.- Una empresa ha pagat per un cert article: 225, 250, 300 i 200 ptes. de preu. Determinau el preu mitjà en les hipòtesis següents:

a) Comprava per valor de 38.250, 47.500, 49.500 i 42.000 ptes., respectivament.

b) Comprava cada vegada un mateix import global.

c) Comprava 174, 186, 192 i 214 unitats, respectivament.

1.6.- D'una mostra de 56 botigues distintes, es varen obtenir els següents preus de venda d'un determinat article:

3260	3510	3410	3180	3300	3540	3320
3450	3840	3760	3340	3260	3720	3430
3320	3460	3600	3700	3670	3610	3910
3610	3610	3620	3150	3520	3430	3330
3370	3620	3750	3220	3400	3520	3360
3300	3340	3410	3600	3320	3670	3420
3320	3290	3550	3750	3710	3530	3500
3290	3410	3100	3860	3560	3440	3620

Es demana:

a) Agrupau la informació en sis intervals d'igual amplada i representació gràfica corresponent.

b) Mitjana geomètrica, quadràtica, aritmètica i harmònica.

c) Desviació de les quatre mitjanes.

1.7.- La següent distribució correspon al capital pagat per les 420 empreses de la construcció amb adreça social en una regió determinada:

Capital (milions de ptes.)	Nombre d'empreses
menys de 5	12
de 5 a 13	66
de 13 a 20	212
de 20 a 30	84
de 30 a 50	30
de 50 a 100	14
més de 100	2

a) Fent servir com a marques de classe del primer i últim interval 4 i 165, respectivament, trobau la mitjana aritmètica i la desviació típica.

b) Calculau la moda i la mediana.

c) Coeficient d'asimetria de Fisher.

Part II

Probabilitats

Capítol 2

Probabilitat elemental

2.1 Resum teòric

Molts de sistemes d'interès presenten fenòmens que mostren variacions impredictibles, degudes a l'atzar. Aquests fenòmens, anomenats **fenòmens aleatoris** o **experiments aleatoris**, es caracteritzen perquè les condicions sota les quals es realitzen no permeten precisar-ne *a priori* el resultat. Per tant, tendrem tot un conjunt de resultats possibles. Per exemple, en el llançament d'un dau, el conjunt de resultats possibles és $\{1, 2, 3, 4, 5, 6\}$. En aquests casos, els models deterministes no són adequats, ja que aquests, cada vegada que es repeteixen les condicions inicials, donen el mateix resultat. Per tant, s'han d'utilitzar els anomenats **models probabilístics**.

D'altra banda, a cada experiment aleatori hi podem associar successos, que són qüestions que es poden contestar amb "sí" o "no" i, una vegada realitzat l'experiment, podrem observar quina és la resposta correcta. En el llançament d'un dau, podem estar interessats, per exemple, en si s'obté un nombre parell. Associat amb "sí" tenim el conjunt de resultats possibles $\{2, 4, 6\}$ i associat amb "no" tenim $\{1, 3, 5\}$.

Finalment, sovint estarem interessats a avançar, abans de realitzar l'experiment, quines opcions té un determinat succés de produir-se (fonament dels jocs d'atzar). Això ho indicam amb frases com "la probabilitat de A és p", on A és un succés com el d'abans, o "el Barça guanyarà la lliga", o "demà farà sol", i p és el nombre o adjectiu que descriu una quantitat tal com "1/2", "alta" o "nul·la".

Així, amb cada experiment aleatori, ens apareixen tres conceptes: el conjunt dels seus possibles resultats, una sèrie de qüestions o successos relacionats amb l'experiment i unes quantitats o adjectius, associats amb cada succés, que anomenam la probabilitat del succés en qüestió i que ens permeten conèixer *a priori* les opcions que té el succés de produir-se. A continuació definirem amb rigor aquests tres conceptes.

Definició 2.1 *Donat un experiment aleatori, direm* **espai mostral** *associat a l'experiment al conjunt de tots els seus resultats possibles. L'indicarem amb* Ω*, i els seus elements, anomenats* **successos elementals**, *amb* $\omega, \omega_i, \ldots$

Definició 2.2 *Donat un conjunt* Ω*, anomenarem* σ**-àlgebra** *de successos sobre* Ω *a tot subconjunt* $\mathcal{F} \subset \Omega$ *tal que:*

1. $\Omega \in \mathcal{F}$.
2. $A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$.
3. $A_1, A_2, \ldots \in \mathcal{F} \Longrightarrow \bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$.

Notes.

1. Anomenarem **succés** a tot element de $\mathcal{F}$. Direm que un succés $A \in \mathcal{F}$ **ha ocorregut** (o ha tengut lloc) si el resultat ω de l'experiment és $\omega \in A$.
2. Observem que els successos sobre Ω són subconjunts de Ω. El fet de dotar la col·lecció de successos de l'estructura de σ-àlgebra és degut al fet que, per poder definir de manera autoconsistent la probabilitat sobre $\mathcal{F}$, necessitam que les operacions de conjunts tals com complementació, unió, intersecció, diferència, etc. siguin operacions internes sobre $\mathcal{F}$.
3. Es pot provar que donada qualsevol col·lecció de successos (les qüestions relacionades amb l'experiment aleatori que ens interessi investigar), existeix la mínima σ-àlgebra que conté la col·lecció, és a dir, que aquesta col·lecció es pot completar afegint el nombre mínim de successos de manera que el nou conjunt de successos obtengut tengui l'estructura de σ-àlgebra.
4. El conjunt buit $\emptyset$ s'anomena el **succés impossible** (mai ocorre) i el conjunt total Ω s'anomena el **succés segur** (sempre ocorre).

Propietats

1. $\emptyset \in \mathcal{F}$.
2. $A_1, A_2, \ldots \in \mathcal{F} \Longrightarrow \bigcap_{i=1}^{\infty} \in \mathcal{F}$.
3. Si $A_1, \ldots, A_n \in \mathcal{F} \Longrightarrow A_1 \cup \cdots \cup A_n \in \mathcal{F}$.
4. Si $A_1, \ldots, A_n \in \mathcal{F} \Longrightarrow A_1 \cap \cdots \cap A_n \in \mathcal{F}$.
5. $A, B \in \mathcal{F} \Longrightarrow A \setminus B = A \cap B^c \in \mathcal{F}$.
6. La σ-àlgebra més petita és $\mathcal{F} = \{\emptyset, \Omega\}$ i la més gran és $\mathcal{F} = P(\Omega)$.

Definició 2.3 *Considerem un conjunt Ω i sigui $\mathcal{F}$ una σ-àlgebra de successos sobre Ω. Una aplicació $p : \mathcal{F} \to [0,1]$ s'anomena una* **mesura de probabilitat** *(o* **probabilitat***) si verifica:*

1. $p\{\Omega\} = 1$.
2. *Si $A_1, A_2, \ldots \in \mathcal{F}$ són disjunts dos a dos ($A_i \cap A_j = \emptyset \quad \forall i \neq j$), aleshores* $p\left\{\bigcup_{n\geq 1} A_n\right\} = \sum_{n\geq 1} p\{A_n\}$.

Propietats

1. $p\{A^c\} = 1 - p\{A\}$.
2. $p\{\emptyset\} = 0$.
3. $A, B \in \mathcal{F} \Longrightarrow p\{A \setminus B\} = p\{A\} - p\{A \cap B\}$.
4. $p\{A \cup B\} = p\{A\} + p\{B\} - p\{A \cap B\}$.
5. $A \subset B \Longrightarrow p\{A\} \leq p\{B\}$.
6. **Continuïtat seqüencial cap a dalt**: Si $A_1 \subset A_2 \subset \cdots$ és una successió creixent de successos de $\mathcal{F}$, definim el seu límit com $A = \lim A_n = \bigcup_{n=1}^{\infty} A_n$. Aleshores resulta que
$$p\{A\} = \lim_{n\to\infty} p\{A_n\}.$$

7. **Continuïtat seqüencial cap a baix**: Si $A_1 \supset A_2 \supset \cdots$ és una successió decreixent de successos de $\mathcal{F}$, definim el seu límit com $A = \lim A_n = \bigcap\limits_{n=1}^{\infty} A_n$. Aleshores resulta que
$$p\{A\} = \lim_{n\to\infty} p\{A_n\}.$$

8. p és **subadditiva**: $A_1, \ldots, A_n \in \mathcal{F} \Longrightarrow p\left\{\bigcup\limits_{i=1}^{n} A_i\right\} \leq \sum\limits_{i=1}^{n} p\{A_i\}$.

Notes.

1. Direm que $A \in \mathcal{F}$ és un **succés nul** si $p\{A\} = 0$. No s'han de confondre els successos nuls amb el succés impossible $\emptyset$: els successos nuls poden ocórrer, encara que tenen probabilitat 0. Per exemple, encertar el centre geomètric d'una diana amb un dard té probabilitat 0, però no és el succés impossible.

2. Quan $\Omega = \{\omega_1, \ldots, \omega_n\}$ és finit i tots els seus elements són successos **equiprobables**: $p\{\omega_1\} = \cdots = p\{\omega_n\} = 1/n$, aleshores donat un succés $A \in \mathcal{F}$, tenim l'anomenada **definició clàssica** de probabilitat :
$$p\{A\} = \frac{card(A)}{n} = \frac{nombre\ de\ casos\ favorables\ a\ A}{nombre\ de\ casos\ possibles}.$$
Aquesta definició falla quan Ω no és finit (quants de llançaments d'una moneda són necessaris per obtenir la primera cara?) i quan no hi ha equiprobabilitat dels successos elementals (en el llançament de dues monedes, la probabilitat d'obtenir una cara és el doble de la d'obtenir-ne dues). De totes formes, quan els requisits es compleixen, és una definició molt útil.

3. D'altra banda, la **definició freqüentista** defineix la probabilitat d'un succés com el límit de les freqüències relatives d'ocurrències del succés en una llarga sèrie de realitzacions de l'experiment:
$$p\{A\} = \lim_{n\to\infty} f_A(n) = \lim_{n\to\infty} \frac{nombre\ de\ vegades\ que\ s'ha\ donat\ A}{nombre\ de\ realitzacions\ de\ l'experiment}.$$
Aquesta definició té un clar problema pràctic i, a més, què passa si l'experiment no és repetible, com a mínim sota les mateixes condicions? De totes formes, en certes situacions serà molt útil.

Definició 2.4 *Direm* **espai de probabilitat** *(o* **model probabilístic***) associat a un experiment aleatori donat a una terna* $(\Omega, \mathcal{F}, p)$*, on:*

- Ω *és l'espai mostral, o conjunt de tots els resultats possibles de l'experiment,*
- $\mathcal{F}$ *és una* σ*-àlgebra de successos sobre* Ω*,*
- p *és una mesura de probabilitat sobre* $\mathcal{F}$*.*

Definició 2.5 *Sigui* B *un succés no nul. Definim la* **probabilitat condicionada de** A **donat** B *com:*

$$p\{A/B\} = \frac{p\{A \cap B\}}{p\{B\}}.$$

Propietats.

1. $P(A \cap B) = p\{B\} \cdot p\{A/B\} = p\{A\} \cdot p\{B/A\}$.
2. En general, donats $A_1, \ldots, A_n \in \mathcal{F}$ tals que $p\{A_1 \cap \cdots \cap A_{n-1}\} > 0$, resulta:

$$p\{A_1 \cap \cdots \cap A_n\} = p\{A_1\} \cdot p\{A_2/A_1\} \cdot p\{A_3/A_1 \cap A_2\} \cdots p\{A_n/A_1 \cap \cdots A_{n-1}\}.$$

3. **Fórmula de les probabilitats totals:**

   Considerem una partició finita (o numerable) de Ω, $\pi = \{A_i\} \subset \mathcal{F}$ ($\Omega = \bigcup\limits_i A_i$, $A_i \cap A_j = \emptyset \ \forall i \neq j$) tal que $p\{A_i\} > 0 \ \forall i$. Donat un succés $A \in \mathcal{F}$, resulta:

$$p\{A\} = \sum_i p\{A \cap A_i\} = \sum_i p\{A_i\} \cdot p\{A/A_i\}.$$

4. **Regla de Bayes:**

   Donada una partició finita (o numerable) de Ω, $\pi = \{A_i\} \subset \mathcal{F}$ de successos no nuls, i $B \in \mathcal{F}$,

$$p\{A_i/B\} = \frac{p\{A_i \cap B\}}{p\{B\}} = \frac{p\{A_i\} \cdot p\{B/A_i\}}{\sum\limits_j p\{A_j\} \cdot p\{B/A_j\}}.$$

En el cas de dos successos A i B, resulta:

$$p\{A/B\} = \frac{p\{A\} \cdot p\{B/A\}}{p\{A\} \cdot p\{B/A\} + p\{A^c\} \cdot p\{B/A^c\}}.$$

Aquesta propietat és útil quan B és un succés observable i A no ho és, però del qual coneixem la seva probabilitat *a priori* $p\{A\}$. Volem conèixer la seva probabilitat *a posteriori* $p\{A/B\}$, una vegada que hem observat B.

5. Si $B \in \mathcal{F}$ és un succés no nul, l'aplicació $p\{\bullet/B\} : \mathcal{F} \to [0,1]$ donada per:

$$p\{\bullet/B\}(A) = p\{A/B\} \quad \forall A \in \mathcal{F}$$

és una nova probabilitat sobre $\mathcal{F}$.

Definició 2.6 *Direm que dos successos* $A, B \in \mathcal{F}$ *són* **independents** *si* $p\{A \cap B\} = p\{A\} \cdot p\{B\}$.

Propietats.

1. Són equivalents:
   (a) A, B independents,
   (b) A, B^c independents,
   (c) A^c, B independents.
2. Ω i $\emptyset$ són independents de qualsevol succés A.
3. Si A és independent d'ell mateix, aleshores o bé $p\{A\} = 0$ o bé $p\{A\} = 1$.

Com ja hem dit, donat un succés no nul B, la probabilitat condicionada $p\{\bullet/B\}$ també és una probabilitat sobre $\mathcal{F}$. Aleshores podem definir el següent tipus d'independència:

Definició 2.7 *Dos successos* $A, B \in \mathcal{F}$ *són* **condicionalment independents** *donat un altre succés no nul* $C \in \mathcal{F}$ *si:*

$$p\{A \cap B/C\} = p\{A/C\} \cdot p\{B/C\}.$$

Nota:

En principi, pareix que hi hauria d'haver relació entre la independència de dos successos i la independència condicionada. Per exemple, si $C = \Omega$, coincideixen les dues independències. De totes formes, en general, no hi ha cap relació entre elles. (Vegeu problemes proposats.)

2.2 Problemes resolts

2.1.- Trobau l'espai mostral per a l'experiment que consisteix a treure dues bolles amb reposició i sense reposició d'una urna que conté 5 bolles (suposau que les bolles estan numerades d'1 a 5).

Resolució. Sigui Ω l'espai mostral. Aleshores tenim que:

- $\Omega = \{(i,j), i = 1,\ldots,5, j = 1,\ldots,5\}$, si l'experiment és amb reposició.
- $\Omega = \{(i,j), i = 1,\ldots,5, j = 1,\ldots,5, j \neq i\}$, si l'experiment és sense reposició.

2.2.- Per al problema anterior, especificau els successos següents:

- $A =$ "Treure exactament una bolla parell"
- $B =$ "Treure la segona bolla parell"

Resolució.

- $A = \{(1,2),(1,4),(2,1),(2,3),(2,5),(3,2),(3,4),(4,1),(4,3),(4,5),(5,2),(5,4)\}$, tant si l'experiment és amb reposició com sense.
-
  - $B = \{(1,2),(1,4),(2,2),(2,4),(3,2),(3,4),(4,2),(4,4),(5,2),(5,4)\}$, si l'experiment és amb reposició.
  - $B = \{(1,2),(1,4),(2,4),(3,2),(3,4),(4,2),(5,2),(5,4)\}$, si l'experiment és sense reposició.

2.3.- Donats $\Omega = \{1, 2, 3\}$, $A = \{1\}$, $B = \{3\}$, $C = \{2\}$ i suposant que $p\{A\} = \frac{1}{3}$, $p\{B\} = \frac{1}{3}$, trobau: $p\{C\}$, $p\{A \cup B\}$, $p\{A^c\}$, $p\{A^c \cap B^c\}$, $p\{A^c \cup B^c\}$ i $p\{B \cup C\}$.

Resolució.

- $p\{C\} = 1 - p\{A\} - p\{B\} = 1 - \frac{1}{3} - \frac{1}{3} = \frac{1}{3}$.
- $p\{A \cup B\} = p\{A\} + p\{B\} = \frac{2}{3}$, ja que $A \cap B = \emptyset$.
- $p\{A^c\} = 1 - p\{A\} = \frac{2}{3}$.
- $p\{A^c \cap B^c\} = p\{(A \cup B)^c\} = 1 - p\{A \cup B\} = \frac{1}{3}$.
- $p\{A^c \cup B^c\} = p\{A^c\} + p\{B^c\} - p\{A^c \cap B^c\} = \frac{2}{3} + \frac{2}{3} - \frac{1}{3} = 1$.
- $p\{B \cup C\} = p\{B\} + p\{C\} = \frac{2}{3}$, ja que $B \cap C = \emptyset$.

2.4.- Es pinta un "1" en una cara d'una moneda i un "2" en l'altra cara. Es tiren tres monedes pintades d'aquesta manera. Trobau la probabilitat que la suma dels tres nombres que surtin sigui 3, que sigui 4, que sigui 5 i que sigui 6.

Resolució. L'espai mostral serà en aquest cas:

$$\begin{aligned} \Omega = \; & \{(1,1,1),(1,1,2),(1,2,1),(1,2,2), \\ & (2,1,1),(2,1,2),(2,2,1),(2,2,2)\} \end{aligned}$$

Fixau-vos que, en aquest cas, tots els successos elementals tenen la mateixa probabilitat de sortir. Així, doncs,

- $p\{\text{“Suma igual a 3”}\} = p\{(1,1,1)\} = \frac{1}{8}$.
- $p\{\text{“Suma igual a 4”}\} = p\{(1,1,2),(1,2,1),(2,1,1)\} = \frac{3}{8}$.
- $p\{\text{“Suma igual a 5”}\} = p\{(1,2,2),(2,2,1),(2,1,2)\} = \frac{3}{8}$.
- $p\{\text{“Suma igual a 6”}\} = p\{(2,2,2)\} = \frac{1}{8}$.

2.5.- En un formiguer hi ha formigues vermelles i negres. En un passadís del formiguer, on l'espai és tan estret que només pot passar una formiga a la vegada, passen 4 formigues una darrera l'altra. Quantes combinacions diferents de color es poden produir suposant que les formigues vermelles no es poden distingir entre elles igualment que les negres?

Resolució. Hem de trobar totes les variacions amb repetició del conjunt {vermelles, negres} presos els elements de 4 en 4. Per tant, el nombre de combinacions serà:

$$VR_2^4 = 2^4 = 16.$$

2.6.- Vint automòbils han de participar en una carrera, 8 dels quals són de la marca A, 7 de la marca B i la resta de la marca C. Si només es mira la marca del cotxe a l'arribada, de quantes maneres poden arribar els cotxes a la meta? De totes les possibles formes d'arribar a la meta, quantes n'hi ha en què un cotxe de la marca A estigui en primer lloc? Quantes en què hi hagi dos cotxes de la marca A en els dos primers llocs?

Resolució.

- Maneres d'arribar a la meta $= PR_{20}^{8,7,5} = \frac{20!}{8! \cdot 7! \cdot 5!} = 99768240$ maneres.
- Maneres en què un cotxe de la marca A arriba en primer lloc $= PR_{19}^{7,7,5} = \frac{19!}{7! \cdot 7! \cdot 5!} = 39907296$ maneres.
- Maneres en què dos cotxes de la marca A arriben en primer lloc $= PR_{18}^{6,7,5} = \frac{18!}{6! \cdot 7! \cdot 5!} = 14702688$ maneres.

2.7.- Trobau la probabilitat que 5 cartes escollides a l'atzar d'una baralla de 52 cartes contenguin:

a) exactament dues parelles.

b) *full* (3 cartes d'un mateix nombre i 2 d'un altre).

c) Flor (5 cartes d'un mateix pal).

d) Correguda (5 cartes en seqüència ascendent).

Resolució.

a) $p\{2 \text{ parelles }\} = \frac{Casos\ favorables}{Casos\ possibles} = \frac{13\cdot\binom{4}{2}\cdot 12\cdot\binom{4}{2}\cdot 11\binom{4}{1}}{\binom{52}{5}} = \frac{247104}{2598960} \approx 0.0951$

L'explicació de la fórmula anterior és la següent:

Primer hem de tenir en compte que hem de pensar en nombres en lloc de cartes.

Així doncs, tenim 13 nombres per formar la primera parella i 4 pals per triar per cada nombre. Per tant, les maneres que tenim per formar la primera parella seran: $13 \cdot \binom{4}{2}$.

Ara ens queden 12 nombres per formar la segona parella i 4 pals per triar per cada nombre. Per tant, les maneres que tenim per formar la segona parella seran: $12 \cdot \binom{4}{2}$.

Finalment, ens queda un nombre per triar d'entre 4 pals. Per tant, les maneres que tenim de triar l'última carta seran: $11 \cdot \binom{4}{1}$.

Combinant els tres resultats anteriors, surt la fórmula.

Tots els altres casos s'han resolt seguint la mateixa filosofia.

b) $p\{\text{“full”}\} = \frac{Casos\ favorables}{Casos\ possibles} = \frac{13\cdot\binom{4}{3}\cdot 12\cdot\binom{4}{2}}{\binom{52}{5}} = \frac{3744}{2598960} \approx 0.00144$

c) $p\{\text{“Flor”}\} = \frac{Casos\ favorables}{Casos\ possibles} = \frac{4\cdot\binom{13}{5}}{\binom{52}{5}} = \frac{1287}{2598960} \approx 0.000495$

d) $p\{\text{“Correguda”}\} = \frac{Casos\ favorables}{Casos\ possibles} = \frac{\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot 9}{\binom{52}{5}} = \frac{9216}{2598960} \approx 0.003546$

L'última fórmula necessita una mica d'explicació.

Fixau-vos que si posam els nombres en ordre ascendent:

`1 2 3 4 5 6 7 8 9 10 J Q K`

podem fer en total 9 corregudes, suposant que no es pot donar la volta, o sigui, no s'admet el cas per exemple `J Q K 1 2`.

2.8.- Una urna conté 2 bolles negres i 5 de blanques. Se selecciona una bolla a l'atzar. Si la bolla és blanca, es torna a l'urna i s'afegeixen dues bolles blanques més. Si la bolla és negra, no es torna a l'urna i no s'afegeix cap bolla més. Es treu una segona bolla. Quina és la probabilitat que la bolla sigui blanca?

Resolució. Siguin els successos:

$B_1 = \{$ La primera bolla és blanca $\}$.

$N_1 = \{$ La primera bolla és negra $\}$.

$B_2 = \{$ La segona bolla és blanca $\}$.

Aplicant el teorema de les probabilitats totals, tenim:

$$\begin{aligned} p\{B_2\} &= p\{B_2/B_1\} \cdot p\{B_1\} + p\{B_2/N_1\} \cdot p\{N_1\} \\ &= \frac{7}{9} \cdot \frac{5}{7} + \frac{5}{6} \cdot \frac{2}{7} = \frac{35}{63} + \frac{10}{42} \approx 0.7936 \end{aligned}$$

2.9.- Dos fabricants, A i B, entreguen la mateixa peça a un altre fabricant, que guarda totes les existències en el mateix lloc. Els antecedents demostren que el 5% de les peces entregades per A estan defectuoses i que el 9% de les peces entregades per B també estan defectuoses. A més a més, A entrega 4 vegades més peces que B. Si es treu a l'atzar una peça i es troba que no està defectuosa, quina és la probabilitat que l'hagi fabricada A.

Resolució. Considerem els successos:

$A = \{$"Peça entregada per A"$\}$.

$B = \{$"Peça entregada per B"$\}$.

$D = \{$"Peça defectuosa"$\}$.

Com que A entrega 4 vegades més peces que B, tenim que: $p\{A\} = \frac{4}{5}$, $p\{B\} = \frac{1}{5}$. A més a més, tenim les probabilitats següents:

$p\{D/A\} = 0.05$, $p\{D/B\} = 0.09$.

Aplicant la fórmula de Bayes, podem trobar la probabilitat demanada:

$$\begin{aligned} p\{A/D^c\} &= \frac{p\{A \cap D^c\}}{p\{D^c\}} = \frac{p\{A\} \cdot p\{D^c/A\}}{p\{A\} \cdot p\{D^c/A\} + p\{B\} \cdot p\{D^c/B\}} \\ &= \frac{0.8 \cdot 0.95}{0.8 \cdot 0.95 + 0.2 \cdot 0.91} \approx 0.8068. \end{aligned}$$

2.10.- Provau que si els successos A i B són independents, també ho són A^c i B^c.

Resolució. Estam suposant que $p\{A \cap B\} = p\{A\} \cdot p\{B\}$.

Per tant, tenim:

$$\begin{aligned} p\{A^c \cap B^c\} &= p\{(A \cup B)^c\} = 1 - p\{A \cup B\} = 1 - (p\{A\} + p\{B\} - p\{A \cap B\}) \\ &= 1 - p\{A\} - p\{B\} + p\{A\} \cdot p\{B\} = (1 - p\{A\}) \cdot (1 - p\{B\}) \\ &= p\{A^c\} \cdot p\{B^c\} \end{aligned}$$

2.11.- Considerem una urna amb N bolles, de les quals n'hi ha n de blanques. Traiem una bolla de l'urna. Si surt blanca, la tornam a posar a l'urna, però si no és blanca no la hi tornam a posar. Traiem una bolla per segona vegada i surt blanca. Trobau la probabilitat p que hagi sortit blanca la primera vegada.
Final. Setembre 94.

Resolució. Considerem els successos:

$B_1 = \{$"Surt blanca la primera vegada"$\}$.

$B_2 = \{$"Surt blanca la segona vegada"$\}$.

$N_1 = \{$"No surt blanca la primera vegada"$\}$.

Aplicant la fórmula de Bayes, tenim:

$$
\begin{aligned}
p &= p\{B_1/B_2\} = \frac{p\{B_1 \cap B_2\}}{p\{B_2\}} = \frac{p\{B_1\} \cdot p\{B_2/B_1\}}{p\{B_1\} \cdot p\{B_2/B_1\} + p\{N_1\} \cdot p\{B_2/N_1\}} \\
&= \frac{\frac{n}{N} \cdot \frac{n}{N}}{\frac{n}{N} \cdot \frac{n}{N} + \frac{N-n}{N} \cdot \frac{n}{N-1}} = \frac{\frac{n}{N}}{\frac{n}{N} + \frac{N-n}{N-1}} \\
&= \frac{\frac{n}{N}}{\frac{n(N-1)+N(N-n)}{N(N-1)}} = \frac{n(N-1)}{nN - n - nN + N^2} = \frac{n(N-1)}{N^2 - n}
\end{aligned}
$$

2.12.- Una quarta part de la població ha estat vacunada contra una malaltia contagiosa. Durant una epidèmia, s'observa que d'entre els malalts n'hi ha un que ha estat vacunat per cada quatre que no hi estan.

a) Ha tengut gens d'eficàcia la vacuna?

b) D'altra banda, se sap que hi ha un malalt entre cada 12 persones vacunades. Quina és la probabilitat que estigui malalta una persona que no s'ha vacunat?

Resolució.

a) Sí que ha tengut eficàcia, ja que

$$p\{\text{estar vacunat}\} = \frac{1}{4} > p\{\text{estar vacunat / malalt}\} = \frac{1}{5}.$$

b) Posem:

$$V : \text{estar vacunat}, \quad M : \text{estar malalt}.$$

Sabem:

$$
\begin{aligned}
p\{V\} &= \frac{1}{4}, \qquad p\{V/M\} = \frac{1}{5}, \\
p\{V^c\} &= \frac{3}{4}, \qquad p\{V^c/M\} = \frac{4}{5}.
\end{aligned}
$$

Aleshores

$$p\{M/V^c\} = \frac{p\{M \cap V^c\}}{p\{V^c\}} = \frac{p\{M\} \cdot p\{V^c/M\}}{p\{V^c\}}.$$

Falta conèixer $p\{M\}$. Sabem però que $p\{M/V\} = \frac{1}{12}$. Per tant

$$p\{V/M\} = \frac{p\{V \cap M\}}{p\{M\}} \Longrightarrow p\{M\} = \frac{p\{V \cap M\}}{p\{V/M\}}.$$

I ara,

$$p\{V \cap M\} = p\{V\} \cdot p\{M/V\} = \frac{1}{4} \cdot \frac{1}{12} = \frac{1}{48},$$

d'on $p\{M\} = \frac{1/48}{1/5} = \frac{5}{48}$. Finalment,

$$p\{M/V^c\} = \frac{5/48 \cdot 4/5}{3/4} = \frac{1/12}{3/4} = \frac{4}{36} = \frac{1}{9}.$$

2.13.- Un llarg missatge s'ha codificat en termes de dos símbols A i B per a transmetre'l a través d'un canal de comunicació. La codificació és tal que A apareix el doble de vegades que B en el missatge codificat. El soroll del canal és tal que, quan A es transmet, es rep com a A amb probabilitat 0.8 i com a B amb probabilitat 0.2; quan B es transmet, es rep com a B amb probabilitat 0.7 i com a A amb probabilitat 0.3.

a) Quina és la freqüència relativa de A en el missatge rebut?

b) Si la darrera lletra del missatge que s'ha rebut és una A, quina és la probabilitat que s'hagi enviat una A?

Resolució.

a) Posem:

$$\begin{array}{ll} A_t : A \text{ transmès}, & A_r : A \text{ rebut}, \\ B_t : B \text{ transmès}, & B_r : B \text{ rebut}. \end{array}$$

Aleshores tenim:

$$p\{A_t\} = 2 \cdot p\{B_t\} \Longrightarrow p\{A_t\} = 2/3,\ p\{B_t\} = 1/3.$$

També,

$$\begin{array}{rlrl} p\{A_r/A_t\} = & 0.8, & p\{B_r/A_t\} = & 0.2, \\ p\{B_r/A_t\} = & 0.3, & p\{B_r/B_t\} = & 0.7. \end{array}$$

Aleshores,

$$\begin{aligned} p\{A_r\} &= p\{A_t\} \cdot p\{A_r/A_t\} + p\{B_t\} \cdot p\{A_r/B_t\} \\ &= \frac{2}{3} \cdot \frac{8}{10} + \frac{1}{3} \cdot \frac{3}{10} = \frac{19}{30} = 0.633. \end{aligned}$$

b)

$$p\{A_t/A_r\} = \frac{p\{A_t\} \cdot p\{A_r/A_t\}}{p\{A_r\}} = \frac{\frac{2}{3} \times \frac{8}{10}}{\frac{19}{30}} = \frac{16}{19} = 0.84.$$

2.14.- Un comerciant ha de viatjar en avió entre Bangkok i Bagdad. Preocupat, demana a la companyia aèria quina és la probabilitat que hi hagi com a mínim una bomba dins l'avió i li diuen que és 0.1. Més preocupat encara, demana quina seria la probabilitat que n'hi hagués com a mínim dues, i li diuen que seria 0.01. Més tranquil·litzat, decideix dur una bomba en el seu equipatge. Quina valoració estadística podem fer de la seva decisió?

Resolució. Posem:

$$\begin{array}{lcl} A & : & \text{com a mínim una bomba}, \\ B & : & \text{com a mínim dues bombes}. \end{array}$$

Sabem que $p\{A\} = 0.1$ i $p\{B\} = 0.01$, i per tant $p\{A \cap B\} = p\{B\} = 0.01$. Aleshores,

$$p\{B/A\} = \frac{p\{A \cap B\}}{p\{A\}} = \frac{0.01}{0.1} = 0.1 = p\{A\}!!$$

Per tant, de res serveix que dugui la seva bomba, com no sigui per augmentar la probabilitat que una bomba exploti per casualitat. El comerciant ha confós $p\{B\}$ amb $p\{B/A\}$.

2.15.- Un auditor ha estat informat de l'existència d'un error deliberat en un dels documents del departament de comptabilitat d'una empresa. La probabilitat que l'error es trobi en un lot de 10 documents és 0.05 i és igualment probable que el defecte estigui en qualsevol d'aquests documents. Si ja n'ha inspeccionat 9 sense trobar l'error, quina és la probabilitat que l'error no es trobi en aquest lot?

Resolució. Posem:

D_i : l'error es troba en el document i-èssim (i = 1, . . . , 10),
L : l'error es troba en el lot.

Aleshores ens demanam

$$p\{D_{10}^c/D_1^c \cap \cdots \cap D_9^c\} = \frac{p\{D_1^c \cap \cdots \cap D_{10}^c\}}{p\{D_1^c \cap \cdots \cap D_9^c\}}.$$

I ara,

$$\begin{aligned} p\{D_1^c \cap \cdots \cap D_{10}^c\} &= p\{L^c\} = 0.95 \\ p\{D_1^c \cap \cdots \cap D_9^c\} &= p\{L\} \cdot p\{D_1^c \cap \cdots \cap D_9^c/L\} + p\{L^c\} \cdot p\{D_1^c \cap \cdots \cap D_9^c/L^c\} \\ &= 0.05 \times \frac{1}{10} + 0.95 \times 1 = 0.955. \end{aligned}$$

Finalment,

$$p\{D_{10}^c/D_1^c \cap \cdots \cap D_9^c\} = \frac{0.95}{0.955} = 0.995.$$

2.3 Problemes proposats

2.1.- Se seleccionen a l'atzar tres cartes sense reposició d'una baralla que conté 3 cartes vermelles, 3 de blaves, 3 de verdes i 3 de negres. Especificau un espai mostral per a aquest experiment i trobau els successos següents:

- $A =$ "Totes les cartes seleccionades són vermelles"
- $B =$ "1 carta és vermella, 1 és verda i l'altra blava"
- $C =$ "Surten tres cartes de colors diferents"

2.2.- Es llancen a l'aire dues monedes iguals. Trobau la probabilitat que surtin dues cares iguals.

2.3.- Suposau que s'ha trucat un dau de tal manera que la probabilitat que surti un determinat nombre sigui proporcional al nombre. Trobau la probabilitat dels successos elementals, que surti nombre parell i que surti nombre senar. Feu el mateix problema però ara suposant que la probabilitat que surti un determinat nombre sigui inversament proporcional al nombre.

2.4.- De quantes maneres diferents es poden col·locar tres llibres distints en una taula?

2.5.- Sis persones es disposen a entrar al cinema. De quantes maneres diferents es poden col·locar en fila?

2.6.- Tres ciutadans destacats han de rebre premis. Si hi ha 4 candidats a aquests premis, de quantes maneres diferents es poden distribuir aquests si un ciutadà només pot rebre com a màxim 1 premi? I si en pot rebre més d'un?

2.7.- Donat un conjunt de 15 punts en el pla, quantes línies es necessiten per juntar totes les possibles parelles de punts?

2.8.- Donada una caixa amb els següents focus: 2 de 25 vats, 4 de 40 vats i 4 de 100 vats, de quantes maneres se'n poden escollir 3?

2.9.- Suposem que les plaques de matrícula dels cotxes es componen de 3 lletres seguides de 3 dígits. Si es poden emprar totes les combinacions possibles, quantes plaques diferents es poden formar?

2.10.- De quantes maneres diferents es poden formar 2 equips d'una lliga que tengui 8 equips?

2.11.- En un magatzem hi ha caixes vermelles i verdes. De quantes maneres es poden col·locar en fila 20 caixes si 15 són vermelles i 5 verdes? I si n'hi ha 10 de cada color?

2.12.- En una presó de 100 presos, es varen seleccionar a l'atzar dues persones per posar-les en llibertat. Quina és la probabilitat que el més vell dels presos sigui un dels elegits? I que se seleccioni la parella formada pel més vell i el més jove?

2.13.- S'apunten A, B i C en una cursa. Quina és la probabilitat que A acabi abans que C, si tots tenen la mateixa habilitat i no hi pot haver empats? Quina és la probabilitat que A acabi abans que B i C?

2.14.- En una sala es troben n persones. Trobau la probabilitat que hi hagi almenys 2 persones amb el mateix mes de naixement. Donau aquest valor per a $n = 3, 4, 5, 6$.

2.15.- Una urna conté 4 bolles numerades amb 1, 2, 3 i 4, respectivament. Es treuen 2 bolles sense reposició. Sigui A el succés que la suma sigui 5, i sigui B_i el succés que la primera bolla treta tengui una i, amb $i = 1, 2, 3, 4$. Trobau $p\{A/B_i\}$, $i = 1, 2, 3, 4$ i $p\{B_i/A\}$, $i = 1, 2, 3, 4$.

2.16.- Es llança a l'aire una moneda no trucada. Quina és la probabilitat que la quarta vegada surti cara, si surt cara en les tres primeres tirades? I si surten 2 cares en les 4 tirades?

2.17.- La urna 1 conté 2 bolles vermelles i 4 de blaves. La urna 2 conté 10 bolles vermelles i 2 de blaves. Si escollim a l'atzar una urna i en traiem una bolla, quina és la probabilitat que la bolla seleccionada sigui blava? I que sigui vermella?

2.18.- Suposem que la ciència mèdica ha desenvolupat una prova per al diagnòstic de càncer que té un 95% d'exactitud, tant en els que tenen càncer com en els que no. Si el 5 per mil de la població realment té càncer, trobau la probabilitat que un determinat individu tengui càncer, si la prova ha donat positiu.

2.19.- Es tira una sola vegada un parell de daus. Si la suma dels dos és com a mínim 7, trobau la probabilitat que sigui igual a i, per a $i = 7, 8, 9, 10, 11, 12$.

2.20.- Una moneda no trucada es tira a l'aire 2 vegades. Considerem els següents successos:

- A: Surt una cara en la primera tirada.
- B: Surt una cara en la segona tirada.

Són els dos successos independents?

2.21.- Una urna conté 4 bolles numerades amb 1, 2, 3 i 4, respectivament. Traiem dues bolles sense reposició. Sigui A el succés que la primera bolla treta tengui un 1 marcat i sigui B el succés que la segona bolla treta tengui un 1 marcat. Es pot dir que A i B són independents? I si l'experiment és amb reposició?

2.22.- Se sap que els $\frac{2}{3}$ dels interns d'una certa presó són menors de 25 anys. També se sap que els $\frac{3}{5}$ són homes i que els $\frac{5}{8}$ dels interns o són dones o majors de 25 anys. Quina és la probabilitat que un presoner escollit a l'atzar sigui dona i de menys de 25 anys?
Primer parcial Curs 92-93.

2.23.- Considerem una urna amb $2n$ bolles numerades de 1 a $2n$. Traiem 2 bolles de l'urna sense reposició. Sabent que la segona bolla és parell, quina és la probabilitat que la primera bolla sigui senar?
Final. Juny 95.

2.24.- Considerem el següent experiment aleatori: traiem 5 nombres a l'atzar sense reposició a partir dels nombres naturals $1, 2, \ldots, 20$. Trobeu la probabilitat p que exactament dos dels nombres siguin múltiples de 3.
Final. Setembre 95.

2.25.- Sigui Ω un espai mostral i A, B i C tres successos. Provau que:

a) Si A i B són independents, també ho són A i B^c.

b) Si A, B i C són independents, també ho són A, B i C^c.

c) És cert que si A, B i C són independents, també ho són A, B^c i C^c? I A^c, B^c i C^c? En el cas que la resposta sigui negativa, donau contraexemples on la propietat falli.

Final. Setembre 95.

2.26.- En una urna hi ha 10 bolles, numerades d'1 fins a 10. Les 4 primeres bolles, o sigui, les bolles 1, 2, 3 i 4 són blanques. Les bolles 5 i 6 són negres i les bolles 7, 8, 9 i 10 són vermelles. Traiem dues bolles sense reposició. Sabent que la segona bolla és negre, trobau la probabilitat p que la primera bolla sigui blanca.
Final. Juny 96.

2.27.- Llançam un dau no trucat 3 vegades. Trobau la probabilitat p que la suma de les 3 cares sigui 10.
Final. Setembre 96.

2.28.- Quatre cartes numerades de l'1 al 4 estan girades cap avall damunt d'una taula. Una persona, suposadament clarivident, anirà endevinant els valors de les 4 cartes una a una. Suposant que és un farsant i que el que fa és dir els quatre nombres a l'atzar, quina és la probabilitat que n'encerti com a mínim un? (Òbviament, no repeteix cap nombre.)

2.29.- Una manera d'augmentar la fiabilitat d'un sistema és mitjançant la introducció d'una còpia dels components en una configuració paral·lela. Suposem que la NASA vol una probabilitat no menor que 0.99999 que el transbordador espacial entri en òrbita al voltant de la Terra amb èxit. Quants de motors s'han de configurar en paral·lel per tal d'assolir aquesta fiabilitat, si se sap que la probabilitat que un qualsevol dels motors funcioni adequadament és 0.95? Suposem que els motors funcionen de manera independent els uns amb els altres.

2.30.- Dues empreses A i B fabriquen el mateix producte. L'empresa A té un 2% de productes defectuosos mentre que l'empresa B en té un 1%. Un client rep una comanda d'una de les empreses (no sap de quina) i comprova que la primera peça funciona. Si suposam que l'estat de les peces de cada empresa és independent, quina és la probabilitat que la segona peça que provi sigui bona? Comprovau que l'estat de les dues peces no és independent, però en canvi és condicionalment independent donada l'empresa que les fabrica.

2.31.- Trobau un exemple de tres successos A, B i C tals que A i B siguin independents però en canvi no siguin condicionalment independents donat C.

Capítol 3

Variables aleatòries

3.1 Resum teòric

Definició 3.1 *Sigui* $(\Omega, \mathcal{F}, p)$ *un espai de probabilitat. Direm* **variable aleatòria** *(real) a tota aplicació* $X : \Omega \to \mathbb{R}$ *tal que* $\forall x \in \mathbb{R}, X^{-1}((-\infty, x]) \in \mathcal{F}$ *(és a dir, és un succés). Aquest conjunt* $X^{-1}((-\infty, x])$ *se sol indicar amb* $\{X \leq x\}$*.*

Definició 3.2 *La* **funció de distribució** *d'una variable aleatòria* X *és la funció* $F : \mathbb{R} \to [0, 1]$ *definida per:*

$$F(x) = p \circ X^{-1}((-\infty, x]) = p\{X \leq x\}.$$

Propietats.

1. F és creixent.
2. F és contínua per la dreta.
3. $\lim\limits_{x \to +\infty} F(x) = 1, \ \lim\limits_{x \to -\infty} F(x) = 0.$

El recíproc també és cert: si $F : \mathbb{R} \to [0, 1]$ és una funció que verifica (1)-(3), aleshores F és la funció de distribució d'una variable aleatòria X.

Notes: Sigui X una variable aleatòria i F la seva funció de distribució.

1. En general, F no és contínua per l'esquerra. De fet, $F(x-) = F(x) - p\{X = x\}$.

2. Mitjançant la funció de distribució es poden expressar les probabilitats que la variable aleatòria X prengui els seus valors en intervals diversos. Per exemple, si $a, b \in \mathbb{R}$,

   (a) $p\{X < a\} = F(a-)$,
   (b) $p\{a \leq X \leq b\} = F(b) - F(a-)$,
   (c) $p\{a < X \leq b\} = F(b) - F(a)$,
   (d) $p\{a \leq X < b\} = F(b-) - F(a-)$,
   (e) $p\{a < X < b\} = F(b-) - F(a)$.

Definició 3.3 *Direm que una variable aleatòria X és* **discreta** *si el conjunt $X(\Omega)$ de tots els valors que pren és finit o numerable sense punts d'acumulació.*

Notes.

1. Per a aquestes variables aleatòries, tenim:

$$\{X \leq t\} \in \mathcal{F} \Longleftrightarrow X^{-1}(x_k) = \{X = x_k\} \in \mathcal{F},\ \forall x_k \in X(\Omega).$$

2. Si Ω és finit, aleshores tota variable aleatòria X sobre $(\Omega, \mathcal{F}, p)$ és discreta.

3. Si X és una variable aleatòria discreta i g és una funció real de variable real és tal que $g(X(\Omega))$ no té punts d'acumulació, aleshores $g(X) : \Omega \to \mathbb{R}$ donada per $g(X)(\omega) = g(X(\omega))$ també és una variable aleatòria discreta.

4. Si X i Y són variables aleatòries discretes, també ho són $X + Y$ i $X \cdot Y$.

Definició 3.4 *Sigui X una variable aleatòria discreta, amb $X(\Omega) = \{x_1, x_2, \ldots\}$. Direm* **funció de probabilitat** *de X a l'aplicació $f_X : \mathbb{R} \to \mathbb{R}$ donada per*

$$f_X(x) = p\{X = x\}\ \ \forall x \in \mathbb{R}.$$

Notes.

1. Si $x \notin X(\Omega),\ f_X(x) = 0$.

2. Si $B \subset \mathbb{R}$ és tal que $\{X \in B\} \in \mathcal{F}$, aleshores $p\{X \in B\} = \sum_{x_i \in B \cap X(\Omega)} f_X(x_i)$

3. $F(x) = \sum_{x_i \le x, x_i \in X(\Omega)} f_X(x_i) \ \forall x \in \mathbb{R}$.

Definició 3.5 *Direm que la variable aleatòria X és* **contínua** *si la seva funció de distribució és contínua.*

Propietat: X és contínua si i només si $p\{X = x\} = 0 \ \forall x \in \mathbb{R}$.

Definició 3.6 *Una funció $f : \mathbb{R} \to \mathbb{R}$ s'anomena* **densitat** *si compleix les condicions següents:*

1. $f \ge 0$,
2. *f és integrable (en el sentit de Riemann) en $\mathbb{R}$,*
3. $\int_{-\infty}^{+\infty} f(x)\, dx = 1$.

Definició 3.7 *Direm que X és una variable aleatòria* **absolutament contínua amb densitat** *f si la seva funció de distribució F se pot escriure com:*

$$F(x) = \int_{-\infty}^{x} f(y)\, dy \quad \forall x \in \mathbb{R},$$

on f és una densitat.

Notes.

1. Tota variable aleatòria absolutament contínua és contínua. El recíproc no és cert en general.
2. Si X és absolutament contínua amb densitat f, aleshores

$$p\{X \in A\} = \int_A f(x)\, dx \quad \forall A \subset \mathbb{R} \text{ tal que } \{X \in A\} \in \mathcal{F}.$$

De vegades, donada una variable aleatòria X, estarem interessats en la distribució d'una nova variable aleatòria Y funció de X, $Y = g(X)$, on $g : \mathbb{R} \to \mathbb{R}$ és una funció donada. El cas discret és trivial; ens centrarem, per tant, en el cas absolutament continu.

Propietats.

1. Si g és creixent,
$$F_Y(y) = F_X(x)\mid_{x=g^{-1}(y)}.$$
2. Si g és decreixent,
$$F_Y(y) = 1 - F_X(x)\mid_{x=g^{-1}(y)}.$$
3. Si, a més, g és diferenciable,
$$f_Y(y) = \left.\frac{f_X(x)}{|g'(x)|}\right|_{x=g^{-1}(y)}.$$
4. En general, si g és diferenciable (no necessàriament monòtona) tal que $\forall y \in \mathbb{R}$, l'equació $g(x) = y$ té un nombre finit de solucions $x_1, \ldots, x_n$,
$$f_Y(y) = \sum_{x_k} \left.\frac{f_X(x)}{|g'(x)|}\right|_{x=x_k}.$$

Per tal de descriure completament el comportament d'una variable aleatòria, s'ha de donar una funció, o bé la de distribució o bé la de densitat (de probabilitat en el cas discret). En certes situacions, estarem interessats en uns quants paràmetres que resumeixin la informació donada per aquestes funcions. Vegem-ne el més importants.

Definició 3.8 *Direm* **esperança** *(***mitjana,** *o* **valor esperat***) d'una variable aleatòria X a:*

- $\mathrm{E}X = \sum_k x_k \cdot p\{X = x_k\}$ *si X és discreta amb $X(\Omega) = \{x_1, x_2, \ldots\}$,*
- $\mathrm{E}X = \int_{-\infty}^{+\infty} x\, f(x)\, dx$ *si X és absolutament contínua amb densitat f.*

L'esperança de X està definida si la sèrie o integral anterior convergeix absolutament, és a dir si $\mathrm{E}|X| < \infty$.

Notes: Si X és una variable aleatòria no negativa, es compleix:

1. $\mathrm{E}X = \sum\limits_{k=0}^{\infty} p\{X > k\}$ (cas discret amb $X(\Omega) = \mathbb{Z}^+$).
2. $\mathrm{E}X = \int\limits_{0}^{+\infty} (1 - F_X(x))\, dx$ (cas absolutament continu).

Si $Y = g(X)$ és absolutament contínua, aleshores, segons la definició, $\mathrm{E}Y = \int\limits_{-\infty}^{+\infty} y\, f_Y(y)\, dy$, però es pot provar que també $\mathrm{E}Y = \int\limits_{-\infty}^{+\infty} g(x)\, f_X(x)\, dx$, on aquesta integral existirà si $\mathrm{E}Y = \int\limits_{-\infty}^{+\infty} |g(x)|\, f_X(x)\, dx < \infty$.

Propietats.

1. $\forall a \in \mathbb{R}$, $\mathrm{E}(a) = a$.
2. $\forall a \in \mathbb{R}$, $\mathrm{E}(aX) = a \cdot \mathrm{E}X$ (si existeix).
3. $\mathrm{E}\left(\sum\limits_{k=1}^{n} a_k g_k(X)\right) = \sum\limits_{k=1}^{n} a_k \cdot \mathrm{E}(g_k(X))$ (si existeixen totes les esperances).
4. Si $\mathrm{E}|X| < \infty$, aleshores $|\mathrm{E}X| \leq \mathrm{E}|X|$.
5. $X \geq 0 \Longrightarrow \mathrm{E}X \geq 0$ (si existeix).
6. Si f_X és simètrica respecte de a ($f_X(a - x) = f_X(a + x)$), aleshores $\mathrm{E}X = a$ (si existeix).

Definició 3.9 *Direm que una variable aleatòria X té* **moment d'ordre n finit** *si X^n té esperança finita i, en aquest cas,* $\mathrm{E}(X^n)$ *s'anomena el moment d'ordre n de X.*

Definició 3.10 *Si X té esperança finita, direm que X té* **moment central d'ordre n finit** *si existeix* $\mathrm{E}((X - \mathrm{E}X)^n)$*, i aquest és el moment central d'ordre n de X.*

Definició 3.11 *Direm* **variància** *d'una variable aleatòria X al moment central de segon ordre:* $\mathrm{Var}X = \sigma_X^2 = \mathrm{E}(X - \mathrm{E}X)^2$. *L'arrel quadrada positiva de* $\mathrm{Var}X$ *s'anomena la* **desviació típica** *(o* **estàndard***) de X:* $\sigma_X = +\sqrt{\mathrm{Var}X}$.

Propietats.

1. $\mathrm{Var}X = \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2$.
2. $\mathrm{Var}(aX) = a^2 \cdot \mathrm{Var}X \quad \forall a \in \mathbb{R}$.
3. $\mathrm{Var}(a) = 0 \quad \forall a \in \mathbb{R}$.
4. $\mathrm{Var}(X+Y) = \mathrm{Var}X + \mathrm{Var}Y + 2(\mathrm{E}(XY) - \mathrm{E}X \cdot \mathrm{E}Y)$.
5. $\mathrm{Var}(aX+b) = a^2 \cdot \mathrm{Var}X \quad \forall a, b \in \mathbb{R}$.
6. $\mathrm{Var}(X - \mathrm{E}X) = \mathrm{Var}X$.

Desigualtat de Txebixef:
Sigui X una variable aleatòria no negativa, i $g : \mathbb{R}^+ \to \mathbb{R}^+$ una funció creixent tal que $\mathrm{E}(g(X)) < \infty$. Sigui $a \in \mathbb{R}$ tal que $g(a) > 0$. Aleshores

$$p\left\{X \geq a\right\} \leq \frac{1}{g(a)}\, \mathrm{E}(g(X))$$

Notes.

1. **Desigualtat de Markov**: En particular, si X té moment de segon ordre finit i $a > 0$,

$$p\left\{|X| \geq a\right\} \leq \frac{\mathrm{E}X^2}{a^2}.$$

2. Si X té moment central de segon ordre finit,

$$p\left\{|X - \mathrm{E}X| \geq a\right\} \leq \frac{\mathrm{Var}X}{a^2}.$$

3. Si X té moment central de segon ordre finit,

$$p\left\{|X - \mathrm{E}X| \geq a \cdot \sigma_X\right\} \leq 1/a^2.$$

   En particular,
$$p\left\{|X - \mathrm{E}X| \geq 3\sigma_X\right\} \leq 1/9 = 0.111,$$
   és a dir que la probabilitat que X prengui valors fora de l'interval $(\mathrm{E}X - 3\sigma_X, \mathrm{E}X + 3\sigma_X)$ és menor que 0.111 (si $a = 4$, la cota és 0.0625).

Per tal d'obtenir més operativitat en el tractament de certs problemes, se solen aplicar transformacions de la distribució d'una variable aleatòria, de manera que, la transformada conservi tota la informació referent a la variable aleatòria i que el procés sigui reversible. Vegem-ne unes quantes d'aquestes transformacions.

Definició 3.12 *Direm que* $X = X_1 + \mathrm{i} \cdot X_2$ *és una* **variable aleatòria complexa** *si* X_1 *i* X_2 *són dues variables aleatòries reals.*

Definició 3.13 *Donada una variable aleatòria real* X*, definim la* **funció característica** *de* X *com la funció* $\phi_X : \mathbb{R} \to \mathbb{C}$ *donada per*

$$\phi_X(t) = \mathrm{E}(\mathrm{e}^{\mathrm{i}tX}) = \mathrm{E}(\cos tX) + \mathrm{i} \cdot \mathrm{E}(\sin tX) \ \forall t \in \mathbb{R}.$$

Propietats.

1. La funció característica sempre existeix.
2. ϕ_X caracteritza la variable aleatòria X, en el sentit que dues variables aleatòries amb la mateixa funció característica tenen la mateixa distribució.
3. Si X té moment d'ordre k finit,
$$\mathrm{E}X^k = \frac{1}{\mathrm{i}^k} \cdot \left. \frac{d^k \phi_X(t)}{dt^k} \right|_{t=0}.$$
4. Si X és absolutament contínua i si $|\phi_X|$ és integrable en $\mathbb{R}$, podem recuperar f_X a partir de ϕ_X:
$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi_X(t)\, \mathrm{e}^{-\mathrm{i}tx}\, dt \ \ \forall x \in \mathbb{R}.$$
5. Si X és discreta amb $X(\Omega) = \mathbb{Z}$, obtenim també una fórmula d'inversió:
$$p\{X = k\} = \frac{1}{2\pi} \int_0^{2\pi} \phi_X(t)\, \mathrm{e}^{-\mathrm{i}tk}\, dt \ \ \forall k \in \mathbb{Z}.$$

Definició 3.14 *Si en la definició de la funció característica substituïm* it *per* t*, obtenim la* **funció generatriu de moments***:*

$$m_X(t) = \mathrm{E}(\mathrm{e}^{tX})$$

Si X *té moment d'ordre* k *finit, resulta:* $\mathrm{E}X^k = m_X(0)$.

Definició 3.15 *La* **funció generatriu cumulativa** *és la funció* $c_X(t) = \ln m_X(t)$.

Definició 3.16 *La* **funció generatriu factorial** *és la funció* $\psi_X(t) = \mathrm{E}(t^X)$.

3.2 Problemes resolts

3.1.- Una urna conté 4 bolles numerades 1, 2, 3 i 4 respectivament. Sigui Y el nombre que surt quan s'extreu una bolla a l'atzar. Quina és la funció de probabilitat per a Y?

Resolució. El rang de la variable aleatòria Y és: $Y(\Omega) = \{1, 2, 3, 4\}$.

Tenint en compte que tots els valors de $Y(\Omega)$ tenen la mateixa probabilitat de sortir, la funció de probabilitat serà per a qualsevol valor de $Y(\Omega)$: $f_Y(i) = p\{Y = i\} = \frac{1}{4}$.

La funció de probabilitat de Y queda esquematitzada en la taula següent:

Y	1	2	3	4
$f_Y(y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

3.2.- Considerem l'urna del problema anterior. Es treuen dues bolles sense reposició i sigui Z la suma dels dos nombres que surten. Trobau la funció de probabilitat per a Z.

Resolució. El rang de valors de Z serà: $Z(\Omega) = \{3, 4, 5, 6, 7\}$ on Ω és l'espai mostral: $\Omega = \{(i, j), i \neq j\}$.

Tenint en compte que els successos elementals de Ω tenen la mateixa probabilitat de sortir, podem escriure els successos $\{Z = i\}$ en funció dels elements de Ω per així poder trobar la funció de probabilitat de Z.

Per tant, observem que:

$$\begin{array}{llll} \{Z=3\} = & \{(1,2),(2,1)\}, & \{Z=4\} = & \{(1,3),(3,1)\}, \\ \{Z=5\} = & \{(1,4),(4,1),(2,3),(3,2)\}, & \{Z=6\} = & \{(2,4),(4,2)\}, \\ \{Z=7\} = & \{(3,4),(4,3)\}. & & \end{array}$$

La funció de probabilitat valdrà:

Z	3	4	5	6	7
$f_Z(z)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$

3.3.- Comprovau que:

$$F_X(t) = \begin{cases} 0, & \text{si } t < -3, \\ \frac{1}{3}, & \text{si } -3 \leq t < -1, \\ \frac{2}{3}, & \text{si } -1 \leq t < 0, \\ 1, & \text{si } t \geq 0, \end{cases}$$

és una funció de distribució i trobau la funció de probabilitat per a X.

Resolució. Comprovem que $F_X(t)$ verifica totes les propietats d'una funció de distribució:

- F_X és creixent ja que a mesura que t creix $F_X(t)$ també.
- F_X és contínua per la dreta ja que $F_X(t)$ és contínua en tot punt distint de $-3, -1$ i 0, i en aquest punts es verifica:
  - $\lim_{t \to -3^+} F_X(t) = \lim_{t \to -3^+} \frac{1}{3} = F_X(-3)$.
  - $\lim_{t \to -1^+} F_X(t) = \lim_{t \to -1^+} \frac{2}{3} = F_X(-1)$.
  - $\lim_{t \to 0^+} F_X(t) = \lim_{t \to 0^+} 1 = F_X(0)$.
- F_X compleix que:
  - $\lim_{t \to -\infty} F_X(t) = 0$.
  - $\lim_{t \to \infty} F_X(t) = 1$.

Com podem observar, es tracta d'una funció de distribució corresponent a una variable aleatòria discreta ja que F_X és una funció esglaonada.

El rang de la variable aleatòria X són els punts on F_X no és contínua. Per tant, tenim que: $X(\Omega) = \{-3, -1, 0\}$.

Per trobar la funció de densitat en aquests punts basta calcular el salt de la discontinuïtat corresponent a cada punt, o sigui:

$$f_X(i) = \lim_{t \to i^+} F_X(t) - \lim_{t \to i^-} F_X(t).$$

La funció de probabilitat serà, doncs:

X	-3	-1	0
$f_X(x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

3.4.- Verificau que:

$$F_Z(t) = \begin{cases} 0, & \text{si } t < 0, \\ t^2, & \text{si } 0 \leq t < \frac{1}{2}, \\ 1 - 3(1-t)^2, & \text{si } \frac{1}{2} \leq t < 1, \\ 1, & \text{si } t \geq 1, \end{cases}$$

és una funció de distribució i trobau la funció de densitat per a Z.

Resolució.

- F_Z és creixent. Comprovem que la derivada és positiva en cada un dels trossos on està definida, ja que en els punts 0,1/2 i 1 els trossos enganxen bé:
  - Si $t < 0$, tenim que $F'_Z(t) = 0$.
  - Si $0 \leq t < \frac{1}{2}$, tenim que $F'_Z(t) = 2t \geq 0$.
  - Si $\frac{1}{2} \leq t < 1$, tenim que $F'_Z(t) = 6(1-t) \geq 0$.
  - Si $t \geq 1$, tenim que $F'_Z(t) = 0$.
- F_Z és trivialment contínua en tot punt distint de $0, \frac{1}{2}$ i 1 ja que $F_Z(t)$ està definit com un polinomi. Vegem que en aquests punts F_Z també és contínua.

- $\lim_{t\to 0^+} F_Z(t) = \lim_{t\to 0^+} t^2 = 0 = \lim_{t\to 0^-} F_Z = \lim_{t\to 0^-} 0.$
- $\lim_{t\to \frac{1}{2}^+} F_Z(t) = \lim_{t\to \frac{1}{2}^+} 1-3(1-t)^2 = \frac{1}{4} = \lim_{t\to \frac{1}{2}^-} F_Z(t) = \lim_{t\to \frac{1}{2}^-} t^2.$
- $\lim_{t\to 1^+} F_Z(t) = \lim_{t\to 1^+} 1 = 1 = \lim_{t\to 1^-} F_Z = \lim_{t\to 1^-} 1-3(1-t)^2.$

- F_Z compleix que:
  - $\lim_{t\to -\infty} F_Z(t) = 0.$
  - $\lim_{t\to \infty} F_Z(t) = 1.$

Com podem observar es tracta d'una funció de distribució d'una variable aleatòria absolutament contínua ja que F_Z és contínua.

Per trobar la funció de densitat, basta fer servir la fórmula $f_Z(t) = F_Z'(t)$ en els punts on $F_Z(t)$ sigui derivable, o sigui, per a tot t distint de $0, \frac{1}{2}$ i 1. En aquests punts, podem definir $f_Z(t)$ com vulguem, ja que la funció de densitat d'una variable aleatòria absolutament contínua està determinada excepte un nombre finit de punts (de fet pot ser un nombre numerable).

Així doncs, $f_Z(t)$ val:

$$f_Z(t) = \begin{cases} 0, & \text{si } t < 0, \\ 2t, & \text{si } 0 \leq t < \frac{1}{2}, \\ 6(1-t), & \text{si } \frac{1}{2} \leq t < 1, \\ 0, & \text{si } t \geq 1. \end{cases}$$

3.5.- Sigui la variable aleatòria X amb funció de probabilitat:

$$f_X(x) = \begin{cases} \frac{1}{4}, & \text{si } x = 2, 4, 8, 16, \\ 0, & \text{en cas contrari.} \end{cases}$$

Calculau $\mathrm{E}X$, $\mathrm{E}X^2$, $\mathrm{E}\left(\frac{1}{X}\right)$, $\mathrm{E}\left(2^{\frac{X}{2}}\right)$ i $\mathrm{Var}\, X$.

Resolució. El rang de la variable aleatòria X és

$$X(\Omega) = \{2^i, i = 1, 2, 3, 4\}.$$

Així doncs:

$$\begin{aligned}
\mathrm{E}X &= \sum_{i=1}^{4} \tfrac{1}{4}\cdot 2^i = 7.5,\\
\mathrm{E}\left(X^2\right) &= \sum_{i=1}^{4} \tfrac{1}{4}\cdot 2^{2i} = 85,\\
\mathrm{E}\left(\tfrac{1}{X}\right) &= \sum_{i=1}^{4} \tfrac{1}{4}\cdot \tfrac{1}{2^i} = 0.234375,\\
\mathrm{E}\left(2^{\frac{X}{2}}\right) &= \sum_{i=1}^{4} \tfrac{1}{4}\cdot 2^{2^{i-1}} = 69.5,\\
\mathsf{Var}\ X &= \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = 85 - 7.5^2 = 28.75
\end{aligned}$$

3.6.- Prenent la funció de densitat següent:

$$f_X(x) = \begin{cases} 2(1-x), & \text{si } 0 < x < 1,\\ 0, & \text{en cas contrari,}\end{cases}$$

calculau $\mathrm{E}X$, $\mathrm{E}\left(X^2\right)$, $\mathrm{E}\left((X+10)^2\right)$, $\mathrm{E}\left(\frac{1}{1-X}\right)$ i Var X.

Resolució. El rang de la variable aleatòria X és l'interval $(0,1)$.

Així doncs:

$$\begin{aligned}
\mathrm{E}X &= \int_0^1 2x(1-x)\,dx = \int_0^1 (2x-2x^2)\,dx = \left[x^2 - \frac{2}{3}x^3\right]_0^1\\
&= 1 - \frac{2}{3} = \frac{1}{3},\\
\mathrm{E}\left(X^2\right) &= \int_0^1 2x^2(1-x)\,dx = \int_0^1 (2x^2-2x^3)\,dx = \left[\frac{2}{3}x^3 - \frac{1}{2}x^4\right]_0^1\\
&= \frac{2}{3} - \frac{1}{2} = \frac{1}{6},\\
\mathrm{E}\left((X+10)^2\right) &= \mathrm{E}\left(X^2 + 20X + 100\right) = \mathrm{E}\left(X^2\right) + 20\mathrm{E}X + 100\\
&= \frac{1}{6} + 20\frac{1}{3} + 100 = \frac{641}{6},\\
\mathrm{E}\left(\frac{1}{1-X}\right) &= \int_0^1 2\frac{1-x}{1-x}\,dx = [2x]_0^1 = 2,\\
\mathsf{Var}\ X &= \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}.
\end{aligned}$$

3.7.- Ens donen la següent funció de distribució d'una variable aleatòria X:

$$F_X(t) = \begin{cases} 0, & \text{si } t < 0, \\ \frac{t}{100}, & \text{si } 0 \le t \le 100, \\ 1, & \text{si } t > 100. \end{cases}$$

Trobau el percentil 10% ($t_{0.10}$), el percentil 20% ($t_{0.20}$) i el percentil 80% ($t_{0.80}$).

Resolució. Fixau-vos que es tracta d'una variable aleatòria contínua.

Per tant, per trobar el percentil t_p hem de resoldre l'equació: $F_X(t_p) = p$. A més a més, tots els valors dels percentils es trobaran en l'interval $(0, 100)$, ja que aquest és el rang de la variable aleatòria X.

L'equació anterior es transformarà, doncs, en $\frac{t_p}{100} = p$. D'on deduïm que el percentil t_p val $t_p = 100 \cdot p$.

Els percentils valen:

$$t_{0.10} = 10, \; t_{0.20} = 20, \; t_{0.80} = 80.$$

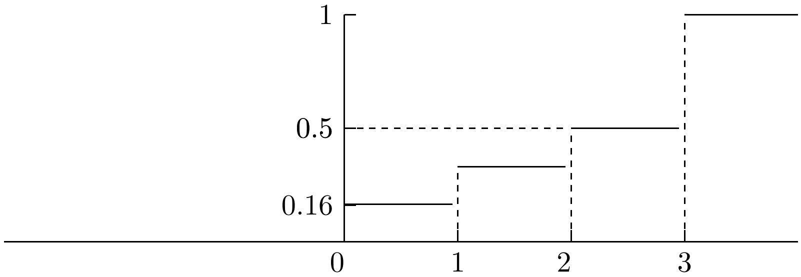

Figura 3.1: Funció de distribució.

3.8.- Sigui $F_X(t)$ la funció de distribució següent:

$$F_X(t) = \begin{cases} 0, & \text{si } t < 0, \\ \frac{1}{6}, & \text{si } 0 \leq t < 1, \\ \frac{1}{3}, & \text{si } 1 \leq t < 2, \\ \frac{1}{2}, & \text{si } 2 \leq t < 3, \\ 1, & \text{si } t \geq 3. \end{cases}$$

Trobau la mediana i el percentil 16% ($t_{0.16}$).

Resolució. Ens demanen trobar el percentil 50% (mediana o $t_{0.50}$) i el percentil 16% o $t_{0.16}$.

La fórmula general per trobar un percentil $100p$% o t_p és

$$t_p = \min\{t \mid F_X(t) \geq p\}. \tag{3.1}$$

En el cas continu, com ja hem comentat en el problema anterior, resoldre l'equació anterior és equivalent a resoldre l'equació $F(t_p) = p$, però en el cas discret, com és el cas que ens ocupa, no podem aplicar la darrera fórmula. En aquest darrer cas, hem d'aplicar la fórmula original (3.1).

En la figura 3.1 hem dibuixat la funció de distribució. Fixau-vos que el percentil 16% serà en aquest cas $t_{0.16} = 0$, ja que el valor 0 és el valor més petit en què la funció de distribució és més gran o igual que 0.16.

De la mateixa manera, tenim que el percentil 50% val $t_{0.50} = 2$.

3.9.- Suposau que X és una variable aleatòria (discreta) amb funció generatriu de moments
$$\psi_X(t) = \frac{(1-t^{n+1})}{(n+1)(1-t)}.$$
Trobau $f_X(x)$ (funció de probabilitat) i $m_X(t)$.
Suposau que el rang de X és $X(\Omega) = \{0, 1, \ldots, n\}$.

Resolució. Abans de resoldre el problema vegem la proposició següent:

Proposició 3.1 *Sigui X una variable aleatòria discreta amb rang $X(\Omega) = \{0, 1, \ldots, n\}$. Aleshores la funció de probabilitat es pot trobar segons la fórmula següent:*

$$f_X(k) = \frac{\psi_X^{(k)}(0)}{k!}, \text{ on } \psi_X(t) \text{ és la funció generatriu factorial de } X \text{ per a } k = 0, \ldots, n. \tag{3.2}$$

Prova de la proposició.

Fixau-vos que:

$$\begin{aligned} \psi_X(t) &= \sum_{j=0}^{n} t^j f_X(j), \\ \psi_X'(t) &= \sum_{j=0}^{n-1} (j+1) t^j f_X(j+1). \end{aligned}$$

Es pot veure per inducció de forma molt fàcil que la derivada k-èssima de la funció generatriu factorial compleix la fórmula per a $k = 1, \ldots, n$:

$$\psi_X^{(k)}(t) = \sum_{j=0}^{n-k} (j+1) \cdot (j+2) \ldots (j+k) t^j f_X(j+k). \tag{3.3}$$

D'aquí, doncs, podem deduir que:

$$\psi_X^{(k)}(0) = k! f_X(k),$$

d'on, aïllant $f_X(k)$, queda vista la proposició □

Resolguem el problema.

Primer simplificam l'expressió de $\psi_X(t)$:

$$\psi_X(t) = \frac{1}{(n+1)}\frac{(t^{n+1}-1)}{(t-1)} = \frac{1}{n+1}\sum_{j=0}^{n} t^j.$$

Raonant de manera semblant a la prova de la proposició anterior, podem deduir la següent fórmula per a la derivada k-èssima:

$$\psi_X^{(k)}(t) = \frac{1}{n+1}\sum_{j=0}^{n-k}(j+1)\dots(j+k)t^j.$$

D'aquí deduïm que $\psi_X^{(k)}(0) = \frac{1}{n+1}k!$ per $k = 0, \dots, n$.

Finalment, aplicant la fórmula (3.2) tenim que la funció de probabilitat de X val:

$$f_X(k) = \frac{1}{n+1}, k = 0, 1, \dots, n.$$

A continuació trobarem $m_X(t)$:

$$m_X(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right) = \psi_X\left(\mathrm{e}^{t}\right) = \frac{\mathrm{e}^{(n+1)t}-1}{(n+1)(\mathrm{e}^t-1)}.$$

3.10.- Donada la funció generatriu cumulativa $c_X(t)$, provau que $c'_X(0) = \mu_X = \mathrm{E}(X)$ i $c''_X(0) = \sigma_X^2 = \mathrm{Var}\ X$.

Resolució. Fent càlculs, podem trobar les dues primeres derivades de $c_X(t)$:

$$\begin{aligned} c'_X(t) &= \frac{m'_X(t)}{m_X(t)}, \\ c''_X(t) &= \frac{m''_X(t)m_X(t) - m'_X(t)^2}{m_X(t)^2}. \end{aligned}$$

Tenint en compte les relacions que hi ha entre les derivades de $m_X(t)$ i els moments,

$$m_X(0) = \mathrm{E}\left(\mathrm{e}^{t0}\right) = \mathrm{E}(1) = 1,\ m'_X(0) = \mathrm{E}X,\ m''_X(0) = \mathrm{E}\left(X^2\right), \tag{3.4}$$

arribam a les conclusions següents:

$$\begin{aligned} c'_X(0) &= \frac{m'_X(0)}{m_X(0)} = \mathrm{E}X, \\ c''_X(0) &= \frac{m''_X(0)m_X(0) - m'_X(0)^2}{m_X(0)^2} = \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = \mathsf{Var}\ X. \end{aligned}$$

3.11.- Si

$$F_Y(t) = \begin{cases} 1 - \mathrm{e}^{-t}, & \text{si } t \geq 0, \\ 0, & \text{si } t < 0, \end{cases}$$

trobau $F_X(t)$ i $f_X(t)$ on $X = 2Y - 7$.

Resolució. Recordem que si tenim una variable aleatòria Y i feim un canvi del tipus $X = g(Y)$ on $g(Y)$ és una funció estrictament monòtona (creixent o decreixent), les relacions que hi ha entre les funcions de distribució i de densitat de Y i X són les següents:

$$\begin{aligned} F_X(t) &= F_Y(g^{-1}(t)), & (3.5) \\ f_X(t) &= f_Y(g^{-1}(t)) \cdot \frac{dg^{-1}}{dt}(t) = \frac{f_Y(g^{-1}(t))}{\left|\frac{dg}{dt}(g^{-1}(t))\right|}. & (3.6) \end{aligned}$$

En el nostre cas, tenim que $g(Y) = 2Y - 7$ i $g^{-1}(Y) = \frac{Y+7}{2}$.

Fixau-vos que la funció de densitat la podem trobar de dues maneres; una, derivant la funció de distribució en tractar-se Y d'una variable aleatòria contínua i, dues, aplicant directament la fórmula (3.6). Ho farem de la primera manera.

Per tant, aplicant (3.5) podem trobar $F_X(t)$ i $f_X(t)$.

$$\begin{aligned} F_X(t) &= F_Y\left(\frac{(t+7)}{2}\right) = \begin{cases} 1 - \mathrm{e}^{-\frac{(t+7)}{2}}, & \text{si } \frac{t+7}{2} \geq 0, \\ 0, & \text{si } \frac{t+7}{2} < 0. \end{cases} = \begin{cases} 1 - \mathrm{e}^{-\frac{(t+7)}{2}}, & \text{si } t \geq -7, \\ 0, & \text{si } t < -7. \end{cases} \\ f_X(t) &= F'_X(t) = \begin{cases} \frac{1}{2}\mathrm{e}^{-\frac{(t+7)}{2}}, & \text{si } t \geq -7, \\ 0, & \text{si } t < -7. \end{cases} \end{aligned}$$

3.12.- Donada la funció de distribució següent

$$F_U(t) = \begin{cases} 0, & \text{si } t < -1, \\ \frac{1}{3}, & \text{si } -1 \leq t < 0, \\ \frac{2}{3}, & \text{si } 0 \leq t < 1, \\ 1, & \text{si } t \geq 1, \end{cases}$$

trobau la funció de distribució per a la forma estàndard de U.

Resolució. La forma estàndard d'una variable aleatòria U és una altra variable Z_U relacionada de la següent forma amb U:

$$Z_U = \frac{U - \mathrm{E}U}{\sqrt{\mathsf{Var}\ U}}$$

Per tant, primer de tot hem de trobar $\mathrm{E}U$ i Var U.

Observem que es tracta d'una variable aleatòria discreta amb rang

$$U(\Omega) = \{-1, 0, 1\}.$$

La funció de probabilitat serà el valor dels salts de discontinuïtat en la funció de distribució:

U	-1	0	1
$f_U(u)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Per tant,

$$\begin{aligned} \mathrm{E}U &= \sum_{i=-1}^{1} i\frac{1}{3} = \frac{1}{3}(-1+0+1) = 0, \\ \mathsf{Var}\ U &= \mathrm{E}\left(U^2\right) - (\mathrm{E}U)^2 = \sum_{i=-1}^{1} i^2\frac{1}{3} = \frac{2}{3}. \end{aligned}$$

Així, doncs, el canvi que hem de fer és:

$$Z_U = \frac{U}{\sqrt{\frac{2}{3}}} := g(U).$$

Tenint en compte que $g^{-1}(U) = \sqrt{\frac{2}{3}}U$, la funció de distribució de la variable Z_U serà:

$$F_{Z_U}(t) = F_U\left(\sqrt{\frac{2}{3}}t\right) = \begin{cases} 0, & \text{si } \sqrt{\frac{2}{3}}t < -1, \\ \frac{1}{3}, & \text{si } -1 \leq \sqrt{\frac{2}{3}}t < 0, \\ \frac{2}{3}, & \text{si } 0 \leq \sqrt{\frac{2}{3}}t < 1, \\ 1, & \text{si } \sqrt{\frac{2}{3}}t \geq 1, \end{cases} = \begin{cases} 0, & \text{si } t < -\sqrt{\frac{3}{2}}, \\ \frac{1}{3}, & \text{si } -\sqrt{\frac{3}{2}} \leq t < 0, \\ \frac{2}{3}, & \text{si } 0 \leq t < \sqrt{\frac{3}{2}}, \\ 1, & \text{si } t \geq \sqrt{\frac{3}{2}}. \end{cases}$$

3.13.- Considerem un examen tipus TEST amb n preguntes on cada pregunta té només k possibles respostes. En cada pregunta només hi ha una sola resposta certa.
Suposem que l'alumne contesta a l'atzar i que la probabilitat de contestar a qualsevol pregunta val $\frac{1}{2}$.

a) Si l'alumne encerta una pregunta li donam 1 punt, si la deixa en blanc, 0 punts i si falla, li restam x punts. Trobau el valor de x perquè l'esperança d'encertar una pregunta a l'atzar sigui 0.

b) Considerem $n = 5$ i $k = 5$. Sigui X la variable aleatòria que ens dóna la puntuació a l'examen tenint en compte a), o sigui, que si l'alumne contesta malament una pregunta, li restam el que li correspongui.
Trobau la funció de probabilitat de X, EX i Var X.
Final. Juny 93.

Resolució.

a) Definim la variable aleatòria discreta:

$$P = \text{“Puntuació d'una pregunta”}.$$

El rang de la variable aleatòria P és: $P(\Omega) = \{-x, 0, 1\}$.

La funció de probabilitat de P és:

p	-x	0	1
$f_X(p)$	$\frac{k-1}{2k}$	$\frac{1}{2}$	$\frac{1}{2k}$

Expliquem una mica els resultats anteriors:

$$\begin{aligned}
p\{P=-x\} &= p\{\text{“fallar la pregunta”}\} \\
&= p\{\text{“contestar la pregunta”} \cap \text{“fallar la pregunta”}\} \\
&= p\{\text{“contestar la pregunta”}\} \cdot p\{\text{“fallar la pregunta”}\} \\
&= \tfrac{1}{2} \cdot \tfrac{k-1}{k} = \tfrac{k-1}{2k}, \\
p\{P=0\} &= p\{\text{“no contestar la pregunta”}\} = \tfrac{1}{2}, \\
p\{P=1\} &= p\{\text{“encertar la pregunta”}\} \\
&= p\{\text{“contestar la pregunta”} \cap \text{“encertar la pregunta”}\} \\
&= p\{\text{“contestar la pregunta”}\} \cdot p\{\text{“encertar la pregunta”}\} \\
&= \tfrac{1}{2} \cdot \tfrac{1}{k} = \tfrac{1}{2k}.
\end{aligned}$$

A continuació, trobem $\mathrm{E}P$:

$$\mathrm{E}P = -\frac{k-1}{2k} \cdot x + \frac{1}{2} \cdot 0 + \frac{1}{2k} \cdot 1 = \frac{1-x(k-1)}{2k}.$$

Si imposam que $\mathrm{E}P = 0$ ens resulta que el valor de x és $x = \frac{1}{k-1}$.

b) Considerem que el nombre de preguntes és $n = 5$ i el nombre de respostes és $k = 5$.

Fixau-vos que si contesta i preguntes bé i j preguntes malament, la nota que té, tenint en compte a), és: Nota$= i - \frac{j}{4}$.

A més a més, tenim que la probabilitat que contesti i preguntes bé i j malament val:

$$\begin{aligned}
p\{i \text{ preguntes bé } \cap j \text{ preguntes malament}\} &= \\
PR_5^{i,j,5-i-j} \cdot \left(\tfrac{1}{2\cdot5}\right)^i \cdot \left(\tfrac{4}{2\cdot5}\right)^j \cdot \left(\tfrac{1}{2}\right)^{5-i-j} &= \tfrac{5!}{i!j!(5-i-j)!} \cdot \tfrac{4^j}{2^5 5^{i+j}}.
\end{aligned} \tag{3.7}$$

Farem una taula on queda indicada la nota, suposant que hi ha i preguntes ben contestades i j de mal contestades:

$j\backslash i$	0	1	2	3	4	5
0	0.	1.	2.	3.	4.	5.
1	-0.25	0.75	1.75	2.75	3.75	–
2	-0.5	0.5	1.5	2.5	–	–
3	-0.75	0.25	1.25	–	–	–
4	-1.	0.	–	–	–	–
5	-1.25	–	–	–	–	–

De la taula anterior podem deduir que el rang de la variable aleatòria $X =$"nota" és el següent:

$$\begin{aligned} X(\Omega) = \{ \quad & -1.25, -1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1, \\ & 1.25, 1.5, 1.75, 2, 2.5, 2.75, 3, 3.75, 4, 5\}. \end{aligned}$$

Per trobar la funció de probabilitat de X fent servir la taula anterior, fixau-vos que, per a cada nota, existeix una única i i una única j tals que:

$$f_X(x) = p\{X = x\} = p\{i \text{ preguntes bé } \cap j \text{ preguntes malament}\}.$$

El valor de la darrera probabilitat el dóna la fórmula (3.7).

L'únic valor del rang de X on falla la fórmula anterior és el valor 0. Fixau-vos que per al valor 0 hi ha dos possibles i (preguntes ben contestades) i dos possibles j (preguntes mal contestades): $i = 0, j = 0$ i $i = 1, j = 4$. Així doncs:

$$p\{X = 0\} = \frac{5!}{0!0!5!} \cdot \frac{4^0}{2^5 5^0} + \frac{5!}{1!4!0!} \cdot \frac{4^4}{2^5 5^5} = 0.04405.$$

Exposam a continuació la taula de la funció de probabilitat de X:

x	-1.25	-1.	-0.75	-0.5	-0.25	0.	0.25
$f_X(s)$	0.01024	0.064	0.16	0.2	0.125	0.04405	0.064
x	0.5	0.75	1.	1.25	1.5	1.75	2.
$f_X(s)$	0.12	0.1	0.03125	0.0064	0.024	0.03	0.0125
x	2.5	2.75	3.	3.75	4.	5.	
$f_X(x)$	0.0016	0.004	0.0025	0.0002	0.00025	0.00001	

Per trobar $\mathrm{E}X$, hem de tenir en compte que $X = \sum_{i=1}^{5} P_i$ on P_i és la variable aleatòria que ens dóna la puntuació de la pregunta i-èssima.

Tenint en compte l'apartat a), $\mathrm{E}(P_i) = 0$. D'aquí, doncs, $\mathrm{E}X = 0$.

La variància de X es podrà trobar com:

$$\begin{aligned}\text{Var } X &= \mathrm{E}\left(X^2\right) = \sum_{x_i \in X(\Omega)} x_i^2 f_X(x_i) = (-1.25)^2 \cdot 0.01024 + \cdots + 5^2 \cdot 0.00001 \\ &= \frac{5}{8} = 0.625.\end{aligned}$$

De fet, quan s'introdueixi el concepte d'independència de variables aleatòries, podrem trobar la variància d'una forma molt semblant al càlcul de l'esperança. Més concretament, tenint en compte que les variables P_i són independents, podem escriure:

$$\text{Var } X = \sum_{i=1}^{5} \text{Var } P_i = 5\text{Var } P,$$

on P és qualsevol variable de les P_i.

El càlcul de Var P és molt senzill:

$$\text{Var } P = \mathrm{E}\left(P^2\right) = (-0.25)^2 \cdot \frac{4}{10} + 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{10} = \frac{1}{8}.$$

Per tant, Var $X = \frac{5}{8}$.

3.14.- S'ha estimat que el temps de vida X, en hores, d'un cert component electrònic segueix una distribució donada per la densitat

$$f(x) = \begin{cases} 0, & \text{si } x < 0, \\ \frac{1}{8}\mathrm{e}^{-x/8}, & \text{en cas contrari.} \end{cases}$$

El departament de control de qualitat rebutja tots els components que fallen durant les 3 primeres hores i comercialitza la resta.

a) Determinau la distribució del temps de vida dels components comercialitzats.

b) Quina és la probabilitat que un component comercialitzat funcioni més de 12 hores?

Resolució.

a) Posem:

Y : temps de vida d'un component comercialitzat.

Per pròpia definició, Y pren valors més grans que 3. Aleshores, si $y \leq 3$, $F_Y(y) = p\{Y \leq y\} = 0$. D'altra banda, si $y > 3$,

$$\begin{aligned} F_Y(y) &= p\{Y \leq y\} = p\{X \leq y/X > 3\} \\ &= \frac{p\{X \leq y, X > 3\}}{p\{X > 3\}} = \frac{p\{3 < X \leq y\}}{1 - p\{X \leq 3\}} = \frac{F_X(y) - F_X(3)}{1 - F_X(3)}. \end{aligned}$$

Tenim

$$F_X(y) = \int_0^y \frac{1}{8}\mathrm{e}^{-x/8}\, dx = \frac{1}{8}\left[\frac{\mathrm{e}^{-x/8}}{-1/8}\right]_0^y = 1 - \mathrm{e}^{-y/8}.$$

Així

$$F_X(3) = 1 - \mathrm{e}^{-3/8},$$

i resulta

$$F_Y(y) = \frac{1 - \mathrm{e}^{-y/8} - 1 + \mathrm{e}^{-3/8}}{1 - 1 + \mathrm{e}^{-3/8}} = 1 - \mathrm{e}^{-\frac{1}{8}(y-3)}.$$

Finalment,

$$F_Y(y) = \begin{cases} 0, & \text{si } y \leq 3, \\ 1 - \mathrm{e}^{-\frac{1}{8}(y-3)}, & \text{en cas contrari.} \end{cases}$$

b)

$$\begin{aligned} p\{Y > 12\} &= 1 - p\{Y \leq 12\} = 1 - F_Y(12) \\ &= 1 - 1 + \mathrm{e}^{-\frac{1}{8}(12-3)} = \mathrm{e}^{-9/8}. \end{aligned}$$

3.15.- La funció de densitat d'una variable aleatòria X és:

$$f_X(x) = \begin{cases} x+1, & \text{si } x \in (-1, 0], \\ -x+1, & \text{si } x \in (0, 1], \\ 0, & \text{si } x \in (-\infty, -1] \cup (1, \infty). \end{cases}$$

Definim la variable aleatòria $Y = g(X)$, on g és la funció

$$g(x) = \begin{cases} 1, & \text{si } x \in (1/2, \infty), \\ 0, & \text{si } x \in (-1/2, 1/2], \\ -1, & \text{si } x \in (-\infty, -1/2]. \end{cases}$$

Determinau la funció de probabilitat i la de distribució de Y.

Resolució. Tenim

$$\begin{aligned} p\{Y = 1\} &= p\{X > 1/2\} = 1 - p\{X \leq 1/2\} = 1 - \int_{-\infty}^{1/2} f_X(x)\, dx \\ &= 1 - \int_{-1}^{0} (x+1)\, dx - \int_{0}^{1/2} (-x+1)\, dx \\ &= 1 - \left[\frac{x^2}{2} + x\right]_{-1}^{0} - \left[\frac{-x^2}{2} + x\right]_{0}^{1/2} \\ &= 1 + \frac{1}{2} - 1 + \frac{1}{8} - \frac{1}{2} = \frac{1}{8}. \end{aligned}$$

També

$$p\{Y = -1\} = p\{X \leq -1/2\} = \int_{-\infty}^{-1/2} f_X(x)\, dx$$

$$= \int_{-1}^{-1/2} (x+1)\,dx = \left[\frac{x^2}{2} + x\right]_{-1}^{-1/2} = \frac{1}{8} - \frac{1}{2} - \frac{1}{2} + 1 = \frac{1}{8}.$$

Finalment,

$$p\{Y=0\} = 1 - [p\{Y=1\} + p\{Y=-1\}] = 1 - \frac{1}{8} - \frac{1}{8} = \frac{3}{4}.$$

Així, la funció de distribució de la Y serà

$$F_Y(x) = \begin{cases} 0, & \text{si } x < -1, \\ 1/8, & \text{si } -1 \leq x < 0, \\ 7/8, & \text{si } 0 \leq x < 1, \\ 1, & \text{si } x \geq 1. \end{cases}$$

3.16.- Un joc es diu que és just si el guany esperat de cada jugador és 0. Dos jugadors A i B tiren un dau per torns i guanya el primer que obté un '5'. Cada jugador aposta una quantitat $c_j (j = 1, 2)$ i el total se'l quedarà el guanyador. Si suposam que comença a jugar A, quina condició han de verificar c_1 i c_2 perquè el joc sigui just?

Resolució. Calculem la probabilitat que guanyi cada jugador. Posem

A_i : el jugador A obté un '5' en la jugada i-èssima,
B_i : el jugador B obté un '5' en la jugada i-èssima,

entenent en cada cas que s'obté el '5' per primera vegada. Aleshores

$$p\{A\} = \sum_{i=1}^{\infty} p\{A_{2i-1}\},$$

on $p\{A\}$ és la probabilitat que guanyi el jugador A.

Ara bé,

$$p\{A_1\} = \frac{1}{6},\ p\{A_3\} = \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}.$$

En general,

$$p\{A_{2i-1}\} = \left(\frac{5}{6}\right)^{2i-2} \cdot \frac{1}{6}.$$

Per tant

$$\begin{aligned} p\{A\} &= \sum_{i=1}^{\infty}\left(\frac{5}{6}\right)^{2i-2}\cdot\frac{1}{6}=\frac{1}{6}\cdot\sum_{i=1}^{\infty}\left(\frac{25}{36}\right)^{i}\cdot\frac{1}{25/36}=\frac{6}{25}\cdot\sum_{i=1}^{\infty}\left(\frac{25}{36}\right)^{i} \\ &= \frac{6}{25}\cdot\frac{\frac{25}{36}}{1-\frac{25}{36}}=\frac{6}{25}\cdot\frac{\frac{25}{36}}{\frac{11}{36}}=\frac{6}{11}. \end{aligned}$$

Anàlogament,

$$p\{B_2\}=\frac{5}{6}\cdot\frac{1}{6},\; p\{B_4\}=\left(\frac{5}{6}\right)^{3}\cdot\frac{1}{6},$$

i, en general,

$$p\{B_{2i}\}=\left(\frac{5}{6}\right)^{2i-1}\cdot\frac{1}{6}.$$

Aleshores

$$p\{B\}=\sum_{i=1}^{\infty}p\{B_{2i}\}=\sum_{i=1}^{\infty}\left(\frac{5}{6}\right)^{2i-1}\cdot\frac{1}{6}=\frac{1}{6}\cdot\sum_{i=1}^{\infty}\left(\frac{25}{36}\right)^{i}\cdot\frac{1}{5/6}=\frac{1}{5}\cdot\frac{\frac{25}{36}}{\frac{11}{36}}=\frac{5}{11},$$

on $p\{B\}$ és la probabilitat que guanyi el jugador B.

Si ara X : guany del jugador A, $X(\Omega)=\{c_2,-c_1\}$ i la funció de probabilitat és:

$$p\{X=c_2\}=p\{A\}=6/11,\; p\{X=-c_1\}=p\{B\}=5/11.$$

Aleshores $\mathrm{E}X=\frac{6}{11}\cdot c_2=\frac{5}{11}\cdot c_1$. Si el joc ha de ser just,

$$\mathrm{E}X=0\Longrightarrow 6c_2=5c_1.$$

Observem que si Y : guany del jugador B, aleshores $Y=-X$ i $\mathrm{E}Y=-\mathrm{E}X$ i, per tant, s'obté la mateixa condició.

3.17.- El preu per estacionament en un aparcament és de 75 pts. per a la primera hora o fracció, i de 60 pts. a partir de la segona hora o fracció. Suposem que el temps, en hores, que un vehicle qualsevol roman a l'aparcament es modelitza segons la funció de densitat

$$f_X(x)=\begin{cases}\mathrm{e}^{-x}, & \text{si } x\geq 0,\\ 0, & \text{en cas contrari.}\end{cases}$$

Calculau l'ingrés mitjà per vehicle

Resolució. Siguin Y : ingrés per aparcament i Z : nombre d'hores (inclosa la fracció de la darrera hora) que un vehicle està aparcat. Aleshores $Y = 75 + 60(Z-1)$. Z pren valors $1, 2, \ldots$ amb probabilitats:

$$\begin{aligned} p\{Z=1\} &= p\{0 \le X \le 1\} = \int_0^1 \mathrm{e}^{-x}\,dx = 1 - \mathrm{e}^{-1} \\ p\{Z=2\} &= p\{1 \le X \le 2\} = \int_1^2 \mathrm{e}^{-x}\,dx = \mathrm{e}^{-1} - \mathrm{e}^{-2} \\ &\vdots \\ p\{Z=k\} &= p\{k-1 \le X \le k\} = \int_{k-1}^k \mathrm{e}^{-x}\,dx = \mathrm{e}^{-k+1} - \mathrm{e}^{-k}. \end{aligned}$$

Per tant

$$\begin{aligned} \mathrm{E}Z &= \sum_{k\ge 1} k \cdot p\{Z=k\} = \sum_{k \ge 1} k \cdot (\mathrm{e}^{-k+1} - \mathrm{e}^{-k}) \\ &= 1 - \mathrm{e}^{-1} + 2\mathrm{e}^{-1} - 2\mathrm{e}^{-2} + 3\mathrm{e}^{-2} - 3\mathrm{e}^{-3} + \cdots \\ &= 1 + \mathrm{e}^{-1} + \mathrm{e}^{-2} + \cdots = \sum_{k \ge 0} \mathrm{e}^{-k} = \frac{1}{1-\mathrm{e}^{-1}} \approx 1.582. \end{aligned}$$

Finalment,

$$\mathrm{E}Y = \mathrm{E}(75 + 60(Z-1)) = 75 + 60(\mathrm{E}Z - 1) \approx 75 + 60 \times 0.582 = 109.92.$$

3.3 Problemes proposats

3.1.- Hi ha 10 estudiants inscrits en una classe d'Estadística, d'entre els quals 3 tenen 19 anys, 4 tenen 20 anys, 1 té 21 anys, 1 té 24 anys i 1 té 26 anys. D'aquesta classe se seleccionen dos estudiants sense reposició. Sigui X la edat mitjana dels dos estudiants seleccionats. Trobau la funció de probabilitat per a X.

3.2.- Verificau que:

$$F_W(t) = \begin{cases} 0, & \text{si } t < 3, \\ \frac{1}{3}, & \text{si } 3 \le t < 4, \\ \frac{1}{2}, & \text{si } 4 \le t < 5, \\ \frac{2}{3}, & \text{si } 5 \le t < 6, \\ 1, & \text{si } t \ge 6, \end{cases}$$

és una funció de distribució i especificau la funció de probabilitat per a W. Trobau també $p\{3 < W \le 5\}$.

3.3.- La variable aleatòria Z té per funció de probabilitat:

$$f_Z(x) = \begin{cases} \frac{1}{3}, & \text{si } x = 0, 1, 2, \\ 0, & \text{en els altres casos.} \end{cases}$$

Quina és la funció de distribució per a Z?

3.4.- Verificau que:

$$F_X(t) = \begin{cases} 0, & \text{si } t < -1, \\ \frac{t+1}{2}, & \text{si } -1 \le t \le 1, \\ 1, & \text{si } t > 1, \end{cases}$$

és una funció de distribució i trobau la funció de densitat per a X. Calculau també $p\left\{-\frac{1}{2} \le X \le \frac{1}{2}\right\}$.

3.5.- Sigui Y una variable contínua amb funció de densitat:

$$f_Y(y) = \begin{cases} 2(1-y), & \text{si } 0 < y < 1, \\ 0, & \text{en els altres casos.} \end{cases}$$

Trobau la funció de distribució $F_Y(t)$.

3.6.- Verificau que:

$$F_Y(t) = \begin{cases} 0, & \text{si } t < 0, \\ \sqrt{t}, & \text{si } 0 \leq t \leq 1, \\ 1, & \text{si } t > 1, \end{cases}$$

és una funció de distribució i especificau la funció de densitat per a Y. Feu servir aquest resultat per trobar $p\left\{-\frac{1}{2} < Y < \frac{3}{4}\right\}$.

3.7.- Considerem la següent funció de distribució per a una variable aleatòria U:

$$F_U(t) = \begin{cases} 0, & \text{si } t < 1, \\ \ln t, & \text{si } 1 \leq t \leq \mathrm{e}, \\ 1, & \text{si } t > \mathrm{e}. \end{cases}$$

Trobau la mediana (el percentil 50%), $\mathrm{E}(U)$ i $\mathrm{Var}\, U$.

3.8.- Es vol rifar un cotxe amb un valor de 3.000$, per la qual cosa es venen 10.000 paperetes que valen 1$ cada una. Si es compra 1 papereta, quin és el benefici esperat? Quin és el benefici esperat si es compren 100 paperetes? Trobau la variància en cada un dels dos casos.

3.9.- Suposau que Y, el nombre de minuts que es necessiten per dinar en qualsevol dia, té la mateixa versemblança (probabilitat) d'estar en l'interval de 30 a 40 minuts. Trobau $m_Y(t)$ i $m_{Y-\mu_Y}(t)$.

3.10.- Se selecciona a l'atzar un nombre del conjunt $\{1, 2, 3, \ldots, n\}$. Sigui V l'enter seleccionat. Trobau $\psi_V(t)$.

3.11.- Suposau que X és una variable aleatòria (discreta) amb funció generatriu

$$\psi_X(t) = \frac{pt}{1-(1-p)t}, \text{ on } 0 < p < 1.$$

Trobau $f_X(x)$ (funció de probabilitat) i $m_X(t)$.

3.12.- Sigui X una variable aleatòria amb funció de distribució $F_X(t)$ i sigui $Y = a + bX$ on $b < 0$. Trobau la funció de distribució per a la variable Y.

3.13.- A partir de

$$F_X(t) = \begin{cases} 0, & \text{si } t < -1, \\ \\ \frac{t+1}{2}, & \text{si } -1 \leq t \leq 1, \\ \\ 1, & \text{si } t > 1, \end{cases}$$

trobau la funció de distribució per a $Y = 15 + 2X$ i la funció de densitat per Y.

3.14.- Si

$$F_X(t) = \begin{cases} 0, & \text{si } t < 0, \\ t, & \text{si } 0 \leq t \leq 1, \\ 1, & \text{si } t > 1, \end{cases}$$

trobau la funció de distribució i la funció de densitat de la forma estàndard de X $(Z = \frac{X-\mu_X}{\sigma_X})$.

3.15.- Per formar un jardí circular, un jardiner talla una corda, la ferma a una estaca i hi marca el perímetre. Suposau que la llargada de la corda té la mateixa versemblança d'estar en l'interval comprès entre $r - 0.1$ i $r + 0.1$. Quina és la distribució de X, l'àrea de la superfície del jardí? Quina és la probabilitat que l'àrea de la superfície sigui més gran que πr^2?

3.16.- En un determinat lloc, el 60% dels conductors respecten els senyals de reducció de velocitat, mentre que el 40% no els respecten. En aquest lloc hi ha un semàfor i uns 100 metres abans d'arribar-hi hi ha un senyal de reducció de velocitat. Per als conductors que respecten l'esmentat senyal el semàfor continua verd però per als conductors que van a una velocitat elevada sense fer cas del senyal, el semàfor es posa vermell.
Considerem la variable aleatòria $X =$"Nombre de cotxes que passen pel semàfor sense aturar-se a partir d'una certa hora".
Trobau la funció de probabilitat de X, EX i Var X.
Indicació: $\sum_{n \geq 1} nx^n = \frac{x}{(1-x)^2}$, $\sum_{n \geq 1} n^2 x^n = \frac{x(x+1)}{(1-x)^3}$.
Final. Setembre 93.

3.17.- Sigui X una variable aleatòria contínua amb funció de densitat $f_X(x)$. Considerem la variable aleatòria $Y = e^X$. Trobau la funció de densitat de la variable aleatòria Y, $f_Y(y)$.
Final. Setembre 94.

3.18.- Sigui

$$f(x) = \begin{cases} ax \sin x, & \text{si } x \in [0, \pi], \\ 0, & \text{en cas contrari.} \end{cases}$$

Trobau el valor de a perquè $f(x)$ sigui una funció de densitat.

Primer parcial. Febrer 95.

3.19.- Sigui X una variable aleatòria $\mathrm{Exp}(\lambda)$ (exponencial amb paràmetre λ). Considerem la variable $Y = \mathrm{e}^X$. Trobau la funció de densitat de Y $f_Y(t)$.

Primer parcial. Febrer 95.

3.20.- Sigui $f(x) = \begin{cases} \frac{C}{1+x}, & \text{si } x \in (0, 2), \\ 0, & \text{en cas contrari,} \end{cases}$ on C es calcula de tal forma que $f(x)$ sigui una densitat. Trobau el valor de l'esperança de X ($\mathrm{E}(X)$).

Final. Juny 96.

3.21.- Sigui X una variable aleatòria amb funció generatriu $m_X(t)$. Trobau la funció generatriu de la variable aleatòria $X + 1$.

Final. Setembre 96.

3.22.- Sigui X una variable aleatòria contínua amb funció de densitat:

$$f(x) = \begin{cases} \frac{1}{2} \sin x, & \text{si } x \in [0, \pi], \\ 0, & \text{en cas contrari.} \end{cases}$$

a) Trobau la funció generatriu de moments, $\mathrm{E}(X)$ i $\mathrm{Var}\ (X)$.

b) Calculau $p\left\{X^2 - 3X + 2 > 0\right\}$.

Final. Setembre 96.

3.23.- La durada de les conferències telefòniques és una variable aleatòria amb funció de distribució:

$$F(x) = \begin{cases} 0, & \text{si } x \leq 0, \\ 1 - \frac{1}{2}\mathrm{e}^{-x/3} - \frac{1}{2}\mathrm{e}^{-E[x/3]}, & \text{en cas contrari,} \end{cases}$$

on $E[x]$ és la part entera de x. Trobau la probabilitat que una conferència duri:

a) més de 6 minuts,

b) menys de 4 minuts,

c) exactament 3 minuts,

d) menys de 9 minuts, sabent que n'ha durat més de 5,

e) més de 5 minuts, sabent que n'ha durat menys de 9.

3.24.- El temps de vida, en anys, d'un cert component d'una màquina es modelitza mitjançant la funció de densitat següent:

$$f_X(x) = \begin{cases} \mathrm{e}^{-x}, & \text{si } x \geq 0, \\ 0, & \text{en cas contrari.} \end{cases}$$

Si el cost Y de funcionament del component, en milions de ptes., és funció del temps de vida: $Y = 2X^2 + 1$, calculau la quantitat que espera gastar l'empresa en concepte de manteniment.

3.25.- Un sistema de transmissió emet els dígits -1, 0 i 1. Quan es transmet el símbol i, es rep el símbol j amb les probabilitats següents: $p\{r_1/t_1\} = 1$, $p\{r_{-1}/t_{-1}\} = 1$, $p\{r_1/t_0\} = 0.1$, $p\{r_{-1}/t_0\} = 0.1$, $p\{r_0/t_0\} = 0.8$. Es diu en aquest cas que s'ha produït una distorsió $(i-j)^2$. Quin és el valor mitjà de la distorsió?

3.26.- Una font binària emet de manera equiprobable i independent un bloc de 3 dígits (0 o 1) cada segon. De cada bloc envia a un canal de transmissió un 0 si al bloc hi ha més 0 que 1, i un 1 en cas contrari. El canal transmet el dígit amb una probabilitat d'error p. El receptor reconstrueix la terna de dígits repetint tres vegades el dígit que ha rebut. Quin és el nombre mitjà de bits erronis per bloc? Quina hauria de ser la probabilitat p per tal que aquest valor mitjà no fos més gran que 1?

Capítol 4

Estudi d'algunes distribucions conegudes

4.1 Resum teòric

1. **Bernoulli,** $B(1, p)$**.**

   Considerem un experiment aleatori amb dos resultats possibles, a_1 i a_2, amb probabilitats respectives $p \in (0, 1)$ i $q = 1 - p$.

   Definim la variable aleatòria $X : \Omega \to \mathbb{R}$ donada per $X(a_1) = 1$, $X(a_2) = 0$. a_1 se sol anomenar **encert** i a_2 **fallada**. La funció de probabilitat de X és $p\{X = 1\} = p$, $p\{X = 0\} = q$, i X s'anomena **variable aleatòria de Bernoulli amb paràmetre** p, indicada amb $B(1, p)$. Cada experiment s'anomena una **prova de Bernoulli.** Resulta: $\mathrm{E}X = p$ i $\mathrm{Var}X = p \cdot q$.

2. **Binomial,** $B(n, p)$**.**

   Considerem una successió de n repeticions independents d'una prova de Bernoulli amb paràmetre p. Sigui X la variable aleatòria que dóna el nombre d'encerts en les n proves. Aleshores $X(\Omega) = \{0, 1, \ldots, n\}$, i

   $$p\{X = k\} = \binom{n}{k} p^k q^{n-k}.$$

   X s'anomena **variable aleatòria binomial amb paràmetres** n **i** p , i s'indica amb $B(n, p)$. Resulta: $\mathrm{E}X = np, \mathrm{Var}X = npq$.

3. **Geomètrica,** $Geom(p)$**.**

   Considerem una successió de repeticions independents d'una prova de Bernoulli amb paràmetre p. Sigui X la variable aleatòria que dóna el nombre de repeticions de l'experiment fins que s'obté el primer encert. Aleshores $X(\Omega) = \{1, 2, \ldots\}$, i

   $$p\{X = k\} = p \cdot q^{k-1}.$$

   X s'anomena **variable aleatòria geomètrica amb paràmetre** p. Resulta: $\mathrm{E}X = 1/p, \mathrm{Var}X = 1/p^2$.

   De vegades, es pren com a X el nombre de repeticions de l'experiment abans d'obtenir el primer encert. Aleshores $X(\Omega) = \{0, 1, 2, \ldots\}$, i $p\{X = k\} = p \cdot q^k$. En aquest cas, $\mathrm{E}X = q/p$, i $\mathrm{Var}X = q/p^2$.

   X també s'anomena el **temps d'espera**. Aquesta variable aleatòria queda caracteritzada dins les variables aleatòries discretes per la propietat anomenada de **falta de memòria** :

   $$p\{X \geq m + n/X \geq m\} = p\{X \geq n + 1\} \ \ \forall m, n.$$

   (La probabilitat que un determinat component superi més de $m + n$ proves sabent que n'ha superat més de m és igual a la que superi més de n proves).

4. **Poisson,** $Poiss(\lambda)$**.**

   Direm que X, una variable aleatòria discreta amb $X(\Omega) = \{0, 1, 2, \ldots\}$, té la **distribució de Poisson amb paràmetre** $\lambda > 0$, indicada amb $Poiss(\lambda)$, si

   $$p\{X = k\} = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \ k \geq 0.$$

   Resulta: $\mathrm{E}X = \lambda, \mathrm{Var}X = \lambda$.

   Aquesta distribució és molt important i pot aparèixer de dues maneres:

   (a) Com a límit d'una $B(n, p)$ quan $n \to \infty$ i $p \to 0$ de manera que $\lambda = np$ es conserva constant. A la pràctica, l'aproximació és bona per a $n \geq 30$ i $p \leq 0.1$.

   (b) Com a model d'un procés aleatori que dóna lloc a successos al llarg de l'espai o del temps amb les condicions següents:

- el nombre de successos que tenen lloc en una unitat arbitrària de temps (espai) és independent del nombre de successos que tenen lloc en una altra unitat de la mateixa mida,
- la probabilitat que ocorri un succés en una unitat molt petita és proporcional a la mida de la unitat,
- la probabilitat que tengui lloc més d'un succés en una unitat molt petita és despreciable.

Si λ és el ritme d'aparició dels successos, resulta que el nombre de successos que apareixen en l'interval $(0, t]$ és una variable aleatòria $Poiss(\lambda t)$.

5. **Hipergeomètrica.**

S'aplica a les extraccions sense reemplaçament. Considerem una urna amb N bolles, de les quals N_1 són negres i N_2 blanques ($N_1 + N_2 = N$). Extraiem n bolles sense reemplaçament. Considerem la variable aleatòria X que dóna el nombre de bolles negres que han sortit. Aleshores resulta que X té una funció de probabilitat donada per:

$$p\{X = k\} = \frac{\binom{N_1}{k}\binom{N_2}{n-k}}{\binom{N}{n}}.$$

X s'anomena una **variable aleatòria hipergeomètrica**. Resulta: $\mathrm{E}X = np$, i $\mathrm{Var}X = npq\frac{N-n}{N-1}$, on $p = \frac{N_1}{N}$ i $q = 1 - p$.

6. **Uniforme,** $U[a, b]$**.**

Siguin $a, b \in \mathbb{R},\ a < b$. Definim $f : \mathbb{R} \to \mathbb{R}$ com

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a, b], \\ 0, & \text{en cas contrari.} \end{cases}$$

f és una funció de densitat i la corresponent funció de distribució és

$$F(x) = \begin{cases} 0, & \text{si } x \leq a, \\ \frac{x-a}{b-a}, & \text{si } a \leq x \leq b, \\ 1, & \text{si } x \geq b. \end{cases}$$

La variable aleatòria X que té per funció de distribució F s'anomena **variable aleatòria uniforme en l'interval** $[a, b]$ i s'indica amb $U[a, b]$. Resulta

$$\mathrm{E}X = \frac{a+b}{2},\ \mathrm{Var}X = \frac{(b-a)^2}{12}.$$

7. **Exponencial, Exp(λ).**

   Fixem $\lambda > 0$. Direm que una variable aleatòria absolutament contínua X té una **distribució exponencial amb paràmetre** λ, indicada amb Exp(λ), si la seva densitat la dóna:

   $$f(x) = \lambda \cdot \mathrm{e}^{-\lambda x}, \ x \geq 0.$$

   La funció de distribució serà aleshores:

   $$F(x) = 1 - \mathrm{e}^{-\lambda x}, \ x \geq 0.$$

   Una manera d'introduir la distribució exponencial és a partir de la de Poisson. Considerem un flux de successos aleatoris com en un procés de Poisson amb ritme λ. Sigui X la variable aleatòria que dóna l'interval de temps entre dos successos consecutius. Aleshores X té una distribució Exp(λ). Resulta: $\mathrm{E}X = 1/\lambda, \mathrm{Var}X = 1/\lambda^2$.

   Dins les variables aleatòries absolutament contínues, les exponencials queden caracteritzades per la propietat de falta de memòria.

8. **Normal,** $N(\mu, \sigma^2)$**.**

   És probablement la distribució absolutament contínua més important. És la que té per densitat la funció:

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

   la gràfica de la qual és la coneguda campana de Gauss. Es diu que X té la **distribució normal** o **gaussiana**. Resulta: $\mathrm{E}X = \mu$ i $\mathrm{Var}X = \sigma^2$.

   Si consideram $\mu = 0$ i $\sigma^2 = 1$, obtenim la **distribució normal estàndard** $N(0, 1)$.

   Donada una variable aleatòria X amb distribució $N(\mu, \sigma^2)$, la variable aleatòria $Y = \dfrac{X - \mu}{\sigma}$ té una distribució $N(0, 1)$. Recíprocament, si X té una distribució $N(0, 1)$, aleshores $Y = \sigma \cdot X + \mu$ té una distribució $N(\mu, \sigma^2)$.

4.2 Problemes resolts

4.1.- En una urna hi ha 8 bolles vermelles i 2 de negres. Es treuen 20 bolles amb reposició. Sigui Y la variable aleatòria que dóna el nombre de bolles vermelles que surten. Calculau μ_Y, σ_Y, $p\{Y = 16\}$, $p\{Y < 14\}$ i $p\{Y > 18\}$.

Resolució. La variable aleatòria Y és una variable binomial amb paràmetres $n = 20$ i $p =$"probabilitat que una bolla sigui vermella"$= \frac{8}{10} = \frac{4}{5}$.

Per tant,

$$\begin{aligned}
\mu_Y &= np = 20 \cdot \frac{4}{5} = 16, \\
\sigma_Y &= \sqrt{np(1-p)} = \sqrt{20 \cdot \frac{4}{5} \cdot \frac{1}{5}} = \sqrt{\frac{80}{25}} \approx 1.7888, \\
p\{Y = 16\} &= \binom{20}{16} \left(\frac{4}{5}\right)^{16} \left(\frac{1}{5}\right)^{4} \approx 0.2182, \\
p\{Y < 14\} &= 1 - p\{Y \geq 14\} = 1 - \sum_{i=14}^{20} \binom{20}{i} \left(\frac{4}{5}\right)^{i} \left(\frac{1}{5}\right)^{20-i} \approx 0.0867, \\
p\{Y > 18\} &= \binom{20}{19} \left(\frac{4}{5}\right)^{19} \left(\frac{1}{5}\right) + \binom{20}{20} \left(\frac{4}{5}\right)^{20} \approx 0.06917.
\end{aligned}$$

4.2.- Trobau la funció generatriu dels moments per a una variable aleatòria binomial amb paràmetres n i p.

Resolució. Si X és una variable aleatòria binomial amb paràmetres n i p, aleshores el rang de X és:

$$X(\Omega) = \{0, 1, \ldots, n\},$$

amb funció de probabilitat:

$$f_X(i) = \binom{n}{i} p^i (1-p)^{n-i} := \binom{n}{i} p^i q^{n-i} \ (q = 1-p).$$

Per tant, la funció generatriu de moments serà:

$$m_X(t) = \mathrm{E}\left(\mathrm{e}^{tx}\right) = \sum_{i=0}^{n} \mathrm{e}^{ti} f_X(i) = \sum_{i=0}^{n} \mathrm{e}^{ti} \binom{n}{i} p^i q^{n-i}$$

$$= \sum_{i=0}^{n} \left(pe^t\right)^i \binom{n}{i} q^{n-i} = \left(pe^t + q\right)^n.$$

4.3.- Es llança una moneda no trucada fins que surti cara. Quina és la probabilitat que es necessitin menys de 3 intents? I la que es necessitin menys de 4 intents?

Resolució. La variable aleatòria $X =$"nombre d'intents necessaris fins que surti cara" és una variable aleatòria geomètrica amb paràmetre $p = \frac{1}{2}$, ja que la moneda no està trucada.

El rang de X és

$$X(\Omega) = \{1, 2, \ldots\} = \{n \in \mathbb{Z} | n \geq 1\},$$

amb funció de probabilitat:

$$f_X(n) = p \cdot (1-p)^{n-1} := pq^{n-1}, \ n \geq 1.$$

Per tant:

$$p\{X \leq 3\} = \sum_{i=1}^{3} p \cdot q^{i-1} = \frac{7}{8}, \quad p\{X \leq 4\} = \sum_{i=1}^{4} p \cdot q^{i-1} = \frac{15}{16}.$$

4.4.- Una caixa conté 5 bolles, 3 de les quals estan fetes malbé. Escollim dues bolles a l'atzar sense reposició. Quina és la funció de probabilitat per al nombre de bolles fetes malbé en la mostra?

Resolució. Considerem la variable aleatòria $X =$"nombre de bolles escollides fetes malbé". Fixau-vos que X és una variable hipergeomètrica amb paràmetres $M = 5$ (nombre total de bolles), $W = 3$ (nombre de bolles d'un determinat tipus; en aquest cas, fetes malbé) i $n = 2$ (nombre de bolles escollides).

El rang de X serà: $X(\Omega) = \{0, 1, 2\}$ ja que com a màxim escollim dues bolles.

La funció de probabilitat serà:

$$f_X(k) = p\{X = k\} = \frac{\binom{3}{k} \cdot \binom{2}{2-k}}{\binom{5}{2}}.$$

La fórmula anterior surt d'aplicar la fórmula $\frac{\text{casos favorables}}{\text{casos possibles}}$ ja que, en aquest cas, tenim en total $\binom{5}{2}$ casos possibles d'escollir dues bolles qualsevol i $\binom{3}{k}\cdot\binom{2}{2-k}$ casos d'escollir k bolles fetes malbé d'un total de 3 i $2-k$ bolles no fetes malbé d'un total de 2.

La funció de probabilitat queda reflectida en la taula següent:

k	0	1	2
$f_X(k)$	0.1	0.6	0.3

4.5.- Es fan proves de Bernoulli amb probabilitat p d'èxit. Sigui X la variable aleatòria que ens dóna el nombre de proves necessàries per tenir èxit r vegades. Trobau la funció de probabilitat de X (en aquest cas direm que la variable aleatòria X té una **distribució de Pascal** o **binomial negativa**).

Resolució. El rang de la variable aleatòria X serà:

$$X(\Omega) = \{r, r+1, \ldots\} = \{n \in \mathbb{Z} \mid n \geq r\},$$

ja que hem de fer com a mínim r proves per poder obtenir èxit r vegades.

De cara a trobar la funció de probabilitat, hem de tenir en compte que si $X = n$, la darrera prova de Bernoulli ha de ser èxit ja que, en cas contrari, hauríem obtengut èxit en un nombre de proves més petit que n. En les $n-1$ proves restants s'ha obtengut èxit $r-1$ vegades. La funció de probabilitat, serà, doncs:

$$f_X(n) = p\{X = n\} = \binom{n-1}{r-1} p^r (1-p)^{n-r} := \binom{n-1}{r-1} p^r q^{n-r},\ n \geq r.$$

4.6.- En una determinada fàbrica han tengut lloc accidents a raó d'1 cada 2 mesos. Suposant que tenen lloc de forma independent, quin és el nombre d'accidents esperats en un any? Quina és la desviació estàndard del nombre d'accidents a l'any? Quina és la probabilitat que no hi hagi cap accident en un determinat mes?

Resolució. Podem dir que en aquesta fàbrica hi ha hagut 0.5 accidents per mes.

Sigui la variable aleatòria $X =$"nombre d'accidents en una any". X és una variable de Poisson amb paràmetre λs, amb $\lambda = 0.5$ i $s = 12$ mesos (1 any). Per tant, X és $Poiss(6)$.

El nombre d'accidents esperats en un any serà

$$\mathrm{E}X = \lambda s = 6 \text{ accidents.}$$

i la desviació estàndard del nombre d'accidents a l'any serà:

$$\sigma = \sqrt{\text{Var } X} = \sqrt{\lambda s} = \sqrt{6} \approx 2.4494 \text{ accidents.}$$

Considerem la variable aleatòria $Y =$"nombre d'accidents en un mes". Y és una variable aleatòria de Poisson amb paràmetre λs, amb $\lambda = 0.5$ i $s = 1$ mes. Per tant, Y és $Poiss(0.5)$.

La probabilitat que no hi hagi cap accident en un determinat mes serà:

$$p\{Y = 0\} = f_Y(0) = \frac{(\lambda s)^0}{0!}\mathrm{e}^{-\lambda s} = \mathrm{e}^{-0.5} \approx 0.6065$$

4.7.- Si X és una variable aleatòria de Poisson amb paràmetre λs, provau que el valor de X més probable és λs.
Indicació: Considerau $\frac{f_X(k)}{f_X(k-1)}$, $k=1,2,\ldots$

Resolució. La funció de probabilitat d'una variable de Poisson amb paràmetre λs és:

$$f_X(k) = \frac{(\lambda s)^k}{k!}\mathrm{e}^{-\lambda s}.$$

Considerem

$$\frac{f_X(k)}{f_X(k-1)} = \frac{\frac{(\lambda s)^k}{k!}}{\frac{(\lambda s)^{k-1}}{(k-1)!}} = \frac{\lambda s}{k}.$$

Per tant:

Si $\lambda s \geq k \;\Rightarrow\; f_X(k) \geq f_X(k-1) \;\Rightarrow f_X(k)$ creix.

Si $\lambda s \leq k \;\Rightarrow\; f_X(k) \leq f_X(k-1) \;\Rightarrow f_X(k)$ decreix.

Per a $k = \lambda s$, $f_X(k)$ pren el màxim.

Nota: si $\lambda s \notin \mathbb{Z}$, aleshores:

$$\max_{k\in\mathbb{Z}} f_X(k) = \max\{f_X([\lambda s]), f_X([\lambda s]+1)\},$$

on $[\lambda s]$ és la part entera de λs.

4.8.- Trobau la funció generatriu de moments per a la variable aleatòria de Poisson de paràmetre λs.

Resolució.

$$\begin{aligned} m_X(t) &= \mathrm{E}\left(\mathrm{e}^{tx}\right) = \sum_{k=0}^{\infty} \mathrm{e}^{tk} \frac{(\lambda s)^k}{k!} \mathrm{e}^{-\lambda s} \\ &= \sum_{k=0}^{\infty} \frac{\left(\lambda s \mathrm{e}^t\right)^k}{k!} \mathrm{e}^{-\lambda s} = \mathrm{e}^{-\lambda s} \cdot \mathrm{e}^{\lambda s \mathrm{e}^t} = \mathrm{e}^{\lambda s\left(\mathrm{e}^t - 1\right)}. \end{aligned}$$

4.9.- Trobau la funció generatriu cumulativa per a una variable aleatòria de Poisson i feu-la servir per trobar μ_X i σ_X^2.

Resolució. Tenim que la funció generatriu d'una variable aleatòria de Poisson val:

$$m_X(t) = \mathrm{e}^{\lambda s\left(\mathrm{e}^t - 1\right)}.$$

Per tant, la funció generatriu cumulativa valdrà:

$$c_X(t) = \ln m_X(t) = \lambda s\left(\mathrm{e}^t - 1\right).$$

Trobem a continuació $\mathrm{E}X$ i Var X:

$$c_X'(t) = \lambda s \mathrm{e}^t, \Rightarrow \mathrm{E}X = c_X'(0) = \lambda s,$$

$$c_X''(t) = \lambda s \mathrm{e}^t, \Rightarrow \text{Var } X = c_X''(0) = \lambda s.$$

4.10.- Suposau que X està distribuïda uniformement en l'interval $(1, 2)$ i que es forma un quadrat amb costats de llargada X. Trobau la funció de densitat de la variable $Y = X^2$ que és l'àrea del quadrat i trobau $p\{Y > 2\}$.

Resolució. Si X és $U(1,2)$, les funcions de distribució i de densitat seran:

$$F_X(t) = \begin{cases} 0, & \text{si } t \leq 1, \\ t-1, & \text{si } 1 < t \leq 2, \\ 1, & \text{si } t \geq 2, \end{cases} \qquad f_X(t) = \begin{cases} 1, & \text{si } t \in (1,2), \\ 0, & \text{en cas contrari.} \end{cases}$$

Observem que el rang de X és l'interval $(1,2)$: $X(\Omega) = (1,2)$. Per tant, el rang de $Y = X^2$ serà l'interval $(1,4)$: $Y(\Omega) = (1,4)$. D'aquí deduïm que si $t \leq 1$, $F_Y(t) = 0$ i si $t \geq 4$, $F_Y(t) = 1$. Trobem $F_Y(t)$ per a $t \in (1,4)$:

$$F_Y(t) = p\{Y \leq t\} = p\left\{X^2 \leq t\right\} = p\left\{1 < X \leq \sqrt{t}\right\} = F_X(\sqrt{t}) - F_X(1) = \sqrt{t} - 1.$$

Les funcions de distribució i de densitat de Y seran:

$$F_Y(t) = \begin{cases} 0, & \text{si } t \leq 1, \\ \sqrt{t}-1, & \text{si } 1 \leq t \leq 4, \\ 1, & \text{si } t \geq 4, \end{cases} \qquad f_Y(t) = \begin{cases} \frac{1}{2\sqrt{t}}, & \text{si } t \in (1,4), \\ \\ 0, & \text{en cas contrari.} \end{cases}$$

Per calcular $p\{Y > 2\}$, ho podem posar en termes de la funció de distribució:

$$p\{Y > 2\} = 1 - F_Y(2) = 1 - \left(\sqrt{2} - 1\right) = 2 - \sqrt{2}.$$

4.11.- X és una variable geomètrica amb paràmetre p i Y és exponencial amb paràmetre λ. Trobau el valor de λ tal que $p\{X > 1\} = p\{Y > 1\}$.

Resolució. La funció de probabilitat de la variable aleatòria X és:

$$f_X(n) = p \cdot (1-p)^{n-1}, \; n \in \mathbb{Z}, \; n \geq 1.$$

Les funcions de distribució i de densitat de la variable aleatòria Y són:

$$F_Y(t) = 1 - \mathrm{e}^{-\lambda t}, \; t \geq 0, \quad f_Y(t) = \lambda \mathrm{e}^{-\lambda t}, \; t \geq 0.$$

Per tant:

$$\begin{aligned} p\{X > 1\} &= 1 - p\{X = 1\} = 1 - p, \\ p\{Y > 1\} &= 1 - F_Y(1) = 1 - \left(1 - \mathrm{e}^{-\lambda}\right) = \mathrm{e}^{-\lambda}. \end{aligned}$$

L'equació que hem de resoldre per trobar λ és:

$$e^{-\lambda} = 1 - p,$$

d'on resulta que $\lambda = \ln\left(\frac{1}{1-p}\right)$.

4.12.- Es dóna un procés de Poisson amb paràmetre λ. Es comença a observar el procés a partir del temps zero. Sigui S el temps passat fins que té lloc la segona incidència. Trobau la funció de densitat per a S.

Resolució. Considerem les variables aleatòries:

$S =$ "temps fins que té lloc la segona incidència en el procés de Poisson",

$X_s =$ "nombre de successos en l'interval $(0, s]$" $\rightarrow Poiss(\lambda s)$.

Observem que els dos successos següents són el mateix:

$$\{S > s\} = \{X_s \leq 1\},$$

ja que si per a $t = s$ no ha tengut lloc la segona incidència en el procés de Poisson, el valor de X_s (nombre d'incidències en el procés de Poisson) és com a màxim 1.

Per tant,

$$p\{S > s\} = p\{X_s = 0\} + p\{X_s = 1\} = e^{-\lambda s}(1 + \lambda s), \text{ per a } s > 0.$$

De la darrera expressió podem trobar les funcions de distribució i densitat de la variable aleatòria S:

$$\begin{aligned} F_S(s) &= 1 - p\{S > s\} = 1 - e^{-\lambda s}(1 + \lambda s), \text{ per a } s > 0, \\ f_S(s) &= F'_S(s) = \lambda e^{-\lambda s}(1 + \lambda s) - \lambda e^{-\lambda s} = \lambda^2 s e^{-\lambda s}, \text{ per a } s > 0. \end{aligned}$$

4.13.- Generalització de l'exercici anterior. Suposau que se dóna un procés de Poisson amb paràmetre λ i sigui T_r el temps passat fins que té lloc la $r-$èssima incidència ($r \geq 1$). Trobau la funció de densitat de la variable aleatòria T_r (en aquest cas direm que la variable T_r té la **distribució gamma.**)
Ind.: El succés $T_r > t$ és equivalent al succés: "observar $r-1$ incidències en l'interval de temps $(0, t)$".

Resolució. Considerem les variables aleatòries:

$T_r =$ "temps fins que té lloc la $r-$èssima incidència en el procés de Poisson",

$X_s = Poiss(\lambda s)$.

Observem que els dos successos següents són el mateix:

$$\{T_r > s\} = \{X_s \leq r-1\},$$

ja que si per a $t = s$ no ha tengut lloc la $r-$èssima incidència en el procés de Poisson, el valor de X_s (nombre d'incidències en el procés de Poisson) és com a màxim $r-1$.

Per tant,

$$p\{T_r > s\} = \sum_{i=0}^{r-1} \mathrm{e}^{-\lambda s} \cdot \frac{(\lambda s)^i}{i!}, \text{ per a } s > 0.$$

De la darrera expressió podem trobar les funció de distribució i densitat de la variable aleatòria T_r:

$$\begin{aligned} F_{T_r}(s) &= 1 - p\{S > s\} = 1 - \sum_{i=0}^{r-1} \mathrm{e}^{-\lambda s} \cdot \frac{(\lambda s)^i}{i!}, \text{ per a } s > 0, \\ f_{T_r}(s) &= F'_{T_r}(s) = \lambda \mathrm{e}^{-\lambda s} \sum_{i=0}^{r-1} \frac{(\lambda s)^i}{i!} - \lambda \mathrm{e}^{-\lambda s} \sum_{i=1}^{r-1} \frac{(\lambda s)^{i-1}}{(i-1)!} \\ &= \lambda \mathrm{e}^{-\lambda s} \left(\sum_{i=0}^{r-1} \frac{(\lambda s)^i}{i!} - \sum_{i=0}^{r-2} \frac{(\lambda s)^i}{i!} \right) = \frac{\lambda^r s^{r-1}}{(r-1)!} \mathrm{e}^{-\lambda s}, \text{ per a } s > 0. \end{aligned}$$

4.14.- Si X és una variable aleatòria normal amb mitjana μ i variància σ^2, trobau la funció de distribució de la variable aleatòria $U = |X|$ en funció de la funció de distribució de la normal estàndard.
Com a aplicació, trobau $p\{|X| \leq 3\}$ si X és $N(\mu = 1, \sigma^2 = 4)$.

Resolució. Siguin les variables aleatòries $X\ N(\mu,\sigma^2)$, $Z\ N(0,1)$ i $U=|X|$. La relació entre la funció de distribució de U i de Z és la següent:

$$\begin{aligned} F_U(t) &= p\{U=|X|\le t\}=p\{-t\le X\le t\} \\ &= F_X(t)-F_X(-t)=F_Z\left(\frac{t-\mu}{\sigma}\right)-F_Z\left(\frac{-t-\mu}{\sigma}\right). \end{aligned}$$

En el cas que $\mu=1$, $\sigma^2=4$ i $t=3$, tendrem que:

$$\begin{aligned} p\{|X|\le 3\} &= F_U(3)=F_Z\left(\frac{3-1}{2}\right)-F_Z\left(\frac{-3-1}{2}\right) \\ &= F_Z(1)-F_Z(-2)\approx 0.8413-(1-0.9772)=0.8185. \end{aligned}$$

4.15.- Se sap que el percentil 90% d'una distribució normal és igual a 50 i que el seu percentil 15% és 25.

a) Trobau μ i σ.

b) Quin és el percentil 40% ?

Resolució.

a) Per trobar μ i σ, hem de resoldre el sistema d'equacions següent:

$$\left.\begin{aligned} p\{X\le 50\} &= p\left\{Z\le \tfrac{50-\mu}{\sigma}\right\}=0.9, \\ p\{X\le 25\} &= p\left\{Z\le \tfrac{25-\mu}{\sigma}\right\}=0.15. \end{aligned}\right\}$$

Mirant a les taules, resulta que:

$$\left.\begin{aligned} \tfrac{50-\mu}{\sigma} &= 1.28, \\ \tfrac{25-\mu}{\sigma} &= -1.04. \end{aligned}\right\}$$

Les solucions del sistema anterior són: $\mu\approx 36.2068$ i $\sigma\approx 10.7758$.

b) Hem de trobar el valor $t_{0.4}$ tal que $p\{X \leq t_{0.4}\} = 0.4$. Si posam la condició anterior en termes de la variable aleatòria $Z = N(0,1)$, podem dir que hem de trobar el valor $t_{0.4}$ tal que

$$p\left\{Z \leq \frac{t_{0.4} - 36.2068}{10.7758}\right\} = 0.4$$

Mirant a les taules, resulta que $\frac{t_{0.4}-36.2068}{10.7758} \approx -0.26$, d'on resulta que $t_{0.4} \approx 33.4051$.

4.16.- Sigui T la variable aleatòria que ens dóna el temps d'arribada d'una persona al seu lloc de treball a partir de les 8.00 h. del matí. Suposem que T és Exp ($\lambda = 1$ h.). Escollim a l'atzar 10 dies. Trobau:

a) La probabilitat que arribi 5 dies més tard de les 8.30 h.

b) La probabilitat que arribi com a màxim un dia més prest de les 9.00 h.

c) El nombre mitjà de dies que arriba més tard de les 8.30 h.

Final. Juny 94.

Resolució.

a) Considerem la variable aleatòria:

$$X_a = \text{“ nombre de dies que arriba més tard de les 8.30 h ”}.$$

Fixau-vos que X_a és una variable aleatòria binomial de paràmetres $n = 10$ dies i $p_a =$ probabilitat que un dia a l'atzar arribi més tard de les 8.30 h. Aquest darrer valor ens el donarà la variable aleatòria T:

$$p_a = p\left\{T \geq \frac{1}{2}\right\} = \mathrm{e}^{-0.5} \approx 0.6065.$$

Per tant, la probabilitat que arribi 5 dies més tard de les 8.30 h val:

$$p\{X_a = 5\} = \binom{10}{5} p_a^5 (1-p_a)^5 \approx 0.1950.$$

b) Considerem la variable aleatòria:

$$X_b = \text{“ nombre de dies que arriba més prest de les 9h.”}.$$

Fixau-vos que X_b és una variable aleatòria binomial de paràmetres $n = 10$ dies i $p_b =$ probabilitat que un dia a l'atzar arribi més prest de les 9h. Aquest darrer valor ens el donarà la variable aleatòria T:

$$p_b = p\{T \leq 1\} = 1 - \mathrm{e}^{-1} \approx 0.6321.$$

Per tant, la probabilitat que arribi com a màxim un dia més prest de les 9.00 h. val:

$$p\{X_b \leq 1\} = \binom{10}{0}(1-p_b)^{10} + \binom{10}{1}p_b(1-p_b)^9 \approx 0.000825.$$

c) El nombre mitjà de dies que arriba més tard de les 8.30 h s'obté amb:

$$\mathrm{E}X_a = 10p_a \approx 6.06 \text{ dies}.$$

4.17.- Un canal de transmissió accepta un voltatge arbitrari V com a *input*, i treu com a *output* un voltatge $Y = V + N$, on N és una variable aleatòria $N(0,1)$. Suposem que el canal s'utilitza per transmetre informació binària de la següent manera:

- per transmetre $0 \rightarrow V = -1$,
- per transmetre $1 \rightarrow V = 1$.

El receptor decideix que s'ha enviat un 0 si el voltatge rebut és negatiu i un 1 si és positiu. Trobau la probabilitat de cometre un error si s'ha enviat un 0; el mateix si s'ha enviat un 1.

Resolució. Tenim:

$$\begin{aligned} p\{\text{error / 0 transmès}\} &= p\{Y > 0/V = -1\} = p\{V + N > 0/V = -1\} \\ &= \frac{p\{-1 + N > 0\}}{P\{V = -1\}} = \frac{p\{N > 1\}}{1/2} = \frac{0.159}{1/2} = 0.318, \end{aligned}$$

i, anàlogament,

$$\begin{aligned} p\{\textsf{error / 1 transmès}\} &= p\{Y<0/V=1\} = p\{V+N<0/V=1\} \\ &= \frac{p\{1+N<0\}}{P\{V=1\}} = \frac{p\{N<-1\}}{1/2} = \frac{0.159}{1/2} = 0.318. \end{aligned}$$

4.3 Problemes proposats

4.1.- Es tiren una sola vegada 5 daus no trucats. Sigui X el nombre d'uns que surten. Calculau l'esperança de X, la variància de X, $p\{1 \leq X < 4\}$ i $p\{X \geq 2\}$.

4.2.- Se sap que el 10% dels tassons fabricats per una determinada màquina té algun defecte. Si se seleccionen a l'atzar 10 dels tassons fabricats per aquesta màquina, quina és la probabilitat que cap sigui defectuós? Quants tassons defectuosos s'esperaria trobar?

4.3.- Se sap que Y és una variable binomial amb mitjana $\mu_Y = 6$ i variància $\sigma_Y^2 = 4$. Trobau la distribució de Y, o sigui, trobau n i p.

4.4.- Un fabricant de peces les envia en paquets de 20 als seus clients. Suposau que la probabilitat que una peça sigui defectuosa és 0.05.

a) Quin és el nombre esperat de peces defectuoses per paquet?

b) Quina és la probabilitat que un determinat paquet no tengui cap peça defectuosa?

4.5.- Suposau que una urna conté 10 bolles, una de les quals és negra. Sigui Z el nombre d'extraccions amb reposició necessàries per treure la bolla negra. Quina és la funció de probabilitat per a Z? I la mitjana de Z?

4.6.- Es llança una moneda a l'aire fins que surt cara. Suposau que les tirades són independents i que la probabilitat que surti cara cada vegada és p.

a) Demostrau que la probabilitat que faci falta un nombre senar de llançaments és $\frac{p}{1-q^2}$ on $q = 1 - p$.

b) Trobau el valor de p tal que la probabilitat que faci falta un nombre senar de vegades sigui 0.6.

c) Es pot trobar un valor de p tal que la probabilitat que faci falta un nombre senar de vegades sigui 0.5?

4.7.- S'ha observat que el trànsit mitjà de cotxes en un determinat punt d'un camí rural és de 3 cotxes/hora. Suposau que els instants en què passen els cotxes són independents. Sigui X el nombre de cotxes que passen per aquest punt en un interval de 20 minuts. Trobau $p\{X = 0\}$ i $p\{X \geq 2\}$.

4.8.- Suposau que un de cada 10.000 nins neix cec. Si un hospital d'una ciutat gran va tenir 5.000 naixements el 1970, aproximau per la distribució de Poisson la probabilitat que cap dels nascuts en aquest any fos cec en néixer. Aproximau també la probabilitat que hagi nascut exactament 1 nin cec i la que hagin nascut almenys 2 nins cecs.

4.9.- Suposau que les vendes que fa un venedor de cotxes usats es realitzen segons un procés de Poisson amb paràmetre $\lambda = 1$ cotxe per setmana.

a) Quina és la probabilitat que hi hagi exactament 3 vendes en un període de 2 setmanes? I com a mínim 3 vendes? I com a màxim 3 vendes?

b) Quina és la probabilitat que passin 3 períodes de 2 setmanes consecutives sense cap venda?

4.10.- Trobau la funció generatriu factorial per a una variable aleatòria X distribuïda uniformement en (a, b).

4.11.- Suposem que X està distribuïda uniformement en $(0, 2)$ i Y és una variable exponencial amb paràmetre λ. Trobau el valor de λ tal que $p\{X < 1\} = p\{Y < 1\}$.

4.12.- Suposem que el temps X que fa falta perquè un corredor de fons recorri una milla és una variable normal amb paràmetres $\mu = 4$ minuts i 1 segon i $\sigma = 2$ segons. Quina és la probabilitat que aquest atleta recorri la milla en menys de 4 minuts? I en més de 3 minuts i 55 segons?

4.13.- L'alçada que salta un atleta de salt d'alçada en cada intent és una variable aleatòria normal amb mitjana de 2 metres i desviació estàndard de 0.8 metres.

a) Quina és la màxima alçada que pot saltar amb probabilitat 0.95?

b) Quina és l'alçada que salta en només el 10% dels intents?

4.14.- Una centraleta telefònica rep per terme mitjà 100 telefonades per hora.

a) Trobau la probabilitat que passin almenys 25 minuts per rebre la primera telefonada.

b) Trobau la probabilitat que passin almenys 30 minuts per rebre la segona telefonada.

c) Si definim com a p_k la probabilitat que passin almenys 30 minuts per rebre la $k-$èssima telefonada, trobau la relació que hi ha entre p_k i p_{k-1}.

(Indicació: feu servir la distribució gamma.) Primer Parcial. Curs 92-93.

4.15.- Sigui X una variable aleatòria normal amb paràmetres $\mu = 1$ i $\sigma^2 = 1$. Trobau el valor de b tal que $p\left\{(X-1)^2 \leq b\right\} = 0.1$.
Final. Juny 96.

4.16.- Sigui Z una variable aleatòria $N(0, 1)$. Trobau $p\left\{\left(Z - \frac{1}{4}\right)^2 > \frac{1}{16}\right\}$.
Final. Setembre 96.

4.17.- Dues persones juguen a cara o creu i han decidit continuar la partida fins que s'hagin obtengut com a mínim 3 cares i 3 creus. Trobau la probabilitat que el joc no s'acabi en 10 tirades.

4.18.- Arriben missatges a un ordinador a un ritme mitjà de 15 missatges per segon. Sabem que el nombre de missatges que arriben en un segon és una variable aleatòria de Poisson.

a) Trobau la probabilitat que no arribi cap missatge en un segon.

b) Trobau la probabilitat que arribin més de 10 missatges en un segon.

4.19.- De dos xips, es planteja utilitzar-ne un en un cert sistema. El temps de vida del xip 1 s'ha modelat segons una $N(20.000, 4.000^2)$ (la probabilitat d'un temps de vida negatiu és despreciable) i el del xip 2, segons una $N(22.000, 1.000^2)$. Quin xip s'hauria de triar si el temps de vida objectiu del sistema és de 20.000 hores? I si és de 24.000 hores?

4.20.- Sigui X el nombre d'encerts en n proves de Bernoulli independents amb probabilitat p d'encert.

a) Si k és el valor més probable de X, provau que

$$(n+1)p - 1 \leq k \leq (n+1)p.$$

b) Si llançam 10 vegades un dau sense biaix, quin és el nombre més probable de vegades que obtendrem un '2'?

Capítol 5

Variables aleatòries vectorials

5.1 Resum teòric

Definició 5.1 *Si* $X_1, \ldots, X_n$ *són variables aleatòries en* $(\Omega, \mathcal{F}, p)$*, el vector* $X = (X_1, \ldots, X_n)$ *s'anomena una* **variable aleatòria vectorial** *(o* **vector aleatori n-dimensional***) en* $(\Omega, \mathcal{F}, p)$*.*

Donat un vector aleatori, estam interessats a determinar la seva distribució conjunta, com també les distribucions de cada component (distribucions marginals). Primer considerarem només vectors aleatoris bidimensionals.

Definició 5.2 *Direm* **funció de distribució conjunta** *de* (X, Y) *a:*

$$F_{XY}(x, y) = p\{X \leq x, Y \leq y\}.$$

Propietats.

1. F_{XY} és creixent en x i y: Si $x_1 \leq x_2$ i $y_1 \leq y_2, F_{XY}(x_1, y_1) \leq F_{XY}(x_2, y_2)$.
2. $\lim\limits_{x \to -\infty} F_{XY}(x, y) = \lim\limits_{y \to -\infty} F_{XY}(x, y) = 0$.
3. $\lim\limits_{(x,y) \to (+\infty, +\infty)} F_{XY}(x, y) = 1$.
4. Les **funcions de distribució marginals** són:
   - $F_X(x) = \lim\limits_{y \to +\infty} F_{XY}(x, y)$,

- $F_Y(y) = \lim_{x\to+\infty} F_{XY}(x,y)$.

5. $\lim_{x\to a^+} F_{XY}(x,y) = F_{XY}(a,y)$, $\lim_{y\to b^+} F_{XY}(x,y) = F_{XY}(x,b)$.

Si X i Y són variables aleatòries discretes, amb $X(\Omega) = \{x_1, x_2, \ldots\}$ i $Y(\Omega) = \{y_1, y_2, \ldots\}$, aleshores el vector aleatori (X,Y) és també discret, amb $(X,Y)(\Omega) = \{(x_j, y_k) : j = 1,2,\ldots, k = 1,2,\ldots\}$. Aleshores podem donar la definició següent:

Definició 5.3 *Direm* **funció de probabilitat conjunta** *del vector aleatori discret (X,Y) a:*

$$f_{XY}(x_j, y_k) = p\left\{X = x_j, Y = y_k\right\}, j = 1,2,\ldots, k = 1,2,\ldots$$

Propietats.

1. Si A és tal que $\{(X,Y) \in A\}$ és un succés, $p\left\{(X,Y) \in A\right\} = \sum_{(x_j,y_k)\in A} f_{XY}(x_j, y_k)$.

2. $\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} f_{XY}(x_j, y_k) = 1$.

3. Les **funcions de probabilitat marginals** s'obtenen com:

   - $f_X(x_j) = p\left\{X = x_j\right\} = \sum_{k=1}^{\infty} f_{XY}(x_j, y_k)$.
   - $f_Y(y_k) = p\left\{Y = y_k\right\} = \sum_{j=1}^{\infty} f_{XY}(x_j, y_k)$.

El recíproc no és cert en general: conegudes f_X i f_Y, no podem determinar f_{XY}.

Definició 5.4 *Direm que dues variables aleatòries X i Y són* **conjuntament absolutament contínues** *(o que (X,Y) és un* **vector aleatori absolutament continu***) si F_{XY} és la integral d'una funció de densitat conjunta, és a dir si*

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u,v)\,dudv, \quad \forall (u,v) \in \mathbb{R}^2.$$

Propietats.

1. Si A és tal que $\{(X,Y) \in A\}$ és un succés, $p\,\{(X,Y) \in A\} = \int\int_A f_{XY}(x,y)dxdy$.
2. $\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x,y)\,dxdy = 1$.
3. $f_{XY}(x,y) = \dfrac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$.
4. Les **funcions de densitat marginals** són:
   - $f_X(x) = \displaystyle\int_{-\infty}^{+\infty} f_{XY}(x,y)\,dy$.
   - $f_Y(y) = \displaystyle\int_{-\infty}^{+\infty} f_{XY}(x,y)\,dx$.

   El recíproc no és cert en general: donades f_X i f_Y, no podem determinar f_{XY}.

Exemples.

1. **Llei uniforme en $\mathbb{R}^2$.**

   (X,Y) té una **llei uniforme sobre el recinte fitat** $B \subset \mathbb{R}^2$ si la seva funció de densitat és:

$$f_{XY}(x,y) = \begin{cases} \frac{1}{|B|}, & \text{si}(x,y) \in B, \\ 0, & \text{en cas contrari,} \end{cases}$$

   on $|B| =$ àrea de B.

2. **Llei gaussiana bidimensional.**

   Suposem que el vector aleatori (X,Y) té la densitat conjunta següent:

$$f(x,y) = \frac{1}{2\pi\sqrt{1-r^2}} \mathrm{e}^{-(x^2+y^2-2rxy)/(2(1-r^2))}, \ |r| \leq 1.$$

   Aleshores direm que X i Y són **conjuntament gaussianes**. Es pot provar que les distribucions marginals són totes dues $N(0,1)$.

En general, direm que dues variables aleatòries són conjuntament gaussianes si la seva funció de densitat conjunta és donada per:

$$\begin{aligned} f(x,y) &= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\exp(g(x,y)), \\ g(x,y) &= \frac{-1}{2(1-r^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2r\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right], \end{aligned}$$

on $|r| \leq 1$.

Definició 5.5 *Donades dues variables aleatòries X i Y, direm que són* **independents** *si*

$$F_{XY}(x,y) = F_X(x)\cdot F_Y(y) \quad \forall (x,y)\in\mathbb{R}^2.$$

Propietats.

1. X i Y són independents si i només si $f_{XY}(x,y) = f_X(x)\cdot f_Y(y)$.
2. Si X i Y són independents, i $g,h:\mathbb{R}\to\mathbb{R}$ són tals que $g(X)$ i $h(Y)$ són dues noves variables aleatòries, aleshores aquestes també són independents.

Definició 5.6 *Si X i Y són discretes, definim la* **funció de probabilitat condicionada de Y donat que** $X=x$ *com*

$$f_{Y/X}(y/x) = \frac{f_{XY}(x,y)}{f_X(x)},$$

per als x tals que $f_X(x)>0$.

$f_{Y/X}$ satisfà totes les propietats d'una probabilitat; en particular, si $A\in\mathcal{F}$,

$$p\{Y\in A/X=x\} = \sum_{y_k\in A\cap Y(\Omega)} f_{Y/X}(y_k/x).$$

Definició 5.7 *Si X i Y són conjuntament absolutament contínues, definim la* **funció de densitat condicionada de Y donat que** $X=x$ *com*

$$f_{Y/X}(y/x) = \frac{f_{XY}(x,y)}{f_X(x)},$$

per als x tals que $f_X(x)>0$.

Propietats. Sigui $A \subset \mathbb{R}$.

1. Si X és discreta,

$$p\{Y \in A\} = \sum_{x_j} p\{Y \in A/X = x_j\} \cdot f_X(x_j).$$

2. Si X és absolutament contínua,

$$p\{Y \in A\} = \int_{-\infty}^{+\infty} p\{Y \in A/X = x\} \cdot f_X(x)\,dx.$$

Definició 5.8 **L'esperança condicionada de** Y **donat que** $X = x$ *es defineix com*

$$\mathrm{E}[Y/X = x] = \int_{-\infty}^{+\infty} y \cdot f_{Y/X}(y/x)\,dy,$$

per al cas conjuntament absolutament continu, i

$$\mathrm{E}[Y/X = x] = \sum_{y_k} y_k \cdot f_{Y/X}(y_k/x),$$

per al cas discret.

$\mathrm{E}[Y/X = x]$ és una funció de x, posem $g(x)$. Aleshores té sentit parlar de la funció de la variable aleatòria X $\hat{Y} = g(X) = \mathrm{E}[Y/X]$ com la variable aleatòria que, quan $X = x$, val $\mathrm{E}[Y/X = x]$. Aquesta variable aleatòria s'anomena l'**esperança condicionada de** Y **donada** X.

Definició 5.9 *La* **variància de** Y **donat que** $X = x$ *és la variància de la distribució condicionada:*

$$\mathrm{Var}(Y/X = x) = \mathrm{E}[(Y - \mathrm{E}[Y/X = x])^2/X = x] = \mathrm{E}[Y^2/X = x] - (\mathrm{E}[Y/X = x])^2.$$

Observem que $\mathrm{Var}(Y/X = x)$ és una funció de x i així podem definir la variable aleatòria $\mathrm{Var}(Y/X) = \mathrm{E}[(Y - \mathrm{E}[Y/X])^2/X]$, anomenada **variància condicionada de** Y **donada** X.

Considerem ara el cas de n variables aleatòries $X_1, \ldots, X_n$. La **funció de distribució conjunta** és $F_{X_1 \ldots X_n}(x_1, \ldots, x_n) = p\{X_1 \leq x_1, \ldots, X_n \leq x_n\}$.

Les **funcions de distribució marginals** s'obtenen posant $+\infty$ en les components adequades. Per exemple, $F_{X_1 X_2}(x_1, x_2) = F_{X_1 \ldots X_n}(x_1, x_2, +\infty, \ldots, +\infty)$.

La funció de probabilitat conjunta (cas discret) i la funció de densitat conjunta (cas absolutament continu) es defineixen exactament igual que el cas bidimensional.

Resulta que $X_1, \ldots, X_n$ són independents si i només si

$$F_{X_1 \ldots X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n),$$

i això equival a

$$f_{X_1 \cdots X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

Exemple: Llei multinomial

Considerem un experiment aleatori amb k resultats possibles $A_1, \ldots, A_k$ mútuament excloents ($A_i \cap A_j = \emptyset \; \forall i \neq j, \; i, j = 1 \ldots, k$) tals que $\Omega = \bigcup\limits_{i=1}^{k} A_i$. Posem $p_i = p\{A_i\}, \; i = 1, \ldots, k$. Repetim l'experiment n vegades. Considerem les variables aleatòries

$$X_i = \textit{nombre de vegades que s'obté el resultat } A_i \; (i = 1, \ldots, k).$$

Aleshores el vector aleatori k-dimensional $(X_1, \ldots, X_k)$ es diu que segueix una **llei multinomial amb paràmetres** $n, p_1, \ldots, p_k$**.** La seva funció de probabilitat és:

$$f_{X_1 \cdots X_k}(n_1, \ldots, n_k) = p\{X_1 = n_1, \ldots, X_k = n_k\} = \frac{n!}{n_1! \cdots n_k!} \, p_1^{n_1} \cdots p_k^{n_k},$$

on $n_1 + \cdots + n_k = n$.

Es pot provar que cada variable aleatòria X_i té una distribució $B(n, p_i)$ $(i = 1, \ldots, k)$.

Estudiem ara el cas de funcions de diverses variables. En el que segueix, $X_1, \ldots, X_n$ seran n variables aleatòries conjuntament absolutament contínues.

1. Sigui $Z = g(X_1, \ldots, X_n)$. Aleshores

$$F_Z(z) = p\{(X_1, \ldots, X_n) \in B_z\} = \int \cdots \int_{B_z} f_{X_1 \cdots X_n}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n,$$

on $B_z = \{(x_1, \ldots, x_n) : g(x_1, \ldots, x_n) \leq z\}$.

2. Considerem $\underline{Z} = (Z_1, \ldots, Z_m) = (g_1(X_1, \ldots, X_n), \ldots, g_m(X_1, \ldots, X_n))$. Aleshores

$$F_{\underline{Z}}(\underline{z}) = p\left\{(X_1, \ldots, X_n) \in B_{\underline{z}}\right\},$$

on $B_{\underline{z}} = \{(x_1, \ldots, x_n) : g_1(x_1, \ldots, x_n) \leq z_1, \ldots, g_m(x_1, \ldots, x_n) \leq z_m\}$.

3. Si $\underline{Y} = (Y_1, \ldots, Y_m) = (g_1(\underline{X}), \ldots, g_m(\underline{X})) = \underline{g}(\underline{X})$ verifica

   (a) $\underline{X} = (X_1, \ldots, X_n)$ és conjuntament absolutament continu amb densitat conjunta $f_{\underline{X}}$,

   (b) $\underline{Y}$ té la mateixa dimensió que $\underline{X}$ $(m = n)$,

   (c) $\underline{g}$ és bijectiva i diferenciable per tot,

   aleshores $\underline{Y}$ és conjuntament absolutament continu amb densitat conjunta

$$f_{\underline{Y}}(\underline{y}) = \left.\frac{f_{\underline{X}}(\underline{x})}{|J(\underline{y}, \underline{x})|}\right|_{\underline{x}=\underline{g}^{-1}(\underline{y})},$$

   on $J(\underline{y}, \underline{x}) = det\left(\dfrac{\partial y_i}{\partial x_j}\right)_{n\times n}$ = jacobià de $\underline{g}$.

4. Si les g_i són lineals, per exemple, $(Z_1, Z_2) = (aX_1 + bX_2, cX_1 + dX_2)$ on $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ és tal que $\det A \neq 0$, aleshores

$$f_{Z_1 Z_2}(z_1, z_2) = \left.\frac{f_{X_1 X_2}(x_1, x_2)}{|\det A|}\right|_{(x_1,x_2)^T = A^{-1}\,(z_1,z_2)^T}.$$

Donarem ara els moments de funcions de dues variables aleatòries més utilitzats. El cas de més de dues variables aleatòries és semblant.

- Si $Z = g(X, Y)$,
  - $\mathrm{E}Z = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x, y) f_{XY}(x, y)\, dxdy$ (cas absolutament continu),

- $\mathrm{E}Z = \sum_j \sum_k g(x_j, y_k) f_{XY}(x_j, y_k)$ (cas discret).

- Si $g(X,Y) = g_1(X) \cdot g_2(Y)$ i X i Y són independents, aleshores

$$\mathrm{E}(g_1(X) \cdot g_2(Y)) = \mathrm{E}(g_1(X)) \cdot \mathrm{E}(g_2(Y)).$$

- Els **moments conjunts** de X i Y són $\mu_{rs} = \mathrm{E}(X^r Y^s)$, $r, s = 1, 2, \ldots$

- Els **moments marginals** són $\mu_{r0} = \mathrm{E}X^r$ i $\mu_{0s} = \mathrm{E}Y^s$.

- Els **moments centrals** són $\mu'_{rs} = \mathrm{E}[(X - \mathrm{E}X)^r (Y - \mathrm{E}Y)^s]$. Si $r = s = 1$, obtenim la **covariància** de X i Y:

$$\mathrm{Cov}(X,Y) = \sigma_{XY} = \mathrm{E}[(X - \mathrm{E}X)(Y - \mathrm{E}Y)] = \mathrm{E}(XY) - \mathrm{E}X \cdot \mathrm{E}Y.$$

  El **coeficient de correlació** entre X i Y és

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y},$$

  i verifica:

  - $|\rho_{XY}| \leq 1$,
  - $|\rho_{XY}| = 1 \Longleftrightarrow Y = aX + b$, amb $a > 0$ si $\rho_{XY} = 1$ i $a < 0$ si $\rho_{XY} = -1$.

  En aquest darrer cas es diu que X i Y estan **relacionades linealment.** En general ρ_{XY} indica la mesura en què Y es pot aproximar com una funció lineal de X. Direm que X i Y estan **incorrelacionades** si $\rho_{XY} = 0$. Si X i Y són independents, $\rho_{XY} = 0$. El recíproc no és cert en general.

Donat un vector aleatori n-dimensional $\underline{X} = (X_1, \ldots, X_n)$, els moments més importants són:

- **vector mitjana**:

$$\mathrm{E}\underline{X} = \begin{pmatrix} \mathrm{E}X_1 \\ \vdots \\ \mathrm{E}X_n \end{pmatrix},$$

- **matriu de covariàncies**:

$$V_{\underline{X}} = \begin{pmatrix} \mathrm{Var}X_1 & \mathrm{Cov}(X_1, X_2) & \dots & \mathrm{Cov}(X_1, X_n) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{Cov}(X_1, X_n) & \mathrm{Cov}(X_2, X_n) & \dots & \mathrm{Var}X_n \end{pmatrix},$$

- **funció característica**:

$$\phi_{\underline{X}}(t_1, \dots, t_n) = \mathrm{E}[\mathrm{e}^{\mathrm{i}(t_1 X_1 + \cdots + t_n X_n)}].$$

5.2 Problemes resolts

5.1.- Es treuen dues cartes sense reposició d'una baralla de 52 cartes. Sigui X el nombre d'asos que surten i sigui Y el nombre d'espases. Trobau la funció de probabilitat conjunta $f_{XY}(x, y)$ i calculau $p\{X > Y\}$.

Resolució. Els rangs de les variables X i Y són, respectivament:

$$X(\Omega) = \{0, 1, 2\}, \quad Y(\Omega) = \{0, 1, 2\}.$$

Trobem a continuació la funció de probabilitat conjunta f_{XY}:

$$\begin{aligned} f_{XY}(0,0) &= p\{X = 0 \cap Y = 0\} = p\{\text{ 0 asos } \cap \text{ 0 espases }\} \\ &= p\{\text{ carta 1 no as ni espasa }\} \cdot \\ &\quad\; p\{\text{ carta 2 no as ni espasa/carta 1 no as ni espasa }\} \\ &= \frac{36}{52} \cdot \frac{35}{51} \approx 0.4751, \\ f_{XY}(1,0) &= p\{X = 1 \cap Y = 0\} = p\{\text{ 1 as } \cap \text{ 0 espases }\} \\ &= p\{\text{ carta 1 as no espasa}\} \cdot \\ &\quad\; p\{\text{ carta 2 no espasa ni as/carta 1 as no espasa }\} \\ &\quad\; + p\{\text{ carta 1 no as ni espasa }\} \cdot \\ &\quad\; p\{\text{ carta 2 as no espasa/carta 1 no as ni espasa }\} \\ &= \frac{3}{52} \cdot \frac{36}{51} + \frac{36}{52} \cdot \frac{3}{51} \approx 0.0814, \\ f_{XY}(2,0) &= p\{X = 2 \cap Y = 0\} = p\{\text{ 2 asos } \cap \text{ 0 espases }\} \\ &= p\{\text{ carta 1 as no espasa }\} \cdot \end{aligned}$$

$$
\begin{aligned}
& p\{\ \text{carta 2 as no espasa/carta 1 as no espasa}\ \} \\
= \;& \frac{3}{52}\cdot\frac{2}{51} \approx 0.0022, \\
f_{XY}(0,1) = \;& p\{X=0\cap Y=1\} = p\{\ \text{0 asos } \cap \text{ 1 espasa}\ \} \\
= \;& p\{\ \text{carta 1 no as ni espasa}\ \}\cdot \\
& p\{\ \text{carta 2 espasa no as/carta 1 no as ni espasa}\ \}+ \\
& p\{\ \text{carta 1 espasa no as}\ \}\cdot \\
& p\{\ \text{carta 2 no as ni espasa/carta 1 espasa no as}\ \} \\
= \;& \frac{36}{52}\cdot\frac{12}{51}+\frac{12}{52}\cdot\frac{36}{51} \approx 0.3257, \\
f_{XY}(1,1) = \;& p\{X=1\cap Y=1\} = p\{\ \text{1 as } \cap \text{ 1 espasa}\ \} \\
= \;& p\{\ \text{carta 1 as no espasa}\ \}\cdot \\
& p\{\ \text{carta 2 espasa no as/carta 1 as no espasa}\ \} \\
& +p\{\ \text{carta 1 espasa no as}\ \}\cdot \\
& p\{\ \text{carta 2 as no espasa/carta 1 espasa no as}\ \} \\
& +p\{\ \text{carta 1 as d'espases}\ \} \\
& \cdot p\{\ \text{carta 2 no as ni espasa/carta 1 as d'espases}\ \} \\
& +p\{\ \text{carta 1 no as ni espasa}\ \}\cdot \\
& p\{\ \text{carta 2 as d'espases/carta 1 no as ni espasa}\ \} \\
= \;& \frac{3}{52}\cdot\frac{12}{51}+\frac{12}{52}\cdot\frac{3}{51}+\frac{1}{52}\cdot\frac{36}{51}+\frac{36}{52}\cdot\frac{1}{51} \approx 0.0542, \\
f_{XY}(2,1) = \;& p\{X=2\cap Y=1\} = p\{\ \text{2 asos } \cap \text{ 1 espasa}\ \} \\
= \;& p\{\ \text{carta 1 as d'espases}\ \}\cdot \\
& p\{\ \text{carta 2 as no espasa/carta 1 as d'espases}\ \} \\
& +p\{\ \text{carta 1 as no espasa}\ \}\cdot \\
& p\{\ \text{carta 2 as d'espases/carta 1 as no espasa}\ \} \\
= \;& \frac{1}{52}\cdot\frac{3}{51}+\frac{3}{52}\cdot\frac{1}{51} \approx 0.0022, \\
f_{XY}(0,2) = \;& p\{X=0\cap Y=2\} = p\{\ \text{0 asos } \cap \text{ 2 espases}\ \} \\
= \;& p\{\ \text{carta 1 espasa no as}\ \}\cdot \\
& p\{\ \text{carta 2 espasa no as/carta 1 espasa no as}\ \} \\
= \;& \frac{12}{52}\cdot\frac{11}{51} \approx 0.0497, \\
f_{XY}(1,2) = \;& p\{X=1\cap Y=2\} = p\{\ \text{1 as } \cap \text{ 2 espases}\ \}
\end{aligned}
$$

$$
\begin{aligned}
&= p\{\text{ carta 1 as d'espases }\}\cdot \\
&\quad p\{\text{ carta 2 espasa no as/carta 1 as d'espases }\} \\
&\quad +p\{\text{ carta 1 espasa no as }\}\cdot \\
&\quad p\{\text{ carta 2 as d'espases/carta 1 espasa no as }\} \\
&= \frac{1}{52}\cdot\frac{12}{51}+\frac{12}{52}\cdot\frac{1}{51}\approx 0.0090, \\
f_{XY}(2,2) &= p\{X=2\cap Y=2\} = p\{\text{ 2 asos } \cap \text{ 2 espases }\} = 0.
\end{aligned}
$$

La funció de probabilitat conjunta queda esquematitzada en la taula següent:

$Y\backslash X$	0	1	2
0	0.4751	0.0814	0.0022
1	0.3257	0.0542	0.0022
2	0.0497	0.0090	0

Per trobar $p\{X>Y\}$, podem posar la probabilitat anterior en termes de la funció de probabilitat conjunta:

$$
\begin{aligned}
p\{X>Y\} &= \sum_{i>j} f_{XY}(i,j) = f_{XY}(1,0)+f_{XY}(2,0)+f_{XY}(2,1) \\
&\approx 0.0814+0.0022+0.0022 = 0.0858.
\end{aligned}
$$

5.2.- Suposem que se selecciona a l'atzar un punt de l'interior del cercle centrat a l'origen amb radi 1. Sigui X la coordenada x i Y la coordenada y del punt elegit. Trobau la funció de densitat conjunta $f_{XY}(x,y)$, la funció de densitat de X, $f_X(x)$ i la funció de densitat de Y, $f_Y(y)$.

Resolució. La variable (X,Y) serà una variable bidimensional contínua ja que el seu rang és:

$$(X,Y)(\Omega) = \{(x,y)\in\mathbb{R}^2 \mid x^2+y^2\le 1\}.$$

La funció de densitat conjunta tendrà la forma:

$$f_{XY}(x,y) = \begin{cases} k, & \text{si } x^2+y^2\le 1, \\ 0, & \text{en cas contrari.} \end{cases}$$

Per trobar k hem de fer servir que:

$$\int\int_{\mathbb{R}^2} f_{XY}(x,y)\,dx\,dy = 1.$$

Per tant:

$$\int\int_{(X,Y)(\Omega)} k\,dx\,dy = k \cdot \text{ Àrea } ((X,Y)(\Omega)) = k\pi.$$

D'on deduïm que $k = \frac{1}{\pi}$.

Si $x \notin (-1,1)$, $f_X(x) = 0$ ja que $f_{XY}(x,y) = 0$ per a tot $y \in \mathbb{R}$.

De la mateixa manera, si $y \notin (-1,1)$, $f_Y(y) = 0$ ja que $f_{XY}(x,y) = 0$ per a tot $x \in \mathbb{R}$.

Per tant, suposem que $x \in (-1,1)$. Fixau-vos que $f_{(X,Y)}(x,y) \neq 0$ per a $y \in (-\sqrt{1-x^2}, \sqrt{1-x^2})$. Per tant,

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy = \frac{2}{\pi}\sqrt{1-x^2}.$$

La funció de densitat marginal de X queda aleshores:

$$f_X(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2}, & \text{si } x \in (-1,1), \\ 0, & \text{en cas contrari.} \end{cases}$$

Fent un raonament semblant podem trobar la funció de densitat marginal de Y:

$$f_Y(y) = \begin{cases} \frac{2}{\pi}\sqrt{1-y^2}, & \text{si } y \in (-1,1), \\ 0, & \text{en cas contrari.} \end{cases}$$

5.3.- Provau que:

$$p\{(a < X \leq b) \cap (Y \leq d)\} = F_{XY}(b,d) - F_{XY}(a,d).$$

Indicació: Feu un dibuix.

Resolució. Farem primer una prova gràfica.

En la figura 5.1 es veu que l'àrea que hi ha entre les tres rectes $x = a$, $x = b$ i $y = d$ (àrea situada per davall), que anomenarem àrea ABD, és la resta entre l'àrea que hi ha entre les dues rectes $x = b$ i $y = d$ (àrea situada per davall i a l'esquerra), que anomenarem àrea BD, i l'àrea entre les rectes $x = a$ i $y = d$, que anomenarem àrea AD.

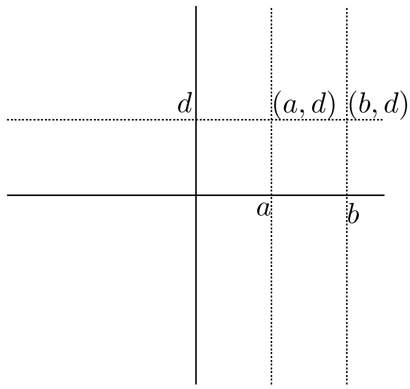

Figura 5.1: Prova gràfica del problema

Fixau-vos que l'àrea ABD són els punts del pla (X, Y) que compleixen:

$$\{a < X \leq b\} \cap \{Y \leq d\},$$

i les àrees BD i AD són els punts del pla (X, Y) que compleixen, respectivament:

$$\{X \leq b \cap Y \leq d\}, \quad \{X \leq a \cap Y \leq d\}.$$

Per tant, de forma gràfica podem escriure que:

Àrea$(\{a < X \leq b\} \cap \{Y \leq d\}) =$ Àrea$(\{X \leq b \cap Y \leq d\}) -$ Àrea$(\{X \leq a \cap Y \leq d\})$.

Farem ara la prova analítica.

Hem de provar que:

$$p\{(a < X \leq b) \cap (Y \leq d)\} = F_{XY}(b, d) - F_{XY}(a, d). \tag{5.1}$$

Vegem primer la següent relació entre els successos que intervenen en la fórmula anterior:

$$\{(a < X \leq b) \cap (Y \leq d)\} = \{X \leq b \cap Y \leq d\} - \{X \leq a \cap Y \leq d\}. \tag{5.2}$$

Si provam (5.2) quedarà vist (5.1), ja que:

$$\begin{aligned} F_{XY}(b, d) &= p\{X \leq b \cap Y \leq d\}, \\ F_{XY}(a, d) &= p\{X \leq a \cap Y \leq d\}. \end{aligned}$$

Vegem, doncs, (5.2).

Inclusió cap a la dreta $\subseteq$.

Sigui $\omega \in \{(a < X \leq b) \cap (Y \leq d)\}$. Aleshores $X(\omega) \leq b$ i $Y(\omega) \leq d$ però $X(\omega) > a$.

Per tant $\omega \in \{X \leq b \cap Y \leq d\}$ i $\omega \notin \{X \leq a \cap Y \leq d\}$. Queda vista la inclusió.

Inclusió cap a l'esquerra $\supseteq$.

Sigui $\omega \in \{X \leq b \cap Y \leq d\} - \{X \leq a \cap Y \leq d\}$. Aleshores $X(\omega) \leq b$, $Y(\omega) \leq d$ i $X(\omega) > a$.

Per tant $\omega \in \{(a < X \leq b) \cap (Y \leq d)\}$. Queda vista la inclusió. □

5.4.- Suposem que (X, Y) té densitat $f_{XY}(x, y) = \frac{1}{2}$ per a (x, y) pertanyent al quadrat de vèrtexs (a, a), $(a, -a)$, $(-a, a)$ i $(-a, -a)$ i que $f_{XY}(x, y)$ val 0 en els altres casos.

a) Trobau el valor de a.

b) Trobau les densitats marginals de X i Y.

Resolució.

a) Per trobar el valor de a, hem de fer servir que:

$$\int \int_{\mathbb{R}^2} f_{X,Y}(x, y)\, dx\, dy = 1.$$

En el nostre cas, $f_{X,Y}(x, y) \neq 0$ si $(x, y) \in (-a, a) \times (-a, a)$. Per tant, la fórmula anterior es redueix a:

$$\int_{-a}^{a} \int_{-a}^{a} \frac{1}{2}\, dx\, dy = \frac{1}{2}(2a)^2 = 1,$$

d'on $a = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}$.

b) Càlcul de $f_X(x)$ i $f_Y(y)$.

Si $x \notin (-a, a)$, $f_X(x) = 0$ ja que $f_{XY}(x, y) = 0$ per a tot $y \in \mathbb{R}$.

De forma semblant si $y \notin (-a, a)$, $f_Y(y) = 0$ ja que $f_{XY}(x, y) = 0$ per a tot $x \in \mathbb{R}$.

Sigui, doncs, $x \in (-a, a)$. En aquest cas $f_{XY}(x, y) \neq 0$ si $y \in (-a, a)$. Per tant:

$$f_X(x) = \int_{-a}^{a} \frac{1}{2}\, dy = \frac{1}{2} \cdot 2a = a = \frac{\sqrt{2}}{2}.$$

Així, doncs, la funció de densitat de X és:

$$f_X(x) = \begin{cases} \frac{\sqrt{2}}{2}, & \text{si } x \in \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \\ 0, & \text{en cas contrari.} \end{cases}$$

Aleshores la variable X està distribuïda uniformement en l'interval $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$.

Fent un raonament semblant, tenim que la variable Y està distribuïda uniformement en el mateix interval.

5.5.- Donada la següent funció de probabilitat de la variable aleatòria bidimensional (W, Z):

$$f_{(W,Z)} = \begin{cases} \frac{1}{n}, & \text{si } (w, z) = (1, 1), (2, 2), \ldots, (n, n), \\ 0, & \text{en els altres casos,} \end{cases}$$

trobau els valors de $\mathrm{E}\left(\frac{W}{Z}\right)$, $\mathrm{E}\left(\frac{Z}{W}\right)$, $\mathrm{E}\left(W^2 + Z^2\right)$ i $\mathrm{E}\left(\frac{W^3}{Z^2}\right)$.

Resolució. El rang de les variables discretes W i Z és:

$$W(\Omega) = Z(\Omega) = \{1, \ldots, n\}.$$

Fixau-vos que $f_{WZ}(w, z) \neq 0$ si i només si $w = z \in W(\Omega) = Z(\Omega)$.

Tenint en compte aquest fet, trobem una fórmula general per calcular $\mathrm{E}(g(W, Z))$:

$$\mathrm{E}(g(W, Z)) = \sum_{i=1}^{n} \sum_{j=1}^{n} g(i, j) f_{WZ}(i, j) = \frac{1}{n} \sum_{k=1}^{n} g(k, k).$$

A continuació trobam les esperances demanades:

$$\begin{aligned} \mathrm{E}\left(\frac{W}{Z}\right) &= \frac{1}{n} \sum_{k=1}^{n} \frac{k}{k} = \frac{1}{n} \cdot n = 1, \\ \mathrm{E}\left(\frac{Z}{W}\right) &= \frac{1}{n} \sum_{k=1}^{n} \frac{k}{k} = \frac{1}{n} \cdot n = 1, \end{aligned}$$

$$
\begin{aligned}
\mathrm{E}\left(W^2+Z^2\right) &= \frac{1}{n}\sum_{k=1}^{n}(k^2+k^2) = \frac{2}{n}\sum_{i=1}^{n}k^2 = \frac{2}{n}\frac{n(n+1)(2n+1)}{6} \\
&= \frac{(n+1)(2n+1)}{3}, \\
\mathrm{E}\left(\frac{W^3}{Z^2}\right) &= \frac{1}{n}\sum_{k=1}^{n}\frac{k^3}{k^2} = \frac{1}{n}\sum_{k=1}^{n}k = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}.
\end{aligned}
$$

5.6.- Donada la següent funció de densitat de la variable aleatòria bidimensional (U,V):

$$
f_{UV} = \begin{cases} 6(1-u-v), & \text{si } 0<u<1,\ 0<v<1-u, \\ 0, & \text{en els altres casos,} \end{cases}
$$

trobau $\mathrm{E}(U)$, $\mathrm{E}(V)$, $\mathrm{E}(UV)$ i σ_{UV}.

Resolució.

$$
\begin{aligned}
\mathrm{E}(U) &= \int_0^1\int_0^{1-u} 6(1-u-v)u\,dv\,du = 6\int_0^1\left[v-uv-\frac{v^2}{2}\right]_0^{1-u} u\,du \\
&= 6\int_0^1 u\left(1-u-u(1-u)-\frac{(1-u)^2}{2}\right)du = 6\int_0^1\left(\frac{u}{2}-u^2+\frac{u^3}{2}\right)du \\
&= 6\left[\frac{u^2}{4}-\frac{u^3}{3}+\frac{u^4}{8}\right]_0^1 = 6\left(\frac{1}{4}-\frac{1}{3}+\frac{1}{8}\right) = \frac{1}{4}, \\
\mathrm{E}(V) &= \int_0^1\int_0^{1-u} 6(1-u-v)v\,dv\,du = 6\int_0^1\left[\frac{v^2}{2}-u\frac{v^2}{2}-\frac{v^3}{3}\right]_0^{1-u} du \\
&= 6\int_0^1 \frac{(1-u)^2}{2}-u\frac{(1-u)^2}{2}-\frac{(1-u)^3}{3}\,du = \int_0^1(1-3u+3u^2-u^3)\,du \\
&= \left[u-\frac{3}{2}u^2+u^3-\frac{u^4}{4}\right]_0^1 = 1-\frac{3}{2}+1-\frac{1}{4} = \frac{1}{4}, \\
\mathrm{E}(UV) &= \int_0^1\int_0^{1-u} 6(1-u-v)uv\,dv\,du = \int_0^1 u(1-3u+3u^2-u^3)\,du \\
&= \int_0^1(u-3u^2+3u^3-u^4)\,du = \left[\frac{u^2}{2}-u^3+\frac{3}{4}u^4-\frac{u^5}{5}\right]_0^1
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2} - 1 + \frac{3}{4} - \frac{1}{5} = \frac{1}{20}, \\
\sigma_{UV} &= \mathrm{E}(UV) - \mathrm{E}(U) \cdot \mathrm{E}(V) = \frac{1}{20} - \frac{1}{4} \cdot \frac{1}{4} = -\frac{1}{80}.
\end{aligned}
$$

5.7.- Suposem que X i Y són variables aleatòries i definim $Z = X + Y$. Demostrau que:

a) $\mathrm{E}(Z) = \mathrm{E}(X) + \mathrm{E}(Y)$.

b) $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}$.

Resolució.

a) Fent servir les propietats de l'esperança, tenim que:

$$\mathrm{E}(Z) = \mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y).$$

b) Per provar la fórmula de la variància farem servir que l'esperança de la suma és la suma d'esperances i que l'esperança d'una constant per una variable és la constant per l'esperança de la variable.

$$
\begin{aligned}
\sigma_Z^2 &= \sigma_{X+Y}^2 = \mathrm{E}(X+Y)^2 - (\mathrm{E}(X+Y))^2 \\
&= \mathrm{E}\left(X^2 + Y^2 + 2XY\right) - (\mathrm{E}X + \mathrm{E}Y)^2 \\
&= \mathrm{E}\left(X^2\right) + \mathrm{E}\left(Y^2\right) + 2\mathrm{E}(XY) - \left((\mathrm{E}X)^2 + (\mathrm{E}Y)^2 + 2\mathrm{E}X \cdot \mathrm{E}Y\right) \\
&= \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 + \mathrm{E}\left(Y^2\right) - (\mathrm{E}Y)^2 + 2(\mathrm{E}(XY) - \mathrm{E}X \cdot \mathrm{E}Y) \\
&= \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}
\end{aligned}
$$

5.8.- Siguin X i Y variables aleatòries discretes amb funció de probabilitat conjunta:

$$f_{XY}(x,y) = \begin{cases} \frac{2}{n(n+1)}, & \text{per } x = 1,2,\ldots,n, \quad y = 1,2,\ldots,x, \\ 0, & \text{en els altres casos.} \end{cases}$$

Demostrau que X i Y no són independents. Calculeu també la regressió de Y sobre X ($\mathrm{E}(Y/X = x)$) i la regressió de X sobre Y ($\mathrm{E}(X/Y = y)$).

Resolució. Escriguem en forma de taula la funció de probabilitat conjunta de les variables X i Y, com també les funcions de densitat de X i Y:

$Y\backslash X$	1	2	$\ldots$	n	$f_Y(y)$
1	$\frac{2}{n(n+1)}$	$\frac{2}{n(n+1)}$	$\ldots$	$\frac{2}{n(n+1)}$	$\frac{2n}{n(n+1)} = \frac{2}{n+1}$
2	0	$\frac{2}{n(n+1)}$	$\ldots$	$\frac{2}{n(n+1)}$	$\frac{2(n-1)}{n(n+1)}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	0	0	$\ldots$	$\frac{2}{n(n+1)}$	$\frac{2}{n(n+1)}$
$f_X(x)$	$\frac{2}{n(n+1)}$	$\frac{4}{n(n+1)}$	$\ldots$	$\frac{2}{n+1}$	1

Com ja es veu, el rang de les variables X i Y és el conjunt $X(\Omega) = Y(\Omega) = \{1,\ldots,n\}$ i les fórmules per a les funcions de densitat de X i Y són les següents:

$$\begin{aligned} f_X(x) &= \frac{2x}{n(n+1)}, \ x = 1,\ldots,n, \\ f_Y(y) &= \frac{2(n-y+1)}{n(n+1)}, \ y = 1,\ldots,n. \end{aligned}$$

Podem deduir, doncs, que X i Y no són independents ja que:

$$f_X(x) \cdot f_Y(y) = \frac{4x(n-y+1)}{n^2(n+1)^2} \neq \frac{2}{n(n+1)}.$$

Basta prendre, per exemple $x = 1$ i $y = 1$ i veiem que en la fórmula anterior es compleix la igualtat només per a $n = 2$. Preneu $x = 1$ i $y = 2$ i no es complirà.

Deduïm, doncs, que X i Y no són independents.

Calculem a continuació la regressió de Y sobre X: $\mathrm{E}(Y/X = x)$.

Primer hem de trobar la funció de probabilitat de la variable aleatòria $Y/X = x$, on $x \in \{1, \ldots, n\}$.

En principi, podem dir que el rang de la variable $Y/X = x$ està inclòs en el conjunt $\{1, \ldots, n\}$. Trobem la funció de probabilitat:

$$f_{Y/X}(y/x) = \frac{f_{XY}(x,y)}{f_X(x)} = \begin{cases} \frac{\frac{2}{n(n+1)}}{\frac{2x}{n(n+1)}} = \frac{1}{x}, & \text{si } y = 1, \ldots, x, \\ 0, & \text{en cas contrari.} \end{cases}$$

Ara podem dir que el rang de la variable $Y/X = x$ és $\{1, \ldots, x\}$.

La regressió de Y sobre X valdrà, doncs:

$$\mathrm{E}(Y/X = x) = \sum_{y=1}^{x} y f_{Y/X}(y/x) = \sum_{y=1}^{x} y \cdot \frac{1}{x} = \frac{1}{x} \sum_{y=1}^{x} y = \frac{1}{x} \cdot \frac{x(x+1)}{2} = \frac{x+1}{2}.$$

Trobem ara la regressió de X sobre Y: $\mathrm{E}(X/Y = y)$.

Trobem primer la funció de probabilitat de la variable $X/Y = y$, on $y \in \{1 \ldots, n\}$:

$$f_{X/Y}(x/y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \begin{cases} 0, & \text{si } x = 1, \ldots, y-1, \\ \frac{\frac{2}{n(n+1)}}{\frac{2(n-y+1)}{n(n+1)}} = \frac{1}{n-y+1}, & \text{si } x = y, \ldots, n. \end{cases}$$

El rang de la variable $X/Y = y$ serà: $X/Y = y(\Omega) = \{y, y+1, \ldots, n\}$.

La regressió de X sobre Y valdrà, doncs:

$$\begin{aligned} \mathrm{E}(X/Y = y) &= \sum_{x=y}^{n} x \frac{1}{n-y+1} = \frac{1}{n-y+1} \sum_{i=1}^{n-y+1} (i + y - 1) \\ &= \frac{1}{n-y+1} \left(\frac{(n-y+2)(n-y+1)}{2} + (y-1)(n-y+1) \right) \\ &= \frac{n-y+2}{2} + y - 1 = \frac{n+y}{2}. \end{aligned}$$

5.9.- Suposant que X i Y són variables aleatòries independents normals amb paràmetres μ_X, σ_X, μ_Y i σ_Y, respectivament, quina és la funció generatriu dels moments conjunts per a (X, Y)?

Resolució. Recordem que si la variable aleatòria X és normal $N(\mu, \sigma^2)$, aleshores la funció generatriu de moments és:

$$m_X(t) = \mathrm{e}^{t\mu + \frac{t^2\sigma^2}{2}}.$$

Per tant, si X és $N(\mu_X, \sigma_X^2)$ i Y és $N(\mu_Y, \sigma_Y^2)$ i són independents, la funció generatriu de moments conjunta serà:

$$\begin{aligned} m_{XY}(t,s) &= \mathrm{E}\left(\mathrm{e}^{tx+sy}\right) = \mathrm{E}\left(\mathrm{e}^{tx}\right) \cdot \mathrm{E}\left(\mathrm{e}^{sy}\right) = m_X(t) \cdot m_Y(s) \\ &= \mathrm{e}^{t\mu_X + \frac{t^2\sigma_X^2}{2}} \cdot \mathrm{e}^{s\mu_Y + \frac{s^2\sigma_Y^2}{2}} = \mathrm{e}^{t\mu_X + s\mu_Y + \frac{t^2\sigma_X^2 + s^2\sigma_Y^2}{2}}. \end{aligned}$$

5.10.- Sigui X l'hora en què una persona s'aixeca del llit el matí (mesurada en fraccions d'hora després de les 7.00 h. del matí) i Y el temps que tarda en arribar a l'oficina (en fraccions d'hora) després d'aixecar-se. Suposem que la densitat condicional de Y per $X = x$ és:

$$f_{Y/X}(y/x) = \begin{cases} \frac{2y}{(1-x)^2}, & \text{per } 0 < x < \frac{2}{3} \text{ i } 0 < y < 1-x, \\ 0, & \text{en els altres casos,} \end{cases}$$

mentre que la marginal de X és:

$$f_X(x) = \begin{cases} \frac{81}{26}(1-x)^2, & \text{si } 0 < x < \frac{2}{3}, \\ 0, & \text{en els altres casos.} \end{cases}$$

Sabem que aquesta persona, un determinat matí, tarda 30 minuts en arribar a l'oficina (per tant, $Y = \frac{1}{2}$). Trobau la probabilitat que s'hagi aixecat del llit més tard de les 7.15 h. Si tarda 50 minuts per arribar a l'oficina, quina és la probabilitat que s'hagi aixecat després de les 7.20 h.?

Resolució. Considerem la variable aleatòria $X/Y = \frac{1}{2}$. Fixau-vos que es tracta d'una variable aleatòria contínua ja que tant X com Y ho són.

El que demanen calcular és $p\left\{\left(X/Y = \frac{1}{2}\right) \geq \frac{1}{4}\right\}$ (probabilitat que s'hagi aixecat després de les 7.15 h. $X = \frac{1}{4}$ d'hora després de les 7.00 h.) suposant que ha tardat 30 minuts en arribar a l'oficina ($Y = \frac{1}{2}$ d'hora).

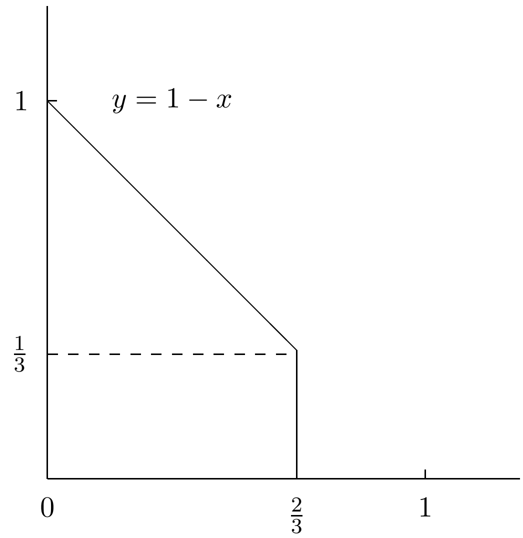

Figura 5.2: Representació gràfica del domini de la funció de densitat conjunta.

Trobarem primer la funció de densitat conjunta de les variables X i Y:

$$f_{XY}(x,y) = f_{Y/X}(y|x) \cdot f_X(x) = \begin{cases} \frac{81}{26} \cdot 2y = \frac{81}{13}y, & \text{si } 0 < x < \frac{2}{3},\ 0 < y < 1-x, \\ 0, & \text{en cas contrari.} \end{cases}$$

En la figura 5.2 està representat gràficament el domini de la funció de densitat conjunta on aquesta no és nul·la.

Trobem a continuació la funció de densitat de la variable aleatòria Y fent servir la fórmula:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx.$$

Fixau-vos en la figura 5.2 que, per a $0 \leq y \leq \frac{1}{3}$, $f_{XY}(x,y)$ no és nul·la per a $0 \leq x \leq \frac{2}{3}$, i, per a $\frac{1}{3} \leq y \leq 1$, $f_{XY}(x,y)$ no és nul·la per a $0 \leq x \leq 1-y$.

La funció de densitat de Y serà, doncs:

$$f_Y(y) = \begin{cases} \int_0^{\frac{2}{3}} \frac{81}{13} y\,dx = \frac{54}{13}y, & \text{si } 0 \leq y \leq \frac{1}{3}, \\ \int_0^{1-y} \frac{81}{13} y\,dx = \frac{81}{13}y(1-y), & \text{si } \frac{1}{3} < y \leq 1. \end{cases}$$

Trobem la funció de densitat de la variable aleatòria $X/Y = \frac{1}{2}$ per fer-la servir per trobar $p\left\{\left(X/Y=\frac{1}{2}\right)\geq\frac{1}{4}\right\}$.

$$f_{X/Y}\left(x/\frac{1}{2}\right)=\frac{f_{XY}\left(x,\frac{1}{2}\right)}{f_Y\left(\frac{1}{2}\right)}=\frac{\frac{81}{13}\cdot\frac{1}{2}}{\frac{81}{13}\cdot\frac{1}{2}\cdot\frac{1}{2}}=2,\ \text{si } 0\leq x\leq 1-\frac{1}{2}=\frac{1}{2}.$$

Podem dir, que la variable aleatòria $X/Y=\frac{1}{2}$ és uniforme en l'interval $\left(0,\frac{1}{2}\right)$.

Recordem que si una variable aleatòria W és uniforme en l'interval (a,b), la seva funció de distribució és:

$$F_W(t)=\begin{cases}0, & \text{si } t<a,\\ \frac{t-a}{b-a}, & \text{si } a\leq t\leq b,\\ 1, & \text{si } t>b.\end{cases}\tag{5.3}$$

Per tant, fent servir la fórmula 5.3 podem trobar $p\left\{\left(X/Y=\frac{1}{2}\right)\geq\frac{1}{4}\right\}$:

$$p\left\{\left(X/Y=\frac{1}{2}\right)\geq\frac{1}{4}\right\}=1-F_{X/Y}\left(\frac{1}{4}/\frac{1}{2}\right)=1-\frac{\frac{1}{4}-0}{\frac{1}{2}-0}=1-\frac{1}{2}=\frac{1}{2}.$$

També es demana trobar $p\left\{\left(X/Y=\frac{5}{6}\right)\geq\frac{1}{3}\right\}$ (probabilitat que s'hagi aixecat després de les 7.20 h., $X=\frac{20}{60}=\frac{1}{3}$ d'hora) suposant que tarda 50 minuts en arribar a l'oficina ($Y=\frac{50}{60}=\frac{5}{6}$ d'hora).

Resoldrem aquesta segona part de forma semblant a la primera.

Trobem primer la funció de densitat de la variable aleatòria $X/Y=\frac{5}{6}$:

$$f_{X/Y}\left(x/\frac{5}{6}\right)=\frac{f_{XY}\left(x,\frac{5}{6}\right)}{f_Y\left(\frac{5}{6}\right)}=\frac{\frac{81}{13}\cdot\frac{5}{6}}{\frac{81}{13}\cdot\frac{5}{6}\cdot\frac{5}{6}}=6,\ \text{si } 0\leq x\leq 1-\frac{5}{6}=\frac{1}{6}.$$

Podem dir que la variable aleatòria $X/Y=\frac{5}{6}$ és uniforme en l'interval $\left(0,\frac{1}{6}\right)$.

Per tant, fent servir la fórmula 5.3,

$$p\left\{\left(X/Y=\frac{5}{6}\right)\geq\frac{1}{3}\right\}=1-F_{X/Y}\left(\frac{1}{3}/\frac{5}{6}\right)=1-1=0.$$

5.11.- Si X_1, X_2, X_3 són variables aleatòries independents de Poisson, cada una amb paràmetre $\lambda_i s$ $(i = 1, 2, 3)$, trobau la funció de probabilitat conjunta $f_{X_1X_2X_3}(x_1, x_2, x_3)$.

Resolució. Recordem que si X és una variable aleatòria de Poisson amb paràmetre λs, la seva funció de probabilitat és:

$$f_X(x) = \frac{(\lambda s)^x}{x!} \cdot \mathrm{e}^{-\lambda s}, \; x = 0, 1, \ldots$$

Si X_1, X_2 i X_3 són independents, la funció de probabilitat conjunta serà el producte de les funcions de probabilitat de les X_i.

Així, doncs:

$$\begin{aligned} f_{X_1X_2X_3}(x_1, x_2, x_3) &= f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot f_{X_3}(x_3) \\ &= \frac{(\lambda_1 s)^{x_1}}{x_1!} \cdot \mathrm{e}^{-\lambda_1 s} \cdot \frac{(\lambda_2 s)^{x_2}}{x_2!} \cdot \mathrm{e}^{-\lambda_2 s} \cdot \frac{(\lambda_3 s)^{x_3}}{x_3!} \cdot \mathrm{e}^{-\lambda_3 s} \\ &= \frac{\lambda_1^{x_1} \cdot \lambda_2^{x_2} \cdot \lambda_3^{x_3}}{x_1! \cdot x_2! \cdot x_3!} \cdot s^{x_1+x_2+x_3} \cdot \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)s}. \end{aligned}$$

5.12.- Si $X_1, X_2, \ldots, X_k$ són variables aleatòries binomials independents amb paràmetres $n_1, p, n_2, p, \ldots, n_k, p$, respectivament, demostrau que $Y = \sum_{i=1}^{k} X_i$ és una variable aleatòria binomial amb paràmetres $\sum_{i=1}^{k} n_i$ i p.

Resolució. Abans de resoldre el problema recordem dues propietats de variables aleatòries:

a) Si X és una variable aleatòria binomial de paràmetres n i p, la funció generatriu de moments val:
$$m_X(t) = (p \cdot \mathrm{e}^t + q)^n, \text{ on } q = 1 - p.$$

b) Si $X_1, \ldots, X_k$ són variables aleatòries independents i definim la variable $Y = \sum_{i=1}^{k} X_i$, aleshores la funció generatriu de moments de la variable Y és el producte de les funcions generatrius de moments de les variables $X_i,\ i = 1, \ldots, k$:

$$m_Y(t) = m_{X_1}(t) \ldots m_{X_k}(t) := \prod_{i=1}^{k} m_{X_i}(t).$$

Tenint en compte aquestes dues propietats, per comprovar que la variable $Y = \sum_{i=1}^{k} X_i$ és una variable aleatòria binomial de paràmetres $\sum_{i=1}^{k} n_i$ i p, basta veure que la seva funció generatriu de moments és:

$$m_Y(t) = (p \cdot \mathrm{e}^t + q)^{\sum_{i=1}^{k} n_i}.$$

Així, doncs:

$$m_Y(t) = \prod_{i=1}^{k} m_{X_i}(t) = \prod_{i=1}^{k} (p \cdot \mathrm{e}^t + q)^{n_i} = (p \cdot \mathrm{e}^t + q)^{\sum_{i=1}^{k} n_i},$$

com volíem veure. □

5.13.- Demostrau que si $X_1, X_2, \ldots, X_n$ són variables aleatòries independents de Poisson amb paràmetres $\lambda_1 s_1, \lambda_2 s_2, \ldots, \lambda_n s_n$, respectivament, aleshores $Y = \sum_{i=1}^{n} X_i$ és una variable aleatòria de Poisson amb paràmetre $\sum_{i=1}^{n} \lambda_i s_i$.

Resolució. Recordem que si X és $Poiss(\lambda s)$, la seva funció generatriu de moments és:

$$m_X(t) = \mathrm{E}\left(\mathrm{e}^{(tX)}\right) = \sum_{k=0}^{\infty} \mathrm{e}^{tk} \frac{(\lambda s)^k}{k!} \mathrm{e}^{-\lambda s} = \mathrm{e}^{-\lambda s} \sum_{k=0}^{\infty} \frac{(\mathrm{e}^t \lambda s)^k}{k!} = \mathrm{e}^{\lambda s (\mathrm{e}^t - 1)}.$$

Suposem ara que X_i és $Poiss(\lambda_i s_i),\ i = 1, \ldots, n$ i sigui $Y = \sum_{i=1}^{n} X_i$. Vegem que la funció generatriu de Y coincideix amb la funció generatriu d'una variable de Poisson amb paràmetre $\sum_{i=1}^{n} \lambda_i s_i$.

$$m_Y(t) = \prod_{i=1}^{n} m_{X_i}(t) = \prod_{i=1}^{n} \mathrm{e}^{\lambda_i s_i (\mathrm{e}^t - 1)} = \mathrm{e}^{\sum_{i=1}^{n} \lambda_i s_i (\mathrm{e}^t - 1)}.$$

5.14.- En una placa de devora la porta d'un ascensor es pot llegir: "Capacitat màxima 6 persones, 500 Kg". Suposem que els pesos de les persones que fan servir aquest ascensor se seleccionen d'una distribució normal amb $\mu = 63.5$ Kg, $\sigma = 13.6$ Kg. Si 6 persones entren en l'ascensor, quina és la probabilitat que el seu pes combinat sigui més gran que la capacitat màxima de 500 Kg?

Resolució. Considerem la variable aleatòria X :"Pes d'una persona". Tenim que X es distribueix segons una llei Normal amb paràmetres $\mu = 63.5$ Kg i $\sigma = 13.6$ Kg. Per tant, tenim que el pes de les 6 persones serà $Y = \sum_{i=1}^{6} X_i$ on X_i representa el pes de la persona i-èssima. La variable Y serà normal amb paràmetres $\mu_Y = \sum_{i=1}^{6} 63.5 = 381$ Kg i $\sigma_Y = \sqrt{\sum_{i=1}^{6} 13.6^2} \approx 33.3131$ Kg. Per tant, el que ens demanen es pot calcular de la manera següent:

$$\begin{aligned} p\{Y \geq 500\} &= p\left\{Z = N(0,1) \geq \frac{500-381}{33.3131}\right\} = p\{Z \geq 3.5721\} \\ &= 1 - F_Z(3.5721) \approx 1 - 0.9998 = 0.0002 \end{aligned}$$

5.15.- Siguin X i Y variables aleatòries independents i uniformes en l'interval $[0,1]$. Considerem la variable aleatòria $U = X + Y$. Trobau la funció de densitat i la de distribució de U. Trobau també l'esperança i la variància de U.
Final. Juny 93.

Resolució. Considerem les variables aleatòries X i Y uniformes en l'interval $(0,1)$ i independents. Recordem que la funció de densitat de cada una val:

$$f_X(t) = f_Y(t) = \begin{cases} 1, & \text{si } x \in (0,1), \\ 0, & \text{en cas contrari.} \end{cases}$$

Com que X i Y són independents, la seva funció de densitat conjunta serà:

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y) = \begin{cases} 1, & \text{si } (x,y) \in (0,1) \times (0,1), \\ 0, & \text{en cas contrari.} \end{cases}$$

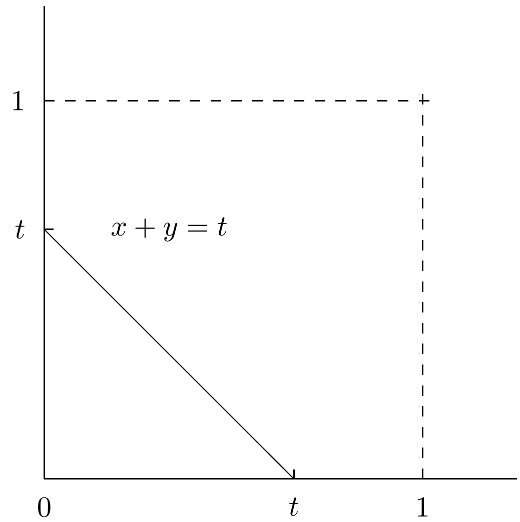

Figura 5.3: Domini d'integració en el cas 1.

Considerem ara la variable $Z = X + Y$. Fixau-vos que el rang de Z serà l'interval $(0, 2)$, ja que tant X com Y prenen valors dins l'interval $(0, 1)$.

Trobem primer la funció de distribució de Z, $F_Z(t)$. Basta considerar $t \in (0, 2)$ ja que, per a $t \leq 0$, $F_Z(t) = 0$ i per a $t \geq 2$, $F_Z(t) = 1$.

Així, doncs:

$$F_Z(t) = p\{Z = X + Y \leq t\} = \int\int_{\{(x,y)\in\mathbb{R}^2 \ \mid \ x+y\leq t\}} f_{XY}(x, y)\, dx\, dy.$$

Només ens interessa integrar en el domini on la funció de densitat conjunta no sigui nul·la. O sigui, basta integrar en el domini que resulta de fer la intersecció entre els conjunts:

$$\{(x, y) \in \mathbb{R}^2 \mid x + y \leq t\} \cap ((0, 1) \times (0, 1)).$$

Considerem dos casos:

Cas 1. $0 \leq t \leq 1$. En aquest cas, la intersecció anterior serà el triangle de vèrtexs $(0, 0)$, $(t, 0)$ i $(0, t)$ (veure figura 5.3).

Per tant, la funció de distribució de Z serà:

$$F_Z(t) = \int_0^t \int_0^{t-x} 1\, dy\, dx = \int_0^t (t - x)\, dx = \left[tx - \frac{x^2}{2}\right]_0^t = t^2 - \frac{t^2}{2} = \frac{t^2}{2}.$$

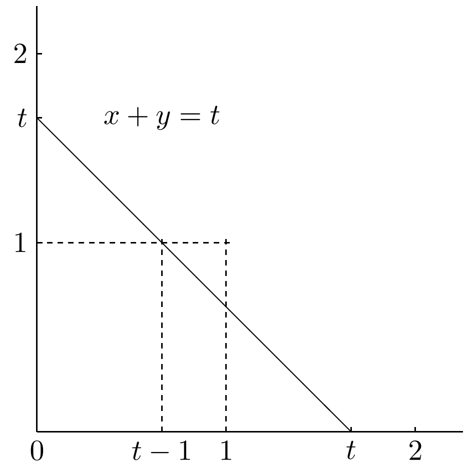

Figura 5.4: Domini d'integració en el cas 2.

Cas 2. $1 \leq t \leq 2$. En aquest cas, la intersecció anterior serà el trapezi de vèrtexs $(0, 0)$, $(1, 0)$, $(1, t-1)$, $(t-1, 1)$ i $(0, 1)$ (vegeu figura 5.4).

Per tant, la funció de distribució de Z serà:

$$\begin{aligned} F_Z(t) &= \int_0^{t-1}\int_0^1 1\,dy\,dx + \int_{t-1}^1\int_0^{t-x} 1\,dy\,dx = \int_0^{t-1} 1\,dx + \int_{t-1}^1 (t-x)\,dx \\ &= t-1+\left[tx-\frac{x^2}{2}\right]_{t-1}^1 = t-1+t-\frac{1}{2}-t(t-1)+\frac{(t-1)^2}{2} \\ &= -\frac{t^2}{2}+2t-1. \end{aligned}$$

La funció de distribució de Z serà, doncs:

$$F_Z(t) = \begin{cases} 0, & \text{si } t \leq 0, \\ \frac{t^2}{2}, & \text{si } 0 \leq t \leq 1, \\ -\frac{t^2}{2}+2t-1, & \text{si } 1 \leq t \leq 2, \\ 1, & \text{si } t \geq 1. \end{cases}$$

Per trobar la funció de densitat de Z, basta derivar la funció anterior:

$$f_Z(t) = F'_Z(t) = \begin{cases} t, & 0 \le t \le 1, \\ 2-t, & \text{si } 1 < t \le 2, \\ 0, & \text{en cas contrari.} \end{cases}$$

Per trobar $\mathrm{E}Z$ i $\mathrm{Var}Z$, ho podem fer de dues formes: integrant directament amb la funció de densitat de Z o fent servir les propietats de la variància i l'esperança tenint en compte que $\mathrm{E}X = \mathrm{E}Y = \frac{1}{2}$ i $\mathrm{Var}X = \mathrm{Var}Y = \frac{1}{12}$. Ho farem de la segona forma. Deixam al lector la comprovació dels resultats fent servir la funció de densitat de Z.

$$\begin{aligned} \mathrm{E}Z &= \mathrm{E}(X+Y) = \mathrm{E}X + \mathrm{E}Y = \frac{1}{2} + \frac{1}{2} = 1, \\ \mathrm{Var}Z &= \mathrm{Var}(X+Y) = \mathrm{Var}X + \mathrm{Var}Y = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}. \end{aligned}$$

5.16.- Siguin X i Y variables aleatòries independents.

a) Provau que $m_{X+Y}(t) = m_X(t)m_Y(t)$, o sigui, la funció generatriu de la suma és el producte de les funcions generatrius.

b) Suposant que X és $N(\mu_1, \sigma_1^2)$, Y és $N(\mu_2, \sigma_2^2)$, i que són independents, i fent sevir l'apartat a), provau que $X+Y$ és $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

c) Suposant que X és $N(\mu, \sigma^2)$, provau que $Y = aX + b$ és $N(a\mu + b, a^2\sigma^2)$.

d) Siguin $X_1, X_2, \ldots, X_n$ variables aleatòries totes $N(\mu, \sigma^2)$ i independents. Provau fent servir b) i c) que la variable aleatòria $\overline{X} = \frac{X_1+\ldots+X_n}{n}$ és $N(\mu, \frac{\sigma^2}{n})$.

e) Suposant $\sigma = 2$, trobau n perquè $p\left\{|\overline{X} - \mu| \le 1\right\} = 0.95$ amb les mateixes suposicions que en l'apartat d).

Primer parcial. Curs 92-93.

Resolució.

a) Suposem que X i Y són independents. Per tant, la funció generatriu de la suma valdrà:

$$m_{X+Y}(t) = \mathrm{E}\left(\mathrm{e}^{t(X+Y)}\right) = \mathrm{E}\left(\mathrm{e}^{tX} \cdot \mathrm{e}^{tY}\right) = \mathrm{E}\left(\mathrm{e}^{tX}\right) \cdot \mathrm{E}\left(\mathrm{e}^{tY}\right) = m_X(t) \cdot m_Y(t).$$

b) Suposem ara que X és $N(\mu_1, \sigma_1^2)$ i Y és $N(\mu_2, \sigma_2^2)$. Vegem que $X+Y$ és $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$. Per provar-ho veurem que la funció generatriu de $X+Y$ coincideix amb la funció generatriu d'una variable aleatòria normal de paràmetres $\mu_1+\mu_2$ i $\sigma_1^2+\sigma_2^2$.

Recordem que si X és $N(\mu, \sigma^2)$, la funció generatriu de moments val:

$$m_X(t) = \mathrm{e}^{t\mu + \frac{t^2}{2}\sigma^2}.$$

Per tant, tenint en compte la fórmula anterior i l'apartat a) podem trobar la funció generatriu de la variable $X+Y$:

$$m_{X+Y}(t) = m_X(t)\cdot m_Y(t) = \mathrm{e}^{t\mu_1+\frac{t^2}{2}\sigma_1^2}\cdot \mathrm{e}^{t\mu_2+\frac{t^2}{2}\sigma_2^2} = \mathrm{e}^{t(\mu_1+\mu_2)+\frac{t^2}{2}(\sigma_1^2+\sigma_2^2)},$$

funció que coincideix amb la funció generatriu d'una variable aleatòria $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

c) Suposem que X és $N(\mu, \sigma^2)$. Per veure que la variable $Y = aX+b$ és $N(a\mu+b, a^2\sigma^2)$, farem servir la mateixa tècnica que en l'apartat b). Per tant:

$$\begin{aligned} m_Y(t) &= \mathrm{E}\left(\mathrm{e}^{t(aX+b)}\right) = \mathrm{e}^{tb}\cdot \mathrm{E}\left(\mathrm{e}^{X(at)}\right) \\ &= \mathrm{e}^{tb}\cdot m_X(at) = \mathrm{e}^{tb}\cdot \mathrm{e}^{at\mu+\frac{a^2t^2}{2}\sigma^2} = \mathrm{e}^{t(a\mu+b)+\frac{a^2t^2}{2}\sigma^2}, \end{aligned}$$

funció que correspon a una variable aleatòria $N(a\mu+b, a^2\sigma^2)$.

d) Sigui $X_1, \ldots, X_n$ variables aleatòries $N(\mu, \sigma^2)$ i independents. Considerem $\overline{X} = \frac{X_1+\cdots+X_n}{n}$. Vegem que $\overline{X}$ és $N\left(\mu, \frac{\sigma^2}{n}\right)$.

Fent servir l'apartat b), podem dir que la variable aleatòria $S = X_1 + \ldots X_n$ és $N(n\mu, n\sigma^2)$. Es pot provar per inducció damunt n. L'apartat b) ens diu que l'afirmació anterior és certa per a $n=2$.

Fent servir l'apartat c), tenim que $\overline{X} = \frac{S}{n}$ és $N\left(\mu, \frac{\sigma^2}{n}\right)$. Basta prendre $a = \frac{1}{n}$ i $b = 0$.

e) Suposem $\sigma = 2$. Aleshores $\overline{X}$ és $N\left(\mu, \frac{4}{n}\right)$. Per tant:

$$\begin{aligned} p\left\{\left|\overline{X}-\mu\right| \le 1\right\} &= p\left\{\left|\frac{\overline{X}-\mu}{\frac{2}{\sqrt{n}}}\right| \le \frac{1}{\frac{2}{\sqrt{n}}}\right\} = p\left\{|Z = N(0,1)| \le \frac{\sqrt{n}}{2}\right\} \\ &= p\left\{-\frac{\sqrt{n}}{2} \le Z \le \frac{\sqrt{n}}{2}\right\} = F_Z\left(\frac{\sqrt{n}}{2}\right) - F_Z\left(-\frac{\sqrt{n}}{2}\right) \\ &= 2F_Z\left(\frac{\sqrt{n}}{2}\right) - 1. \end{aligned}$$

Per tant:

$$F_Z\left(\frac{\sqrt{n}}{2}\right) = \frac{0.95+1}{2} = 0.975.$$

Mirant les taules, $\frac{\sqrt{n}}{2} = 1.96$. D'on deduïm que $n \approx (1.96)^2 \cdot 4 = 15.36$. n val aproximadament 16.

5.17.- El nombre de clients que arriben a una estació de servei durant un temps t és una variable aleatòria de Poisson amb paràmetre βt. El temps necessari per servir cada client és una variable aleatòria exponencial amb paràmetre α. Determinau la densitat de la variable aleatòria que dóna el nombre de clients que arriben durant el temps de servei T d'un determinat client. Se suposa que les arribades de clients són independents del temps de servei dels clients. Ind.: $\int_0^\infty r^k\, \mathrm{e}^{-r}\, dr = \Gamma(k+1) = k!$

Resolució. Posem:

$$\begin{aligned}
N &: \text{ nombre de clients que arriben durant } t \to Poiss(\beta t),\\
T &: \text{ temps necessari per servir un client} \to Exp(\alpha),\\
X &: \text{ nombre de clients que arriben durant el temps de servei d'un client.}
\end{aligned}$$

Sabem que $X/T = t$ és una $Poiss(\beta t)$. Aleshores

$$p\{X = k/T = t\} = \frac{(\beta t)^k}{k!}\,\mathrm{e}^{-\beta t},\ k \geq 0.$$

També

$$f_T(t) = \alpha \cdot \mathrm{e}^{-\alpha t},\ t \geq 0.$$

Aleshores

$$\begin{aligned}
p\{X = k\} &= \int_{-\infty}^{\infty} p\{X = k/T = t\} \cdot f_T(t)\, dt\\
&= \int_0^\infty \frac{(\beta t)^k}{k!}\,\mathrm{e}^{-\beta t} \cdot \alpha \cdot \mathrm{e}^{-\alpha t}\, dt = \frac{\alpha}{k!} \cdot \int_0^\infty (\beta t)^k \cdot \mathrm{e}^{-(\alpha+\beta)t}\, dt
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\alpha\cdot\beta^k}{k!\cdot(\alpha+\beta)^{k+1}}\cdot\int_0^\infty [(\alpha+\beta)t]^k\cdot \mathrm{e}^{-(\alpha+\beta)t}\cdot(\alpha+\beta)\,dt\\
&= \frac{\alpha}{\alpha+\beta}\cdot\left(\frac{\beta}{\alpha+\beta}\right)^k\cdot\frac{1}{k!}\cdot k! = \frac{\alpha}{\alpha+\beta}\cdot\left(\frac{\beta}{\alpha+\beta}\right)^k,\ k\geq 0.
\end{aligned}
$$

Aleshores X segueix una distribució geomètrica amb paràmetre $\frac{\alpha}{\alpha+\beta}$.

5.3 Problemes proposats

5.1.- Una moneda no trucada té un 1 pintat en una cara i un 2 en l'altra cara. Es llança a l'aire dues vegades la moneda. Sigui X la suma dels dos nombres que surten i sigui Y la diferència dels dos nombres (el primer menys el segon). Trobau la funció de probabilitat conjunta $f_{XY}(x,y)$, la funció de probabilitat de X, $f_X(x)$ i la funció de probabilitat de Y, $f_Y(y)$.

5.2.- Què ha de valer A si es vol que la funció següent

$$
f_{XY}(x,y) = \begin{cases} A\frac{x}{y}, & \text{si } 0<x<1,\ 1<y<2,\\ 0, & \text{en els altres casos,}\end{cases}
$$

sigui una funció de densitat per a la variable aleatòria conjunta (X,Y).

5.3.- Provau que:

$$
p\left\{(a<X\leq b)\cap(c<Y\leq d)\right\} = F_{XY}(b,d)-F_{XY}(a,d)-F_{XY}(b,c)+F_{XY}(a,c).
$$

Indicació: feu un dibuix.

5.4.- Suposem que (X,Y) és una variable aleatòria bidimensional contínua amb funció de densitat:

$$
f_{XY}(x,y) = \begin{cases} \frac{1}{x}, & \text{si } 0<y<x,\ 0<x<1,\\ 0, & \text{en els altres casos.}\end{cases}
$$

Trobau les funcions de densitat marginals per a X i Y.

5.5.- Suposem que es pinta un "+1" en una cara d'una moneda no trucada i un "-1" en l'altra cara. La moneda es llança a l'aire dues vegades. Sigui X el nombre que surt quan la tiram la primera vegada i Y el nombre que surt quan la tiram la segona vegada. Trobau $f_{XY}(x,y)$, EX, EY i E$\left(\frac{X}{Y}\right)$.

5.6.- Es llança 3 vegades una moneda no trucada. Sigui X el nombre de cares que s'obtenen i Y el nombre de creus. Trobau la funció de probabilitat conjunta per a (X,Y) i trobau σ_{XY}.

5.7.- Suposem que (X,Y) té la densitat $f_{XY} = c$ per a (x,y) en el quadrilàter de vèrtexs $(0,0)$, $(1,1)$, $(a, 1-a)$ i $(1-a, a)$ on $0 \leq a \leq \frac{1}{2}$.

a) Trobau el valor de c.

b) Trobau ρ_{XY} si $a = 0$ i $a = \frac{1}{2}$.

5.8.- Siguin X i Y variables aleatòries discretes amb funció de probabilitat conjunta:

$$f_{XY}(x,y) = \begin{cases} \frac{1}{n^2}, & \text{per } x = 1,2,\ldots,n, \quad y = 1,2,\ldots,n, \\ 0, & \text{en els altres casos.} \end{cases}$$

Comprovau que X i Y són independents.

5.9.- Siguin X i Y variables aleatòries contínues conjuntament distribuïdes amb funció de densitat:

$$f_{(X,Y)}(x,y) = \begin{cases} 4, & \text{per } 0 < x < 1, \text{ i } 0 < y < \frac{1}{4}, \\ 0, & \text{en els altres casos.} \end{cases}$$

Comprovau que X i Y són independents.

5.10.- Si la probabilitat conjunta per a (X,Y) és no nul·la en exactament 3 punts, què s'ha de complir per què X i Y siguin independents?

5.11.- Suposem que X_1 i X_2 són variables aleatòries independents, cada una amb mitjana 0 i variància σ^2. Definim

$$\begin{aligned} Y &= a_1X_1 + a_2X_2, \\ Z &= b_1X_1 + b_2X_2. \end{aligned}$$

Calculau μ_Y, μ_Z, σ_Y^2, σ_Z^2 i σ_{YZ}.

5.12.- Suposem que dos amics aposten per saber quin dels dos paga el cafè durant 20 dies. Definim:

$$X_i = \begin{cases} 1, & \text{si el primer guanya el } i\text{-èssim dia,} \\ 0, & \text{si el primer perd en el } i\text{-èssim dia,} \end{cases}$$

per a $i = 1, 2, \ldots, 20$. Aleshores $Y = \sum_{i=1}^{20} X_i$ és el nombre total de vegades que el primer guanya el segon en els 20 dies. Suposant que el joc és aleatori, això és, que la probabilitat de guanyar cada dia és $\frac{1}{2}$ i que els resultats són independents dia a dia, calculau la probabilitat que $Y \geq 10$ i la que $Y \geq 15$.

5.13.- Els pacients amb membres romputs arriben a un hospital d'una universitat en forma de successos de Poisson a raó d'un per dia.

a) Quina és la probabilitat que arribin a l'hospital 7 pacients en un període de 7 dies?

b) Donat que arribaran 7 pacients en un període de 7 dies, quina és la probabilitat que n'arribi un en cada un dels 7 dies?

5.14.- Siguin X_1 i X_2 variables aleatòries binomials independents amb paràmetres $n_1 = 3$, p_1 i $n_2 = 2$, p_2, respectivament. Trobau la distribució per a $X_1 + X_2$. Té forma binomial?

5.15.- De la successió dels nombres naturals, en triam a l'atzar dos. Sigui X_1 la variable aleatòria amb valors 0 si els dos nombres són tots dos parells o tots dos senars i 1 en cas contrari. Sigui X_2 la variable aleatòria amb valors $0, 1, 2$, segons que dels dos nombres triats no sigui parell cap, ho sigui un o ho siguin els dos, respectivament. Trobau:

a) La funció de probabilitat conjunta de X_1 i X_2, com també les funcions de densitat marginals.

b) $\mathrm{E}(X_1)$, $\mathrm{E}(X_2)$, $\mathrm{Var}X_1$, $\mathrm{Var}X_2$.

5.16.- En una petita comunitat de 10 parelles en les quals ambdós membres treballen, l'ingrés anual (en milers de dòlars) té la distribució següent:

Parella	Ingrés de l'home	Ingrés de la dona
1	10	5
2	15	15
3	15	10
4	10	10
5	10	10
6	15	5
7	20	10
8	15	10
9	20	15
10	20	10

Es pren una parella a l'atzar perquè representi aquesta comunitat en una convenció. Sigui X l'ingrés aleatori de l'home i Y, el de la dona. Trobau:

a) La funció de probabilitat conjunta de X i Y.

b) La distribució de X, la seva mitjana i la seva variància.

c) La distribució de Y, la seva mitjana i la seva variància.

d) La covariància de X i Y.

e) $\mathrm{E}(X/Y = 10)$ i $\mathrm{E}(Y/X = 20)$.

Final. Setembre 91.

5.17.- Considerem la funció:

$$f(x,y) = \begin{cases} 3x, & \text{si } 0 \leq x \leq 1 \text{ i } x \leq y \leq 1, \\ 3y, & \text{si } 0 \leq x \leq 1 \text{ i } 0 \leq y \leq x, \\ 0, & \text{en cas contrari.} \end{cases}$$

a) Comprovau que és una funció de densitat.

b) Trobau la funció de distribució.

c) Trobau la funció de densitat de X, Y, X/Y i Y/X.

Final. Setembre 93.

5.18.- Sigui (X, Y) la variable aleatòria contínua amb funció de densitat conjunta:

$$f(x,y) = \begin{cases} 6xy(2-x-y), & \text{si } 0 \leq x, y \leq 1, \\ 0, & \text{en cas contrari.} \end{cases}$$

Trobau $\mathrm{E}(Y/X)$.
Final. Setembre 94.

5.19.- Sigui (X, Y) la variable aleatòria bidimensional amb funció de probabilitat conjunta:

$Y \backslash X$	-1	0	1
-1	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{9}$
0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{18}$
1	$\frac{2}{9}$	0	$\frac{2}{9}$

Trobau $\mathrm{E}(Y/X=1)$.
Primer parcial. Febrer 95.

5.20.- Un dau amb 4 cares està numerat amb $1, 2, 3, 4$. Llançam el dau 5 vegades. Trobau la probabilitat que totes les cares surtin una vegada com a mínim.
Primer parcial. Febrer 95.

5.21.- Sigui

$$f(x,y) = \begin{cases} \alpha xy, & \text{si } (x,y) \text{ pertany al triangle de} \\ & \text{vèrtexs } (0,0),\ (a,0) \text{ i } (0,a), \\ 0, & \text{en cas contrari,} \end{cases}$$

on $a > 0$ és un paràmetre conegut. Què ha de valer α perquè f sigui la funció de densitat conjunta d'una variable aleatòria (X, Y)?
Final. Juny 95.

5.22.- Sigui

$$f(x,y)=\begin{cases} k, & \text{si } (x,y)\in C=\{(x,y)\in\mathbb{R}^2:\ x\geq 0, y\geq 0, x^2+y^2\leq 4, x+y\geq 2\}, \\ 0, & \text{en cas contrari.}\end{cases}$$

Trobau el valor de k perquè f sigui una densitat.

Final. Setembre 96.

5.23.- Siguin X i Y dues variables aleatòries uniformes en l'interval $(0,1)$ i independents. Considerem la variable aleatòria $P=XY$. Trobau la funció de densitat de P, $\mathrm{E}(P)$ i $\mathrm{Var}P$.

Primer parcial. Febrer 95.

5.24.- Trobau $\mathrm{E}(Y/X)$ si la funció de densitat conjunta de la variable aleatòria bidimensional (X,Y) val:

$$f_{XY}(x,y)=\begin{cases} \frac{1}{8}\left(y^2-x^2\right)\mathrm{e}^{-y}, & \text{si } 0<|x|\leq y<\infty, \\ \\ 0, & \text{en cas contrari.}\end{cases}$$

Primer parcial. Febrer 95.

5.25.- La variable (X,Y) està distribuïda uniformement en el cercle $x^2+y^2\leq 4$. Calculau:

a) $p\{Y>kX\}$.

b) Densitat marginal de la variable aleatòria X.

c) Densitat condicionada $f_{X/Y}(x/1)$.

d) $p\{|X|<1/Y=0.5\}$.

Final. Juny 95.

5.26.- Escollim dues cartes d'una baralla espanyola de 48 cartes. Considerem les variables:

X : "nombre d'ors".

Y : "nombre de figures" (les figures són les cartes marcades amb un 10, 11 o 12).

Trobau la funció de probabilitat conjunta de (X, Y) i $\mathrm{E}(Y/X)$.
Examen extraordinari de febrer 95.

5.27.- Considerem una variable aleatòria bidimensional (X, Y) amb funció de densitat conjunta:

$$f(x,y) = \begin{cases} \frac{3y^2}{2}, & \text{si } 0 \leq x \leq 2,\ 0 \leq y \leq 1, \\ 0, & \text{en cas contrari.} \end{cases}$$

Calculau $p\left\{X \leq 1/Y \leq \frac{1}{2}\right\}$.
Final. Setembre 95.

5.28.- Sigui (X, Y) una variable aleatòria bidimensional contínua amb funció de densitat conjunta:

$$f(x,y) = \begin{cases} c, & \text{si } (x,y) \in T, \text{ on } T \text{ és el triangle de} \\ & \quad \text{vèrtexs } (0,3),\ (1,1) \text{ i } (2,3), \\ 0, & \text{en cas contrari.} \end{cases}$$

Es demana:

a) El valor de la constant c.

b) Funcions de densitat marginals de les variables aleatòries X i Y.

c) $p\{X \leq 1/Y > 2\}$.

Final. Febrer 96.

5.29.- Sigui (X, Y) una variable aleatòria bidimensional amb funció de densitat conjunta:

$$f(x,y) = \frac{\sqrt{3}}{2\pi} \mathrm{e}^{-(x^2+xy+y^2)}, \ (x,y) \in \mathbb{R}^2.$$

Trobau $\mathrm{E}(Y/X)$.
Final. Juny 96.

5.30.- Les variables aleatòries X_1 i X_2 són i.i.d. amb densitat comuna

$$f(x) = \begin{cases} 1, & \text{si } 0 \leq x \leq 1, \\ 0, & \text{en cas contrari.} \end{cases}$$

Determinau la densitat de $Y = X_1 - X_2$.

5.31.- Es llancen a l'aire dos daus de diferent color, un de blanc i l'altre de vermell. Sigui X la variable aleatòria que dóna el nombre de punts obtenguts amb el dau blanc i Y la variable aleatòria que dóna el nombre més gran dels punts obtenguts amb els dos daus.

a) Determinau la funció de probabilitat conjunta de X i Y.

b) Obteniu les funcions de probabilitat marginals.

c) Són X i Y independents?

d) Obteniu la distribució de X condicionada a $Y = 4$.

e) Quina és l'esperança de X condicionada a $Y = 4$?

f) Determinau $\mathrm{E}[X/Y]$.

g) Calculau $\mathrm{E}\left(\mathrm{E}(X/Y)\right)$ i comprovau que coincideix amb $\mathrm{E}X$.

5.32.- Suposem que la variable aleatòria X se selecciona a l'atzar de l'interval unitat i, aleshores, la variable aleatòria Y se selecciona a l'atzar de l'interval $(0, X)$. Determinau la distribució de Y.

5.33.- Es llancen a l'aire dos daus sense biaix. Siguin N_1 i N_2 els valors obtenguts en els dos daus. Posem $X = N_1 + N_2$ i $Y = |N_1 - N_2|$. Obteniu la distribució conjunta de X i Y i verificau que estan incorrelacionades. Són independents?

Capítol 6

Lleis dels grans nombres i teorema del límit central

6.1 Resum teòric

Considerem una successió $\{X_n\}_{n\geq 1}$ de variables aleatòries independents i idènticament distribuïdes. Les lleis dels grans nombres estudien el comportament asimptòtic de la successió de mitjanes mostrals $\left\{\dfrac{S_n}{n} : n \geq 1\right\}$, on $S_n = X_1 + \cdots + X_n$. Com a conseqüència de la llei forta s'obté la coneguda propietat que diu que la freqüència relativa d'un succés convergeix cap a la seva probabilitat quan el nombre de repeticions de l'experiment tendeix cap a infinit.

El teorema del límit central afirma que, sota certes condicions, la distribució de la suma normalitzada de n variables aleatòries s'acosta quan n creix cap a la d'una variable aleatòria gaussiana. A diferència de les lleis dels grans nombres, el teorema del límit central no estudia límits de successions de variables aleatòries, sinó límits de les funcions de distribució; més concretament, estudia la distribució límit de la successió de distribucions de les sumes parcials de les variables aleatòries donades.

6.1.1 Lleis del grans nombres

Sigui X una variable aleatòria de la qual desconeixem la seva esperança $\mathrm{E}X = \mu$, que suposarem que és finita. Siguin $X_1, \ldots, X_n$ n còpies repetides independentment de X; això vol dir que les X_i són n **variables aleatòries i.i.d.** (independents i idènticament distribuïdes) amb la mateixa distribució que X. Per tal d'estimar $\mathrm{E}X$ s'utilitza la **mitjana mostral** de la successió:

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{S_n}{n}.$$

La **llei feble dels grans nombres** afirma que

$$\forall \epsilon > 0, \ \lim_{n\to\infty} p\left\{|M_n - \mu| < \epsilon\right\} = 1.$$

Això ens diu que per a un valor de n fixat prou gran, la mitjana mostral usant n mostres estarà pròxima al valor mitjà real amb una probabilitat molt alta.

La **llei forta dels grans nombres** afirma que, donada una successió de variables aleatòries $X_1, X_2, \ldots$ i.i.d. amb mitjana comuna $\mathrm{E}X = \mu$ finita i variància comuna finita,

$$p\left\{\lim_{n\to\infty} M_n = \mu\right\} = 1,$$

és a dir que, amb probabilitat 1, la successió de mitjanes mostrals,

$$M_n = \frac{X_1 + \cdots + X_n}{n},$$

tendeix cap al valor mitjà real μ quan augmentam el nombre n de mostres.

En particular, donat un experiment aleatori, sigui $N_A(n)$ el nombre de vegades que apareix el succés A en n repeticions de l'experiment. Aleshores la freqüència relativa de A en les n repeticions és:

$$f_A(n) = \frac{N_A(n)}{n}.$$

Observem que $N_A(n)$ es pot escriure també com la suma

$$N_A(n) = X_1 + \cdots + X_n$$

on les variables aleatòries $X_1, \ldots, X_n$ són i.i.d. amb distribució comuna de Bernoulli amb paràmetre $p = p\{A\}$ (és a dir, $\mathrm{E}X_i = p \ \forall i = 1, \ldots, n$).

Per tant $f_A(n) = M_n$, la mitjana mostral, i aleshores

$$p\left\{\lim_{n\to\infty} f_A(n) = p\{A\}\right\} = 1,$$

és a dir que la freqüència relativa del succés A convergeix cap a la seva probabilitat amb probabilitat 1.

6.1.2 Teorema del límit central

Sigui $X_1, X_2, \ldots$ una successió de variables aleatòries i.i.d. amb mitjana μ i variància σ^2 finites. Considerem $S_n = X_1 + \cdots + X_n$, i sigui

$$Z_n = \frac{S_n - \mathrm{E}S_n}{\sqrt{\mathrm{Var}S_n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}.$$

Aleshores

$$\lim_{n\to\infty} p\{Z_n \le z\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-x^2/2}\, dx,$$

és a dir que la funció de distribució de Z_n convergeix cap a la funció de distribució de la $N(0,1)$. També resulta que la funció de distribució de S_n convergeix cap a la de la $N(n\mu, n\sigma^2)$.

Com a cas particular tenim el **teorema de De Moivre-Laplace:**

Sigui X una variable aleatòria $B(n,p)$. Aleshores $X = X_1 + \cdots + X_n$, on les X_i són $B(1,p)$ i.i.d. Aleshores la funció de distribució de $\dfrac{X - np}{\sqrt{npq}}$ convergeix cap a la de la $N(0,1)$. També resulta que la funció de distribució de X convergeix cap a la de la $N(np, npq)$.

Notes.

1. De fet, històricament, aquesta va ser la primera versió del teorema del límit central.

2. Com que X és discreta i una normal és absolutament contínua, per aproximar la funció de probabilitat de X es fa la correcció següent (d'altra manera l'aproximació seria nul·la, a causa de la continuïtat de la distribució normal):

$$p\{X = k\} \simeq p\{k - 1/2 \le N(np, npq) \le k + 1/2\}.$$

   També:

$$p\{k_1 \le X \le k_2\} \simeq p\{k_1 - 1/2 \le N(np, npq) \le k_2 + 1/2\}.$$

6.2 Problemes resolts

6.1.- Si X està distribuïda uniformement en l'interval $(0,1)$, comparau $p\{|X - \mu_X| < k\sigma_X\}$ amb els valors donats per la desigualtat de Txebixef per a $k = \frac{5}{4}, \frac{3}{2}, \frac{7}{4}$ i 2.

Resolució. Recordem que si la variable X està distribuïda uniformement en l'interval $(0,1)$, les funcions de densitat i de distribució són:

$$f_X(t) = \begin{cases} 1, & \text{si } t \in (0,1), \\ 0, & \text{en cas contrari}, \end{cases} \qquad F_X(t) = \begin{cases} 0, & \text{si } t < 0, \\ t, & \text{si } 0 \leq t \leq 1, \\ 1, & \text{si } t > 1. \end{cases}$$

L'esperança i la variància valdran, doncs:

$$\begin{aligned} \mu_X &= \int_0^1 x\,dx = \frac{1}{2}, \\ \sigma_X^2 &= \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = \int_0^1 x^2\,dx - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}. \end{aligned}$$

Per tant el valor de $p\{|X - \mu_X| < k\sigma_X\}$ serà:

$$\begin{aligned} p\{|X-\mu_X| < k\sigma_X\} &= p\{\mu_X - k\sigma_X < X < \mu_X + k\sigma_X\} \\ &= p\left\{\frac{1}{2} - k\frac{\sqrt{3}}{6} < X < \frac{1}{2} + k\frac{\sqrt{3}}{6}\right\} = \int_{\max\left\{0, \frac{1}{2} - k\frac{\sqrt{3}}{6}\right\}}^{\min\left\{1, \frac{1}{2} + k\frac{\sqrt{3}}{6}\right\}} 1\,dx \\ &= \min\left\{1, \frac{1}{2} + k\frac{\sqrt{3}}{6}\right\} - \max\left\{0, \frac{1}{2} - k\frac{\sqrt{3}}{6}\right\}. \qquad (6.1) \end{aligned}$$

A continuació exposam els resultats per als diferents k indicant el valor exacte fent servir la fórmula (6.1).

k	Valor exacte	Valor donat per la desigualtat de Txebixef $(1 - \frac{1}{k^2})$
$\frac{5}{4}$	$0.8608 - 0.1391 = 0.7216$	$1 - \frac{16}{25} = 0.36$
$\frac{3}{2}$	$0.9330 - 0.0669 = 0.8660$	$1 - \frac{4}{9} = 0.5555$
$\frac{7}{4}$	$1 - 0 = 1$	$1 - \frac{16}{49} = 0.6734$
2	$1 - 0 = 1$	$1 - \frac{1}{4} = 0.75$

6.2.-

a) Considerem la funció $g : \mathbb{R} \to \mathbb{R}$ donada per $g(p) = p(1-p)\ \forall\, p \in [0,1|$. Demostrau que $g(p)$ té un màxim en $p = \frac{1}{2}$. Deduïu que $g(p) \leq \frac{1}{4}$ per a tot $p \in [0,1]$.

b) Si X és una variable aleatòria binomial amb paràmetres n i p, aleshores $\mu_X = np$, $\sigma_X^2 = np(1-p)$. A més, si $Y = \frac{X}{n}$, tenim que $\mu_Y = p$ i $\sigma_Y^2 = \frac{p(1-p)}{n}$. Demostrau que:

$$p\left\{|Y-p| < \delta\right\} \geq 1 - \frac{p(1-p)}{n\delta^2}.$$

c) Per a la variable Y definida en b), demostrau que:

$$p\left\{|Y-p| < \delta\right\} \geq 1 - \frac{1}{4}\delta^2.$$

d) Suposem que volem trobar un valor de n tal que:

$$p\left\{|Y-p| < \delta\right\} \geq 0.9,$$

on δ és qualque constant positiva. Demostrau que si $n \geq \frac{2.5}{\delta^2}$, aleshores se satisfà el que volíem.

Resolució.

a) Per trobar els extrems relatius, hem de derivar i igualar a 0. Per tant,

$$g'(p) = 1 - 2p = 0,$$

d'on g té un únic extrem relatiu en $p = \frac{1}{2}$. Per veure que és un màxim relatiu, observem que $g''(p) = -2 < 0$.

Per tant, podem dir:

$$\begin{aligned} \max_{p\in[0,1]} g(p) &= \max\left\{g(0), g(1), g\left(\frac{1}{2}\right)\right\} \\ &= \max\left\{g(0) = g(1) = 0, g\left(\frac{1}{2}\right) = \frac{1}{4}\right\} = g\left(\frac{1}{2}\right) = \frac{1}{4}. \end{aligned}$$

Aleshores g té un màxim absolut per a $p \in [0,1]$ en $p = \frac{1}{2}$.

Podem dir, per tant, que $g(p) \leq \frac{1}{4}$, $\forall p \in [0,1]$.

b) Considerem la variable $Y = \frac{X}{n}$ on X és una variable aleatòria binomial amb paràmetres n i p. Hem de provar que:

$$p\{|Y-p| < \delta\} \geq 1 - \frac{p(1-p)}{n\delta^2}.$$

Observem que:

$$\begin{aligned} \mathrm{E}Y &= \mu_Y = \frac{1}{n}\mathrm{E}X = \frac{1}{n}np = p, \\ \mathrm{Var}Y &= \sigma_Y^2 = \frac{1}{n^2}\sigma_X^2 = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n}. \end{aligned}$$

Aplicant la desigualtat de Txebixef, tenim que:

$$\begin{aligned} p\{|Y-\mu_Y| < \delta\} &\geq 1 - \frac{\mathrm{Var}Y}{\delta^2}, \\ p\{|Y-p| < \delta\} &\geq 1 - \frac{\frac{p(1-p)}{n}}{\delta^2} = 1 - \frac{p(1-p)}{n\delta^2}\Box \end{aligned}$$

c) Fent servir l'apartat a) sabem que $p(1-p) \leq \frac{1}{4}$. Per tant, podem escriure, fent servir l'apartat b):

$$p\{|Y-p| < \delta\} \geq 1 - \frac{p(1-p)}{n\delta^2} \geq 1 - \frac{1}{4n\delta^2} \geq 1 - \frac{1}{4\delta^2} \text{ ja que } n \geq 1.\Box$$

d) Hem de veure que si

$$n \geq \frac{2.5}{\delta^2}, \tag{6.2}$$

aleshores:

$$p\{|Y-p| < \delta\} \geq 0.9 \tag{6.3}$$

Sabem que:

$$p\{|Y-p| < \delta\} \geq 1 - \frac{1}{4n\delta^2}.$$

Considerem un valor n tal que $1 - \frac{1}{4n\delta^2} \geq 0.9$. Observem que per a aquest n es complirà (6.3). Vegem que la condició que hem imposat per a n és equivalent a (6.2):

$$1 - \frac{1}{4n\delta^2} \geq 0.9 \Leftrightarrow \frac{1}{4n\delta^2} \leq 0.1 \Leftrightarrow 4n\delta^2 \geq 10 \Leftrightarrow n \geq \frac{10}{4\delta^2} = \frac{2.5}{\delta^2}\Box$$

6.3.- En una taula de daus d'un casino es llancen un parell de daus 180 vegades per hora (aproximadament). Quina és la probabilitat (aproximada) que es llancin entre 700 i 750 sumes de 7 en 24 hores?

Resolució. Considerem la variable aleatòria X :"Nombre de sumes de 7 en 24 hores".

Podem considerar que X és una variable aleatòria de Poisson amb paràmetre λs, on λ és el nombre mitjà de vegades que surt 7 en 1 hora i $s = 24$ hores. Trobem λ:

$$\lambda = 180 \times \text{ probabilitat de sortir 7 una vegada} = 180 \cdot \frac{6}{36} = 30,$$

ja que dels 36 resultats possibles n'hi ha 6 en què la suma dels dos daus val 7:

$$\{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}.$$

Deduïm així que X és $Poiss(\lambda s = 30 \cdot 24 = 720)$.

Ens demanen calcular la probabilitat següent: $p\{700 \leq X \leq 750\}$.

Com que els nombres que surten són molt grans, la probabilitat anterior no es pot trobar aplicant directament la fórmula de la funció de probabilitat d'una variable aleatòria de Poisson. Trobarem la probabilitat anterior aplicant el teorema del límit central, o sigui, aproximarem X per una variable aleatòria normal X_N de paràmetres $\mu = \lambda s = 720$ i $\sigma^2 = \lambda s = 720$.

$$\begin{aligned}
p\{700 \leq X \leq 750\} &\approx p\{699.5 \leq X_N \leq 750.5\} \\
&= p\left\{\frac{699.5-720}{\sqrt{720}} \leq Z = N(0,1) \leq \frac{750.5-720}{\sqrt{720}}\right\} \\
&= p\{-0.764 \leq Z \leq 1.136\} = F_Z(1.136) - F_Z(-0.764) \\
&= F_Z(1.136) - 1 + F_Z(0.764) \approx 0.8708 - 1 + 0.7764 = 0.6472.
\end{aligned}$$

6.4.- Suposem que la probabilitat que té un jugador de bàsquet d'encertar un tir lliure és una constant p (desconeguda) i que els seus tirs són independents. Quants de tirs ha d'intentar si es vol que la proporció de bàsquets estigui com a màxim a distància 0.1 de p amb probabilitat mínima de 0.8?

Resolució. Considerem la variable aleatòria X_i que val 1 si encerta en l'i-èssim intent i 0 en cas contrari. Tenim per tant que X_i és una variable aleatòria de Bernoulli amb paràmetre p.

Considerem la variable aleatòria $X = \sum_{i=1}^{n} X_i$ que ens dóna els nombre de bàsquets en n intents. Fixau-vos que X és una variable aleatòria binomial amb paràmetres n i p.

La proporció de bàsquets la donarà la variable $Y = \frac{X}{n}$. Per tant, ens demanen trobar n (nombre d'intents) tal que:

$$p\left\{|Y-p| \leq 0.1\right\} \geq 0.8$$

En general, si volem trobar n tal que:

$$p\left\{|Y-p| \leq \delta\right\} \geq \gamma,$$

aproximam la variable Y per una normal aplicant el teorema del límit central i la probabilitat anterior es pot aproximar per:

$$\begin{aligned}
p\left\{|Y-p| \leq \delta\right\} &= p\left\{p-\delta \leq Y \leq p+\delta\right\} = p\left\{n(p-\delta) \leq X \leq n(p+\delta)\right\} \\
&\approx p\left\{n(p-\delta) - 0.5 \leq X_N = N(np, np(1-p) \leq (p+\delta)n + 0.5\right\} \\
&= p\left\{\frac{-n\delta - 0.5}{\sqrt{npq}} \leq Z = N(0,1) \leq \frac{n\delta + 0.5}{\sqrt{npq}}\right\} \\
&= F_Z\left(\frac{n\delta + 0.5}{\sqrt{npq}}\right) - F_Z\left(\frac{-n\delta - 0.5}{\sqrt{npq}}\right) = 2F_Z\left(\frac{n\delta + 0.5}{\sqrt{npq}}\right) - 1,
\end{aligned}$$

on $q = 1-p$.

Per tant, volem trobar n tal que:

$$2F_Z\left(\frac{n\delta + 0.5}{\sqrt{npq}}\right) - 1 \geq \gamma,$$

d'on deduïm que:

$$\frac{n\delta + 0.5}{\sqrt{npq}} \geq z_{\frac{\gamma+1}{2}}, \tag{6.4}$$

on $z_{\frac{\gamma+1}{2}}$ és el percentil $\frac{\gamma+1}{2}$% de la variable $N(0,1)$.

Fent càlculs en la condició (6.4), tenim que n ha de complir:

$$n\delta + 0.5 \geq z_{\frac{\gamma+1}{2}} \sqrt{npq} \tag{6.5}$$

Per simplificar càlculs, prendrem n tal que

$$n\delta \geq z_{\frac{\gamma+1}{2}} \sqrt{npq} \tag{6.6}$$

Fixau-vos que si n compleix (6.6) també complirà (6.5). Finalment, simplificant (6.6) tenim que n ha de complir:

$$n \geq \frac{{z_{\frac{\gamma+1}{2}}}^2 pq}{\delta^2} \tag{6.7}$$

Fent servir el problema 6.2, podem dir que $pq \leq \frac{1}{4}, \ \forall p \in [0,1]$. Perquè la condició (6.7) no depengui de p (paràmetre desconegut) prendrem n que compleixi:

$$n \geq \frac{{z_{\frac{\gamma+1}{2}}}^2}{4\delta^2} \tag{6.8}$$

En el nostre cas $\gamma = 0.8$ i $\delta = 0.1$ Mirant a les taules, resulta $z_{\frac{\gamma+1}{2}} = z_{0.9} = 1.28$. Així doncs, n complirà:

$$n \geq \frac{1.28^2}{4(0.1)^2} \approx 41.$$

S'hauran d'intentar com a mínim 41 tirs per tenir la seguretat que la proporció de bàsquets estigui com a màxim a distància 0.1 de p amb probabilitat mínima de 0.8.

6.5.- Sigui X una variable aleatòria tal que $\mathrm{E}X = 0$ i $\mathrm{Var}X = 1$. Trobau el tant per cent mínim de valors de X que estan en l'interval $(-2,2)$?
Final. Juny 1994.

Resolució. Ens demanen la quantitat següent:

$$p\{-2 < X < 2\} \times 100.$$

Aplicant la desigualtat de Txebixef, tenim:

$$p\{|X - \mu_X| < 2\} = p\{|X| < 2\sigma_X\} = p\{-2 < X < 2\} \geq 1 - \frac{1}{4} = 0.75$$

Hi ha com a mínim un 75% de valors de la variable X en l'interval $(-2,2)$.

6.6.- Suposem que el 10% dels votants d'una certa comunitat està a favor d'una certa legislació. Es fa un enquesta entre la població i s'obté una freqüència relativa $f_A(n)$ com a estimació de la proporció anterior (n és el nombre d'enquestats). Determinau, aplicant la desigualtat de Txebixef, quants de votants s'haurien d'enquestar per tal que la probabilitat que $f_A(n)$ difereixi de 0.1 menys de 0.02 sigui d'almenys 0.95. Què podem dir si no coneixem la proporció 0.1? (situació més real)

Resolució. Considerem les variables aleatòries:

$$I_i = \begin{cases} 1, & \text{si la persona } i \text{ està a favor de la lesgislació,} \\ 0, & \text{en cas contrari,} \end{cases}$$

$\forall i = 1, \ldots, n$. Aleshores les I_i són variables aleatòries i.i.d. amb distribució comuna $B(1, p)$.

La freqüència relativa serà:

$$f_A(n) = \frac{\sum_{i=1}^{n} I_i}{n}.$$

Aleshores:

$$\mathrm{E}(f_A(n)) = \frac{\sum_{i=1}^{n} \mathrm{E} I_i}{n} = \frac{n \cdot p}{n} = p = 0.1,$$

i

$$\mathrm{Var} f_A(n) = \frac{\sum_{i=1}^{n} \mathrm{Var} I_i}{n^2} = \frac{npq}{n^2} = \frac{pq}{n} = \frac{0.09}{n}.$$

Aleshores

$$p\{|f_A(n) - 0.1| < 0.02\} = 1 - p\{|f_A(n) - 0.1| \geq 0.02\} \geq 1 - \frac{\frac{0.09}{n}}{0.02^2} = 0.95,$$

i d'aquí obtenim n:

$$\frac{0.09}{n \cdot 0.02^2} = 0.05 \Longrightarrow n = \frac{0.09}{0.05 \times 0.02^2} = 4500.$$

Si no coneixem $p = \mathrm{E}(f_A(n))$, tampoc coneixerem $\mathrm{Var} f_A(n) = \frac{pq}{n}$, però de totes formes sabem que $p \cdot q \leq 1/4$, i aleshores

$$p\{|f_A(n) - 0.1| < 0.02\} \geq 1 - \frac{pq}{n \times 0.02^2} \geq 1 - \frac{1}{4n \times 0.02^2} = 0.95,$$

d'on obtenim

$$\frac{1}{4n \times 0.02^2} = 0.05 \Longrightarrow n = \frac{1}{4 \times 0.05 \times 0.02^2} = 12500.$$

6.7.- Es llança a l'aire un dau regular 100 vegades. Aplicau la desigualtat de Txebixef per tal d'obtenir una cota de la probabilitat que el nombre total de punts obtenguts estigui entre 300 i 400. Quina probabilitat s'obté aplicant el teorema del límit central?

Resolució. Considerem les variables aleatòries:

X_i : resultat en el llançament i-èssim, $i = 1, \ldots, 100$.

Tenim:

$$\mathrm{E}X_i = \sum_{k=1}^{6} k \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2},$$

i

$$\mathrm{Var}X_i = \mathrm{E}X_i^2 - (\mathrm{E}X_i)^2 = \sum_{k=1}^{6} k^2 \cdot \frac{1}{6} - \left(\frac{21}{6}\right)^2 = \frac{91}{6} - \frac{441}{36} = \frac{35}{12}.$$

El nombre total de punts obtenguts en els 100 llançaments serà:

$$S_{100} = \sum_{i=1}^{100} X_i.$$

Aleshores:

$$\mathrm{E}S_{100} = \frac{2100}{6} = 350, \quad \mathrm{Var}S_{100} = \frac{3500}{12} = \frac{875}{3}.$$

Per tant

$$\begin{aligned} p\{300 < S_{100} < 400\} &= p\{|S_{100} - \mathrm{E}S_{100}| < 50\} = 1 - p\{|S_{100} - \mathrm{E}S_{100}| \geq 50\} \\ &\geq 1 - \frac{\frac{875}{3}}{50^2} = \frac{53}{60} = 0.883. \end{aligned}$$

Si aplicam el teorema del límit central, $\dfrac{S_{100} - \mathrm{E}S_{100}}{\sqrt{\mathrm{Var}S_{100}}}$ es pot aproximar per una $N(0,1)$. Aleshores,

$$p\{300 < S_{100} < 400\} \simeq p\left\{\frac{299.5 - 350}{\sqrt{875/3}} < \frac{S_{100} - \mathrm{E}S_{100}}{\sqrt{\mathrm{Var}S_{100}}} < \frac{400.5 - 350}{\sqrt{875/3}}\right\}$$

$$\begin{aligned} &\simeq p\{-2.95 < N(0,1) < 2.95\} = 2 \times p\{N(0,1) < 2.95\} - 1 \\ &= 2 \times 0.998411 - 1 = 0.99682. \end{aligned}$$

6.8.- El nombre d'errors d'impremta per pàgina d'un llibre segueix una distribució de Poisson, amb un nombre mitjà d'errors per pàgina igual a 2. En un llibre de 300 pàgines, quina és la probabilitat que en una o més pàgines hi hagi més de 5 errors? Calculau-la directament, aproximant per una Poisson i aproximant per una normal.

Resolució. Posem:

X : nombre d'errors per pàgina.

Aleshores X segueix una distribució $Poiss(2)$. Per tant

$$p\{X > 5\} = 1 - p\{X \leq 5\} = 1 - 0.9834 = 0.0166.$$

Si ara consideram:

Y : nombre de pàgines amb més de 5 errors,

aleshores Y segueix una distribució binomial amb paràmetres $n = 300$ i $p = 0.0166$. Per tant

$$\begin{aligned} p\{Y \geq 1\} &= 1 - p\{Y < 1\} = 1 - p\{Y = 0\} \\ &= 1 - \binom{300}{0} \times 0.0166^0 \times 0.9834^{300} = 1 - 0.0066 = 0.9934. \end{aligned}$$

També es pot calcular aproximant la $B(n,p)$ per una $Poiss(np)$. En el nostre cas,

$$\lambda = n \cdot p = 300 \times 0.0166 = 4.98.$$

Aleshores

$$p\{Y \geq 1\} = 1 - p\{Y = 0\} = 1 - \frac{4.98^0}{0!}\,\mathrm{e}^{-4.98} = 1 - \mathrm{e}^{-4.98} = 0.9931.$$

Si aplicam el teorema de De Moivre-Laplace,

$$\begin{aligned}
p\{Y \geq 1\} &= 1 - p\{Y = 0\} \simeq 1 - p\{-0.5 < Y \leq 0.5\} \\
&\simeq 1 - p\left\{\frac{-0.5 - np}{\sqrt{npq}} < Z \leq \frac{0.5 - np}{\sqrt{npq}}\right\} \\
&= 1 - p\left\{Z < \frac{0.5 - 4.98}{\sqrt{4.897}}\right\} + p\left\{Z < \frac{-0.5 - 4.98}{\sqrt{4.897}}\right\} \\
&= 1 - p\{Z < -2.02\} + p\{Z < -2.47\} = p\{Z < 2.02\} + p\{Z < -2.47\} \\
&= 0.97831 + 0.006756 = 0.985066,
\end{aligned}$$

on Z té la distribució $N(0,1)$.

6.9.- S'ha de computar la suma de 100 nombres reals. Suposem que els nombres s'han aproximat per l'enter més pròxim de manera que cada nombre té un error uniformement distribuït en $(-1/2, 1/2)$. Utilitzau el teorema de límit central per estimar la probabilitat que l'error total en la suma dels 100 nombres superi 6.

Resolució. Sigui

$$X_i : \text{error total del nombre } i\text{-èssim}.$$

Aleshores X_i té una distribució uniforme en l'interval $(-1/2, 1/2)$ i, per tant, $\mathrm{E}X_i = 0$, $\mathrm{Var}X_i = \frac{1}{12}$. L'error total en la suma dels 100 nombres el donarà

$$S_{100} = X_1 + \cdots + X_{100}.$$

Aleshores $\mathrm{E}S_{100} = 0$ i $\mathrm{Var}S_{100} = \frac{100}{12}$. Així resultarà

$$\begin{aligned}
p\{S_{100} > 6\} &= 1 - p\{S_{100} \leq 6\} = 1 - p\left\{Z_{100} \leq \frac{6}{\sqrt{\frac{100}{12}}}\right\} \\
&= 1 - p\{Z_{100} \leq 2.08\} \simeq 1 - p\{Z \leq 2.08\} = 1 - 0.98124 = 0.01876,
\end{aligned}$$

on Z té la distribució $N(0,1)$.

6.10.- La mitjana de bolígrafs que es venen diàriament en una papereria és 30 i la desviació típica, 5. Aquests valors són 20 i 4 per al nombre de quaderns venuts. Se sap, a més, que el coeficient de correlació entre les vendes d'ambdós productes és 0.7. Quina és la probabilitat que el nombre total dels dos articles venuts durant un trimestre estigui comprès entre 4.300 i 4.600 unitats?

Resolució. Posem:

X : nombre de bolígrafs venuts en un dia,
Y : nombre de quaderns venuts en un dia,
D : nombre total d'aquests dos articles venuts en un dia $= X + Y$.

Sabem que $\mathrm{E}X = 30, \mathrm{E}Y = 20, \mathrm{Var}X = 5^2 = 25, \mathrm{Var}Y = 4^2 = 16,$ i $\rho_{XY} = 0.7$. Aleshores $\mathrm{E}D = 50$, i

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y} \quad \Longrightarrow 0.7 = \frac{\mathrm{Cov}(X,Y)}{5 \times 4} \quad \Longrightarrow \mathrm{Cov}(X,Y) = 14.$$

Aleshores $\mathrm{Var}D = \mathrm{Var}X + \mathrm{Var}Y + 2\,\mathrm{Cov}(X,Y) = 25 + 16 + 28 = 69$. El nombre total dels dos articles venuts cada trimestre el dóna

$$S_{90} = D_1 + \cdots + D_{90},$$

on cada $D_i (i = 1, \ldots, 90)$ té la mateixa distribució que D. Aleshores $\mathrm{E}S_{90} = 90 \times \mathrm{E}D = 4500$ i $\mathrm{Var}S_{90} = 90 \times \mathrm{Var}D = 6210$, d'on $\sigma_{S_{90}} = 78.8$. Finalment,

$$\begin{aligned}
p\{4300 < S_{90} \le 4600\} &= p\{S_{90} \le 4600\} - p\{S_{90} \le 4300\} \\
&\simeq p\{S_{90} \le 4600.5\} - p\{S_{90} \le 4300.5\} \\
&= p\left\{Z_{90} \le \frac{4600.5 - 4500}{78.8}\right\} - p\left\{Z_{90} \le \frac{4300.5 - 4500}{78.8}\right\} \\
&= p\{Z_{90} \le 1.28\} - p\{Z_{90} \le -2.53\} \\
&\simeq p\{Z \le 1.28\} - p\{Z \le -2.53\} \\
&= 0.8997 - 0.0057 = 0.894,
\end{aligned}$$

on Z té la distribució $N(0,1)$.

6.3 Problemes proposats

6.1.- Suposem que $X_1, X_2, X_3, \ldots,$ són variables aleatòries independents distribuïdes idènticament, cada una amb funció de densitat:

$$f_X(x) = \begin{cases} \frac{1}{2}, & \text{per a } x = -1, 1, \\ 0, & \text{per als altres casos.} \end{cases}$$

Definim la seqüència de les mitjanes:

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Demostrau que la funció de densitat per a $\overline{X}_n$ és:

$$f_{\overline{X}_n}(x) = \begin{cases} \frac{\binom{n}{\frac{n}{2}(x+1)}}{2^n}, & \text{per a } x = \pm\frac{1}{n}, \pm\frac{3}{n}, \ldots, \pm\frac{n-2}{n}, \pm 1, \\ \\ 0, & \text{en els altres casos.} \end{cases}$$

(Suposau que n és un nombre senar.)

6.2.- El nombre d'accidents en un tros de 10 Km de carretera de 2 carrils és una variable aleatòria de Poisson amb mitjana de 2 accidents per setmana. Quina és la probabilitat (aproximada) que hi hagi menys de 100 accidents en aquest tros durant 1 any?

6.3.- La llargada que es pot estirar un fil de niló és una variable aleatòria exponencial amb mitjana de 1524 metres. Quina és la probabilitat (aproximada) que la llargada mitjana de 100 fils estigui entre 1447.8 m i 1691.64 m?

6.4.- Les telefonades que es reben en un commutador d'una indústria arriben com a successos de Poisson a raó de 120 per hora. Quina és la probabilitat que arribin entre 110 i 125 telefonades entre les 9.00 i les 10.00 del matí de qualsevol dia?

6.5.- Considerem la variable aleatòria $X =$"Nombre de cotxes arribats a un supermercat en un determinat mes". Suposem que l'arribada de cotxes al supermercat és un procés de Poisson i que per terme mitjà arriben 5.000 cotxes per mes. Aplicant el teorema del límit central, calculau la probabilitat p que arribin entre 4.900

i 5.200 cotxes, ambdós inclosos, al supermercat en un mes escollit a l'atzar.
Final. Setembre de 1994.

6.6.- La probabilitat que un jugador de bàsquet encistelli és p.

Quants de llançaments ha de fer com a mínim (aproximadament) per tal que la probabilitat que la mitjana de bàsquets estigui a distància 0.01 de p sigui de 0.99?
Primer parcial. Febrer 95.

6.7.- Sigui $X_1, \ldots, X_n$ amb $n = 48$, una mostra aleatòria simple d'una variable aleatòria uniforme en l'interval $(0, a)$. Aplicant el teorema del límit central, trobau la probabilitat aproximada $p\left\{\sum_{i=1}^{n} X_i > a\right\}$.
Final. Juny 96.

6.8.- Un radiofar està alimentat per una bateria amb un temps de vida útil T governat per una distribució exponencial amb una esperança d'un mes. Trobau el nombre mínim de bateries que s'han de subministrar al radiofar per tal que sigui operatiu amb probabilitat 0.99 almenys un any.

6.9.- Es llança l'aire una moneda sense biaix n vegades. Calculau n de manera que la freqüència relativa del '6' disti menys de 0.01 de la probabilitat teòrica 1/6?

6.10.- Se sap que en una població, la talla dels individus barons adults és una variable aleatòria X amb mitjana $\mu_X = 170$ cm i desviació típica $\sigma_X = 7$ cm. S'elegeixen 140 individus a l'atzar en condicions independents. Calculau la probabilitat que la mitjana mostral $\bar{x}$ difereixi de μ_X menys de 1 cm.

Bibliografia

[AM96] ALBERICH, R.; MIR, A. *Introducció a l'Estadística Descriptiva.* Publicacions UIB, 1996.

[Can92] CANAVOS, G.C. *Probabilidad y Estadística. Aplicaciones y Métodos.* McGraw-Hill, 1992.

[GS88] GRIMMETT, G.; STIRZAKER, D. *Probability and Random Processes.* Oxford Science Publications, 1988.

[iS94] SANZ-SOLÉ, M. *Lliçons de Càlcul de Probabilitats.* Publicacions Universitat de Barcelona, 1994.

[Lar90] LARSON, H.J. *Introducción a la Teoria de las Probabilidades e Inferencia Estadística.* Limusa, 1990.

[LG94] LEON-GARCIA, A. *Probability and Random Processes for Electrical Engineering.* Addison-Wesley, 1994.

[Lli88] BARÓ-LLINÀS, J. *Estadística Descriptiva.* Parramón, 1988.

[Mey92] MEYER, P.L. *Probabilidad y Aplicaciones Estadística.* Addison-Wesley Iberoamericana, 1992.

[MR96] MONTGOMERY, D.C.; RUNGER, G.C. *Probabilidad y Estadística aplicadas a la Ingeniería.* McGraw-Hill, 1996.

[QIL90] QUESADA, V.; ISIDORO, A.; LÓPEZ, L.A. *Curso y Ejercicios de Estadística.* Alhambra Universidad, 1990.

Índex alfabètic