

Matemàtiques II
1r curs de Grau en Química

David Zorío Ventura

Índex

1	Teoria del mostreig estadístic	5
1.1	Estadística descriptiva	6
1.1.1	Mesures de centralització	9
1.1.2	Mesures de dispersió	10
1.2	Teoria de la probabilitat	12
1.2.1	Introducció al concepte de probabilitat	15
1.2.2	Variables aleatòries	24
1.2.3	Distribucions	34
1.3	Inferència estadística	48
1.3.1	Intervals de confiança	50
1.3.2	Contrast d'hipòtesis	56
2	Ajustament i regressió	73
2.1	Ajustament lineal	75
2.1.1	Intervals de confiança	79
2.1.2	Prediccions	82
2.2	Altres ajustaments	82
2.2.1	Ajustament parabòlic	83
2.2.2	Ajustament hiperbòlic	84
2.2.3	Ajustament potencial	86
2.2.4	Ajustament exponencial	88
3	Interpolació polinòmica	91
3.1	Error d'interpolació	94
3.2	Fórmules d'interpolació	95
3.2.1	Fórmula de Lagrange	96
3.2.2	Fórmula de Newton	97
3.2.3	Fórmula de les diferències dividides	100
3.2.4	Mètodes d'Aitken i Neville	102
3.3	Interpolació a trossos	104
3.3.1	Splines cúbics	106
4	Integració numèrica	109
4.1	Fórmules de Newton-Cotes	110
4.1.1	Fórmula dels trapezis	112
4.1.2	Fórmula de Simpson 1/3	116

4.1.3	Fórmula de Simpson 3/8	118
4.2	Integració numèrica d'equacions diferencials	120
4.2.1	Mètode d'Euler	121
4.2.2	Mètode d'Euler millorat	123
4.2.3	Mètode de Runge-Kutta de quart ordre	124

Capítol 1

Teoria del mostreig estadístic

L'estadística és la ciència, amb base matemàtica, que s'encarrega de la recollecció, ordenació, representació, anàlisi i interpretació de dades.

En aquest capítol es proporcionen un seguit de ferramentes d'anàlisi estadístic de dades. Es divideix en tres seccions principals, la descripció de les quals es resumeix a continuació:

- **Estadística descriptiva:** Recollecció, ordenació, representació i anàlisi d'un conjunt de dades.
- **Teoria de la probabilitat:** Estudi de la freqüència d'un esdeveniment aleatori mitjançant un experiment aleatori.
- **Inferència estadística:** Obtenció de propietats d'una població estadística a partir d'una mostra d'aquesta.

1.1 Estadística descriptiva

En aquesta secció es presenten una sèrie de tècniques elementals per al tractament estadístic de dades. Comencem amb una sèrie de definicions bàsiques amb les quals treballarem habitualment.

- **Població:** Conjunt d'individus que es pretén estudiar.
- **Variable:** Característica de la població que pren un valor per a cadascun dels individus d'aquesta. Les variables poden ser de dos tipus:
 - **Numèriques:** Aquelles el valor de les quals ve representat per nombres.
 - **Categòriques:** Aquelles que no prenen valors numèrics.
- **Dades:** Valors que pren una variable.
- **Mostra:** Subconjunt de la població de què es disposa de dades. S'anomena **grandària** de la mostra al seu nombre d'elements.

Exemple Se seleccionen 15 persones a l'atzar dins d'un campus universitari i es registra el gènere i l'alçada (en cm) de cadascun d'ells.

Individu	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Gènere	H	D	D	H	D	H	H	D	D	H	H	D	H	D	D
Alçada	172	165	168	167	160	177	184	157	160	177	172	177	180	165	170

En aquest cas, la població és el conjunt d'estudiants del campus, on les variables que s'estudien són el gènere (categòrica) i l'alçada (numèrica) a partir d'una mostra de grandària 15.

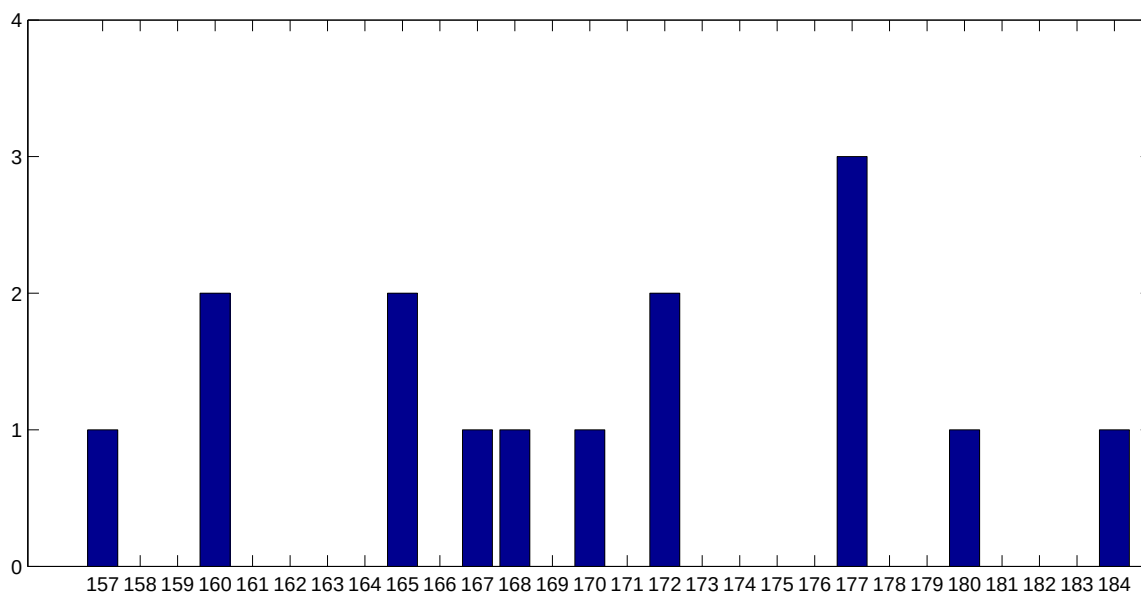
Una vegada obtingudes, les dades poden ser ordenades i representades de diferents maneres per a una millor comprensió de la informació que proporcionen. Seguint amb l'exemple anterior, les dades ordenades (de menor a major) quedarien de la següent manera:

157	160	160	165	165	167	168	170	172	172	177	177	177	180	184
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Per tal d'evitar escriure valors repetits, podem agrupar les dades per **freqüència**, indicant el nombre de vegades que apareix cada valor:

Valor	157	160	165	167	168	170	172	177	180	184
Freqüència	1	2	2	1	1	1	2	3	1	1

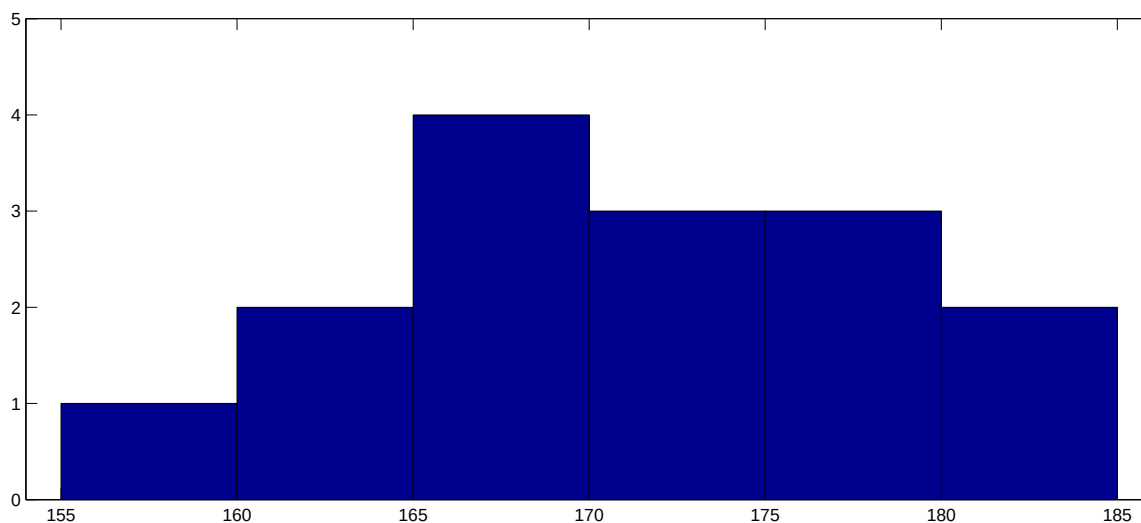
Una manera de representar gràficament els valors de la mostra és mitjançant un **histograma**, que és un diagrama de barres on l'abscissa representa el valor de la mostra i l'ordenada la seua freqüència. Per a l'exemple que s'està analitzant, es mostra el corresponent histograma a la figura següent:



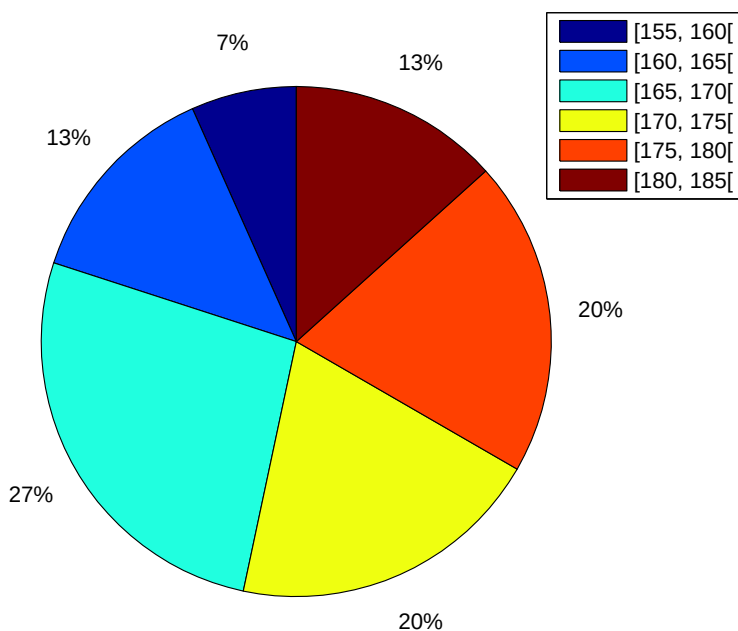
També podem ordenar les dades per **classes**, amb rangs de valors i la freqüència assignada, que en aquest cas representa el nombre de valors que estan dins del rang establert. Ordenem les dades de l'exemple en classes amb un rang de cinc unitats:

Rang	[155, 160[	[160, 165[	[165, 170[	[170, 175[	[175, 180[	[180, 185[
Freqüència	1	2	4	3	3	2

En aquest cas, l'histograma queda de la següent manera:



Una altra possible manera de representar les dades és mitjançant un **diagrama de sectors**, que consisteix a representar sobre una circumferència els valors o classes amb les seues respectives proporcions d'aparició en forma de sectors d'aquesta. En el cas de l'exemple anterior amb l'ordenació per classes prèvia, tenim:

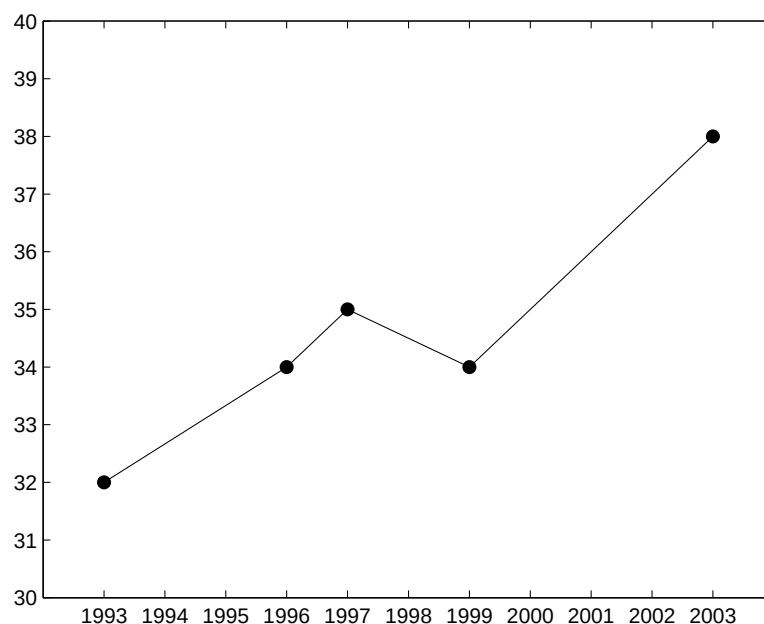


Finalment, quan les dades obtingudes depenen d'una evolució temporal, aquestes es poden representar mitjançant un **gràfic temporal**, on el temps ve indicat a les abscisses i les dades a les ordenades, com mostra el següent exemple.

Exemple La següent taula recull dades sobre el nombre d'accidents de trànsit anuals a una determinada carretera nacional.

Any	1993	1996	1997	1999	2003
Accidents	32	34	35	34	38

El corresponent gràfic temporal és el següent:



1.1.1 Mesures de centralització

A continuació es proporcionen un seguit d'eines elementals que permeten descriure els trets més bàsics d'una mostra estadística. D'ara endavant treballarem tant amb mostres genèriques $\{x_i\}_{i=1}^n$ com amb mostres amb valors agrupats per freqüències $\{x_i\}_{i=1}^r$, amb $\{n_i\}_{i=1}^r$ les freqüències de repetició del valor corresponent, indicant en tot moment amb quina de les dues modalitats es treballa.

- **Mitjana:** Es defineix la mitjana $\bar{x}$ d'una mostra $\{x_i\}_{i=1}^n$ com es segueix:

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

Per a mostres amb valors agrupats per freqüència, aquesta definició es pot reescriure com

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

- **Mediana:** Es defineix la mediana M d'una mostra ordenada de la següent manera, segons la grandària de la mostra n :

- Si n és senar, aleshores la mediana és el valor central de la mostra, ço és:

$$M = x_{\frac{n+1}{2}}.$$

- Si n és parell, aleshores la mediana és la mitjana entre els dos valors que queda enmig de la mostra:

$$M = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

- **Moda:** La moda m és el valor de la mostra que més vegades es repeteix (és a dir, aquell amb major freqüència). En cas d'haver més d'un valor amb la màxima freqüència, la moda és el conjunt format per tots ells.

Exemple Considerem la següent mostra:

$$\{2.3, 2.5, 2.4, 2.4, 2.8, 2.3, 2.6, 2.4, 2.2, 2.6\}.$$

Aleshores la seua mitjana és

$$\bar{x} = \frac{2.3 + 2.5 + 2.4 + 2.4 + 2.8 + 2.3 + 2.6 + 2.4 + 2.2 + 2.6}{10} = 2.45.$$

Alternativament, si agrupem les dades segons les seues freqüències:

i	1	2	3	4	5	6
x_i	2.2	2.3	2.4	2.5	2.6	2.8
n_i	1	2	3	1	2	1

$$\bar{x} = \frac{2.2 + 2 \cdot 2.3 + 3 \cdot 2.4 + 2.5 + 2 \cdot 2.6 + 2.8}{10} = 2.45.$$

D'altra banda, la mostra ordenada queda de la següent manera:

$$2.2, 2.3, 2.3, 2.4, \mathbf{2.4}, \mathbf{2.4}, 2.5, 2.6, 2.6, 2.8.$$

Com que n és parell, la mediana és la mitja entre els valors de la cinquena i la sisena posició (en negreta).

$$M = \frac{2.4 + 2.4}{2} = 2.4$$

Donat que 2.4 és també al mateix temps el valor amb major freqüència (3), la moda és

$$m = 2.4.$$

1.1.2 Mesures de dispersió

Com a complement de les mesures de centralització per a obtindre informació a partir de la mostra, tenim les mesures de dispersió, que proporcionen el grau de disparitat que hi ha a les dades.

- **Variància:** Definim la variància s^2 d'una mostra com el valor:

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

Alternativament, si tenim la mostra amb valors agrupats per freqüències:

$$s^2 = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{x})^2.$$

- **Desviació típica:** La desviació típica s és l'arrel quadrada positiva de la variància, és a dir:

$$s = +\sqrt{s^2}.$$

- **Rang:** El rang d'una mostra és la diferència entre el seu valor màxim i mínim.
 - **Quartils:** El primer quartil Q_1 d'una mostra és el valor que deixa a la seua esquerra la quarta part dels valors d'aquesta, el segon quartil Q_2 el que deixa la meitat d'ells (per tant, $Q_2 = M$) i el tercer quartil Q_3 és el que deixa a la seua esquerra les tres quartes parts dels valors de la mostra.
- La manera de calcular els quartils d'una mostra ordenada és la següent:

- Si la grandària n de la mostra és un nombre parell, aquesta es pot dividir en dues submostres de la mateixa grandària, $\frac{n}{2}$. El primer quartil Q_1 és la mediana de la submostra que queda a l'esquerra i el segon quartil Q_3 la mediana de la que queda a la dreta.
- Si la grandària n de la mostra és un nombre senar, aleshores s'agafa com a submostra esquerra tots els valors que queden per l'esquerra del valor central, $x_{\frac{n+1}{2}}$, (mediana de la mostra), sense incloure aquest, i com a submostra dreta tots els valors que queden per la dreta d'ell (sense incloure'l tampoc). Aleshores Q_1 i Q_3 es calculen com la mediana de les respectives submostres obtingudes.

Exercici 1 Demostra que la variància també es pot expressar de la següent manera:

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

o, alternativament, si els valors de la mostra estan agrupats per freqüències:

$$\frac{1}{n} \sum_{i=1}^r n_i x_i^2 - \bar{x}^2.$$

Exemple Treballem amb la mateixa mostra que a l'exemple anterior, on recordem que la mitjana obtinguda és $\bar{x} = 2.45$. Per tant, emprant la fórmula obtinguda a l'exercici anterior, la variància de la mostra resulta en aquest cas:

$$s^2 = \frac{(2.2 - 2.45)^2 + 2(2.3 - 2.45)^2 + 3(2.4 - 2.45)^2 + (2.5 - 2.45)^2 + 2(2.6 - 2.45)^2 + (2.8 - 2.45)^2}{10} = 0.0285.$$

Alternativament, emprant la fórmula obtinguda a l'exercici anterior, la variància també es pot calcular com:

$$s^2 = \frac{2.2^2 + 2 \cdot 2.3^2 + 3 \cdot 2.4^2 + 2.5^2 + 2 \cdot 2.6^2 + 2.8^2}{10} - 2.45^2 = 0.0285.$$

La desviació típica és per tant:

$$s = +\sqrt{s^2} = \sqrt{0.0285} \approx 0.1688.$$

El rang de la mostra és $2.8 - 2.2 = 0.6$.

Finalment, quant al càlcul dels quartils, com que la grandària de la mostra és 10, un nombre parell, la mostra es pot dividir en dues parts iguals.

Pel que fa a la part esquerra, tenim:

$$2.2, 2.3, \mathbf{2.3}, 2.4, 2.4.$$

On el tercer valor de la submostra, assenyalat en negreta, resulta ser el valor central d'aquesta, i per tant el valor del primer quartil és $Q_1 = 2.3$.

Quant a la part dreta, es té:

$$2.4, 2.5, \mathbf{2.6}, 2.6, 2.8.$$

I per tant el valor del tercer quartil és $Q_3 = 2.6$.

1.2 Teoria de la probabilitat

Per tal de fer un estudi més avançat o plantejar-se qüestions més profundes sobre una mostra donada i les seues propietats, resulta imprescindible fer servir els fonaments de la teoria de la probabilitat. En aquesta secció tractarem les qüestions més remarcables, necessàries per als estudis estadístics que abordarem més endavant.

A continuació es presenten els conceptes més elementals amb els quals treballarem al llarg d'aquesta secció.

Un **experiment** consisteix en la realització d'un conjunt d'accions, el qual produeix un **resultat** determinat i que pot repetir-se indefinidament.

Depenent de la seua naturalesa, poden de ser dos tipus:

- **Experiments deterministes:** Són aquells que sota les mateixes causes i condicions produeixen els mateixos efectes (per exemple, l'experiment consistent en avaluar la funció real de variable real $f(x) = x^2 - 1$ en el punt $x = 2$ sempre donarà com a resultat $f(2) = 3$).
- **Experiments aleatoris:** En les mateixes causes i condicions poden produir-se efectes diferents (per exemple, llançar una moneda a l'aire i observar el resultat).

La realització d'un seguit d'experiments produeix una **mostra**, corresponent al conjunt de resultats obtinguts. En el cas dels experiments aleatoris, rep el nom de **mostra aleatòria**.

Una **variable aleatòria (V.A.)** és aquella que quantifica els resultats d'un experiment aleatori, dins d'un conjunt de possibles resultats anomenat **espai mostral (E.M.)**, que denotarem d'ara endavant per Ω . Aquest pot ser de dos tipus:

- **Discret:** Els resultats possibles són una quantitat finita o infinita *numerable*¹ (indexable mitjançant els nombres naturals). Exemples:
 - Llançament d'una moneda a l'aire: Anomenem Y a la variable aleatòria, que en aquest cas pot prendre el valor $Y = C$ (cara) o $Y = X$ (creu), per la qual cosa l'espai mostral és $\Omega = \{C, X\}$.
 - Llançament de dues monedes a l'aire: $\Omega = \{CC, CX, XC, XX\}$.
 - Llançaments d'una moneda fins obtenir una cara: $\Omega = \{C, XC, XXC, XXXC, XXXXC, \dots\}$ (infinit numerable).
 - Resultat en llançar un dau: $\Omega = \{1, 2, 3, 4, 5, 6\}$
 - Nombre de llançaments d'un dau necessaris fins obtindre un 5: $\Omega = \{1, 2, 3, \dots\} = \mathbb{N}$ (infinit numerable).
- **Continu:** Els resultats possibles són una quantitat infinita *no numerable* (no indexables a través dels nombres naturals). Exemples:
 - Valor de l'expressió $x^2 + 1$ per a un valor de $x \in [0, 2]$ triat a l'atzar: $\Omega = [1, 5]$.
 - Temperatura en un punt de l'espai triat a l'atzar (en °C): $\Omega =] - 273, +\infty[$.
 - Temps d'execució d'un algoritme informàtic (en segons): $\Omega =]0, +\infty[$.
 - Tria d'un punt a l'atzar en el plànol real: $\Omega = \mathbb{R}^2$.
 - Tria d'un punt a l'atzar en el segon quadrant del plànol real: $\Omega =] - \infty, 0] \times [0, +\infty[$.

¹Més informació: https://ca.wikipedia.org/wiki/Conjunt_numerable

Dins d'un espai mostral, Ω , podríem estar interessats no només en estudiar un resultat per separat, sinó un conjunt específic d'ells. Per exemple, en el cas del llançament d'un dau, on l'espai mostral és $\Omega = \{1, 2, 3, 4, 5, 6\}$, ens podria interessar, en lloc d'un valor específic, si el resultat obtingut és parell. En eixe cas, el conjunt de resultats és $A = \{2, 4, 6\}$, que és un **subconjunt** de Ω (és a dir, tots els elements de A formen part de Ω), i es denota per $A \subseteq \Omega$.

Dins de tots els subconjunts d' Ω possibles trobem dos casos especials. D'una banda, parlarem del **conjunt buit** (format per zero elements), que denotarem per $\emptyset$, i, d'altra banda, també es pot considerar el conjunt total, Ω , com a subconjunt de si mateix (ja que tots els elements de Ω formen part d' Ω). Denotarem per $\mathcal{P}(\Omega)$ al conjunt format per tots els subconjunts possibles de Ω .

Exemple *Agafem una carta a l'atzar d'una baralla i ens fixem en el seu pal. Denotem per R el resultat "obtenir rombe", P "obtenir pica", T "obtenir trèvol" i C "obtenir cor". El nostre espai mostral és per tant $\Omega = \{R, P, T, C\}$. Aleshores el conjunt format per tots els seus subconjunts possibles és:*

$$\mathcal{P}(\Omega) = \{\emptyset, \{R\}, \{P\}, \{T\}, \{C\}, \{R, P\}, \{R, T\}, \{R, C\}, \{P, T\}, \{P, C\}, \{T, C\}, \{R, P, T\}, \{R, P, C\}, \{R, T, C\}, \{P, T, C\}, \Omega\}.$$

Exercici 2 *Troba el nombre de subconjunts de l'espai mostral Ω associat a l'experiment "seleccionar aleatòriament un nombre natural entre 1 i 3", és a dir, el nombre d'elements de $\mathcal{P}(\Omega)$. A partir de l'exemple anterior i el resultat obtingut en aquest exercici, pots induir quin és, en general, el nombre d'elements de $\mathcal{P}(\Omega)$ si Ω en té n ?*

A continuació es presenten les operacions bàsiques que es poden realitzar entre conjunts. Donats A , B dos subconjunts qualsevol, tenim:

- **Unió:** La unió entre A i B és el conjunt format per tots els elements de A i B . Es denota per $A \cup B$. La definició formal és $A \cup B = \{x \mid x \in A \vee x \in B\}$, on $\vee$ és l'operador lògic que representa la **disjunció**. L'expressió anterior es llig com: "els elements x tals que x pertany a A o x pertany a B ".

Exemple *Si $A = \{a, b, c\}$ i $B = \{b, c, d\}$ aleshores*

$$A \cup B = \{a, b, c, d\}.$$

- **Intersecció:** La intersecció entre A i B és el conjunt format per tots els elements que tenen en comú A i B . Es denota per $A \cap B$. La definició formal és $A \cap B = \{x \mid x \in A \wedge x \in B\}$, on $\wedge$ és l'operador lògic que representa la **conjunció**. L'expressió anterior es llig com: "els elements x tals que x pertany a A i x pertany a B ".

Exemple *Si $A = \{a, b, c\}$ i $B = \{b, c, d\}$ aleshores*

$$A \cap B = \{b, c\}.$$

*Si canviem l'enunciat de l'exemple i redefinim B com $\{d, e, f\}$ aleshores A i B no tenen elements en comú. En aquest cas la intersecció és **buida**, i per tant el conjunt resultant és el buit:*

$$A \cap B = \emptyset.$$

Direm que dos conjunts A, B són **disjunts** o **mútuament excloents** si $A \cap B = \emptyset$.

- **Diferència:** La diferència o **complementari** de B sobre A , denotat per $A \setminus B$, és el conjunt format per tots els elements de A que no pertanyen a B . Formalment, $A \setminus B = \{x \in A \mid x \notin B\}$, que es pot llegir com “els elements x del conjunt A que **no** pertanyen a B ”.

Exemple Si $A = \{a, b, c, d\}$ i $B = \{c, d, e\}$ aleshores tenim

$$A \setminus B = \{a, b\}.$$

A diferència dels casos anteriors, aquesta operació en general no és commutativa, ja que en aquest exemple fer l'operació en sentit invers resulta:

$$B \setminus A = \{e\}.$$

Exercici 3 Si A i B són dos conjunts tals que $A \setminus B = A$, què podem dir de $B \setminus A$? I de $A \cap B$?

D'ara endavant, si es treballa amb un cert espai mostral Ω , donat $A \subseteq \Omega$ un subconjunt d' Ω denotarem per $\bar{A}$ al complementari de A sobre Ω (o **contrari** de A), és a dir $\bar{A} = \Omega \setminus A$.

Propietats Siguen A , B i C tres conjunts qualsevol. Aleshores:

1. $\bar{\bar{A}} = A$.
2. $\overline{A \cup B} = \bar{A} \cap \bar{B}$.
3. $\overline{A \cap B} = \bar{A} \cup \bar{B}$.
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
5. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
6. Si $A \subseteq B$ i $B \subseteq A \longrightarrow A = B$.

Amb les ferramentes que hem introduït, podem presentar ara el concepte d'àlgebra, amb el qual es recolza la teoria de la probabilitat moderna.

Definició (Àlgebra) Direm que una col·lecció o família $\mathcal{A}$ de subconjunts d'un espai mostral Ω ($\mathcal{A} \subseteq \mathcal{P}(\Omega)$) és una **àlgebra** si es verifica:

1. $\emptyset \in \mathcal{A}$ (el conjunt buit forma part de la col·lecció $\mathcal{A}$).
2. Si $A \in \mathcal{A} \rightarrow \bar{A} \in \mathcal{A}$ (si un cert conjunt està en $\mathcal{A}$, aleshores el seu contrari també hi és).
3. Si $A, B \in \mathcal{A} \rightarrow A \cup B \in \mathcal{A}$ (donats dos conjunts de $\mathcal{A}$, la seua unió també és un element de $\mathcal{A}$).

Els elements de $\mathcal{A}$ s'anomenen **successos**. D'altra banda, $\emptyset$ rep el nom de **succés impossible** i Ω el de **succés segur**.

Des del punt de vista de l'espai mostral (que recordem que és el conjunt de tots els resultats possibles d'un experiment) es pot entendre el concepte de succés com un conjunt format per una sèrie de resultats específics de l'espai mostral, o com un subconjunt d'aquest.

Exemple Considerem l'espai mostral associat a l'experiment “llançar una moneda a l'aire”, $\Omega = \{C, X\}$. Aleshores:

- Si $\mathcal{A} = \{\{C\}, \{X\}, \Omega\}$ aleshores $\mathcal{A}$ no és una àlgebra, ja que $\emptyset \notin \mathcal{A}$.
- Si $\mathcal{A} = \{\emptyset, \{C\}, \Omega\}$ llavors $\mathcal{A}$ no és una àlgebra perquè $\overline{\{C\}} = \{X\} \notin \mathcal{A}$.
- Si $\mathcal{A} = \{\emptyset, \{C\}, \{X\}\}$ aleshores $\mathcal{A}$ tampoc és una àlgebra, ja que $\{C\} \cup \{X\} = \{C, X\} = \Omega \notin \mathcal{A}$.
- Per a qualsevol espai mostral Ω , les famílies $\mathcal{A} = \{\emptyset, \Omega\}$ i $\mathcal{A} = \mathcal{P}(\Omega)$ són àlgebres, en el primer cas la menor possible (**àlgebra trivial**) i en el segon la major possible (**àlgebra total**). En aquest exemple, tant $\mathcal{A} = \{\emptyset, \Omega\} = \{\emptyset, \{C, X\}\}$ com $\mathcal{A} = \mathcal{P}(\Omega) = \{\emptyset, \{C\}, \{X\}, \{C, X\}\}$ són exemples d'àlgebres.

De la definició d'àlgebra es dedueixen les següents propietats:

Propietats Siga $\mathcal{A}$ una àlgebra i $A, B \in \mathcal{A}$. Aleshores:

1. $\Omega \in \mathcal{A}$.
2. Si $A, B \in \mathcal{A} \longrightarrow A \cap B \in \mathcal{A}$.

Exercici 4 Demuestra les dues propietats anteriors.

Indicació: Per a la segona propietat, fes servir el primer i tercer apartat de les propietats de les operacions entre conjunts.

Exercici 5 Considerem l'espai mostral associat a l'experiment "llançar un dau", $\Omega = \{1, 2, 3, 4, 5, 6\}$ i els successos "obtenir parell", $P = \{2, 4, 6\}$, i "obtenir senar", $S = \{1, 3, 5\}$.

- (a) Demuestra que la família $\mathcal{A} = \{\emptyset, P, S, \Omega\}$ és una àlgebra.
- (b) Raona per què la família $\mathcal{A} = \{\emptyset, \{1\}, P, S, \Omega\}$ no és una àlgebra.
- (c) Troba un altre exemple d'una àlgebra sobre aquest espai mostral que no siga la trivial, la total ni la plantejada a l'apartat (a) i que tinga més de quatre elements.

En els casos en què treballarem al llarg d'aquest curs, donat un espai mostral Ω sempre treballarem amb l'àlgebra total, $\mathcal{A} = \mathcal{P}(\Omega)$. És a dir, treballarem amb tots els successos possibles derivats de l'espai mostral, o el que és el mateix, amb els conjunts formats per qualsevol combinació de resultats possible.

1.2.1 Introducció al concepte de probabilitat

L'estudi de la probabilitat de successos naix del desig de conèixer la freqüència d'ocurrència d'un succés determinat, principalment en l'àmbit dels jocs d'atzar. Segons es va anar desenvolupant, va anar estenent-se a uns usos més genèrics i diferents camps de les matemàtiques, fins evolucionar a la teoria de la probabilitat moderna, tot convertint-se en els fonaments de la inferència estadística.

La manera més directa d'estudiar la freqüència d'ocurrència d'un succés és mitjançant l'experimentació. Així doncs, donat un succés A , sotmés a estudi en un experiment, realitzem N repeticions d'aquest i observem el nombre de vegades que ocorre A , n_A . Aleshores es defineix la **freqüència relativa** del succés A com

$$\text{fr}(A) = \frac{n_A}{N}.$$

Notem que, no obstant això, aquesta noció depèn de cada experiment, és a dir, si es realitzen dos tandes d' N experiments el nombre d'ocurrències de A , n_A , no necessàriament ha de ser el mateix en els dos, i en conseqüència tampoc $\text{fr}(A)$; per tant, matemàticament parlant no és un concepte *ben definit*.

La utilitat d'aquest concepte és més bé com a aproximació empírica de la probabilitat real del succés A . És que, si denotem per $p(A)$ a la probabilitat que es doni el succés A , la *Llei de Regularitat Estadística* indica que la freqüència relativa d'un succés s'estabilitza quan el nombre de repeticions de l'experiment creix indefinidament, i ho fa al valor de la probabilitat, és a dir:

$$\text{fr}(A) \xrightarrow{N \rightarrow \infty} p(A).$$

No obstant això, a més de ser un fet empíric, l'expressió anterior és incongruent des del punt de vista formal, ja que, com s'ha indicat abans, $\text{fr}(A)$ no és un concepte ben definit, i per tant el límit anterior tampoc ho és. Cal cercar per tant noves maneres d'abordar l'estudi amb rigor total.

Definició (Fórmula de Laplace) Donat un succés A , si l'espai mostral està format per N resultats possibles, amb n_A casos favorables a l'ocurrència de A , i tots ells tenen la mateixa probabilitat (equiprobables), es pot definir la probabilitat del succés A com

$$p(A) = \frac{n_A}{N} = \frac{\text{nombre de casos favorables}}{\text{nombre de casos possibles}}.$$

Exemples Anem a veure alguns exemples on la fórmula de Laplace és aplicable:

1. Llancem una moneda a l'aire i observem el resultat. Es vol estudiar el succés

$$A = \{C\} = \{\text{obtindre cara}\}.$$

L'espai mostral en aquest cas és $\Omega = \{C, X\}$. Per tant, $n_A = 1$, que correspon al resultat "obtindre cara", mentre que $N = 2$ (els dos resultats possibles, "obtindre cara" o "obtindre creu"). Així doncs:

$$p(A) = \frac{n_A}{N} = \frac{1}{2} = 0.5.$$

2. Llancem un dau i observem el resultat. Es desitja conèixer la probabilitat del succés

$$A = \{\text{obtenir parell}\} = \{2, 4, 6\}.$$

Com que l'espai mostral és $\Omega = \{1, 2, 3, 4, 5, 6\}$, es té $N = 6$ i $n_A = 3$, per tant:

$$p(A) = \frac{n_A}{N} = \frac{3}{6} = \frac{1}{2} = 0.5.$$

3. Llancem dues monedes a l'aire i es vol conèixer la probabilitat del succés

$$A = \{\text{obtenir alguna cara}\}.$$

En aquest cas, es pot escriure l'espai mostral com $\Omega = \{CC, CX, XC, XX\}$. Amb aquesta nomenclatura, el succés A és específicament $A = \{CC, CX, XC\}$, és a dir, el conjunt de resultats en què apareix alguna cara. Així doncs, $N = 4$ i $n_A = 3$, i per tant:

$$p(A) = \frac{n_A}{N} = \frac{3}{4} = 0.75.$$

4. Se selecciona una carta d'una baralla de cartes a l'atzar (amb quatre pals i 13 cartes per pal, numerades de l'1 al 10 i les figures J,Q,K) i es vol saber la probabilitat del succés

$$A = \{\text{ser de piques o una figura}\}.$$

Per tant, l'espai mostrat està format per $N = 4 \cdot 13 = 52$ elements. Pel que fa als casos favorables, tenim els corresponents a l'extracció d'una pica, amb 13 possibilitats i les d'obtenir alguna de les tres figures dels tres pals restants (notem que les figures de piques estan incloses dins dels tretze casos comptats anteriorment); per tant $n_A = 13 + 3 \cdot 3 = 22$. Així doncs,

$$p(A) = \frac{n_A}{N} = \frac{22}{52} = \frac{11}{26} \approx 0.4231.$$

Exemple Considerem ara un dau trucat en què els successos associats a obtenir un valor específic ja no són equiprobables. Per exemple, es truca el dau de manera que:

$$p(\{1\}) = 0.3, \quad p(\{2\}) = 0.1, \quad p(\{3\}) = 0.3, \quad p(\{4\}) = 0.15, \quad p(\{5\}) = 0.1, \quad p(\{6\}) = 0.05.$$

Si ara considerem el succés $A = \{\text{obtenir cara}\}$, aleshores la definició de Laplace ja no és operativa, ja que els diferents resultats que constitueixen l'espai mostrat no són equiprobables.

Tot i així, intuïtivament es pot determinar que la probabilitat associada hauria de ser:

$$p(A) = p(\{2, 4, 6\}) = p(\{2\}) + p(\{4\}) + p(\{6\}) = 0.1 + 0.15 + 0.05 = 0.3.$$

Tot seguit introduïrem el concepte de probabilitat en la seua formulació axiomàtica actual, que permet abordar problemes en un context general, com en el cas de l'exemple anterior.

Definició (Probabilitat) Una probabilitat p sobre una àlgebra de successos $\mathcal{A}$ associada a un espai mostrat Ω és una aplicació

$$p : \mathcal{A} \longrightarrow [0, 1]$$

que verifica:

1. $p(\Omega) = 1$.
2. Si A i B són disjunts, aleshores $p(A \cup B) = p(A) + p(B)$.

La terna $(\Omega, \mathcal{A}, p)$ rep el nom d'**espai probabilístic**.

En altres paraules, una probabilitat és una correspondència que assigna successos (elements de l'àlgebra) a un valor numèric entre 0 i 1 (probabilitat assignada al succés), i que a més ha de verificar que la probabilitat del succés segur, Ω , siga 1, i d'altra banda, donats dos successos A i B sense resultats en comú, la probabilitat del succés generat per la seua unió (és a dir, que ocorrega A o B) és la suma de les probabilitats de cadascun d'ells.

Definició Un succés elemental d'una àlgebra $\mathcal{A}$ es defineix com aquell succés $A \in \mathcal{A}$ no buit ($A \neq \emptyset$) tal que si $B \in \mathcal{A}$ amb $B \subseteq A$, aleshores $B = \emptyset$ o $B = A$.

Per tant, un succés elemental és aquell que no conté cap altre succés (com a subconjunt), tret d'ell mateix o el buit.

Exemples

- Els successos elementals de l'espai mostral associat a l'experiment de llançar un dau,

$$\Omega = \{1, 2, 3, 4, 5, 6\},$$

sota l'àlgebra total, $\mathcal{A} = \mathcal{P}(\Omega)$, són: $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$.

- En general, per a un espai mostral qualsevol, Ω , si l'àlgebra associada és la total, $\mathcal{A} = \mathcal{P}(\Omega)$, aleshores els successos elementals són els conjunts formats per un únic element, és a dir, els successos que contenen un únic resultat. En termes formals, la família de successos elementals és:

$$\{\{x\} \mid x \in \Omega\}.$$

- Si canviem l'àlgebra del primer exemple per $\mathcal{A} = \{\emptyset, P, S, \Omega\}$, on $P = \{2, 4, 6\}$ i $S = \{1, 3, 5\}$, aleshores els successos elementals són en aquest cas P i S .

Teorema Tot succés $A \in \mathcal{A}$ és unió (disjunta) de successos elementals:

$$A = \bigcup_i A_i, \quad A_i \text{ succés elemental.}$$

A més, la probabilitat de A és

$$p(A) = \sum_i p(A_i).$$

Exemple Tornant a l'enunciat de l'exemple del dau trucat considerant l'àlgebra total, com que podem escriure $A = \{2, 4, 6\} = \{2\} \cup \{4\} \cup \{6\}$ (unions disjunes de successos elementals), es té

$$p(A) = p(\{2\}) + p(\{4\}) + p(\{6\}) = 0.1 + 0.15 + 0.05 = 0.3,$$

que es correspon a la idea intuïtiva que es tenia abans d'introduir els axiomes de la probabilitat.

A partir de la definició dels axiomes de la probabilitat es poden deduir les següents propietats:

Propietats

1. $p(\overline{A}) = 1 - p(A)$. En particular, $p(\emptyset) = 0$.
2. Si $A \subseteq B \longrightarrow p(A) \leq p(B)$.
3. Si A i B són successos qualsevol, aleshores $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

Exemple En un cert espai de probabilitat hi ha dos successos A i B complint les següents condicions:

- $p(A) = 0.3$.
- $p(\overline{B}) = 0.6$.
- $p(A \cap B) = 0.1$.

Es demana calcular $p(A \cup B)$.

Fent ús de les propietats anteriors, tenim $p(B) = 1 - p(\overline{B}) = 1 - 0.6 = 0.4$. D'altra banda,

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = 0.3 + 0.4 - 0.1 = 0.6.$$

Probabilitat condicionada

Donats dos successos A i B , hi ha vegades en què l'ocurrència de B altera o *condiciona* la probabilitat d'ocurrència de A si es té en compte la nova situació que genera el fet d'haver succeït B . El següent exemple il·lustra un cas en què es dona eixa situació.

Exemple

- *Suposem que en una rifa consistent en 50 nombres numerats de l'1 al 50, Joan ha comprat el número 35 i Pere el 36. Quina probabilitat té Joan de guanyar? I Pere?*

En aquest cas, el nostre espai mostral és el conjunt dels 50 primers nombres naturals, és a dir, $\Omega = \{1, 2, 3, \dots, 50\}$. Donat que en aquest cas tots els resultats són equiprobables, anomenant els successos $J = \{35\}$ i $P = \{36\}$, tant la probabilitat que guanye Joan com que guanye Pere és

$$p(J) = p(P) = \frac{1}{50}.$$

- *Suposem ara que es filtra que el resultat de la rifa acaba en 5. Quina probabilitat té ara Joan de guanyar? I Pere?*

Anomenem B al succés "eixir en nombre acabat en 5". Aquest succés restringeix els resultats possibles a 10:

$$\{5, 15, 25, 35, 45\},$$

que novament són equiprobables. Com que el nombre de Joan acaba en 5, la seua probabilitat de guanyar coneixent aquesta informació és $p(J) = \frac{1}{5}$, mentre que, per contra, el número de Pere no verifica les noves condicions, i per tant $p(P) = 0$.

Tot seguit veurem com formalitzar aquesta idea, que ens ajudarà a abordar problemes amb condicionament més complexos.

Definició (Probabilitat condicionada) *Siga B un succés amb probabilitat $p(B) \neq 0$ i A un altre succés amb probabilitat $p(A)$. Es defineix la **probabilitat de A condicionada a B** com*

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

Aquesta probabilitat es pot entendre com la probabilitat de A *sabent que s'ha donat B* .

Exemple *Tornant a l'exemple anterior, si assumim que el número del resultat de la rifa acaba en 5, aleshores se sap que s'ha donat el succés $B\{\text{nombre acabat en } 5\} = \{5, 15, 25, 35, 45\}$. Aleshores:*

$$p(B) = \frac{5}{50} = \frac{1}{10}.$$

Pel que fa al cas de Joan, cal calcular $J \cap B$, o el que és el mateix, el succés "guanya Joan i el resultat de la rifa acaba en 5", que és $J \cap B = \{35\} \cap \{5, 15, 25, 35, 45\} = \{35\}$, complint-se per tant que

$$p(J \cap B) = \frac{1}{50}.$$

Així doncs:

$$p(J|B) = \frac{p(J \cap B)}{p(B)} = \frac{\frac{1}{50}}{\frac{1}{10}} = \frac{1}{5}.$$

D'altra banda, quant al cas de Pere, es requereix el càlcul de $P \cap B$, és a dir, el succés "guanya Pere i el resultat de la rifa acaba en 5", essent $P \cap B = \{36\} \cap \{5, 15, 25, 35, 45\} = \emptyset$, que és el succés impossible, ja que efectivament és impossible que isca el número 36 i al mateix temps el resultat de la rifa acabe en 5, quan de fet acaba en 6. Per tant $p(P \cap B) = 0$ i es té:

$$p(P|B) = \frac{0}{\frac{1}{10}} = 0.$$

Propietats Siga $B \subseteq \Omega$ un succés tal que $p(B) \neq 0$ i $A \subseteq \Omega$ un succés qualsevol. Aleshores:

1. $p(\emptyset|B) = 0$.
2. Si $B \subseteq A \longrightarrow p(A|B) = 1$. En particular, $p(\Omega|B) = 1$.
3. $p(\bar{A}|B) = 1 - p(A|B)$.
4. Si $A_1, A_2 \subseteq \Omega \longrightarrow p(A_1 \cup A_2|B) = p(A_1|B) + p(A_2|B) - p(A_1 \cap A_2|B)$.
5. $p(A) = p(A|B)p(B) + p(A|\bar{B})p(\bar{B})$.

Definició (Independència) Donats dos successos A i B , direm que A és **independent** de B si

$$p(A|B) = p(A).$$

En cas contrari, es diu que A és dependent de B .

Nota Si A és independent de B , amb $p(A), p(B) > 0$ com que $p(A|B) = p(A)$, emprant la definició de probabilitat condicionada, es té que

$$p(A \cap B) = p(A)p(B).$$

Al mateix temps, de l'expressió anterior es pot deduir que $p(B|A) = p(B)$, i per tant B també és independent de A .

En resum, les següents afirmacions són equivalents:

1. A és independent de B ($p(A|B) = p(A)$).
2. B és independent de A ($p(B|A) = p(B)$).
3. $p(A \cap B) = p(A)p(B)$.

Així doncs, com que la independència sempre és mútua, parlarem de " A i B són independents" en lloc de " A és independent de B " o " B és independent de A ".

A més, es pot donar la següent definició equivalent d'independència entre successos:

$$A \text{ i } B \text{ són independents si } p(A \cap B) = p(A)p(B).$$

En general, si A i B són successos qualsevol (dependents o independents), la fórmula que es verifica és: $p(A \cap B) = p(A|B)p(B)$ o, alternativament, $p(A \cap B) = p(A)p(B|A)$.

En cas de tenir tres successos A_1, A_2, A_3 , podem escriure

$$\begin{aligned} p(A_1 \cap A_2 \cap A_3) &= p(A_3 \cap (A_1 \cap A_2)) = p(A_3|(A_1 \cap A_2))p(A_1 \cap A_2) = p(A_2 \cap A_1)p(A_3|(A_1 \cap A_2)) \\ &= p(A_2|A_1)p(A_1)p(A_3|(A_1 \cap A_2)) = p(A_1)p(A_2|A_1)p(A_3|(A_1 \cap A_2)). \end{aligned}$$

Aquesta propietat es pot generalitzar per a un nombre arbitrari de successos.

Propietat Donat Ω un espai mostral i $\{A_i\}_{i=1} \subseteq \Omega$ una col·lecció de successos, aleshores

$$p(\cap_i A_i) = \prod_{i=1} p(A_i | \cap_{j=1}^{i-1} A_j).$$

O el que és el mateix,

$$p(A_1 \cap A_2 \cap A_3 \cap A_4 \cap \dots) = p(A_1)p(A_2|A_1)p(A_3|(A_1 \cap A_2))p(A_4|(A_1 \cap A_2 \cap A_3)) \dots$$

És a dir, la probabilitat que es donen simultàniament els successos $A_1, A_2, A_3, A_4, \dots$ és el producte de la probabilitat que es done A_1 , per la probabilitat que es done A_2 sabent que s'ha donat A_1 , per la probabilitat que es done A_3 sabent que s'han donat A_1 i A_2 , per la probabilitat que es done A_4 sabent que s'han donat A_1, A_2 i A_3 , etc.

Exemple

- Una urna conté un total de 9 boles, 5 de les quals són blanques i 4 són negres. Si s'extrauen tres boles amb reemplaçament (és a dir, es tornen a dipositar a l'urna després de cada extracció), quina és la probabilitat de traure, per ordre, blanca, després negra i finalment negra?

Denotem per B_1 al succés "traure una bola blanca en la primera extracció", N_2 a "traure una bola negra en la segona extracció" i N_3 "traure una bola negra en la tercera extracció"; el problema demana la probabilitat que els tres successos anteriors es donen simultàniament, i per tant s'ha de calcular $p(B_1 \cap N_2 \cap N_3)$.

Com que hi ha reemplaçament, l'urna es troba en les mateixes condicions en cadascuna de les tres extraccions, per tant els successos són independents entre si i es té:

$$p(B_1 \cap N_2 \cap N_3) = p(B_1)p(N_2)p(N_3) = \frac{5}{9} \cdot \frac{4}{9} \cdot \frac{4}{9} = \frac{80}{729} \approx 0.1097.$$

- Si s'extrauen tres boles sense devolució, quina és la probabilitat del succés anterior?

Donat que en aquest cas no hi ha reemplaçament, l'estat de l'urna canvia després de cada extracció, amb una bola menys del color corresponent a la bola extreta amb anterioritat. Això vol dir per tant que ara els successos són dependents entre si. En conseqüència, hem d'emprar l'expressió general:

$$p(B_1 \cap N_2 \cap N_3) = p(B_1)p(N_2|B_1)p(N_3|(B_1 \cap N_2)).$$

Per tal de calcular-la, hem de trobar per separat cadascuna de les probabilitats del segon membre de la igualtat.

La primera és la d'obtenir blanca en la primera extracció, que és $p(B_1) = \frac{5}{9}$.

La segona és la d'obtenir negra en la segona extracció sabent que la primera ha sigut blanca; això vol que l'urna ara té una bola blanca menys, i per tant aquesta té en eixe moment 8 boles, 4 de les quals són blanques i altres 4 negres, així doncs $p(N_2|B_1) = \frac{4}{8}$.

Finalment, el tercer factor correspon a la probabilitat d'obtenir negra en la tercera extracció sabent que en la primera ha eixit blanca y en la segona negra. Això vol dir que en eixa situació l'urna té una bola blanca i una bola negra menys, ço és, un total de 7 boles, 4 blanques i 3 negres. Per tant $p(N_3|(B_1 \cap N_2)) = \frac{3}{7}$.

Aleshores:

$$p(B_1 \cap N_2 \cap N_3) = p(B_1)p(N_2|B_1)p(N_3|(B_1 \cap N_2)) = \frac{5}{9} \cdot \frac{4}{8} \cdot \frac{3}{7} = \frac{5}{42} \approx 0.1190.$$

Un altre resultat de gran utilitat en algunes situacions és el teorema de la probabilitat total, que es correspon amb una generalització de la darrera propietat de la probabilitat condicionada i diu el següent.

Teorema (Teorema de la probabilitat total) *Suposem que Ω es descomposa en una partició de successos mútuament excloents, és a dir, $\Omega = \cup_i H_i$, amb $H_i \cap H_j = \emptyset$ si $i \neq j$.*

Siga $A \subseteq \Omega$ un succés. Aleshores la seua probabilitat es calcula per

$$p(A) = \sum_i p(A|H_i)p(H_i).$$

Exemple *Efectuem un llançament d'una moneda i, a continuació, un dau.*

- *Si ix cara, es guanya si s'obté un nombre parell en el llançament del dau.*
- *Si ix creu, es guanya si s'obté un 1 o un 6 en el dau.*

Quina és la probabilitat de guanyar?

Anomenem $A = \{\text{guanyar}\}$. Descomposem el nostre espai mostral en funció del resultat en llançar una moneda: $\Omega = C \cup X$, on $C = \{\text{obtenir cara}\}$ i $X = \{\text{obtenir creu}\}$. Aleshores tenim el següent: D'una banda, si ha eixit cara, la probabilitat de guanyar és $p(A|C) = \frac{3}{6}$ (obtenir parell) i, d'altra, si ha eixit creu, la probabilitat de guanyar és $p(A|X) = \frac{2}{6}$ (obtenir un 1 o un 6).

$$p(A) = p(A|C)p(C) + p(A|X)p(X) = \frac{3}{6} \cdot \frac{1}{2} + \frac{2}{6} \cdot \frac{1}{2} = \frac{5}{12}.$$

Exemple *Llancem dues monedes i observem el resultat. Aleshores, en funció d'aquest, robem un cert nombre de cartes d'una baralla:*

- *Si no ix cap cara, robem una carta.*
- *Si ix una cara, en robem dues.*
- *Si ixen dos cares, n'agafem tres.*

Es desitja conèixer la probabilitat d'obtenir almenys una pica en finalitzar el procediment.

Siga $A = \{\text{obtenir almenys una pica}\}$. Ara, podem descomposar el nostre espai mostral en funció dels resultats obtinguts en llançament de la moneda com $\Omega = H_0 \cup H_1 \cup H_2$, on $H_i = \{\text{eixir } i \text{ cares}\}$, que són mútuament excloents, ja que no es poden donar simultàniament.

Per tant, pel teorema de la probabilitat total es té que

$$p(A) = p(A|H_0)p(H_0) + p(A|H_1)p(H_1) + p(A|H_2)p(H_2).$$

D'una banda, tenim $p(H_0) = p(\{\text{obtenir 0 cares}\}) = \frac{1}{4}$, $p(H_1) = p(\{\text{obtenir 1 cara}\}) = \frac{2}{4} = \frac{1}{2}$ i $p(H_2) = p(\{\text{obtenir 2 cares}\}) = \frac{1}{4}$.

D'altra banda, $p(A|H_0)$ implica el càlcul de la probabilitat d'obtenir una pica si se sap que no s'ha obtingut cap cara. Això vol dir que solament es roba una carta, i per tant, com que hi ha 4 pals, $p(A|H_0) = \frac{1}{4}$.

Pel que fa a $p(A|H_1)$, ara s'ha de calcular la probabilitat d'obtenir almenys una pica si ha eixit una cara, és a dir, robant dues cartes. En casos com aquest, resulta més fàcil plantejar el problema a la inversa: calculem primer la probabilitat de no obtenir cap pica en les dues extraccions (succés contrari), és a dir, $p(\bar{A}|H_1)$, i després calculem la que volem com $p(A|H_1) = 1 - p(\bar{A}|H_1)$. La probabilitat de no

obtenir una pica en la primera extracció és $\frac{3}{4}$. Pel que fa a la segona, com que en la primera no hem obtingut pica, continuen havent a la baralla 13 piques, mentre que ara queden a la baralla 51 cartes. Així doncs, la probabilitat de no obtenir una pica en la segona extracció és $\frac{51-13}{51} = \frac{38}{51}$. Aleshores, $p(\bar{A}|H_1) = \frac{3}{4} \cdot \frac{38}{51} = \frac{19}{34}$, i per tant $p(A|H_1) = 1 - p(\bar{A}|H_1) = \frac{15}{34}$.

Finalment, $p(A|H_2)$ fa referència a la probabilitat d'obtenir almenys una pica si han eixit dues cares, ço és, si es roben tres cartes. Raonant de la mateixa manera que abans, s'obté $p(\bar{A}|H_2) = \frac{3}{4} \cdot \frac{38}{51} \cdot \frac{37}{50} = \frac{703}{1700}$ i, conseqüentment, $p(A|H_2) = \frac{997}{1700}$.

Amb açò es completa el càlcul de les probabilitats requerides, i es té per tant:

$$p(A) = p(A|H_0)p(H_0) + p(A|H_1)p(H_1) + p(A|H_2)p(H_2) = \frac{1}{4} \cdot \frac{1}{4} + \frac{15}{34} \cdot \frac{1}{2} + \frac{997}{1700} \cdot \frac{1}{4} = \frac{1461}{3400} \approx 0.4297.$$

Finalment, el resultat següent relaciona els dos possibles sentits de la probabilitat condicionada, molt útil en alguns casos com veurem més endavant.

Teorema (Teorema de Bayes) Supposem que ha ocorregut el succés A . Aleshores, la probabilitat que vinga donat per la causa H ve donada per

$$p(H|A) = \frac{p(A|H)p(H)}{p(A)}.$$

Exemple Supposem que hem realitzat l'experiment de l'exemple anterior i hem obtingut una pica, és a dir, s'ha donat A . Ara es desitja conèixer la probabilitat que aquest resultat s'haja donat havent eixit una cara en el primer pas.

Denotant $H = H_1$, es té pel teorema de Bayes:

$$p(H|A) = \frac{p(A|H)p(H)}{p(A)} = \frac{\frac{15}{34} \cdot \frac{1}{2}}{\frac{1461}{3400}} = \frac{250}{487} \approx 0.5133.$$

Exemple Supposem que, donat un grup de persones, un 1% d'elles pateix una malaltia determinada. Es disposa d'una prova de detecció d'aquesta, la qual té un 1% de probabilitat d'obtenir un fals positiu, així com un 1% d'obtenir un fals negatiu. Es demana obtenir la probabilitat que un individu patisca la malaltia si el resultat de la prova ha sigut positiu.

Siguen els successos $M = \{\text{patir la malaltia}\}$ i $P = \{\text{obtenir positiu}\}$. Aleshores la interpretació de les dades del problema és: $p(M) = 0.01$ (probabilitat de patir la malaltia), $p(P|\bar{M}) = 0.01$ (probabilitat d'obtenir un fals positiu) i $p(\bar{P}|M) = 0.01$ (probabilitat d'obtenir un fals negatiu). I es demana calcular $p(M|P)$.

Pel teorema de Bayes, es té

$$p(M|P) = \frac{p(P|M)p(M)}{p(P)}.$$

D'una banda, tenim $p(P|M) = 1 - p(\bar{P}|M) = 1 - 0.01 = 0.99$, i d'altra banda hem de calcular $p(P)$, cosa que podem fer partint l'espai mostral en els dos successos mútuament excloents M i $\bar{M}$:

$$\begin{aligned} p(P) &= p(P|M)p(M) + p(P|\bar{M})p(\bar{M}) = (1 - p(\bar{P}|M))p(M) + p(P|\bar{M})(1 - p(M)) \\ &= (1 - 0.01) \cdot 0.01 + 0.01 \cdot (1 - 0.01) = 2 \cdot 0.99 \cdot 0.01. \end{aligned}$$

En conseqüència:

$$p(M|P) = \frac{p(P|M)p(M)}{p(P)} = \frac{0.99 \cdot 0.01}{2 \cdot 0.99 \cdot 0.01} = \frac{1}{2} = 0.5.$$

1.2.2 Variables aleatòries

El concepte de variable aleatòria naix de la necessitat de fer més manejables matemàticament els resultats dels experiments aleatoris. Per exemple, considerem els següents dos experiments:

1. Llançar una moneda defectuosa tal que $p(C) = 0.9$.

Denotem per C al succés {obtenir cara} = $\{C\}$ i X al succés {obtenir creu} = $\{X\}$. Aleshores $\Omega = \{C, X\}$, $p(C) = 0.9$, $p(X) = 0.1$.

2. Observar una peça fabricada en un procés que produeix un 10% de peces defectuoses.

Denotem per B al succés {peça bona} = $\{B\}$ i D al succés {peça defectuosa} = $\{D\}$. Aleshores $\Omega = \{B, D\}$, $p(B) = 0.9$, $p(D) = 0.1$.

Com es pot comprovar, les úniques diferències que hi ha entre els dos experiments és la naturalesa del problema en qüestió i la notació emprada per a representar els successos, però des del punt de vista d'objectes matemàtics són *essencialment* el mateix. Les variables aleatòries permeten treballar d'una manera més genèrica, sense dependre del context del problema sinó únicament de les dades d'aquest, a més de ser una ferramenta de gran utilitat per a tota allò en què treballarem més endavant.

Definició (Variable aleatòria) Siga $(\Omega, \mathcal{A}, p)$ un espai probabilístic i $\mathbb{R}$ el cos dels nombres reals.

Es diu que una aplicació

$$X : \Omega \longrightarrow \mathbb{R}$$

és una **variable aleatòria (V.A.)** si

$$\forall x \in \mathbb{R}, \quad A_x = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{A},$$

és a dir, si A_x és un succés $\forall x \in \mathbb{R}$.

Nota En el cas de l'àlgebra total, $\mathcal{A} = \mathcal{P}(\Omega)$, la condició especificada en la definició sempre es verifica, i per tant en els casos en què treballarem una variable aleatòria serà tota aplicació $X : \Omega \longrightarrow \mathbb{R}$. D'ara endavant, assumirem sempre que estem sota aquestes circumstàncies.

Exemple Prenem nota del resultat obtingut en llançar dues monedes.

En aquest cas, l'espai mostral és $\Omega = \{CC, CX, XC, XX\}$. Considerem l'àlgebra total, $\mathcal{A} = \mathcal{P}(\Omega)$.

Siga $Y : \Omega \longrightarrow \mathbb{R}$ la V.A. que compta el nombre de cares obtingudes: $Y(CC) = 2$, $Y(CX) = 1$, $Y(XC) = 1$, $Y(XX) = 0$.

Definició Donat $I \subseteq \mathbb{R}$, es defineix

$$p(X \in I) = p(A_I),$$

on

$$A_I = \{\omega \in \Omega \mid X(\omega) \in I\}.$$

En el cas que I siga un interval, se sol denotar de la següent manera:

- $I =]-\infty, x]$: $p(X \leq x)$.
- $I =]-\infty, x[$: $p(X < x)$.
- $I = [x, +\infty[$: $p(X \geq x)$.

- $I =]x, +\infty[: p(X > x).$
- $I = \{x\} : p(X = x).$

Exemple Sota les mateixes circumstàncies i notació que a l'exemple anterior, podem ara considerar probabilitats associades a la variable aleatòria Y .

Així doncs, plantejar el càlcul del valor de $p(Y = 1)$ és plantejar-se la probabilitat que el nombre de cares siga igual a 1, i per tant

$$p(Y = 1) = p(\{CX, XC\}) = \frac{2}{4} = \frac{1}{2},$$

o, per exemple, $p(Y \leq 1)$ fa referència a la probabilitat que el nombre de cares siga menor o igual que 1 (per tant, en aquest cas, que isquen 0 cares o 1 cara), de manera que

$$p(Y \leq 1) = p(\{XX, CX, XC\}) = \frac{3}{4}.$$

Definició Donades dues variables aleatòries X, Y , definides en el mateix espai probabilístic, es pot definir la suma $X + Y$, la diferència $X - Y$, el producte XY i, sempre que Y no prenga valors nuls, el quocient $\frac{X}{Y}$.

Propietat Totes les operacions definides anteriorment generen una nova variable aleatòria.

Exemple Seguint amb els exemples anteriors, introduïm ara una nova variable aleatòria Z , que compta el nombre de creus obtingudes: $Z(CC) = 0$, $Z(CX) = 1$, $Z(XC) = 1$, $Z(XX) = 2$.

Podem considerar ara diferents operacions entre Y i Z . Es presenten a continuació alguns exemples.

- $p(Z - Y = -2)$. Representa la probabilitat que el nombre de creus menys el nombre de cares siga -2 , o el que és el mateix, que el nombre de creus siga igual al nombre de cares menys 2 ($Z = Y - 2$). L'única possibilitat en aquest cas és $Y = 2$ i $Z = 0$, corresponent al resultat CC , i per tant

$$p(Z - Y = -2) = p(\{CC\}) = \frac{1}{4}.$$

- $p(Z - Y \geq 0)$. La condició de la probabilitat es tradueix com $Z \geq Y$, o el que és el mateix, que el nombre de creus siga major o igual que el de cares. Els resultats que verifiquen aquesta condició són CX , XC i XX , i per tant

$$p(Z - Y \geq 0) = p(\{CX, XC, XX\}) = \frac{3}{4}.$$

- $p(YZ = 0)$. Quan un producte és zero, algun dels dos factors ho ha de ser, per tant $Y = 0$ (0 cares) o bé $Z = 0$ (0 creus), essent XX i CC , respectivament, els únics resultats que verifiquen aquesta condició. Aleshores

$$p(YZ = 0) = p(\{CC, XX\}) = \frac{2}{4} = \frac{1}{2}.$$

- $p(Y + Z = 2)$. La variable $Y + Z$ representa el nombre de cares més el nombre de creus. Com que es fan un total de 2 llançaments, la suma anterior sempre val 2, i en conseqüència la igualtat de la condició es verifica en qualsevol cas. Així doncs:

$$p(Y + Z = 2) = p(\Omega) = 1.$$

Exercici 6 Amb les dades de l'exemple anterior, troba el valor de $p(YZ = 1)$, $p(YZ = 2)$, $p(Y + Z = 1)$, $p(Y + Z \geq 1)$ i $p(Y - Z \leq 3)$.

Definició (Suport) Anomenem **suport** d'una V.A. X al conjunt de valors que pren amb probabilitat no nul·la i es denota per S_X .

$$S_X = \{x \in \mathbb{R} \mid p(X = x) \neq 0\}.$$

Definició (Variable aleatòria discreta) Una V.A. X es diu **discreta** si el seu suport és un conjunt finit o infinit numerable.

Exemple Les variables aleatòries definides als exemples anteriors, Y, Z , són discretes ja que el seu suport és finit:

$$S_Y = S_Z = \{0, 1, 2\}.$$

Exercici 7 Amb la notació dels exemples anteriors, calcula S_{Y+Z} , S_{Y-Z} , S_{Z-Y} i S_{YZ} .

Definició (Funció de densitat discreta) Donada X una V.A. discreta, es defineix la **funció de densitat de probabilitat** associada a X , f_X , per

$$f_X(x) = \begin{cases} p(X = x) \neq 0 & x \in S_X \\ 0 & x \notin S_X \end{cases}.$$

Exemple Tornant a l'experiment dels exemples anteriors, es té que $f_Y(0) = p(Y = 0) = \frac{1}{4}$, $f_Y(1) = p(Y = 1) = \frac{1}{2}$, $f_Y(2) = p(Y = 2) = \frac{1}{4}$ i $f_Y(x) = 0$ si $x \notin \{0, 1, 2\}$. Pel que fa a la V.A. Z , la seua funció de densitat coincideix amb la de Y , és a dir, $f_Z = f_Y$.

Propietats Sigui X una V.A. discreta. Aleshores:

1. $0 \leq f_X(x) \leq 1$.
2. $\sum_{x \in S_X} f(x) = 1$.

Exemple Es realitza un experiment consistent en llançar una moneda fins obtenir una cara. Sigui Y la variable aleatòria que quantifica el nombre de llançaments necessaris fins obtenir-la. Es demana obtenir la funció de densitat associada a Y , f_Y .

Si $n \in \mathbb{N}$, $p(Y = n)$ és la probabilitat d'haver necessitat n llançaments fins obtenir una cara, és a dir, d'haver obtingut creu als $n - 1$ primers llançaments i cara al llançament que fa n . Si denotem per C_i al succés "obtenir cara al llançament i " i X_i al succés "obtenir creu al llançament i ", el que es demana és $p(X_1 \cap X_2 \cap \dots \cap X_{n-1} \cap C_n)$. Com que cada llançament és independent del resultat dels anteriors, es té

$$p(X = n) = p(X_1 \cap X_2 \cap \dots \cap X_{n-1} \cap C_n) = p(X_1)p(X_2) \cdots p(X_{n-1})p(C_n) = \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2^n}.$$

D'altra banda, si $n \notin \mathbb{N}$ aleshores $p(X = n) = 0$, ja que el nombre de llançaments necessaris és sempre un nombre enter. Per tant, es conclou que $S_X = \mathbb{N}$. La funció de densitat associada és aleshores la següent:

$$f_Y(n) = \begin{cases} \frac{1}{2^n} & n \in \mathbb{N} \\ 0 & n \notin \mathbb{N} \end{cases}.$$

A més, d'una banda tenim

$$\sum_{n \in S_Y} f_Y(n) = \sum_{n \in \mathbb{N}} f_Y(n) = \sum_{n=1}^{\infty} \frac{1}{2^n}.$$

Ara bé, d'altra banda, la segona propietat de les funcions de densitat garanteix que

$$\sum_{n \in S_Y} f_Y(n) = 1.$$

Per tant, i com a curiositat, el que acabem demostrar, fent servir únicament la teoria de la probabilitat que hem vist fins ara, és que

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

Definició (Funció de distribució) Una funció de distribució de probabilitat d'una V.A. X és una funció real de variable real que es defineix per

$$F_X(x) = p(X \leq x), \quad \forall x \in \mathbb{R}.$$

Propietats Siga X una V.A. amb funció de distribució F_X . Aleshores:

1. $0 \leq F_X(x) \leq 1 \quad \forall x \in \mathbb{R}$.
2. Si $f_X(x)$ és la funció de densitat de X , llavors

$$F_X(x) = \sum_{x \in S_X} f_X(x).$$

3. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ i $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
4. $F_X(x)$ és no decreixent, ço és,

$$x_1 < x_2 \longrightarrow F_X(x_1) \leq F_X(x_2), \quad \forall x_1, x_2 \in \mathbb{R}.$$

5. $F(x)$ és contínua per la dreta:

$$\lim_{x \rightarrow a^+} F_X(x) = F_X(a), \quad \forall a \in \mathbb{R}.$$

6. $p(a < X \leq b) = F_X(b) - F_X(a) \quad \forall a, b \in \mathbb{R}, a < b$.

Exemple Calculem la funció de distribució associada a la V.A. Y que quantifica el nombre de cares obtingudes en llançar dues monedes. Siga $y \in \mathbb{R}$:

- Si $y < 0$ aleshores $F_Y(y) = p(Y \leq y) = 0$ (és impossible que isca un nombre negatiu de cares).
- Si $0 \leq y < 1$ llavors $F_Y(y) = p(Y \leq y) = p(\{XX\}) = \frac{1}{4}$ (dir que el nombre de cares és menor o igual que un valor menor que 1 és el mateix que dir que el nombre de cares és 0).
- Si $1 \leq y < 2$ aleshores $F_Y(y) = p(Y \leq y) = p(\{XX, XC, CX\}) = \frac{3}{4}$ (si el nombre de cares és menor o igual que un valor menor que 2, aleshores és que o han eixit 0 o cares o ha eixit 1 cara).
- Si $y \geq 2$ llavors $F_Y(y) = p(Y \leq y) = p(\{XX, XC, CX, CC\}) = 1$ (és segur que el nombre de cares obtingudes serà menor o igual que un valor major o igual que 2, ja que el nombre màxim de cares que poden eixir és 2).

Així doncs, la funció de distribució associada a la V.A. Y és:

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{4} & 0 \leq y < 1 \\ \frac{3}{4} & 1 \leq y < 2 \\ 1 & y \geq 2 \end{cases}.$$

Observem que açò està en concordància amb totes les propietats de les funcions de distribució.

Definició (Variable aleatòria contínua) Una V.A. es diu que és **contínua** si la seua funció de distribució associada, $F_X(x) = p(X \leq x)$, no presenta discontinuïtats. En aquest cas, aquesta funció de distribució també es diu que és **contínua**.

Un cas especial (i amb el qual treballarem generalment) de funcions de distribució contínues són les absolutament contínues, que són aquelles que es poden escriure de la següent manera:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt, \quad \forall x \in \mathbb{R}.$$

Definició La funció $f_X(t)$ de l'integrand anterior rep el nom de **funció de densitat de probabilitat** de la V.A. X .

Propietats

1. $f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$.
2. $\int_{-\infty}^{+\infty} f_X(t) dt = 1$.
3. Si f_X és contínua en x , aleshores $\exists F'_X(x) = f_X(x)$.
4. $p(a < X \leq b) = \int_a^b f_X(t) dt = F_X(b) - F_X(a) \quad \forall a, b \in \mathbb{R}, a < b$.

Exemple Considerem l'experiment "triare un nombre real a l'atzar entre 0 i 1" i X la seua variable associada, corresponent als valors aleatoris corresponents a aquesta tria. Aleshores, donat $x \in \mathbb{R}$, tenim el següent:

- Si $x < 0$, llavors $F_X(x) = p(X \leq x) = 0$ (ja que no és possible triar valors negatius).
- Si $0 \leq x \leq 1$ llavors $F_X(x) = p(X \leq x) = x$, corresponent a la proporció de nombres que hi ha a l'interval $[0, x]$ (amb una àrea de x unitats) respecte de l'interval total $[0, 1]$ (amb una àrea d'una unitat).
- Si $x > 1$ llavors $F_X(x) = p(X \leq x) = 1$ (ja que és segur que el resultat estarà per baix d'un valor major que 1).

Així doncs, la funció de distribució queda com es segueix:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}.$$

Podem definir per tant la funció de densitat associada com:

$$f_X(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases}.$$

Complint-se efectivament que

$$F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

Notem que, a diferència del cas discret, en el cas continu la funció de densitat no indica una probabilitat; tanmateix, es pot interpretar com el *ritme* o *magnitud* en què aquesta es va acumulant en cada punt (en termes infinitesimals).

A continuació anem a introduir els conceptes d'**esperança** i **variància**, que estan íntimament relacionats amb els de mitjana i variància d'una mostra aleatòria.

Definició Sigui X una V.A. amb funció de densitat f_X . Es defineix l'**esperança** o **valor mitjà** de X de la següent manera:

- Si X és discreta:

$$E(X) = \sum_{x \in S_x} x f_X(x).$$

- Si X és contínua:

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx.$$

Notem que en el cas discret, la definició d'esperança guarda certa dualitat amb la definició de mitjana d'una mostra aleatòria. En efecte, si $\{x_i\}_{i=1}^k$ és una mostra aleatòria agrupada per freqüències $\{n_i\}_{i=1}^k$, amb $n = \sum_{i=1}^k n_i$ el nombre total de valors, llavors la mitjana mostral és

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

Note's que el valor $\frac{n_i}{n}$ és la freqüència relativa associada al valor x_i (és a dir, la proporció de vegades que aquest apareix), que es pot interpretar com a una aproximació de la seua probabilitat real d'aparició.

D'altra banda, si X és una V.A. discreta

$$E(X) = \sum_{x \in S_x} x f_X(x) = \sum_{x \in S_x} x p(X = x),$$

que efectivament guarda una clara dualitat amb les consideracions anteriors sobre la mitjana mostral, on ara apareix la probabilitat exacta en lloc de l'aproximació per freqüència relativa mostral.

Exemple

- Siga X el nombre de cares obtingudes en llançar dues vegades una moneda equilibrada. Aleshores tenim $S_X = \{0, 1, 2\}$ i $p(X = 0) = \frac{1}{4}$, $p(X = 1) = \frac{1}{2}$ i $p(X = 2) = \frac{1}{4}$. Aleshores es té

$$f_X(x) = \begin{cases} \frac{1}{4} & x = 0 \\ \frac{1}{2} & x = 1 \\ \frac{1}{4} & x = 2 \\ 0 & x \notin S_X \end{cases}.$$

Per tant

$$E(X) = \sum_{x \in S_X} x f_X(x) = \sum_{x=0}^2 x f_X(x) = \sum_{x=0}^2 x p(X = x) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1.$$

És a dir, el promig de cares que cal esperar obtenir en repetir aquest experiment (llançar dues vegades la moneda i observar el nombre de cares obtingudes) és 1.

- Definim ara X com la V.A. que compta el nombre d'uns obtinguts en llançar dues vegades un dau equilibrat. Novament, tenim que $S_X = \{0, 1, 2\}$, amb les següents probabilitats.

Denotem per A_1 al succés "obtenir un 1 al primer llançament" i A_2 al succés "obtenir un 1 al segon llançament". Llavors:

- $p(X = 0) = p(\overline{A_1} \cap \overline{A_2})$, és a dir, la probabilitat de no obtenir cap 1 als dos llançaments és la probabilitat de no obtenir 1 al primer i no obtenir tampoc 1 al segon. Com que els dos llançaments són independents, es té

$$p(X = 0) = p(\overline{A_1} \cap \overline{A_2}) = p(\overline{A_1})p(\overline{A_2}) = \frac{5}{6} \cdot \frac{5}{6} = \frac{25}{36}.$$

- $p(X = 1) = p((A_1 \cap \overline{A_2}) \cup (\overline{A_1} \cap A_2))$, ço és, la probabilitat d'obtenir un 1 després de llançar els daus és la probabilitat que isca al primer llançament i no isca al segon o bé que no isca en el primer però sí al segon. Emprant que $(A_1 \cap \overline{A_2})$ i $(\overline{A_1} \cap A_2)$ són mútuament excloents i novament la independència entre llançaments es té

$$\begin{aligned} p(X = 1) &= p((A_1 \cap \overline{A_2}) \cup (\overline{A_1} \cap A_2)) = p(A_1 \cap \overline{A_2}) + p(\overline{A_1} \cap A_2) \\ &= p(A_1)p(\overline{A_2}) + p(\overline{A_1})p(A_2) = \frac{1}{6} \cdot \frac{5}{6} + \frac{5}{6} \cdot \frac{1}{6} = \frac{10}{36} = \frac{5}{18}. \end{aligned}$$

- $p(X = 2) = p(A_1 \cap A_2)$, és a dir, la probabilitat d'obtenir dos 1 correspon a la probabilitat associada a obtenir un 1 en les dues tirades, tot obtenint-se, emprant la independència dels llançaments

$$p(X = 2) = p(A_1 \cap A_2) = p(A_1)p(A_2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

Així doncs,

$$f_X(x) = \begin{cases} \frac{25}{36} & x = 0 \\ \frac{5}{18} & x = 1 \\ \frac{1}{36} & x = 2 \\ 0 & x \notin S_X \end{cases}.$$

Per tant

$$E(X) = \sum_{x \in S_X} x f_X(x) = \sum_{x=0}^2 x f_X(x) = \sum_{x=0}^2 x p(X = x) = 0 \cdot \frac{25}{36} + 1 \cdot \frac{5}{18} + 2 \cdot \frac{1}{36} = \frac{1}{3}.$$

És a dir, el que s'espera a llarg termini és que, de cada 3 experiments, cadascun d'ells consistent en llançar dues vegades un dau equilibrat, en un d'ells isca un 1. Açò és coherent amb la intuïció ja que tres experiments impliquen un total de 6 llançaments, que és el nombre de resultats possibles, front a un únic cas favorable (eixir 1).

- Finalment, considerem X la V.A. corresponent al resultat en triar un nombre real a l'atzar a l'interval $[0, 1]$. Com ja s'ha vist, aquesta variable aleatòria és contínua i la seua funció de densitat és

$$f_X(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases}.$$

Per tant, en aquest cas l'esperança es calcula com

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_0^1 x dx = \left[\frac{x^2}{2} \right]_{x=0}^{x=1} = \frac{1^2}{2} - \frac{0^2}{2} = \frac{1}{2}.$$

Que resulta ser el valor mitjà de l'interval considerat. És a dir, el valor promig que s'espera en triar valors reals a l'atzar compresos entre 0 i 1 és $\frac{1}{2}$.

En general, l'esperança és una **mitjana ponderada** dels valors que pren una variable aleatòria, tot atorgant més pes als valors més probables.

Propietats Siguen X, Y variables aleatòries definides sobre un mateix espai mostral i $\lambda \in \mathbb{R}$. Aleshores es compleix:

- $E(\lambda X) = \lambda E(X)$.
- $E(X + Y) = E(X) + E(Y)$.
- Si X, Y són independents, aleshores $E(XY) = E(X)E(Y)$.
- Si g és una funció real de variable real, aleshores:

- Si X és discreta, $E[g(X)] = \sum_{x \in S_X} g(x)f_X(x)$.
- Si X és contínua, $E[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)dx$.

Exercici 8

(a) En un examen de tipus test, amb quatre opcions per pregunta, contestar correctament atorga 1 punt a la nota total, mentre que respondre erròniament resta 0.75 punts. Suposem que hi ha un total de 100 preguntes, amb una puntuació total sobre 100, i que un alumne es troba en la següent situació:

1. Coneix la resposta de 40 de les preguntes.
2. En 30 d'elles, sap que dues de les opcions no són correctes.
3. Hi ha 20 en què sap que una de les opcions no és correcta.
4. No sap res respecte de les 10 preguntes restants.

Suposant que tot el que creu saber l'alumne és correcte, si aquest decideix contestar aleatòriament entre les 30 en què dubta entre dues opcions, quina puntuació addicional caldria esperar que obtinga? És esperable que aprobe?

Fes el mateix per a aquelles en què dubta entre tres opcions o totes elles. Interpreta els resultats i digues si convé o no que responga aquest grup de preguntes.

(b) S'ha comprovat que la funció de densitat associada als beneficis nets diaris d'una empresa privada és, per a cada $x \in \mathbb{R}$:

$$f_X(x) = \frac{1}{\pi(1 + (x + 1)^2)}.$$

Comprova que efectivament f_X és una funció de densitat de probabilitat ($\int_{-\infty}^{+\infty} f_X(x)dx = 1$) i calcula $E(X)$, on X és la V.A. associada als beneficis diaris obtinguts.

Interpreta els resultats i digues si l'empresa resulta rentable a llarg termini.

Definició (Variància) Siga X una V.A. amb funció de densitat $f_X(x)$. Es defineix la **variància** d'aquesta variable com

$$\text{var}(X) = E((X - E(X))^2).$$

L'arrel quadrada no negativa de la variança, $+\sqrt{\text{var}(X)}$, es denomina **desviació típica**.

En el cas discret, aquesta definició també guarda una relació estreta amb la variància d'una mostra aleatòria, que recordem que, donada $\{x_i\}_{i=1}^k$ una mostra aleatòria agrupada per freqüències $\{n_i\}_{i=1}^k$, amb $n = \sum_{i=1}^k n_i$ el nombre total de valors, es defineix com

$$s^2 = \frac{1}{n} \sum_{i=1}^k n_i(x_i - \bar{x})^2 = \sum_{i=1}^k (x_i - \bar{x})^2 \frac{n_i}{n}.$$

D'altra banda, si X és una V.A. discreta, llavors

$$\text{var}(X) = E((X - E(X))^2) = \sum_{x \in S_X} (x - E(X))^2 f_X(x) = \sum_{x \in S_X} (x - E(X))^2 p(X = x).$$

Exemple

- Siga X el nombre de cares obtingudes en llançar dues vegades una moneda equilibrada. Emprant els resultats obtinguts anteriorment:

$$\begin{aligned}\text{var}(X) &= E((X - E(X))^2) = E((X - 1)^2) = \sum_{x \in S_X} (x - 1)^2 f_X(x) = \sum_{x=0}^2 (x - 1)^2 p(X = x) \\ &= (0 - 1)^2 \cdot \frac{1}{4} + (1 - 1)^2 \cdot \frac{1}{2} + (2 - 1)^2 \cdot \frac{1}{4} = \frac{1}{2} = 0.5.\end{aligned}$$

I la desviació típica és per tant $\sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2} \approx 0.7071$.

- Considerem X la V.A. que compta el nombre d'uns obtinguts en llançar dues vegades un dau equilibrat.

$$\begin{aligned}\text{var}(X) &= E((X - E(X))^2) = E\left(\left(X - \frac{1}{3}\right)^2\right) = \sum_{x \in S_X} \left(x - \frac{1}{3}\right)^2 f_X(x) \\ &= \sum_{x=0}^2 \left(x - \frac{1}{3}\right)^2 p(X = x) = \left(0 - \frac{1}{3}\right)^2 \cdot \frac{25}{36} + \left(1 - \frac{1}{3}\right)^2 \cdot \frac{10}{36} + \left(2 - \frac{1}{3}\right)^2 \cdot \frac{1}{36} \\ &= \frac{5}{18} \approx 0.2778.\end{aligned}$$

Tot resultant la desviació típica $\sqrt{\frac{5}{18}} \approx 0.5270$.

- Per acabar, siga X la V.A. corresponent al resultat en triar un nombre real a l'atzar a l'interval $[0, 1]$.

$$\begin{aligned}\text{var}(X) &= E((X - E(X))^2) = E\left(\left(X - \frac{1}{2}\right)^2\right) = \int_{-\infty}^{+\infty} \left(x - \frac{1}{2}\right)^2 f_X(x) dx = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx \\ &= \left[\frac{\left(x - \frac{1}{2}\right)^3}{3}\right]_{x=0}^{x=1} = \frac{\left(1 - \frac{1}{2}\right)^3}{3} - \frac{\left(0 - \frac{1}{2}\right)^3}{3} = \frac{1}{12} \approx 0.0833.\end{aligned}$$

La desviació típica és llavors $\sqrt{\frac{1}{12}} = \frac{\sqrt{3}}{6} \approx 0.2887$.

Propietats Siguen X, Y variables aleatòries definides sobre un mateix espai mostral i $a, b \in \mathbb{R}$. Aleshores es compleix:

1. $\text{var}(X) = E(X^2) - E(X)^2$.
2. $\text{var}(a + bX) = b^2 \text{var}(X)$.
3. Si X, Y són independents, aleshores $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$.

Nota La primera propietat és dual amb el fet que la variància d'una mostra aleatòria agrupada per freqüències es pot escriure com

$$s^2 = \sum_{i=1}^k x_i^2 \frac{n_i}{n} - \bar{x}^2.$$

Exercici 9 Calcula novament les variàncies de l'exemple anterior emprant la primera propietat i comprova que el resultat és el mateix.

1.2.3 Distributions

Tot seguit veurem alguns casos especials de distribucions l'estudi de les quals resulta interessant. Aquestes es divideixen en dos blocs o classes: discretes i contínues.

En tots els casos treballarem sobre un espai de probabilitat $(\Omega, \mathcal{A}, p)$.

Distribucions discretes

• Distribució de Bernoulli

Un experiment aleatori rep el nom de Bernoulli si consisteix en observar i classificar els elements d'una població en dues categories: èxit i fracàs.

Definició Siga $A \in \mathcal{A}$ un succés amb probabilitat $p = p(A)$ coneguda, que associarem amb l'èxit de l'experiment. Direm que una V.A. X és de **Bernoulli** quan ve donada per

$$X(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}.$$

En aquest cas, es denota $X \sim Be(p)$.

$S_X = \{0, 1\}$. Així doncs, la seua funció de densitat és

$$f_X(x) = \begin{cases} p & x = 1 \\ q & x = 0 \\ 0 & x \notin \{0, 1\} \end{cases},$$

on $q = 1 - p$, i es compleix

$$E(X) = \sum_{x \in S_X} x f_X(x) = 0 \cdot f_X(0) + 1 \cdot f_X(1) = 0 \cdot q + 1 \cdot p = p.$$

$$\text{var}(X) = E(X^2) - E(X)^2 = \sum_{x \in S_X} x^2 f_X(x) - p^2 = (0^2 \cdot f_X(0) + 1^2 \cdot f_X(1)) - p^2 = p - p^2 = pq.$$

Exemple Considerem el llançament d'un dau i siga el succés $A = \{5, 6\}$.

Siga la variable aleatòria X tal que

$$X(1) = 0, X(2) = 0, X(3) = 0, X(4) = 0, X(5) = 1, X(6) = 1.$$

Aleshores $p = p(A) = \frac{1}{3}$, $q = 1 - p = \frac{2}{3}$, amb funció de densitat

$$f_X(x) = \begin{cases} \frac{1}{3} & x = 1 \\ \frac{2}{3} & x = 0 \\ 0 & x \notin \{0, 1\} \end{cases},$$

complint

$$E(X) = p = \frac{1}{3}.$$

$$\text{var}(X) = pq = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}.$$

• Distribució binomial

S'efectuen n proves Bernoulli i es pren nota del total d'èxits obtinguts.

Definició $A \in \mathcal{A}$ un succés amb probabilitat $p = p(A)$ coneguda. Direm que una V.A. X és una **binomial** si es pot descomposar com una suma de variables Bernoulli, és a dir, $X = X_1 + \dots + X_n$, on $X_i \sim Be(p)$.

En aquest cas, es denota $X \sim B(n, p)$.

$S_X = \{0, 1, \dots, n\}$, amb funció de densitat

$$f_X(k) = \begin{cases} \binom{n}{k} p^k q^{n-k} & k \in \{0, 1, \dots, n\} \\ 0 & k \notin \{0, 1, \dots, n\} \end{cases},$$

on $q = 1 - p$ i $\binom{n}{k}$ és el nombre combinatori n sobre k , que compta el nombre de combinacions de k èxits que poden repartir-se entre n posicions corresponents al total d'experiments efectuats, i ve donat per la fórmula

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Pel que fa a l'esperança i la variància, d'una banda emprant la propietat 1 (linealitat) de l'esperança, tenim:

$$E(X) = E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n) = p + \dots + p = np.$$

D'altra banda, fent servir ara la independència entre les variables Bernoulli X_i :

$$\text{var}(X) = \text{var}(X_1 + \dots + X_n) = \text{var}(X_1) + \dots + \text{var}(X_n) = pq + \dots + pq = npq.$$

Exemple Llancem una moneda 5 vegades.

- Quina és la probabilitat que isquen un total de 3 cares?
- Quina és la probabilitat que isquen menys de dues cares?
- Calcula $E(X)$ i $\text{var}(X)$.

Solució Siga X la V.A. que compta el total de cares obtingudes i el succés $A = \{\text{obtenir cara}\}$. Llavors $p = p(A) = \frac{1}{2} = 0.5$, i per tant $X \sim B(5, 0.5)$. Aleshores:

- La probabilitat d'obtenir un total de 3 cares és

$$p(X = 3) = \binom{5}{3} \cdot 0.5^3 \cdot (1 - 0.5)^{5-3} = \frac{5!}{3! \cdot (5-3)!} \cdot 0.5^3 \cdot 0.5^2 = 10 \cdot 0.5^5 = 0.3125.$$

- La probabilitat que isquen menys de dues cares és

$$p(X < 2) = p((X = 0) \cup (X = 1)) = p(X = 0) + p(X = 1) = 0.03125 + 0.15625 = 0.1875.$$

- $E(X) = 5 \cdot 0.5 = 2.5$ i $\text{var}(X) = 5 \cdot 0.5 \cdot 0.5 = 1.25$.

• Distribució de Poisson

Per a cada $n \in \mathbb{N}$, considerem una tanda n assajos de Bernouilli amb probabilitat p_n (que varia en funció del nombre d'assajos, n), complint

$$\lim_{n \rightarrow \infty} np_n = \lambda > 0.$$

D'acord amb el que hem vist, si denotem per X_n la V.A. corresponent al nombre d'èxits obtinguts en la tanda de n experiments, llavors $X_n \sim B(n, p_n)$ i per tant es té, denotant $q_n = 1 - p_n$,

$$p(X_n = k) = \binom{n}{k} p_n^k q_n^{n-k}.$$

Es pot demostrar que

$$\lim_{n \rightarrow \infty} p(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Definició Direm que una V.A. X segueix una **distribució de Poisson** de paràmetre $\lambda > 0$ si la seua funció de densitat és

$$f_X(k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda} & k \in \{0, 1, 2, \dots\} \\ 0 & k \notin \{0, 1, 2, \dots\} \end{cases}.$$

Es denota per $X \sim Po(\lambda)$.

En aquest cas, l'esperança i la variància coincideixen amb el valor λ :

$$E(X) = \text{var}(X) = \lambda.$$

A més, es compleix el següent.

Propietat Si $X_1 \sim Po(\lambda_1)$ i $X_2 \sim Po(\lambda_2) \longrightarrow X_1 + X_2 \sim Po(\lambda_1 + \lambda_2)$.

Exemple El nombre d'accidents de trànsit setmanals d'una determinada carretera nacional segueix una distribució de Poisson de paràmetre $\lambda = 2$. Calcula la probabilitat que en quatre setmanes es donen més de 3 accidents de trànsit.

Solució Siga X_i la V.A. corresponent al nombre d'accidents de trànsit ocorreguts a la setmana i , $1 \leq i \leq 4$. Llavors $X_i \sim Po(2)$, $1 \leq i \leq 4$. Hem de treballar amb la variable que recull el total d'accidents, és a dir, $X = X_1 + X_2 + X_3 + X_4$. Per la propietat anterior, es té que $X \sim Po(2 + 2 + 2 + 2) = Po(8)$. Aleshores la probabilitat que en quatre setmanes s'esdevinguin més de 3 accidents de trànsit és

$$\begin{aligned} p(X > 3) &= 1 - p(X \leq 3) = 1 - (p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3)) \\ &= 1 - \left(\frac{8^0}{0!} e^{-8} + \frac{8^1}{1!} e^{-8} + \frac{8^2}{2!} e^{-8} + \frac{8^3}{3!} e^{-8} \right) = 1 - \frac{379}{3} e^{-8} \approx 0.9576. \end{aligned}$$

D'altra banda, el promig esperat d'accidents de trànsit en les quatre setmanes és

$$E(X) = 8,$$

amb una variància

$$\text{var}(X) = 8.$$

• Distribució hipergeomètrica

Aquesta distribució es fa servir en casos en què es fan un total d' n experiments consistents a seleccionar **sense reemplaçament** un element dins d'un grup d'un total d' N elements, entre els quals hi ha K possibles èxits.

Definició Direm que una V.A. aleatòria X és **hipergeomètrica** si, sota les circumstàncies indicades anteriorment, aquesta compta el nombre d'èxits obtinguts sobre el total dels n experiments.

Es denota per $X \sim H(N, K, n)$.

La seua funció de densitat (probabilitat d'obtenir k èxits en n experiments) ve donada per

$$f_X(k) = \begin{cases} \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}} & k \in \{0, 1, \dots, n\} \\ 0 & k \notin \{0, 1, \dots, n\} \end{cases}.$$

Amb esperança i variància:

$$E(X) = n \frac{K}{N}.$$

$$\text{var}(X) = n \frac{K}{N} \frac{N-K}{N} \frac{N-n}{N-1}.$$

Exemple El temari d'una oposició consta d'un total de 75 temes. El dia de l'examen, es posen en una urna 75 boles numerades de l'1 al 75, corresponents a cada tema, i s'extrauen a l'atzar 5 d'elles, de les quals l'examinat ha de triar un dels cinc temes corresponents per al seu desenvolupament a l'examen.

Donís s'ha estudiat 33 dels 75 temes possibles i es presenta a l'examen. Quina és la probabilitat que a Donís li toque algun tema dins del grup dels que s'ha estudiat?

Solució En un grup d' $N = 75$ elements (temes totals), hi ha un total de $K = 33$ èxits (temes estudiats). Es realitzen un total d' $n = 5$ experiments (extracció de boles) i es demana la probabilitat d'obtenir almenys un èxit, és a dir, $p(X > 0)$, on X és la V.A. que quantifica el nombre d'èxits, o el que és el mateix, el contrari de no obtenir-ne cap; per tant, es té que $X \sim H(75, 33, 5)$, complint-se que

$$p(X > 0) = 1 - p(X = 0) = 1 - \frac{\binom{33}{0} \binom{75-33}{5-0}}{\binom{75}{5}} = 1 - \frac{\binom{33}{0} \binom{42}{5}}{\binom{75}{5}} = 1 - \frac{850668}{17259390} = \frac{2734787}{2876565} \approx 0.9507.$$

D'altra banda, s'espera que el promig d'èxits (nombre de temes que s'ha estudiat que ixen) siga

$$E(X) = 5 \cdot \frac{33}{75} = \frac{33}{15} = 2.2,$$

amb variància

$$\text{var}(X) = 5 \cdot \frac{33}{75} \cdot \frac{42}{75} \cdot \frac{75-5}{75-1} = \frac{1078}{925} \approx 1.1654.$$

• Distribució binomial negativa

Siga A un succés amb probabilitat p i X el nombre de vegades que s'ha de repetir una experiència, en condicions independents, fins que es done r vegades A .

Definició Una V.A. X amb les condicions anteriors es diu que segueix una **distribució binomial negativa**. En el cas particular que $r = 1$, rep el nom de **distribució geomètrica**.

Es denota per $X \sim BN(r, p)$.

Siga k el nombre d'absències, és a dir, el nombre de vegades que no s'ha donat A en el procés. Aleshores el nombre total d'experiments és $k + r$ (k en què no s'ha donat A i r on s'ha donat A) i la funció de densitat (associada a la probabilitat que s'hagen donat k absències) és

$$f_X(k) = \begin{cases} \binom{k+r-1}{r-1} p^r q^k & k \in \mathbb{N} \\ 0 & k \notin \mathbb{N} \end{cases},$$

on $q = 1 - p$ i amb esperança i variància

$$E(X) = \frac{rq}{p}.$$

$$\text{var}(X) = \frac{rq}{p^2}.$$

Exemple Llancem una moneda trucada tal que $p(C) = 0.6$ i $p(X) = 0.4$, on $C = \{\text{obtenir cara}\}$ i $X = \{\text{obtenir creu}\}$. Es demana obtenir la probabilitat de necessitar 5 o més llançaments fins obtenir dues creus.

Solució L'experiment finalitza en obtenir $r = 2$ creus, amb probabilitat $p = p(X) = 0.4$. Si k és el nombre d'absències del succés X , aleshores el nombre total d'experiments és $k + r = k + 2$. La condició de l'enunciat és per tant $k + 2 \geq 5$, o el que és el mateix, $k \geq 3$ (3 o més absències). Així doncs, anomenant Y a la V.A. que quantifica el nombre d'absències, llavors $Y \sim BN(2, 0.4)$ i es té

$$\begin{aligned} p(Y \geq 3) &= 1 - p(Y < 3) = 1 - (p(Y = 0) + p(Y = 1) + p(Y = 2)) \\ &= 1 - \left[\binom{0+2-1}{2-1} \cdot 0.4^2 \cdot 0.6^0 + \binom{1+2-1}{2-1} \cdot 0.4^2 \cdot 0.6^1 + \binom{2+2-1}{2-1} \cdot 0.4^2 \cdot 0.6^2 \right] \\ &= 1 - \left[\binom{1}{1} \cdot 0.4^2 \cdot 0.6^0 + \binom{2}{1} \cdot 0.4^2 \cdot 0.6^1 + \binom{3}{1} \cdot 0.4^2 \cdot 0.6^2 \right] \\ &= 1 - [0.16 + 0.192 + 0.1728] = 0.4752. \end{aligned}$$

Pel que fa a l'esperança i la variància:

$$E(X) = \frac{2 \cdot 0.6}{0.4} = 3.$$

$$\text{var}(X) = \frac{2 \cdot 0.6}{0.4^2} = 7.5.$$

Distribucions contínues

• Distribució uniforme

Siguen $a, b \in \mathbb{R}$, $a < b$ i siga X la V.A. aleatòria corresponent al resultat de seleccionar un nombre a l'atzar dins de l'interval $[a, b]$.

Definició Una V.A. aleatòria amb les condicions anteriors es diu que segueix una **distribució uniforme**.

Es denota per $X \sim U(a, b)$.

La funció de densitat corresponent és

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}.$$

L'esperança i la variància són per tant

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \left[\frac{x^2}{2(b-a)} \right]_{x=a}^{x=b} = \frac{b^2 - a^2}{2(b-a)} = \frac{(b+a)(b-a)}{2(b-a)} \\ &= \frac{a+b}{2}. \\ \text{var}(X) &= E(X^2) - E(X)^2 = \int_{-\infty}^{+\infty} x^2 f_X(x) dx - \left(\frac{a+b}{2} \right)^2 = \int_a^b \frac{x^2}{b-a} dx - \left(\frac{a+b}{2} \right)^2 \\ &= \left[\frac{x^3}{3(b-a)} \right]_{x=a}^{x=b} - \left(\frac{a+b}{2} \right)^2 = \frac{b^3 - a^3}{3(b-a)} - \left(\frac{a+b}{2} \right)^2 \\ &= \frac{(b-a)(a^2 + ab + b^2)}{3(b-a)} - \frac{a^2 + 2ab + b^2}{4} = \frac{4(a^2 + ab + b^2)}{12} - \frac{3(a^2 + 2ab + b^2)}{12} \\ &= \frac{a^2 - 2ab + b^2}{12} = \frac{(b-a)^2}{12}. \end{aligned}$$

En aquest cas concret, es pot calcular la seua funció de distribució $F_X(x) = p(X \leq x)$, $x \in \mathbb{R}$ de manera explícita, que és

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}.$$

Exercici 10 Comprova que al funció de distribució d'una V.A. aleatòria uniforme té efectivament l'expressió anterior.

Indicació: Recorda la relació entre la funció de distribució i la funció de densitat en el cas de les V.A. contínues:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt, \quad \forall x \in \mathbb{R}.$$

Cal tenir en compte també que $f_X(x)$ s'anul·la fora de l'interval $[a, b]$.

- **Distribució exponencial**

Definició Siga $\lambda > 0$. Direm que una V.A. X segueix una **distribució exponencial** si la seua funció de distribució és

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

Es denota per $X \sim \text{Exp}(\lambda)$.

La seua esperança, variància i desviació típica són

$$\begin{aligned} E(X) &= \frac{1}{\lambda}. \\ \text{var}(X) &= \frac{1}{\lambda^2}. \\ +\sqrt{\text{var}(X)} &= +\sqrt{\frac{1}{\lambda^2}} = \frac{1}{\lambda}. \end{aligned}$$

Exercici 11 Demuestra les dues primeres igualtats anteriors.

Indicació: Integra per parts en els dos casos.

La funció de distribució corresponent és

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

Exercici 12 Comprova que efectivament la funció de distribució té l'expressió anterior.

Exemple Un estudi sobre obsolescència programada indica que la vida útil (en anys) d'un cert electrodomèstic segueix una distribució exponencial de paràmetre $\frac{1}{2}$. Calcula la probabilitat que aquest aparell dure més de 3 anys i mig.

Solució Siga X la V.A. que indica el moment (en anys) en què l'electrodomèstic deixa de funcionar. Aleshores $X \sim \text{Exp}(\frac{1}{2})$.

Per tant, la probabilitat que dure més de 3 anys i mig (és a dir, el contrari que dure 3 anys i mig o menys) és

$$p(X > 3.5) = 1 - p(X \leq 3.5) = 1 - F_X(3.5) = 1 - (1 - e^{-\frac{1}{2} \cdot 3.5}) = e^{-1.75} \approx 0.1738.$$

D'altra banda, es té

$$\begin{aligned} E(X) &= \frac{1}{\lambda} = \frac{1}{\frac{1}{2}} = 2. \\ \text{var}(X) &= \frac{1}{\lambda^2} = \frac{1}{(\frac{1}{2})^2} = 4. \\ +\sqrt{\text{var}(X)} &= \frac{1}{\frac{1}{2}} = 2 \end{aligned}$$

Per tant, la vida útil mitjana per a l'aparell és de 2 anys, amb una desviació típica de també 2 anys.

• Distribució normal

La distribució normal és, dins de les distribucions contínues, la més important i sovint present en la naturalesa. Constitueix també els fonaments de la inferència estadística, com veurem en la propera secció.

Definició Una V.A. X es diu que segueix una **distribució normal** de mitjana $\mu \in \mathbb{R}$ i desviació típica $\sigma > 0$ si la seua funció de densitat és

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$$

on $\exp(y) = e^y$.

Es denota per $X \sim N(\mu, \sigma)$.

En el cas que $\mu = 0$ i $\sigma = 1$, és a dir, $X \sim N(0, 1)$, la distribució rep el nom de **normal tipificada**.

En general, donada $Z \sim N(0, 1)$ i un valor de probabilitat $0 \leq \alpha \leq 1$, es denota per $z_\alpha \in \mathbb{R}$ al punt tal que $p(Z \leq z_\alpha) = \alpha$.

La seua esperança, variància i desviació típica són

$$\begin{aligned} E(X) &= \mu. \\ \text{var}(X) &= \sigma^2. \\ +\sqrt{\text{var}(X)} &= +\sqrt{\sigma^2} = \sigma. \end{aligned}$$

L'inconvenient principal que suposa treballar amb aquesta distribució és que **la funció f_X no té primitiva elemental coneguda**; en altres paraules, no es pot calcular analíticament expressions com, donada $X \sim N(\mu, \sigma)$,

$$p(X \leq x) = F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

Per a fer-ho, ens valdrem d'una taula d'aproximacions per a la normal tipificada (adjunta al final del formulari) i del següent resultat, que permet *tipificar* una V.A. normal de paràmetres qualsevol.

Propietat Siga $\mu \in \mathbb{R}$, $\sigma > 0$ i $X \sim N(\mu, \sigma)$. Aleshores la nova variable aleatòria

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

compleix $Z \sim N(0, 1)$.

Exemple L'alçada dels homes d'un determinat país amb edats compreses entre 20 i 25 anys segueix una distribució normal de mitjana 175 cm i desviació típica 10 cm. Si se selecciona un home en aquest rang d'edat a l'atzar, quina és la probabilitat que la seua alçada siga superior a 177.5 cm?

Solució Donada X la V.A. corresponent a l'edat d'un individu triat en el grup esmentat a l'enunciat, es té que $X \sim N(175, 10)$ i es demana $p(X > 177.5)$.

El que anem a fer primer és convertir la condició que apareix a l'expressió de la probabilitat en una altra equivalent que involucre la normal tipificada, ço és, la variable

$$Z = \frac{X - 175}{10} \sim N(0, 1).$$

Per a fer-ho, anem a restar el valor 175 i posteriorment dividir per 10 a ambdós membres de la desigualtat que apareix a l'expressió de probabilitat:

$$p(X > 177.5) = p\left(\frac{X - 175}{10} > \frac{177.5 - 175}{10}\right) = p(Z > 0.25) = 1 - p(Z \leq 0.25) = 1 - F_Z(0.25).$$

Mirant a la taula per a la normal tipificada, identifiquem z_α amb el valor que apareix al costat dret de la desigualtat, en aquest cas, $z_\alpha = 0.25$ i α com el valor de la probabilitat, que és el que volem calcular.

Els valors de la primera columna corresponen als dígit de z_α i la seua primera xifra decimal, per tant en aquest cas hem de cercar el valor 0.2, que es troba a la quarta fila. D'altra banda, els valors de la primera fila es corresponen amb la segona xifra decimal, que en aquest cas és 5, i per tant triem la setena columna, corresponent al valor 0.05.

Així doncs, el valor de la probabilitat corresponent amb la quarta fila i setena columna és $\alpha = 0.598706$, i per tant

$$p(X > 177.5) = 1 - F_Z(0.25) = 1 - 0.598706 = 0.401294.$$

Si ens fixem en la taula de la distribució normal tipificada, aquesta només proporciona probabilitats $p(Z \leq x)$ majors o iguals que 0.5, corresponents a valors $z \geq 0$ (és a dir, a la dreta del seu eix de simetria). Per a abordar probabilitats menors que 0.5, o el que és el mateix, associades a valors $z < 0$, podem fer ús del següent resultat, que es beneficia de la simetria d'aquesta distribució.

Propietat Siga $Z \sim N(0, 1)$ i $z \in \mathbb{R}$. Aleshores

$$p(Z \leq -z) = p(Z \geq z) = 1 - p(Z \leq z).$$

Exemple Amb les mateixes condicions de l'exemple anterior, suposem ara que es desitja conèixer la probabilitat que un home triat a l'atzar tinga una alçada inferior a 172.5 cm.

Llavors tenim, emprant el resultat anterior:

$$\begin{aligned} p(X < 172.5) &= p\left(\frac{X - 175}{10} < \frac{172.5 - 175}{10}\right) = p(Z < -0.25) = p(Z > 0.25) \\ &= 1 - p(Z \leq 0.25) = 1 - F_Z(0.25) = 1 - 0.598706 = 0.401294. \end{aligned}$$

Propietat Siga $\{X_i\}_{i=1}^n$ una col·lecció de V.A. normals de mitjana μ i variància σ^2 ($X_i \sim N(\mu, \sigma)$). Aleshores la variable resultat del promig d'aquestes,

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

és normal amb mitjana μ i variància $\frac{\sigma^2}{n}$ ($\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$).

Exemple Novament sota les condicions de l'exemple anterior, suposem que s'agafa una mostra de $n = 100$ individus. Es demana obtenir la probabilitat que la mitjana de les seues alçades supere els 177.5 cm.

Solució Siga $\bar{X} = \frac{1}{100} \sum_{i=1}^{100} X_i$. Aleshores, pel resultat anterior, es té que $\bar{X} \sim N(175, \frac{10}{\sqrt{100}}) = N(175, 1)$. Definim

$$Z = \frac{\bar{X} - 175}{1} \sim N(0, 1).$$

Lavors tenim el següent:

$$\begin{aligned} p(X > 177.5) &= p\left(\frac{X - 175}{1} > \frac{177.5 - 175}{1}\right) = p(Z > 2.5) = 1 - p(Z \leq 2.5) = 1 - F_Z(2.5) \\ &= 1 - 0.993790 = 0.00621. \end{aligned}$$

Teorema (Teorema central del límit) Siga $\{X_i\}_{i=1}^n$ una col·lecció de V.A. independents, idènticament distribuïdes (no necessàriament normals), amb esperances $E(X_i) = \mu$ i variàncies $\text{var}(X_i) = \sigma^2$. Siga $Z \sim N(0, 1)$ i definim

$$Y_n = \sum_{i=1}^n X_i.$$

Aleshores, donat $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} p\left(\frac{Y_n - n\mu}{\sqrt{n}\sigma} \leq x\right) = p(Z \leq x).$$

En altres paraules, la V.A. $\sum_{i=1}^n X_i$, té un comportament cada vegada més semblant al d'una normal de mitjana $n\mu$ i desviació típica $\sqrt{n}\sigma$ segons el nombre d'aquestes variables va augmentant.

Exemple Es llança una moneda equilibrada 400 vegades i es desitja conèixer la probabilitat que isquen 210 cares o menys.

Solució Siga Y_{400} la V.A. que quantifica el nombre de cares obtingudes en 400 llançaments. D'acord amb el que ja hem vist, es compleix que $Y_{400} \sim B(400, 0.5)$. D'altra banda, les variables binomials es construeixen a partir de sumes de binomials, que en aquest cas prenen el valor 1 en cas d'eixir cara (èxit) i 0 en cas d'eixir creu (fracàs), per tant la suma és en efecte el nombre total de cares obtingudes:

$$Y_{400} = \sum_{i=1}^{400} X_i,$$

on $X_i \sim Be(0.5)$ es correspon al resultat obtingut al llançament i (1 si és cara i 0 si és creu). Com que $E(X_i) = 0.5$ i $\text{var}(X_i) = 0.5 \cdot (1 - 0.5) = 0.25$ (i per tant desviació típica $\sqrt{\text{var}(X_i)} = \sqrt{0.25} = 0.5$), aplicant el teorema central del límit es té que Y_{400} és aproximadament una normal de mitjana $400 \cdot 0.5 = 200$ i desviació típica $\sqrt{400} \cdot 0.5 = 100$. Per tant, la variable $\tilde{Z} = \frac{Y_{400} - 200}{100}$ és aproximadament una normal tipificada i, en conseqüència, si $Z \sim N(0, 1)$, aleshores

$$p(Y_{400} \leq 210) = p\left(\frac{Y_{400} - 200}{100} \leq \frac{210 - 200}{100}\right) = p(\tilde{Z} \leq 0.1) \approx p(Z \leq 0.1) \approx 0.539828.$$

Exercici 13 Repeteix l'exemple anterior calculant ara la probabilitat que el nombre de cares siga major o igual que 190. Interpreta el resultat.

• Distribucions deduïdes de la normal

Les distribucions que es presenten a continuació són també d'especial importància des del punt de vista de la inferència estadística.

– χ^2 de Pearson

Definició Siga $\{Z_i\}_{i=1}^n$ una col·lecció d' n V.A. independents tals que $Z_i \sim N(0, 1)$. La V.A. obtenida a partir de la suma dels quadrats, $X_n^2 = \sum_{i=1}^n Z_i^2$, es coneix com a distribució χ^2 de Pearson amb n graus de llibertat, i es denota per $X_n^2 \sim \chi_n^2$. En general, donat n un grau de llibertat i $0 \leq \alpha \leq 1$ un valor de probabilitat, es denota per $x_{n,\alpha} \geq 0$ al punt tal que $p(X_n^2 \leq x_{n,\alpha}) = \alpha$.

Igual que en el cas de la normal, per calcular les probabilitats associades a aquesta distribució cal emprar taules d'aproximacions. En el cas d'aquesta distribució, a la primera columna venen donats diferents graus de llibertat i a la primera fila diferents probabilitats. Així, fixats un grau de llibertat n i una probabilitat α , el valor de fila i columna corresponent donen lloc al punt $x_{n,\alpha}$ tal que $p(X_n^2 \leq x_{n,\alpha}) = \alpha$, on $X_n^2 \sim \chi_n^2$.

Exemple Siga $X \sim \chi_{10}^2$. Calcula el valor x tal que $p(X \leq x) = 0.25$.

Solució Mirant a les taules, fixant la fila corresponent a 10 graus de llibertat i la columna corresponent a una probabilitat de 0.25, el valor que obtenim és $x \approx 6.7372$.

Notem que a les taules només hi apareixen valors fins a 30 graus de llibertat. En el cas que $n > 30$, ens podem valdre del següent resultat.

Propietat Siga $X_n^2 \sim \chi_n^2$, $Z \sim N(0, 1)$ i definim la variable

$$Z_n = \sqrt{2X_n^2} - \sqrt{2n-1}.$$

Aleshores

$$\lim_{n \rightarrow \infty} p(Z_n \leq z) = p(Z \leq z) \quad \forall z \in \mathbb{R}.$$

En altres paraules, $\sqrt{2X_n^2}$ es comporta asimptòticament com una $N(\sqrt{2n-1}, 1)$.

Exemple Siga $X \sim \chi_{41}^2$. Calcula el valor x tal que $p(X \leq x) = 0.77$.

Solució Siga $Y = \sqrt{2X}$ i $Z \sim N(0, 1)$. Aleshores Y és aproximadament una $N(\sqrt{2 \cdot 41 - 1}, 1) = N(9, 1)$, i per tant tenim

$$0.77 = p(X \leq x) = p(\sqrt{2X} \leq \sqrt{2x}) = p(Y \leq y),$$

on $y = \sqrt{2x}$. D'altra banda, com que $Y \approx N(9, 1)$,

$$p(Y \leq y) = p\left(\frac{Y-9}{1} \leq \frac{y-9}{1}\right) = p(Y-9 \leq y-9) \approx p(Z \leq y-9) = F_Z(y-9) = F_Z(z),$$

on $z = y - 9$.

Així doncs, si resollem $F_Z(z) = 0.77$ obtindrem una aproximació del valor de z .

En aquest cas, hem de mirar a les taules de la normal tipificada el valor z_α que té una probabilitat $\alpha = 0.77$, cercant el valor més proper a aquest en la taula, observant a quina fila i columna es troba i, a partir d'elles, deduint el valor amb les dues xifres decimals corresponents. En aquest cas, el valor més proper a 0.77 que apareix a la taula és 0.770350 a la fila on apareix a la primera columna el valor 0.7 i a la columna on apareix a la primera fila el valor 0.04. Per tant, $z \approx 0.74$.

Desfent els canvis de variables, obtenim, d'una banda,

$$y = z + 9 \approx 9.74,$$

i d'altra banda, com que $y = \sqrt{2x}$, llavors

$$x = \frac{y^2}{2} \approx \frac{9.74^2}{2} = 47.4338.$$

– Distribució t de Student

Definició Siguen X, Y dues V.A. independents tals que $X \sim N(0, 1)$ i $Y \sim \chi_n^2$, per a un cert $n \in \mathbb{N}$. La V.A.

$$T_n = \frac{X}{\sqrt{\frac{Y}{n}}}$$

es diu que segueix una **distribució t de Student** amb n graus de llibertat, i es denota per $T_n \sim t_n$. En general, donat n un grau de llibertat i $0 \leq \alpha \leq 1$ un valor de probabilitat, es denota per $t_{n,\alpha} \in \mathbb{R}$ al punt tal que $p(T_n \leq t_{n,\alpha}) = \alpha$.

Igual com ocorre en els dos casos anteriors, per calcular probabilitats relacionades amb aquesta distribució cal fer servir taules d'aproximacions, l'ús de la qual és idèntic a la del cas anterior.

Exemple Calcula el valor t tal que $p(T \leq t) = 0.95$, on $T \sim t_{20}$.

Solució Mirant a la taula de la distribució t , el valor que cerquem es troba a la fila corresponent a $n = 20$ graus de llibertat i la columna associada a una probabilitat de 0.95, i per tant $t \approx 1.72472$.

Igual que en el cas de la normal tipificada, aquesta distribució disposa de simetria respecte del 0, fet del qual es dedueix la següent propietat anàlega a aquesta.

Propietat Siga $n \in \mathbb{N}$, $T \sim t_n$ i $t \in \mathbb{R}$. Aleshores

$$p(T \leq -t) = p(T \geq t) = 1 - p(T \leq t).$$

Exemple Calcula el valor t tal que $p(T \leq t) = 0.05$, on $T \sim t_{20}$.

Solució Per la propietat anterior, es té

$$p(T \leq -t) = 1 - p(T \leq t) = 1 - 0.05 = 0.95.$$

Així doncs, es dedueix de l'exemple anterior que $-t \approx 1.72472$, i per tant $t \approx -1.72472$.

Finalment, quan el grau llibertat supera el valor de 30, és a dir, $n > 30$, podem fer ús del següent resultat.

Propietat Siga $n \in \mathbb{N}$, $T_n \sim t_n$, $Z \sim N(0, 1)$ i $t \in \mathbb{R}$. Aleshores

$$\lim_{n \rightarrow \infty} p(T_n \leq t) = p(Z \leq t).$$

Dit d'una altra manera, una distribució t de Student es comporta asimptòticament com una normal tipificada segons van augmentant els graus de llibertat.

A la pràctica, quan en un problema tinguem $n > 30$ graus de llibertat, emprarem al seu lloc la taula de la normal tipificada.

Exemple Calcula el valor t tal que $p(T \leq t) = 0.975$, on $T \sim t_{50}$.

Solució Pel resultat anterior, $T \approx N(0, 1)$, i per tant, si $Z \sim N(0, 1)$, es té

$$0.975 = p(T \leq t) \approx p(Z \leq t).$$

Per tant, hem de trobar el valor de la taula de la normal tipificada que done el valor de probabilitat més proper a 0.975. Aquest és $z = 1.96$, amb un probabilitat aproximada de 0.975002, i per tant es té $t \approx 1.96$.

– Distribució F de Snedecor

Definició Siga $X \sim \chi_m^2$ i $Y \sim \chi_n^2$. Aleshores la V.A.

$$F_{m,n} = \frac{\frac{X}{m}}{\frac{Y}{n}}$$

es diu que segueix una **distribució F de Snedecor** amb (m, n) graus de llibertat i es denota per $F_{m,n} \sim \mathcal{F}_{m,n}$. En general, donat (m, n) dos graus de llibertat i $0 \leq \alpha \leq 1$ un valor de probabilitat, es denota per $f_{m,n,\alpha} \geq 0$ al punt tal que $p(F_{m,n} \leq f_{m,n,\alpha}) = \alpha$.

Per calcular les probabilitats d'aquesta distribució fem servir les taules que apareixen al formulari, un total de dues per a les probabilitats 0.95 i 0.99. En elles, el grau de llibertat m va variant en funció de la fila i n en funció de la columna.

Exemple Siga $F \sim \mathcal{F}_{20,30}$. Obtén el punt $f \in \mathbb{R}$ tal que $p(F \leq f) = 0.95$.

Solució Mirant a les taules, el valor corresponent a la fila $m = 20$ i columna $n = 30$ és 2.03909, i per tant $f \approx 2.03909$.

En alguns casos podem necessitar calcular valors de probabilitat petits (0.05 o 0.01). Per a fer-ho, podem fer servir el següent resultat.

Propietat Siga $F_{m,n} \sim \mathcal{F}_{m,n}$, $0 \leq \alpha \leq 1$ i $f_{m,n}$ tal que $p(F_{m,n} \leq f_{m,n}) = \alpha$. Aleshores

$$p(F_{n,m} \leq \frac{1}{f_{m,n}}) = 1 - \alpha,$$

on $F_{n,m} \sim \mathcal{F}_{n,m}$.

Dit d'una altra manera,

$$f_{m,n,\alpha} = \frac{1}{f_{n,m,1-\alpha}}.$$

Exemple Troba el valor f tal que $p(F \leq f) = 0.05$, on $F \sim \mathcal{F}_{30,20}$.

Solució El valor f que cerquem és

$$f = f_{30,20,0.05} = \frac{1}{f_{20,30,1-0.05}} = \frac{1}{f_{20,30,0.95}} \approx \frac{1}{2.03909} \approx 0.49041,$$

on $f_{20,30,0.95}$ és el valor calculat a l'exemple anterior.

Exercici 14

(a) La probabilitat de curació d'una determinada malaltia en sotmetre al pacient a un tractament determinat és d'un 70%.

1. Si se sotmeten al tractament 10 persones, quina és la probabilitat que es curen més de 7?
2. Si se sotmeten al tractament 1000 persones, quina és la probabilitat que es curen més de 700?

(b) Siga $X \sim N(114.3, 5.4)$. Calcula:

1. $p(X \leq 115)$.
2. $p(X \geq 117)$.
3. $p(X \leq 112.5)$.
4. $p(X \geq 110)$.
5. El valor x tal que $p(X \leq x) = 0.95$.
6. El valor x tal que $p(X \geq x) = 0.05$.
7. El valor x tal que $p(X \leq x) = 0.2$.
8. El valor x tal que $p(X \geq x) = 0.82$.

(c) Siga $X \sim \chi_n^2$. Es demana calcular:

1. Per a $n = 12$, el valor x tal que $p(X \leq x) = 0.9$.
2. Per a $n = 12$, el valor x tal que $p(X \geq x) = 0.9$.
3. Per a $n = 17$, $p(X \leq 10)$.
4. Per a $n = 19$, $p(X \geq 14.5)$.
5. Per a un n qualsevol, $p(X \geq 0)$.

(d) Siga $T \sim t_n$. Es demana:

1. Per a $n = 25$, el valor t tal que $p(T \geq t) = 0.75$.
2. Per a $n = 25$, $p(T \leq 2.06)$.
3. Per a $n = 40$, $p(T \leq 2.06)$.
4. Per a $n = 40$, el valor t tal que $p(T \leq t) = 0.01$.

(e) Siga $F \sim \mathcal{F}_{15,12}$. Obtén:

1. El valor f tal que $p(F \leq f) = 0.99$.
2. El valor f tal que $p(F \geq f) = 0.99$.
3. $p(F \leq 2.47531)$.
4. $p(F \geq 0.40399)$.

1.3 Inferència estadística

En aquesta secció s'estudien procediments que permeten prendre decisions en problemes de naturalesa aleatòria. L'objectiu principal és l'obtenció d'informació d'un conjunt d'elements (població) solament es pot (o vol) accedir parcialment (mostra).

Si un determinat model probabilístic es considera adequat per descriure les dades sotmeses a l'estudi, es desitja estimar els seus paràmetres a partir de la informació disponible, és a dir, la mostra. Aquests paràmetres, que són quantitats relacionades amb la població, reben el nom de **paràmetre poblacional**.

Per tal d'estimar un paràmetre poblacional, es fa ús de la **teoria del mostreig**, amb dues maneres generals de procedir:

- **Estimació puntual:** Aproximació d'un paràmetre poblacional mitjançant un valor calculat a partir de la mostra. Per exemple, la mitjana i la variància mostrals, ja estudiades a la secció d'estadística descriptiva.
- **Intervals de confiança:** Obtenció d'un interval, a partir d'una mostra, on se situa el paràmetre poblacional amb una certa probabilitat. Aquesta secció es centrarà en aquest punt.

Exemple Siga X la V.A. que quantifica l'altura d'una població formada per nens de 12 anys i suposem que $X \sim N(\mu, \sigma)$. Seleccionem una mostra formada per 15 individus d'aquesta població, amb valors mostrals $\{x_i\}_{i=1}^{15}$, corresponents a les seues altures.

La nostra població és per tant el conjunt de tots els nens de 12 anys, mentre que la mostra que hem seleccionat està formada per 15 individus d'eixa població, és a dir, 15 nens de 12 anys.

Definició (Estimadors no esbiaixats) Un estimador $\hat{\theta}$ d'un paràmetre θ d'una població s'anomena **no esbiaixat** quan la seua esperança matemàtica és igual al paràmetre en qüestió, és a dir:

$$E(\hat{\theta}) = \theta.$$

En cas contrari s'anomena **esbiaixat**.

Tot seguit analitzarem un seguit d'estimadors per a la mitjana i la variància d'una població normal.

Siguen $\{X_i\}_{i=1}^n$ tals que $X_i \sim N(\mu, \sigma)$, $1 \leq i \leq n$. Es té que $E(X_i) = \mu$ i $\text{var}(X_i) = \sigma^2$, $1 \leq i \leq n$. Considerem la variable aleatòria consistent en fer la mitjana de les V.A. inicials:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Aleshores $\bar{X}$ és un estimador no esbiaixat de la mitjana poblacional μ . En efecte,

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \mu.$$

No obstant això, l'estimador corresponent a la V.A. consistent en la variància de les V.A. originals:

$$D^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

és un estimador esbiaixat de la variància poblacional σ^2 , ja que es pot comprovar que

$$E(D^2) = \frac{n-1}{n} \sigma^2 \neq \sigma^2.$$

Exercici 15 *Demostra la igualtat anterior.*

Una manera de solucionar el problema anterior és, fent benefici de la linealitat de l'esperança (vista en les seues propietats) és considerar la nova V.A.

$$S^2 = \frac{n}{n-1} D^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2,$$

que és un estimador no esbiaixat de σ^2 ja que

$$E(S^2) = E\left(\frac{n}{n-1} D^2\right) = \frac{n}{n-1} E(D^2) = \frac{n}{n-1} \cdot \frac{n-1}{n} \sigma^2 = \sigma^2.$$

Açò motiva la introducció de la següent definició.

Definició Donada una mostra aleatòria $\{x_i\}_{i=1}^n$ es defineix la seua **quasivariància** com

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

on

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

és la mitjana de la mostra.

Es defineix la **quasidesviació típica** com l'arrel quadrada positiva de la quasivariància

$$S = +\sqrt{S^2}.$$

Es recopilen a continuació un seguit de resultats que emprarem per al càlcul d'interval de confiança, algunes de les quals ja hem vist o són senzilles de deduir.

Propietats Siga $\{X_i\}_{i=1}^n$ una col·lecció de V.A. complint $X_i \sim N(\mu, \sigma)$ i definim les següents variables aleatòries:

$$\begin{aligned} \bar{X} &= \frac{1}{n} \sum_{i=1}^n X_i, \\ Y &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2}, \\ Z &= \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}, \\ T &= \frac{Z}{\sqrt{\frac{Y}{n-1}}} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{S}. \end{aligned}$$

Aleshores es compleix:

- $\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$.
- $Y \sim \chi_{n-1}^2$.
- $\bar{X}, Y$ són independents.
- $Z \sim N(0, 1)$.
- $T \sim t_{n-1}$.

1.3.1 Intervalls de confiança

Fins ara hem estat treballant en casos en què es coneixen els paràmetres poblacionals, és a dir, en el cas de V.A. normals $N(\mu, \sigma)$, la mitjana μ i la variància σ . No obstant això, en la pràctica aquesta informació manca i cal fer ús de procediments d'inferència estadística per tal d'obtenir informació sobre aquests a partir d'una **mostra** obtinga de la població en qüestió.

Els **intervalls de confiança** associats a un paràmetre determinat consisteixen en un rang de valors on dit paràmetre es troba amb una certa probabilitat $\alpha = 1 - \varepsilon$, que anomenarem **nivell de confiança**. El marge o probabilitat d'error és per tant $\varepsilon = 1 - \alpha$, que rep el nom de **significació**.

Suposem que agafem dades de les V.A. $\{X_i\}_{i=1}^n$ d'una població normal $N(\mu, \sigma)$, amb valors mostrals $\{x_i\}_{i=1}^n$. Aleshores podem fer diferents anàlisis en funció del paràmetre que es desitja estimar i de la informació disponible.

Intervalls de confiança per a la mitjana

Suposem que en una població que segueix una distribució $N(\mu, \sigma)$ es desitja estimar la mitjana μ . Llavors tenim les possibilitats que es presenten a continuació.

- **La variància σ^2 és coneguda.** En aquest cas, pel que hem vist anteriorment es compleix que

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1).$$

Donat un nivell de significació $0 < \varepsilon < 1$, ens plantejem trobar el valor $K > 0$ tal que

$$p(|Z| \leq K) = 1 - \varepsilon.$$

Descomposant el primer membre de la igualtat, obtenim

$$\begin{aligned} p(|Z| \leq K) &= p(-K \leq Z \leq K) = p(Z \leq K) - p(Z \leq -K) = p(Z \leq K) - p(Z \geq K) \\ &= p(Z \leq K) - (1 - p(Z \leq K)) = 2p(Z \leq K) - 1. \end{aligned}$$

Així doncs, es té

$$2p(Z \leq K) - 1 = 1 - \varepsilon \iff p(Z \leq K) = 1 - \frac{\varepsilon}{2}.$$

Obtenim per tant que K resulta ser el valor complint que la probabilitat que deixa a la seua esquerra una normal tipificada és $1 - \frac{\varepsilon}{2}$, és a dir, $K = z_{1-\frac{\varepsilon}{2}}$.

Substituint ara pels corresponents valor mostrals, la condició inicial de la probabilitat es tradueix com

$$\left| \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \right| \leq z_{1-\frac{\varepsilon}{2}},$$

ço és

$$-z_{1-\frac{\varepsilon}{2}} \leq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{1-\frac{\varepsilon}{2}}.$$

Aïllant la mitjana μ s'obté per tant

$$\bar{x} - z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}}.$$

I per tant, l'interval de confiança per a la mitjana amb un nivell de confiança $\alpha = 1 - \varepsilon$ (és a dir, que conté la mitjana poblacional amb una probabilitat α) és

$$\left[\bar{x} - z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}} \right].$$

Exemple Considerem la següent mostra resultant d'obtenir les alçades de 7 nens de 12 anys d'un cert país, procedent d'una població $N(\mu, 2)$.

$$\{143.2, 140.3, 141.4, 139.9, 142.7, 144.1, 142.7\}.$$

Es demana calcular un interval de confiança amb un 90% de nivell de confiança per a l'altura mitjana.

Solució Ens trobem en un cas en què es demana un interval de confiança per a la mitjana μ amb variància coneguda ($\sigma^2 = 2^2 = 4$); per tant, estem en condicions d'aplicar la fórmula anterior.

Com que el nivell de confiança és $\alpha = 0.9$, es té que el nivell de significació és $\varepsilon = 1 - \alpha = 1 - 0.9 = 0.1$. Calculant els paràmetres necessaris

$$\bar{x} = \frac{143.2 + 140.3 + 141.4 + 139.9 + 142.7 + 144.1 + 142.4}{7} = 142$$

$$z_{1-\frac{\varepsilon}{2}} = z_{0.95} \approx 1.645$$

$$\sigma = 2$$

$$\sqrt{n} = \sqrt{7} \approx 2.646$$

s'obté que l'interval de confiança per a la mitjana amb $\alpha = 0.9$ és aproximadament

$$\left[142 - 1.645 \cdot \frac{2}{2.646}, \quad 142 + 1.645 \cdot \frac{2}{2.646} \right] \approx [140.76, 143.24].$$

En altres paraules, a partir de les dades de la mostra, s'infereix que l'altura mitjana de la població formada pels nens de 12 anys del país objecte d'estudi es troba, amb una probabilitat d'un 90%, dins de l'interval $[140.76, 143.24]$.

- **La variància σ^2 és desconeguda.** Per tal de calcular l'interval de confiança en aquest cas fem ús de la propietat

$$T = \frac{(\bar{X} - \mu)}{\frac{S}{\sqrt{n}}} \sim t_{n-1}.$$

Donat un nivell de significació $0 < \varepsilon < 1$, ens plantegem trobar el valor $K > 0$ tal que

$$p(|T| \leq K) = 1 - \varepsilon.$$

Raonant exactament de la mateixa manera que abans, obtenim que en aquest cas $K = t_{n-1, 1-\frac{\varepsilon}{2}}$ i l'interval de confiança per a la mitjana resulta

$$\left[\bar{x} - t_{n-1, 1-\frac{\varepsilon}{2}} \frac{S}{\sqrt{n}}, \quad \bar{x} + t_{n-1, 1-\frac{\varepsilon}{2}} \frac{S}{\sqrt{n}} \right].$$

Exemple Amb les dades del problema anterior, suposant que en aquest cas la desviació típica poblacional σ no ve donada, es demana calcular un interval de confiança per a la mitjana amb un nivell de confiança d'un 90%.

Solució La mitjana mostral, ja calculada a l'exemple anterior, és $\bar{x} = 142$. D'altra banda, es té

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

$$= \frac{143.2^2 + 140.3^2 + 141.4^2 + 139.9^2 + 142.7^2 + 144.1^2 + 142.4^2}{6} - \frac{7}{6} \cdot 142^2 = 2.36$$

$$S = \sqrt{S^2} \approx 1.536$$

$$t_{n-1, 1-\frac{\varepsilon}{2}} = t_{6, 0.95} \approx 1.94318.$$

Així doncs, l'interval de confiança aproximat resulta ser

$$\left[142 - 1.94318 \cdot \frac{1.536}{2.646}, \quad 142 + 1.94318 \cdot \frac{1.536}{2.646} \right] \approx [140.87, 143.13].$$

- **Intervals de confiança per a la diferència de dues mitjanes.**

Siguen $X \sim N(\mu_X, \sigma_X)$ i $Y \sim N(\mu_Y, \sigma_Y)$, amb respectives mostres generades $\{x_i\}_{i=1}^m$, $\{y_i\}_{i=1}^n$. Per tal de veure si hi ha diferències significatives en el compartament del valor promig que pren cada variable de les dues poblacions, ens podem plantejar el càlcul d'un interval de confiança per a la diferència entre les seues dues mitjanes, $\mu_X - \mu_Y$. Depenent de si coneixem o no les seues variàncies, tenim diferents possibilitats.

- **Les variàncies σ_X^2, σ_Y^2 són conegudes.**

En aquest cas, l'interval de confiança resulta

$$\left[(\bar{x} - \bar{y}) - z_{1-\frac{\varepsilon}{2}} \sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}, \quad (\bar{x} - \bar{y}) + z_{1-\frac{\varepsilon}{2}} \sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}} \right].$$

Podem denotar-lo d'una manera més abreviada com

$$(\bar{x} - \bar{y}) \pm z_{1-\frac{\varepsilon}{2}} \sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}},$$

on el signe “menys” fa referència a l'extrem esquerre de l'interval i el signe “més” a l'extrem dret.

Exemple Es desitja comparar les altures de la població dels exemples anteriors, que ve donada per 7 valors d'una V.A. $X \sim N(\mu, 2)$, corresponent a les altures dels nens de 12 anys d'un determinat país A, amb les altures d'un país B, procedents d'una V.A. $Y \sim N(\mu, 1.5)$ a partir de les següents dades mostrals

$$\{144.5, 142.3, 145.7, 143.1, 142.6\}.$$

Es desitja obtenir un interval amb un nivell de confiança d'un 90% per a la mitjana.

Solució Identificant els paràmetres que cal calcular i aprofitant els ja obtinguts als exemples anteriors, obtenim:

$$\begin{aligned} m &= 7 \\ \bar{x} &= 142 \\ \sigma_X^2 &= 4 \\ n &= 5 \\ \bar{y} &= \frac{144.5 + 142.3 + 145.7 + 143.1 + 142.6}{5} = 143.64 \\ \sigma_Y^2 &= 1.5^2 = 2.25 \\ z_{1-\frac{\epsilon}{2}} &= z_{0.95} \approx 1.645. \end{aligned}$$

Així doncs, l'interval de confiança resulta ser

$$(142 - 143.64) \pm 1.645 \cdot \sqrt{\frac{4}{7} + \frac{2.25}{5}} \approx -1.64 \pm 1.662 = [-3.302, 0.022].$$

– Les variàncies σ_X^2, σ_Y^2 són desconegudes, però iguals ($\sigma_X^2 = \sigma_Y^2$).

En aquest cas l'interval de confiança es pot calcular per

$$(\bar{x} - \bar{y}) \pm t_{m+n-2, 1-\frac{\epsilon}{2}} \sqrt{\left(\frac{1}{m} + \frac{1}{n}\right) \left(\frac{(m-1)S_X^2 + (n-1)S_Y^2}{m+n-2}\right)}.$$

Exemple Calcular un interval de confiança amb $\alpha = 0.9$ per a l'exemple anterior, suposant que ara les variàncies són desconegudes però iguals.

Solució Aprofitant totes les dades calculades anteriorment, solament és necessari el càlcul de S_Y^2 i $t_{m+n-2, 1-\frac{\epsilon}{2}}$, on

$$\begin{aligned} S_X^2 &= 2.36 \\ S_Y^2 &= \frac{144.5^2 + 142.3^2 + 145.7^2 + 143.1^2 + 142.6^2}{4} - \frac{5}{4} \cdot 143.64^2 = 2.038 \\ t_{m+n-2, 1-\frac{\epsilon}{2}} &= t_{10, 0.95} \approx 1.81246 \end{aligned}$$

Així doncs, l'interval de confiança és aproximadament

$$\begin{aligned} (142 - 143.64) \pm 1.81246 \sqrt{\left(\frac{1}{7} + \frac{1}{5}\right) \left(\frac{6 \cdot 2.36 + 4 \cdot 2.038}{10}\right)} &\approx -1.64 \pm 1.585 \\ &= [-3.225, -0.055]. \end{aligned}$$

– Les variàncies σ_X^2, σ_Y^2 són desconegudes i m, n són grans.

En cas que es desconeixi si les variàncies són iguals i els tamany de les respectives mostres, m, n , siguin *suficientment* grans, l'interval de confiança es pot aproximar pel corresponent al cas que les variàncies siguin conegudes, tot substituint les variàncies poblacionals σ_X^2, σ_Y^2 per les quasivariàncies mostrals S_X^2, S_Y^2 , és a dir,

$$(\bar{x} - \bar{y}) \pm z_{1-\frac{\epsilon}{2}} \sqrt{\frac{S_X^2}{m} + \frac{S_Y^2}{n}}.$$

Intervals de confiança per a la variància

Suposem ara que en una població que segueix una distribució $N(\mu, \sigma)$ el que es vol estimar és σ . Llavors tenim les possibilitats que es presenten tot seguit.

- **La mitjana μ és coneguda.** Siguen $\{X_i\}_{i=1}^n$ tals que $X_i \sim N(\mu, \sigma)$, $1 \leq i \leq n$. Llavors és conegut que $Z_i = \frac{X_i - \mu}{\sigma} \sim N(0, 1)$. Aleshores, per la definició de la χ^2 de Pearson es té

$$\sum_{i=1}^n Z_i^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi_n^2.$$

Donat un nivell de significació $\varepsilon > 0$, es té per tant que

$$p\left(x_{n, \frac{\varepsilon}{2}} \leq \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \leq x_{n, 1 - \frac{\varepsilon}{2}}\right) = 1 - \varepsilon,$$

o el que és el mateix,

$$p\left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{x_{n, 1 - \frac{\varepsilon}{2}}} \leq \sigma^2 \leq \frac{\sum_{i=1}^n (X_i - \mu)^2}{x_{n, \frac{\varepsilon}{2}}}\right) = 1 - \varepsilon.$$

Així doncs, l'interval de confiança que conté a la variància poblacional σ^2 amb una probabilitat d' $1 - \varepsilon$ (és a dir, amb un nivell de confiança $\alpha = 1 - \varepsilon$) és

$$\left[\frac{\sum_{i=1}^n (x_i - \mu)^2}{x_{n, 1 - \frac{\varepsilon}{2}}}, \frac{\sum_{i=1}^n (x_i - \mu)^2}{x_{n, \frac{\varepsilon}{2}}} \right] = \left[\frac{(n-1)S^2 + n(\bar{x} - \mu)^2}{x_{n, 1 - \frac{\varepsilon}{2}}}, \frac{(n-1)S^2 + n(\bar{x} - \mu)^2}{x_{n, \frac{\varepsilon}{2}}} \right].$$

Exercici 16 *Demostra que*

$$\sum_{i=1}^n (x_i - \mu)^2 = (n-1)S^2 + n(\bar{x} - \mu)^2.$$

Exemple *Continuem treballant amb les dades associades a la mostra consistent en les altures dels 7 nens de 12 anys del país A. Suposem ara que les dades procedeixen d'una població $N(141.8, \sigma)$ i es demana un interval de confiança amb un nivell de confiança d'un 90% per a la variància.*

Solució *Com que $\mu = 141.8$ i $\varepsilon = 0.1$, emprant les dades calculades amb anterioritat es té*

$$\sum_{i=1}^n (x_i - \mu)^2 = (n-1)S^2 + n(\bar{x} - \mu)^2 = 6 \cdot 2.36 + 7 \cdot (142 - 141.8)^2 = 14.44$$

$$x_{n, \frac{\varepsilon}{2}} = x_{7, 0.05} \approx 2.16735$$

$$x_{n, 1 - \frac{\varepsilon}{2}} = x_{7, 0.95} \approx 14.0671$$

Així doncs, l'interval de confiança que conté la variància amb un 90% de probabilitat és aproximadament

$$\left[\frac{14.44}{14.0671}, \frac{14.44}{2.16735} \right] \approx [1.027, 6.663].$$

• **La mitjana μ és desconeguda.**

Per les propietats anteriors, es compleix que

$$Y = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Raonant com abans, s'arriba a que

$$p \left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{x_{n-1, 1-\frac{\varepsilon}{2}}} \leq \sigma^2 \leq \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{x_{n-1, \frac{\varepsilon}{2}}} \right) = 1 - \varepsilon$$

i per tant l'interval de confiança resultant, amb un nivell de confiança $\alpha = 1 - \varepsilon$, és

$$\left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{x_{n-1, 1-\frac{\varepsilon}{2}}}, \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{x_{n-1, \frac{\varepsilon}{2}}} \right] = \left[\frac{(n-1)S^2}{x_{n-1, 1-\frac{\varepsilon}{2}}}, \frac{(n-1)S^2}{x_{n-1, \frac{\varepsilon}{2}}} \right].$$

Exemple Amb les mateixes condicions de l'exemple anterior suposem que la mitjana poblacional no ens ve donada i es demana el càlcul de l'interval de confiança per a $\alpha = 0.9$.

Solució Es calcula

$$x_{n-1, \frac{\varepsilon}{2}} = x_{6, 0.05} \approx 1.63538$$

$$x_{n-1, 1-\frac{\varepsilon}{2}} = x_{6, 0.95} \approx 12.5916$$

i l'interval resultant és aproximadament

$$\left[\frac{6 \cdot 2.36}{12.5916}, \frac{6 \cdot 2.36}{1.63538} \right] \approx [1.125, 8.659].$$

• **Interval de confiança per al quocient entre dues variàncies.**

Siguen $X \sim N(\mu_X, \sigma_X)$ i $Y \sim N(\mu_Y, \sigma_Y)$, amb respectives mostres generades $\{x_i\}_{i=1}^m$, $\{y_i\}_{i=1}^n$. Ara anem a comparar les variàncies de les respectives poblacions, σ_X, σ_Y , a través del seu quocient. Per a fer-ho, emprarem el fet que

$$\frac{S_X}{S_Y} \sim \mathcal{F}_{m-1, n-1},$$

d'on es dedueix que l'interval de confiança per al quocient $\frac{\sigma_X}{\sigma_Y}$ amb una significació ε és

$$\left[f_{m-1, n-1, \frac{\varepsilon}{2}} \frac{S_X^2}{S_Y^2}, f_{m-1, n-1, 1-\frac{\varepsilon}{2}} \frac{S_X^2}{S_Y^2} \right] = \left[\frac{1}{f_{n-1, m-1, 1-\frac{\varepsilon}{2}}} \frac{S_X^2}{S_Y^2}, f_{m-1, n-1, 1-\frac{\varepsilon}{2}} \frac{S_X^2}{S_Y^2} \right].$$

Exemple Es demana calcular l'interval de confiança per al quocient de variàncies de les dues mostres corresponents a les diferents altures dels 7 nens del país A i els 5 nens del país B, amb una significació d'un 10%.

Solució L'interval de confiança resultant per al quocient de les variàncies amb una significació $\varepsilon = 0.1$ és

$$\left[\frac{1}{f_{4, 6, 0.95}} \frac{2.36}{2.038}, f_{6, 4, 0.95} \frac{2.36}{2.038} \right] \approx \left[\frac{1}{4.53368} \frac{2.36}{2.038}, 6.16313 \cdot \frac{2.36}{2.038} \right] \approx [0.255, 7.137].$$

1.3.2 Contrast d'hipòtesis

Un contrast d'hipòtesi consisteix en prendre una decisió sobre el valor d'un paràmetre θ (desconegut) d'una certa població a partir de dades d'una mostra $\{x_i\}_{i=1}^n$ generades a partir de la població anteriorment esmentada. També és generalitzable a valors de diferents paràmetres a partir de dades de diferents mostres procedents de diverses poblacions. Es basa en contrastar una **hipòtesi nul·la**, H_0 , front a una **hipòtesi alternativa**, H_a , que es pren com a vertadera si es conclou que H_0 és falsa.

En prendre una decisió, és a dir, pronunciar-nos sobre H_0 o H_a , podem incórrer en una resolució incorrecta:

- Rebutjar H_0 quan aquesta és certa (error del primer tipus).
- Acceptar H_0 quan aquesta és falsa (error del segon tipus).

La postura que prendrem en tot moment és conservadora respecte de la hipòtesi nul·la, és a dir, serà rebutjada únicament quan existisquen indicis suficients de veracitat de la hipòtesi alternativa. Ens centrarem per tant en els errors del primer tipus i anomenem ε a la probabilitat de cometre un error del primer tipus, que anomenarem **nivell de significació**.

En els casos en què anem a treballar, la hipòtesi nul·la sempre serà una igualtat, és a dir, de la forma $H_0 : \theta = \theta_0$, on θ_0 és un valor donat, mentre que l'alternativa pot presentar les tres formes següents:

$$H_a : \begin{cases} \theta \neq \theta_0 \\ \theta > \theta_0 \\ \theta < \theta_0 \end{cases}$$

En el primer cas el contrast rep el nom de **bilateral** o en els dos darrers **unilateral**.

Contrast d'una mitjana

Tot seguit veurem una sèrie de contrastos d'hipòtesi per a la mitjana μ (desconeguda) d'una població normal $N(\mu, \sigma)$ a partir d'una mostra $\{x_i\}_{i=1}^n$ obtinguda d'aquesta. Donat un valor μ_0 , es proposa un contrast amb hipòtesi nul·la

$$H_0 : \mu = \mu_0$$

i les tres possibilitats següents per a la hipòtesi alternativa:

$$H_a : \begin{cases} \mu \neq \mu_0 \\ \mu > \mu_0 \\ \mu < \mu_0 \end{cases}$$

En funció de si es coneix o no la variància σ^2 , tenim les següents possibilitats.

- **La variància σ^2 és coneguda.**

En funció de la hipòtesi alternativa, tenim tres possibilitats, que es detallen a continuació.

- $H_a : \mu \neq \mu_0$.

L'interval de confiança per a la mitjana amb variància coneguda per a una significació ε és

$$\left[\bar{x} - z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}} \right].$$

Recordem que l'interval anterior conté amb una probabilitat $1 - \varepsilon$ la mitjana μ de la població d'on procedeix la mostra donada. Així doncs, si

$$\mu_0 \in \left[\bar{x} - z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

s'accepta H_0 , mentre que es rebutja en favor de H_a si

$$\mu_0 \notin \left[\bar{x} - z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}} \right].$$

La condició d'acceptació de H_0 és per tant

$$\bar{x} - z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + z_{1-\frac{\varepsilon}{2}} \frac{\sigma}{\sqrt{n}},$$

que operant tradueix en

$$\left| \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \right| \leq z_{1-\frac{\varepsilon}{2}}.$$

D'altra banda, es rebutja H_0 si es verifica el contrari, ço és,

$$\left| \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \right| > z_{1-\frac{\varepsilon}{2}}.$$

– $H_a : \mu > \mu_0$.

En aquest cas calculem l'interval de confiança per a la mitjana amb variància coneguda amb un nivell de significació 2ε :

$$\left[\bar{x} - z_{1-\varepsilon} \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + z_{1-\varepsilon} \frac{\sigma}{\sqrt{n}} \right].$$

Notem que amb aquesta significació, per arguments de simetria, la probabilitat que μ quede per fora de l'interval és ε per a cada extrem, sumant un total de 2ε .

Es considera que la desigualtat de la hipòtesi alternativa guanya una versemblança suficient quan μ_0 queda fora i al costat esquerre de l'interval de confiança, és a dir,

$$\mu_0 < \bar{x} - z_{1-\varepsilon} \frac{\sigma}{\sqrt{n}},$$

criteri que prendrem per a determinar el rebuig d' H_0 en favor d' H_a , i que operant es pot escriure com

$$\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > z_{1-\varepsilon}.$$

Per tant, la condició d'acceptació d' H_0 és

$$\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq z_{1-\varepsilon}.$$

Notem que la probabilitat de rebutjar incorrectament H_0 es correspon amb la probabilitat que μ realment estiga fora de l'interval i a la part esquerra, que ja hem raonat anteriorment que és ε .

- $H_a : \mu < \mu_0$. Es treballa amb el mateix interval que en el cas anterior:

$$\left[\bar{x} - z_{1-\varepsilon} \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + z_{1-\varepsilon} \frac{\sigma}{\sqrt{n}} \right].$$

En aquest cas, es considera que hi ha indicis suficients per rebutjar H_0 en favor d' H_a si μ_0 queda més enllà de l'extrem dret de l'interval, és a dir,

$$\mu_0 > \bar{x} + z_{1-\varepsilon} \frac{\sigma}{\sqrt{n}},$$

condició que es pot reescriure com

$$\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} < -z_{1-\varepsilon} = z_\varepsilon.$$

Així doncs, la condició d'acceptació de H_0 és

$$\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \geq -z_{1-\varepsilon} = z_\varepsilon.$$

Els tres casos anteriors es poden condensar com es segueix.

Definim

$$K_{\text{obs}}^0 = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma}$$

i, en funció de la hipòtesi alternativa, tenim la següent taula de decisió.

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\mu \neq \mu_0$	$ K_{\text{obs}}^0 \leq z_{1-\frac{\varepsilon}{2}}$	$ K_{\text{obs}}^0 > z_{1-\frac{\varepsilon}{2}}$
$\mu > \mu_0$	$K_{\text{obs}}^0 \leq z_{1-\varepsilon}$	$K_{\text{obs}}^0 > z_{1-\varepsilon}$
$\mu < \mu_0$	$K_{\text{obs}}^0 \geq z_\varepsilon$	$K_{\text{obs}}^0 < z_\varepsilon$

Exemple Se sap que els pesos dels homes adults d'un país determinat segueix una distribució normal amb desviació típica $\sigma = 12.5$. Es desitja contrastar l'afirmació que indica que el pes mitjà de la població anterior és 70.7 amb un nivell de significació d'un 10% a partir de les dades mostrals següents, obtingudes analitzant el pes de 9 individus d'aquesta població:

$$\{74.3, 68.5, 62.3, 79.4, 77.2, 88.9, 81.7, 73.4, 62.1\}$$

Solució L'enunciat del problema demana fer un contrast d'hipòtesi per a la mitjana de la població (és a dir, el pes mitjà), on la desviació típica poblacional ve donada ($\sigma = 12.5$). Com que simplement es proposa comprovar si realment és acceptable que la mitjana siga 70.7 o no, el contrast és bilateral, és a dir, de la forma

$$\left. \begin{array}{l} H_0 : \mu = 70.7 \\ H_a : \mu \neq 70.7 \end{array} \right\}$$

Hem de calcular per tant

$$K_{\text{obs}}^0 = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma},$$

on

$$n = 9,$$

$$\bar{x} = \frac{74.3 + 68.5 + 62.3 + 79.4 + 77.2 + 88.9 + 81.7 + 73.4 + 62.1}{9} = 74.2,$$

$$\mu_0 = 70.7,$$

$$\sigma = 12.5.$$

Aleshores

$$K_{\text{obs}}^0 = \frac{\sqrt{9} \cdot (74.2 - 70.7)}{12.5} = 0.84.$$

D'altra banda, com que $\varepsilon = 0.1$,

$$z_{1-\frac{\varepsilon}{2}} = z_{0.95} \approx 1.645.$$

Com que $|K_{\text{obs}}^0| = 0.84 \leq 1.645 \approx z_{1-\frac{\varepsilon}{2}}$, acceptem la hipòtesi nul·la H_0 ; per tant, no hi ha indicis suficients, amb un nivell de significació d'un 10%, per a refutar la hipòtesi que afirma que el pes mitjà és 70.7.

- **La variància σ^2 és desconeguda.**

Raonant exactament de la mateixa manera que adès, fent servir ara l'expressió de l'interval de confiança per a la mitjana amb variància desconeguda, s'obté que la manera de procedir és calcular

$$K_{\text{obs}}^0 = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{x} - \mu_0)}{S}$$

i aplicar la següent taula de decisió.

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\mu \neq \mu_0$	$ K_{\text{obs}}^0 \leq t_{n-1, 1-\frac{\varepsilon}{2}}$	$ K_{\text{obs}}^0 > t_{n-1, 1-\frac{\varepsilon}{2}}$
$\mu > \mu_0$	$K_{\text{obs}}^0 \leq t_{n-1, 1-\varepsilon}$	$K_{\text{obs}}^0 > t_{n-1, 1-\varepsilon}$
$\mu < \mu_0$	$K_{\text{obs}}^0 \geq t_{n-1, \varepsilon}$	$K_{\text{obs}}^0 < t_{n-1, \varepsilon}$

Exemple En els fullets de propaganda, una empresa assegura que les bombetes que fabrica tenen una durada mitjana de 1600 hores. A fi de contrastar aquesta dada per tal de prevenir una possible estafa, es va prendre una mostra aleatòria de 25 bombetes obtenint-se una durada mitjana de 1570 hores, amb una desviació típica de 80 hores. Pot acceptar-se la informació dels fullets amb un nivell de confiança del 95%?

Solució En aquest cas, el problema no ofereix cap tipus de dada poblacional (μ, σ) i proposa fer un contrast per a la durada mitjana, μ , al voltant del valor 1600, amb un nivell de significació d'un 95%.

D'acord amb l'enunciat, es posa a prova l'afirmació d'un fullet de propaganda, per la qual cosa hem de contrastar l'afirmació consistent en que la informació que proporciona és correcta ($\mu = 1600$) front al cas que ens trobem amb un cas d'estafa ($\mu < 1600$). Notem que aquesta consideració s'ha deduït del context del problema, ja que no és esperable que una empresa anuncie menys quantitat de la que realment proporciona ($\mu > 1600$) i, en qualsevol cas, si es donara hipotèticament eixa situació seria favorable per al consumidor, i en conseqüència no ens trobaríem en el cas d'una estafa. Per tant, el contrast d'hipòtesi que cal fer és

$$\left. \begin{array}{l} H_0 : \mu = 1600 \\ H_a : \mu < 1600 \end{array} \right\}$$

D'altra banda, les dades mostrals proporcionades són $n = 25$ bombetes, $\bar{x} = 1570$ hores de durada mitjana i $s = 80$ hores de desviació típica. A partir d'aquestes, hem de calcular

$$K_{\text{obs}}^0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s}.$$

Recordem que

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{n-1} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{n-1} s^2,$$

i per tant

$$S = \sqrt{\frac{n}{n-1}} s = \sqrt{\frac{25}{24}} \cdot 80 \approx 81.64966.$$

Aleshores es té

$$K_{\text{obs}}^0 \approx \frac{\sqrt{25} \cdot (1570 - 1600)}{81.64966} \approx -1.83712.$$

D'altra banda, com que $\varepsilon = 0.05$, i emprant les propietats de simetria de la t de Student,

$$t_{n-1, \varepsilon} = -t_{n-1, 1-\varepsilon} = -t_{24, 0.95} \approx -1.71088.$$

Així doncs, es compleix

$$K_{\text{obs}}^0 \approx -1.83712 < -1.71088 \approx t_{n-1, \varepsilon},$$

amb la qual cosa rebutgem H_0 i acceptem H_a . Llavors, podem afirmar amb un nivell de confiança d'un 95% que l'empresa falta a la veritat als fulls de propaganda, tot anunciant una duració menor que la que realment ofereix, i per tant estafant als possibles compradors.

Contrast d'una variància

Partim novament d'una població normal $N(\mu, \sigma)$, on ara fem un contrast al voltant del paràmetre σ^2 mitjançant l'anàlisi d'una mostra $\{x_i\}_{i=1}^n$ obtinguda a partir d'aquesta població. Donat un valor σ_0^2 , es proposa un contrast amb hipòtesi nul·la

$$H_0 : \sigma^2 = \sigma_0^2$$

i les tres possibilitats següents per a la hipòtesi alternativa:

$$H_a : \begin{cases} \sigma^2 \neq \sigma_0^2 \\ \sigma^2 > \sigma_0^2 \\ \sigma^2 < \sigma_0^2 \end{cases}$$

En funció de si es coneix o no la mitjana μ , tenim les següents possibilitats.

• La mitjana μ és coneguda.

Fent el raonament amb l'interval de confiança per a la variància amb μ coneguda, s'obté que s'ha de calcular

$$K_{\text{obs}}^0 = \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma_0} \right)^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_0^2} = \frac{(n-1)S^2 + n(\bar{x} - \mu)^2}{\sigma_0^2}$$

i aplicar la següent taula de decisió.

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\sigma^2 \neq \sigma_0^2$	$x_{n, \frac{\varepsilon}{2}} \leq K_{\text{obs}}^0 \leq x_{n, 1-\frac{\varepsilon}{2}}$	$K_{\text{obs}}^0 < x_{n, \frac{\varepsilon}{2}}$ o $K_{\text{obs}}^0 > x_{n, 1-\frac{\varepsilon}{2}}$
$\sigma^2 > \sigma_0^2$	$K_{\text{obs}}^0 \leq x_{n, 1-\varepsilon}$	$K_{\text{obs}}^0 > x_{n, 1-\varepsilon}$
$\sigma^2 < \sigma_0^2$	$K_{\text{obs}}^0 \geq x_{n, \varepsilon}$	$K_{\text{obs}}^0 < x_{n, \varepsilon}$

Exemple Se sap que una determinada màquina produeix un cert tipus de peça amb una llargària mitjana de 45 cm. Per requeriments de precisió, es considera que una màquina d'aquesta classe treballa defectuosament quan la desviació típica de les peces produïdes supera els 0.5 cm.

Per tal de comprovar el funcionament d'una màquina específica, s'ha observat la següent mostra formada per les llargàries de 5 peces produïdes per aquesta

$$\{45.02, 44.72, 45.9, 43.72, 46.07\}.$$

Es pot considerar, amb una significació d'un 1%, que la màquina és defectuosa?

Solució Aquest problema demana la realització d'un contrast d'hipòtesi per a la variància, on ens ve donada la mitjana mostral, $\mu = 45$, corresponent a la llargària mitjana (en centímetres) de les peces produïdes. Com que hem de contrastar el bon funcionament front al mal funcionament ($\sigma > 0.5$), el contrast que cal plantejar és

$$\left. \begin{array}{l} H_0 : \sigma^2 = 0.5^2 = 0.25 \\ H_a : \sigma^2 > 0.5^2 = 0.25 \end{array} \right\}$$

Hem de calcular

$$\begin{aligned} K_{\text{obs}}^0 &= \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_0^2} \\ &= \frac{(45.02 - 45)^2 + (44.72 - 45)^2 + (45.9 - 45)^2 + (43.72 - 45)^2 + (46.07 - 45)^2}{0.5^2} \\ &= 14.6884. \end{aligned}$$

D'altra banda, $\varepsilon = 0.01$ i es té

$$x_{n, 1-\varepsilon} = x_{5, 0.99} \approx 15.0863.$$

Com que

$$K_{\text{obs}}^0 = 14.6884 \leq 15.0863 \approx x_{n, 1-\varepsilon},$$

acceptem H_0 . Per tant, no hi ha indicis suficients, amb una significació d'un 1%, per a determinar que la màquina és defectuosa.

- **La mitjana μ és desconeguda.** En aquest cas, raonant al voltant dels extrems de l'interval de confiança per a la variància amb mitjana desconeguda, s'obté que s'ha de calcular

$$K_{\text{obs}}^0 = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma_0} \right)^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma_0^2} = \frac{(n-1)S^2}{\sigma_0^2}$$

i prendre la decisió d'acord amb la següent taula.

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\sigma^2 \neq \sigma_0^2$	$x_{n-1, \frac{\varepsilon}{2}} \leq K_{\text{obs}}^0 \leq x_{n-1, 1-\frac{\varepsilon}{2}}$	$K_{\text{obs}}^0 < x_{n-1, \frac{\varepsilon}{2}} \text{ o } K_{\text{obs}}^0 > x_{n-1, 1-\frac{\varepsilon}{2}}$
$\sigma^2 > \sigma_0^2$	$K_{\text{obs}}^0 \leq x_{n-1, 1-\varepsilon}$	$K_{\text{obs}}^0 > x_{n-1, 1-\varepsilon}$
$\sigma^2 < \sigma_0^2$	$K_{\text{obs}}^0 \geq x_{n-1, \varepsilon}$	$K_{\text{obs}}^0 < x_{n-1, \varepsilon}$

Exemple Una empresa de jocs d'atzar necessita analitzar les estadístiques relacionades amb els gols que marquen i encaixen els diferents equips de les lligues professionals per tal d'ajustar les apostes. Aquesta desitja conèixer si la desviació típica dels gols encaixats per temporada d'un equip determinat és de 0.125 gols amb un nivell de significació d'un 5%. Per a fer-ho, recull les dades del promig de gols encaixats en les cinc temporades anteriors, tot obtenint-se les dades següents

$$\{2.3, 1.95, 2.12, 2.57, 2.46\}.$$

Solució En aquest cas es proposa un contrast d'hipòtesi bilateral per a la variància sense proporcionar informació de la mitjana mostral.

$$\left. \begin{array}{l} H_0 : \sigma^2 = 0.125^2 = 0.015625 \\ H_a : \sigma^2 \neq 0.125^2 = 0.015625 \end{array} \right\}$$

Per tant, hem de calcular

$$K_{\text{obs}}^0 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma_0^2},$$

on

$$\bar{x} = \frac{2.3 + 1.95 + 2.12 + 2.57 + 2.46}{5} = 2.28.$$

Aleshores

$$\begin{aligned} K_{\text{obs}}^0 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma_0^2} \\ &= \frac{(2.3 - 2.28)^2 + (1.95 - 2.28)^2 + (2.12 - 2.28)^2 + (2.57 - 2.28)^2 + (2.46 - 2.28)^2}{0.125^2} \\ &= 16.0896. \end{aligned}$$

D'altra banda, tenint en compte que $\varepsilon = 0.05$, hem de calcular

$$\begin{aligned} x_{n-1, \frac{\varepsilon}{2}} &= x_{4, 0.025} \approx 0.48442, \\ x_{n-1, 1-\frac{\varepsilon}{2}} &= x_{4, 0.975} \approx 11.14330. \end{aligned}$$

Com que K_{obs}^0 no es troba entre els dos valors anteriors; concretament,

$$K_{\text{obs}}^0 = 16.0896 > 11.14330 \approx x_{n-1, 1-\frac{\epsilon}{2}},$$

es rebutja H_0 en favor d' H_a . Conseqüentment, es pot afirmar amb un 95% de confiança que la desviació típica no és 0.125.

Altres contrastos

A banda dels esmentats anteriorment, tenim altres tipus de contrastos, que poden implicar més d'un paràmetre o bé paràmetres diferents als d'una normal.

• Contrast de dues mitjanes.

En aquest cas, es disposa de dues mostres, $\{x_i\}_{i=1}^m$ i $\{y_i\}_{i=1}^n$, procedents de poblacions $N(\mu_X, \sigma_X)$ i $N(\mu_Y, \sigma_Y)$, respectivament. A partir d'aquesta informació, es poden contrastar afirmacions que involucren els diferents paràmetres d'aquestes dues distribucions. Una de les possibilitats és comparar les dues mitjanes, en un contrast de la forma

$$H_0 : \mu_X = \mu_Y$$

i amb les tres possibilitats següents per a la hipòtesi alternativa:

$$H_a : \begin{cases} \mu_X \neq \mu_Y \\ \mu_X > \mu_Y \\ \mu_X < \mu_Y \end{cases}$$

Segons es coneguen o no les respectives variàncies, σ_X^2, σ_Y^2 , o si aquestes són o no iguals, tenim els següents casos.

– **Es coneixen σ_X^2 i σ_Y^2 .**

Sota aquestes condicions, es calcula

$$K_{\text{obs}}^0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}}$$

i es pren la decisió segons la taula següent.

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\mu_X \neq \mu_Y$	$ K_{\text{obs}}^0 \leq z_{1-\frac{\epsilon}{2}}$	$ K_{\text{obs}}^0 > z_{1-\frac{\epsilon}{2}}$
$\mu_X > \mu_Y$	$K_{\text{obs}}^0 \leq z_{1-\epsilon}$	$K_{\text{obs}}^0 > z_{1-\epsilon}$
$\mu_X < \mu_Y$	$K_{\text{obs}}^0 \geq z_{\epsilon}$	$K_{\text{obs}}^0 < z_{\epsilon}$

– **σ_X^2 i σ_Y^2 desconegudes.**

En aquesta situació, el procediment es divideix en els dos subcasos següents.

* **Tot i no conèixer σ_X^2 i σ_Y^2 , se sap que $\sigma_X^2 = \sigma_Y^2$.**

En aquest cas calculem

$$K_{\text{obs}}^0 = \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{\bar{x} - \bar{y}}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$= \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{\bar{x} - \bar{y}}{\sqrt{(m-1)S_X^2 + (n-1)S_Y^2}}$$

i es pren la decisió en funció de la taula següent:

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\mu_X \neq \mu_Y$	$ K_{\text{obs}}^0 \leq t_{m+n-2, 1-\frac{\epsilon}{2}}$	$ K_{\text{obs}}^0 > t_{m+n-2, 1-\frac{\epsilon}{2}}$
$\mu_X > \mu_Y$	$K_{\text{obs}}^0 \leq t_{m+n-2, 1-\epsilon}$	$K_{\text{obs}}^0 > t_{m+n-2, 1-\epsilon}$
$\mu_X < \mu_Y$	$K_{\text{obs}}^0 \geq t_{m+n-2, \epsilon}$	$K_{\text{obs}}^0 < t_{m+n-2, \epsilon}$

* **Els respectius tamany de la mostra, m, n , són grans**

En aquest cas les variàncies mostrals es poden prendre com una aproximació de les variàncies poblacionals i aplicar el primer cas fent les substitucions oportunes:

$$K_{\text{obs}}^0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S_X^2}{m} + \frac{S_Y^2}{n}}},$$

aplicant la següent taula de decisió:

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\mu_X \neq \mu_Y$	$ K_{\text{obs}}^0 \leq z_{1-\frac{\epsilon}{2}}$	$ K_{\text{obs}}^0 > z_{1-\frac{\epsilon}{2}}$
$\mu_X > \mu_Y$	$K_{\text{obs}}^0 \leq z_{1-\epsilon}$	$K_{\text{obs}}^0 > z_{1-\epsilon}$
$\mu_X < \mu_Y$	$K_{\text{obs}}^0 \geq z_{\epsilon}$	$K_{\text{obs}}^0 < z_{\epsilon}$

Exemple Es desitja comparar la durada mitjana de dos aparells de marca diferent i comprovar si hi ha una diferència significativa en la seua vida útil amb un nivell de significació d'un 5%. Per a fer-ho, es recullen una sèrie de dades mostrals, que venen donades en la taula següent.

	Tamany de la mostra	Mitjana mostral	Desviació típica mostral
Aparell marca A	120	7.24	1.82
Aparell marca B	140	7.65	1.21

Es pot concloure, a partir de la taula anterior, que existeixen diferències significatives entre la durada mitjana dels aparells de les dues marques amb el nivell de significació anterior?

Solució Es tracta d'un contrast d'hipòtesi bilateral de dues mitjanes:

$$\left. \begin{array}{l} H_0 : \mu_X = \mu_Y \\ H_a : \mu_X \neq \mu_Y \end{array} \right\}$$

A l'enunciat problema no s'aporta cap dada referent a les desviacions típiques mostrals, però m i n es poden considerar suficientment grans, de manera que calculem

$$K_{\text{obs}}^0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S_X^2}{m} + \frac{S_Y^2}{n}}},$$

on

$$m = 120,$$

$$\bar{x} = 7.24,$$

$$S_X^2 = \frac{m}{m-1} s_X^2 = \frac{120}{119} \cdot 1.82^2 \approx 3.3402,$$

$$n = 140,$$

$$\bar{y} = 7.65,$$

$$S_Y^2 = \frac{n}{n-1} s_Y^2 = \frac{140}{139} \cdot 1.21^2 \approx 1.4746.$$

Amb açò, es té

$$K_{\text{obs}}^0 \approx \frac{7.24 - 7.65}{\sqrt{\frac{3.3402}{120} + \frac{1.4746}{140}}} \approx -2.0931.$$

D'altra banda, el nivell de significació és $\varepsilon = 0.05$, amb $z_{1-\frac{\varepsilon}{2}} = z_{0.975} \approx 1.96$ Com que

$$|K_{\text{obs}}^0| = |-2.0931| = 2.0931 > 1.96 \approx z_{0.975}$$

es rebutja H_0 i s'accepta H_a . I per tant, amb un nivell de confiança d'un 95%, es pot afirmar que sí que existeixen diferències significatives en la vida útil dels productes de les dues marques.

• Contrast de dues variàncies.

Sota les mateixes hipòtesis que en el cas del contrast anterior, ara volem contrastar relacions entre les variàncies poblacionals.

La hipòtesi nul·la és en aquest cas

$$H_0 : \sigma_X^2 = \sigma_Y^2,$$

mentre que les hipòtesis alternatives possibles són

$$H_a : \left\{ \begin{array}{l} \sigma_X^2 \neq \sigma_Y^2 \\ \sigma_X^2 > \sigma_Y^2 \\ \sigma_X^2 < \sigma_Y^2 \end{array} \right.$$

En aquest cas es calcula

$$K_{\text{obs}}^0 = \frac{S_X^2}{S_Y^2}$$

i es decideix d'acord amb els criteris establits a la taula següent.

H_a	Acceptar H_0 si	Rebutjar H_0 si
$\sigma_X^2 \neq \sigma_Y^2$	$f_{m-1,n-1,\frac{\varepsilon}{2}} \leq K_{\text{obs}}^0 \leq f_{m-1,n-1,1-\frac{\varepsilon}{2}}$	$K_{\text{obs}}^0 < f_{m-1,n-1,\frac{\varepsilon}{2}} \text{ o } K_{\text{obs}}^0 > f_{m-1,n-1,1-\frac{\varepsilon}{2}}$
$\sigma_X^2 > \sigma_Y^2$	$K_{\text{obs}}^0 \leq f_{m-1,n-1,1-\varepsilon}$	$K_{\text{obs}}^0 > f_{m-1,n-1,1-\varepsilon}$
$\sigma_X^2 < \sigma_Y^2$	$K_{\text{obs}}^0 \geq f_{m-1,n-1,\varepsilon}$	$K_{\text{obs}}^0 < f_{m-1,n-1,\varepsilon}$

On, recordem, es pot escriure

$$f_{m-1,n-1,\varepsilon} = \frac{1}{f_{n-1,m-1,1-\varepsilon}}.$$

Exemple El departament de zoologia de la Universitat de Virgínia va fer un estudi per a estimar la diferència en la quantitat d'ortofòsfor químic mesurat en dues estacions diferents del riu James. L'ortofòsfor es mesura en miligrams per litre i es distribueix de forma normal. Les mostres obtingudes en cada estació es proporcionen a la taula següent

Estació 1	$m = 41$	$\bar{x} = 3.84$	$s_X = 3.07$
Estació 2	$n = 61$	$\bar{y} = 1.49$	$s_Y = 1.82$

Es demana justificar la suposició que les observacions provenen de poblacions normals amb variàncies diferents mitjançant el corresponent contrast d'hipòtesi amb un nivell de significació $\varepsilon = 0.02$.

Solució Es demana realitzar el següent contrast d'hipòtesi

$$\left. \begin{array}{l} H_0 : \sigma_X^2 = \sigma_Y^2 \\ H_a : \sigma_X^2 \neq \sigma_Y^2 \end{array} \right\}$$

Tenim

$$S_X^2 = \frac{m}{m-1} s_X^2 = \frac{41}{40} \cdot 3.07^2 \approx 9.6605,$$

$$S_Y^2 = \frac{m}{m-1} s_Y^2 = \frac{61}{60} \cdot 1.82^2 \approx 3.3676.$$

Per tant

$$K_{\text{obs}}^0 = \frac{S_X^2}{S_Y^2} \approx \frac{9.6605}{3.3676} \approx 2.8687.$$

D'altra banda, tenint en compte que $\varepsilon = 0.02$,

$$f_{m-1,n-1,\frac{\varepsilon}{2}} = \frac{1}{f_{n-1,m-1,1-\frac{\varepsilon}{2}}} = \frac{1}{f_{60,40,0.99}} \approx \frac{1}{2.01941} \approx 0.49519,$$

$$f_{m-1,n-1,1-\frac{\varepsilon}{2}} = f_{40,60,0.99} \approx 1.93602.$$

Com que

$$K_{\text{obs}}^0 \approx 2.8687 > 1.93602 = f_{m-1,n-1,1-\frac{\varepsilon}{2}}$$

rebutgem H_0 en favor d' H_a , i per tant sí que existeixen diferències significatives entre les variàncies amb un nivell de confiança d'un 98%.

• **Contrast del paràmetre d'una binomial.**

Siga A un succés tal que $p(A) = p$, amb el valor de la probabilitat p desconegut. Realitzem N proves Bernoulli, amb N suficientment gran, i siga n el nombre de vegades que s'ha donat A .

Suposem que es vol contrastar l'afirmació que la probabilitat siga un determinat valor p_0 conegut. Tenim com a hipòtesi nul·la

$$H_0 : p = p_0$$

i com a possibles hipòtesis alternatives

$$H_a : \begin{cases} p \neq p_0 \\ p > p_0 \\ p < p_0 \end{cases}$$

La manera de procedir en aquest cas és calcular

$$K_{\text{obs}}^0 = \frac{\frac{n}{N} - p_0}{\sqrt{\frac{p_0(1-p_0)}{N}}} = \frac{n - Np_0}{\sqrt{Np_0(1-p_0)}}.$$

I es pren la decisió seguint la taula següent

H_a	Acceptar H_0 si	Rebutjar H_0 si
$p \neq p_0$	$ K_{\text{obs}}^0 \leq z_{1-\frac{\varepsilon}{2}}$	$ K_{\text{obs}}^0 > z_{1-\frac{\varepsilon}{2}}$
$p > p_0$	$K_{\text{obs}}^0 \leq z_{1-\varepsilon}$	$K_{\text{obs}}^0 > z_{1-\varepsilon}$
$p < p_0$	$K_{\text{obs}}^0 \geq z_{\varepsilon}$	$K_{\text{obs}}^0 < z_{\varepsilon}$

Exemple Es vol saber si una moneda determinada està o no trucada. Per a fer-ho, s'efectuen 270 llançaments, amb els quals s'obtenen un total de 148 cares. Es pot considerar que la moneda està trucada amb un nivell de significació d'un 10%?

Solució Siga el succés $A = \{\text{obtenir cara}\}$. Definim $p = p(A)$ i es demana fer el següent contrast d'hipòtesi bilateral

$$\begin{cases} H_0 : p = 0.5 \\ H_a : p \neq 0.5 \end{cases}$$

on $p_0 = 0.5$ és la probabilitat d'obtenir cara en una moneda no trucada. Tenim que s'han efectuat un total de $N = 270$ llançaments, tot obtenint-se un total de $n = 122$ cares.

Es té per tant

$$K_{\text{obs}}^0 = \frac{n - Np_0}{\sqrt{Np_0(1-p_0)}} = \frac{122 - 270 \cdot 0.5}{\sqrt{270 \cdot 0.5 \cdot (1-0.5)}} \approx -1.5823.$$

D'altra banda, $\varepsilon = 0.1$ i

$$z_{1-\frac{\varepsilon}{2}} = z_{0.95} \approx 1.645.$$

Així doncs,

$$|K_{\text{obs}}^0| \approx 1.5823 \leq 1.645 \approx z_{0.95},$$

amb la qual cosa acceptem H_0 , i per tant no podem considerar que la moneda està trucada amb un 10% de significació.

- **Contrast de dos percentatges.** Sota les condicions i notació anterior, suposem que es realitzen dos experiments per a registrar l'ocurrència del succés A sota condicions diferents, i es desitja estudiar si hi ha variació de la probabilitat sota les diferents circumstàncies en què s'ha realitzat l'experiment.

Anomenem p_X i p_Y a les probabilitats que es done A sota les primeres i les segones condicions, respectivament i suposem que es realitzem, sota les primeres condicions, un total de M experiments, tot obtenint m ocurrencies de A i, sota les segones condicions, N experiments, tot obtenint-se n ocurrencies de A , amb M, N suficientment grans. Aleshores es planteja el següent contrast d'hipòtesi, on la hipòtesi nul·la és

$$H_0 : p_X = p_Y$$

i les possibles hipòtesis alternatives són

$$H_a : \begin{cases} p_X \neq p_Y \\ p_X > p_Y \\ p_X < p_Y \end{cases}$$

Llavors, calculem

$$K_{\text{obs}}^0 = \frac{\frac{m}{M} - \frac{n}{N}}{\sqrt{p^* (1 - p^*) \left(\frac{1}{M} + \frac{1}{N} \right)}},$$

on p^* és la probabilitat mostral, donada per

$$p^* = \frac{m + n}{M + N},$$

i es decideix d'acord amb la taula següent

H_a	Acceptar H_0 si	Rebutjar H_0 si
$p_X \neq p_Y$	$ K_{\text{obs}}^0 \leq z_{1-\frac{\epsilon}{2}}$	$ K_{\text{obs}}^0 > z_{1-\frac{\epsilon}{2}}$
$p_X > p_Y$	$K_{\text{obs}}^0 \leq z_{1-\epsilon}$	$K_{\text{obs}}^0 > z_{1-\epsilon}$
$p_X < p_Y$	$K_{\text{obs}}^0 \geq z_{\epsilon}$	$K_{\text{obs}}^0 < z_{\epsilon}$

Exemple Es desitja comparar l'efectivitat de dos medicaments per a pal·liar una certa enfermetat. Es proporcionen els resultats en la taula següent:

	Pacients tractats	Curacions
Medicament A	170	122
Medicament B	150	118

Es pot afirmar que el medicament B és més efectiu que el medicament A amb una significació d'un 10%?

Solució Donat que el fet que B siga més efectiu que A es tradueix com que la probabilitat de curació aplicant el medicament B, p_Y , siga major que la probabilitat de curació aplicant el medicament A, p_X , l'enunciat del problema proposa realitzar el següent contrast que es planteja a continuació:

$$\left. \begin{array}{l} H_0 : p_X = p_Y \\ H_a : p_X < p_Y \end{array} \right\}$$

Tenim d'una banda el següent:

$$\begin{aligned} M &= 170, \\ m &= 122, \\ N &= 150, \\ n &= 118, \\ p^* &= \frac{m+n}{M+N} = \frac{122+118}{170+150} = \frac{240}{320} = 0.75, \end{aligned}$$

amb la qual cosa obtenim

$$K_{\text{obs}}^0 = \frac{\frac{m}{M} - \frac{n}{N}}{\sqrt{p^* (1-p^*) \left(\frac{1}{M} + \frac{1}{N}\right)}} = \frac{\frac{122}{170} - \frac{118}{150}}{\sqrt{0.75 \cdot (1-0.75) \left(\frac{1}{170} + \frac{1}{150}\right)}} \approx -1.4229$$

D'altra banda, $\varepsilon = 0.1$ i es té

$$z_\varepsilon = -z_{1-\varepsilon} = -z_{0.9} \approx -1.28.$$

Amb açò, tenim

$$K_{\text{obs}}^0 \approx -1.4229 < -1.28 \approx z_\varepsilon.$$

Conseqüentment, rebutgem H_0 en favor d' H_a , i per tant podem afirmar amb una significació d'un 10% que el medicament B és més efectiu que el medicament A.

• Test χ^2 de bondat d'ajustament.

Siguen $\{A_i\}_{i=1}^k$ k successos mútuament excloents i tals que la probabilitat teòrica és $p(A_i) = p_i$, amb $\sum_{i=1}^k p_i = 1$. Denotem per $e_i = Np_i$ la **frequència esperada**.

Suposem que fem un total d' N proves i observem el nombre d'ocurrències de cadascun dels successos A_i , obtenint-se θ_i ocurrències del succés A_i , $1 \leq i \leq k$, amb $\sum_{i=1}^k \theta_i = N$. θ_i rep el nom de **frequència observada**.

Aquest contrast d'hipòtesi consisteix en comprovar la versemblança del model teòric d'acord amb els resultats obtinguts a la pràctica, és a dir, si les freqüències observades estan en concordància amb les freqüències esperades. La hipòtesi nul·la H_0 es correspon amb un resultat favorable a la concordància, mentre que l'alternativa H_a fa referència a l'existència de discrepàncies en el model teòric d'acord amb els resultats observats.

Siga ε el nivell de significació del contrast i m el nombre de paràmetres que han d'estimar-se per tal de calcular els valors esperats². Aleshores es calcula

$$K_{\text{obs}}^0 = \sum_{i=1}^k \frac{(\theta_i - e_i)^2}{e_i}$$

i s'aplica el criteri de la taula següent:

Acceptar H_0 si	Rebutjar H_0 si
$K_{\text{obs}}^0 \leq x_{k-1-m, 1-\varepsilon}$	$K_{\text{obs}}^0 > x_{k-1-m, 1-\varepsilon}$

Exemple En un estudi de natalitat, el nombre de fills i filles en 480 famílies amb 5 descendents és el que es dona a la taula.

Fills	0	1	2	3	4	5
Filles	5	4	3	2	1	0
Famílies	27	84	165	132	60	12

Són aquestes dades consistents, amb una significació d'un 99%, amb la hipòtesi que la probabilitat de tindre fill o filla és la mateixa?

Solució Siga X la V.A. que compta el nombre de fills homes d'una família amb 5 descendents. Aleshores, si assumim que el model probabilístic està d'acord amb la hipòtesi que tindre fills o filles és equiprobable, aleshores $X \sim B(5, 0.5)$. Per tant, es té

$$p_1 = p(X = 0) = \binom{5}{0} 0.5^0 \cdot 0.5^5 = 0.03125,$$

$$p_2 = p(X = 1) = \binom{5}{1} 0.5^1 \cdot 0.5^4 = 0.15625,$$

$$p_3 = p(X = 2) = \binom{5}{2} 0.5^2 \cdot 0.5^3 = 0.3125,$$

$$p_4 = p(X = 3) = \binom{5}{3} 0.5^3 \cdot 0.5^2 = 0.3125,$$

$$p_5 = p(X = 4) = \binom{5}{4} 0.5^4 \cdot 0.5^1 = 0.15625,$$

$$p_6 = p(X = 5) = \binom{5}{5} 0.5^5 \cdot 0.5^0 = 0.03125.$$

El nombre total de famílies amb què s'ha recopilat les dades és $N = 480$, per tant les freqüències esperades, corresponents a les famílies que s'espera que tinguin el nombre de fills a partir del qual

²En els casos en què treballarem, no necessitarem cap paràmetre addicional, i per tant sempre tindrem $m = 0$.

s'ha calculat cada probabilitat amb el model ideal són

$$e_1 = Np_1 = 480 \cdot 0.03125 = 15$$

$$e_2 = Np_2 = 480 \cdot 0.15625 = 75$$

$$e_3 = Np_3 = 480 \cdot 0.3125 = 150$$

$$e_4 = Np_4 = 480 \cdot 0.3125 = 150$$

$$e_5 = Np_5 = 480 \cdot 0.15625 = 75$$

$$e_6 = Np_6 = 480 \cdot 0.03125 = 15$$

D'altra banda, les dades del problema indiquen les freqüències observades:

$$\theta_1 = 27$$

$$\theta_2 = 84$$

$$\theta_3 = 165$$

$$\theta_4 = 132$$

$$\theta_5 = 60$$

$$\theta_6 = 12$$

Així doncs, es té

$$\begin{aligned} K_{\text{obs}}^0 &= \sum_{i=1}^k \frac{(\theta_i - e_i)^2}{e_i} \\ &= \frac{(27 - 15)^2}{15} + \frac{(84 - 75)^2}{75} + \frac{(165 - 150)^2}{150} + \frac{(132 - 150)^2}{150} + \frac{(60 - 75)^2}{75} + \frac{(12 - 15)^2}{15} \\ &= 17.94. \end{aligned}$$

D'altra banda, $k = 6$; a més, no hem necessitat cap paràmetre addicional per a obtenir les freqüències esperades, de manera $m = 0$. Com que $\varepsilon = 0.01$, hem de calcular

$$x_{k-1-m, 1-\varepsilon} = x_{5, 0.99} \approx 15.0863.$$

Com que

$$K_{\text{obs}}^0 = 17.94 > 15.0863 \approx x_{k-1-m, 1-\varepsilon},$$

rebutgem H_0 en favor d' H_a . Es conclou per tant que els naixements de fills i filles no són equiprobables amb un nivell de confiança d'un 99%.

Capítol 2

Ajustament i regressió

En aquest capítol s'estudia la dependència funcional d'una determinada quantitat, Y , que s'expressa en funció d'una altra, X . Treballarem específicament sota la suposició següent. Donada una funció $f : \mathbb{R} \rightarrow \mathbb{R}$, considerem X i Y dues variables aleatòries, relacionades entre si, complint

$$E(Y|X = x) = f(x).$$

Així doncs, fixat un valor $x \in \mathbb{R}$, el valor observat $y = y(x)$ de Y (és a dir, generat a partir d'aquesta variable aleatòria) diferirà, en general, del valor promig anterior, $f(x) = E(Y|X = x)$, en un terme desconegut, $\varepsilon = \varepsilon(x)$.

Aquest terme correspon a un valor generat a partir d'una V.A. ε (depenent de X) tal que

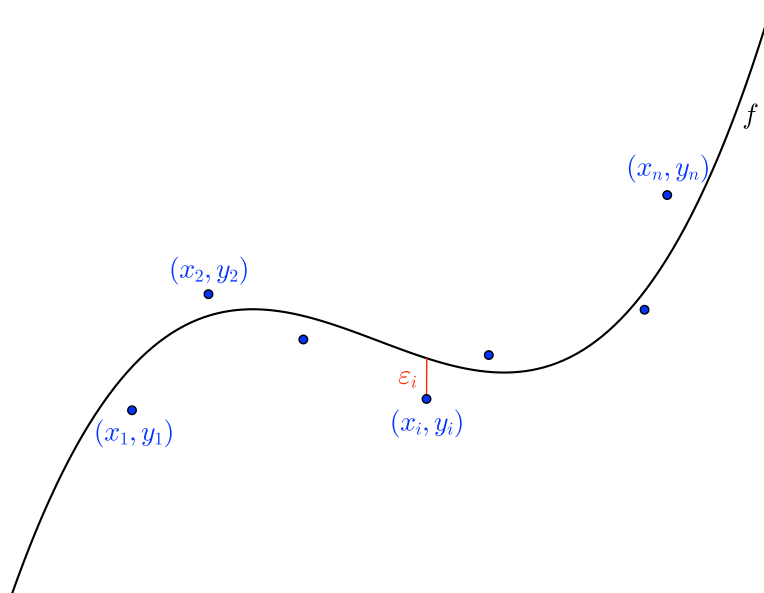
$$Y = f(X) + \varepsilon,$$

complint $E(\varepsilon) = 0$. Així doncs, donada una mostra aleatòria de n valors generats per X , $\{x_i\}_{i=1}^n$, per als quals tenim, en correspondència, els corresponents valors generats per Y , $\{y_i\}_{i=1}^n$, es té que

$$y_i = f(x_i) + \varepsilon_i,$$

on ε_i és l'error aleatori comés en cadascuna de les observacions, generats a partir de la V.A. ε .

S'obté d'eixa manera una mostra de n punts en el pla, $\{(x_i, y_i)\}_{i=1}^n$, corresponents a avaluacions de la funció f que han estat afectades per un error aleatori. La seua representació mitjançant punts del plànol XY rep el nom de **diagrama de dispersió**.



A la pràctica, la situació habitual en què ens trobem és amb el fet que la funció f no es coneix, sinó que es disposa d'una sèrie de dades nodals procedents d'aquesta, afectades per un error aleatori. El que veurem al llarg d'aquest tema és un seguit de procediments que permeten *reconstruir* o aproximar aquesta funció f per una altra, a partir de l'**ajustament** òptim de paràmetres.

Siga $g : \mathbb{R} \times \mathbb{R}^k \rightarrow \mathbb{R}$ una funció, que fa correspondre a cada parell $(x, a_1, \dots, a_k) \in \mathbb{R} \times \mathbb{R}^k$ un valor real determinat $g(x, a_1, \dots, a_k) \in \mathbb{R}$, on $a = (a_1, \dots, a_k) \in \mathbb{R}^k$ és un vector format per k paràmetres, que són els que cal ajustar. Suposem que es disposa d'una sèrie de dades mostrals, $\{(x_i, y_i)\}_{i=1}^n$, que

es desitja ajustar a partir de la funció g . El problema que plantegem a continuació és l'obtenció dels k paràmetres $a_1, \dots, a_k$ que s'ajusten el *millor* possible a aquestes dades.

Un bon criteri que es pot imposar per tal d'obtenir-los és minimitzar la suma del quadrat dels errors mostrals $e_i : \mathbb{R}^k \rightarrow \mathbb{R}$ definits per

$$e_i = e_i(a_1, \dots, a_k) = y_i - g(x_i, a_1, \dots, a_k),$$

és a dir, trobar el mínim de la funció $\phi : \mathbb{R}^k \rightarrow \mathbb{R}$ donada per

$$\phi(a_1, \dots, a_k) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n e_i(a_1, \dots, a_n)^2$$

o el que es coneix també com un procediment d'ajustament per **mínims quadrats**.

Com que el problema consisteix en cercar un possible mínim local de la funció ϕ , la condició que cal imposar és la que compleix qualsevol extrem relatiu, que és, assumint que g disposa de les propietats de diferenciabilitat suficients:

$$\frac{\partial \phi}{\partial a_j}(a_1, \dots, a_k) = 0, \quad 1 \leq j \leq k.$$

Açò genera un sistema de k equacions amb k incògnites (els paràmetres $a_1, \dots, a_k$ que volem ajustar). Concretament, per a cada $j \in \{1, \dots, k\}$, l'equació corresponent resulta

$$\begin{aligned} 0 &= \frac{\partial \phi}{\partial a_j}(a_1, \dots, a_k) = \frac{\partial}{\partial a_j} \left[\sum_{i=1}^n e_i(a_1, \dots, a_n)^2 \right] = 2 \sum_{i=1}^n e_i(a_1, \dots, a_n) \frac{\partial e_i}{\partial a_j}(a_1, \dots, a_n) \\ &= -2 \sum_{i=1}^n (y_i - g(x_i, a_1, \dots, a_k)) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_k). \end{aligned}$$

Així doncs, el sistema d'equacions que cal plantejar, en termes de la funció d'ajustament g , és

$$\sum_{i=1}^n (y_i - g(x_i, a_1, \dots, a_k)) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_k) = 0, \quad 1 \leq j \leq k. \quad (2.1)$$

Aquest plantejament, no obstant, és massa general i el que anem a fer tot seguit és limitar-nos als casos més senzills.

2.1 Ajustament lineal

El primer cas que estudiarem amb profunditat serà el cas en què g és una funció lineal, és a dir, $g(x, a, b) = a + bx$ ($k = 2$ paràmetres). Aquest ajustament és especialment útil sota la suposició que les dades procedeixen d'una funció f lineal, $f(x) = \alpha + \beta x$, però desconeguda, la qual volem ajustar.

En les condicions anteriors, i emprant ara l'expressió de g , el sistema d'equacions anterior és 2×2 i, tenint en compte que

$$\begin{aligned} \frac{\partial g}{\partial a}(x, a, b) &= 1, \\ \frac{\partial g}{\partial b}(x, a, b) &= x, \end{aligned}$$

es té

$$\sum_{i=1}^n (y_i - g(x_i, a, b)) \frac{\partial g}{\partial a}(x_i, a, b) = \sum_{i=1}^n (y_i - (a + bx_i)) \cdot 1 = \sum_{i=1}^n (y_i - (a + bx_i))$$

$$\sum_{i=1}^n (y_i - g(x_i, a, b)) \frac{\partial g}{\partial b}(x_i, a, b) = \sum_{i=1}^n (y_i - (a + bx_i)) x_i = \sum_{i=1}^n (y_i - (a + bx_i)) x_i$$

i per tant aquest presenta el següent aspecte

$$\left. \begin{aligned} \sum_{i=1}^n (y_i - (a + bx_i)) &= 0 \\ \sum_{i=1}^n (y_i - (a + bx_i)) x_i &= 0 \end{aligned} \right\}$$

Desenvolupant la primera equació, s'obté

$$0 = \sum_{i=1}^n (y_i - (a + bx_i)) = \sum_{i=1}^n y_i - na - b \sum_{i=1}^n x_i.$$

Dividint per n a ambdós costats, s'obté

$$\bar{y} - a - b\bar{x} = 0, \quad (2.2)$$

on

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

és la **mitjana de x** i

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

és la **mitjana de y** . Quant a la segona equació, tenim

$$0 = \sum_{i=1}^n (y_i - (a + bx_i)) x_i = \sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i - b \sum_{i=1}^n x_i^2.$$

Dividint de nou per n a ambdós membres, resulta

$$(s_{xy} + \bar{x}\bar{y}) - a\bar{x} - b(s_x^2 + \bar{x}^2) = 0, \quad (2.3)$$

on

$$s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n x_i (y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) y_i = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x}\bar{y}$$

rep el nom de **covariància de x i y** i

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

és la **variància de x** .

Multiplicant per $\bar{x}$ a l'Equació (2.2), el sistema queda com se segueix

$$\left. \begin{aligned} \bar{x}a + \bar{x}^2b &= \bar{x}\bar{y} \\ \bar{x}a + (s_x^2 + \bar{x}^2)b &= (s_{xy} + \bar{x}\bar{y}) \end{aligned} \right\}$$

Restant la primera equació a la segona, obtenim

$$s_x^2 b = s_{xy} \longrightarrow b = \frac{s_{xy}}{s_x^2}.$$

D'altra banda, emprant novament l'Equació (2.2), s'obté

$$a = \bar{y} - b\bar{x} = \bar{y} - \frac{s_{xy}}{s_x^2} \bar{x}.$$

Així doncs, l'expressió de la **recta de regressió de Y sobre X** resulta

$$y = a + bx = \bar{y} - \frac{s_{xy}}{s_x^2} \bar{x} + \frac{s_{xy}}{s_x^2} x,$$

que es pot escriure de manera condensada com

$$y - \bar{y} = \frac{s_{xy}}{s_x^2} (x - \bar{x}).$$

De manera anàloga, i repetint el procediment intercanviant els papers de les variables x i y , la **recta de regressió de X sobre Y** és

$$x - \bar{x} = \frac{s_{xy}}{s_y^2} (y - \bar{y}),$$

on

$$s_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}^2$$

és la **variància de y** .

Definició (Coeficient de correlació) *Es defineix el coeficient de correlació poblacional entre les variables X i Y mitjançant*

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)} \sqrt{\text{var}(Y)}},$$

essent

$$\begin{aligned} \text{cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] = E[E(X)(Y - E(Y))] = E[(X - E(X))E(Y)] \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

la **covariància poblacional**.

Donada una mostra amb n observacions, $\{(x_i, y_i)\}_{i=1}^n$, es defineix el **coeficient de correlació mostral** com

$$r = \frac{s_{xy}}{s_x s_y},$$

on $s_x = +\sqrt{s_x^2}$ i $s_y = +\sqrt{s_y^2}$ són les corresponents **desviacions típiques de x i de y** , respectivament.

Propietat

Es compleix que $|r| \leq 1$.

D'altra banda, si $|r| = 1$ llavors es té que l'ajustament és perfecte, ço és, la recta de regressió passa per tots els punts de la mostra:

$$y_i - \bar{y} = \frac{s_{xy}}{s_x^2}(x_i - \bar{x}), \quad 1 \leq i \leq n.$$

Exercici 17 *Demostra la segona part de la propietat.*

El coeficient de correlació és un indicador de la qualitat de l'ajustament. Com més proper a la unitat siga el seu mòdul millor serà aquest, mentre que com més proper a zero, pitjor. El signe d'aquest també proporciona informació; un signe positiu indica que la tendència és que y augmente en augmentar x (i per tant que x augmente en augmentar y), mentre que un signe negatiu vol dir que la tendència és a una disminució de y en augmentar x (i, conseqüentment, una disminució de x en augmentar y).

Exemple Considerem la següent mostra formada per 5 punts de $\mathbb{R}^2$.

$$\{(1, 0.5), (1.5, 1.3), (2.5, 2.1), (3, 2), (4, 2.7)\}.$$

Es demana calcular la recta de regressió de Y sobre X , de X sobre Y i el coeficient de correlació.

Solució Podem identificar cadascun dels elements de la mostra per components:

$$X = \{1, 1.5, 2.5, 3, 4\},$$

$$Y = \{0.5, 1.3, 2.1, 2, 2.7\}.$$

Aleshores es té

$$\begin{aligned}\bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i = \frac{1 + 1.5 + 2.5 + 3 + 4}{5} = 2.4, \\ \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i = \frac{0.5 + 1.3 + 2.1 + 2 + 2.7}{5} = 1.72, \\ s_x^2 &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \frac{1^2 + 1.5^2 + 2.5^2 + 3^2 + 4^2}{5} - 2.4^2 = 1.14, \\ s_y^2 &= \frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}^2 = \frac{0.5^2 + 1.3^2 + 2.1^2 + 2^2 + 2.7^2}{5} - 1.72^2 = 0.5696, \\ s_{xy} &= \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y} = \frac{1 \cdot 0.5 + 1.5 \cdot 1.3 + 2.5 \cdot 2.1 + 3 \cdot 2 + 4 \cdot 2.7}{5} - 2.4 \cdot 1.72 = 0.772.\end{aligned}$$

Per tant, els coeficients de la recta de regressió són

$$b = \frac{s_{xy}}{s_x^2} = \frac{0.772}{1.14} \approx 0.6772, \quad a = \bar{y} - b\bar{x} \approx 1.72 - 0.6772 \cdot 2.4 \approx 0.0947.$$

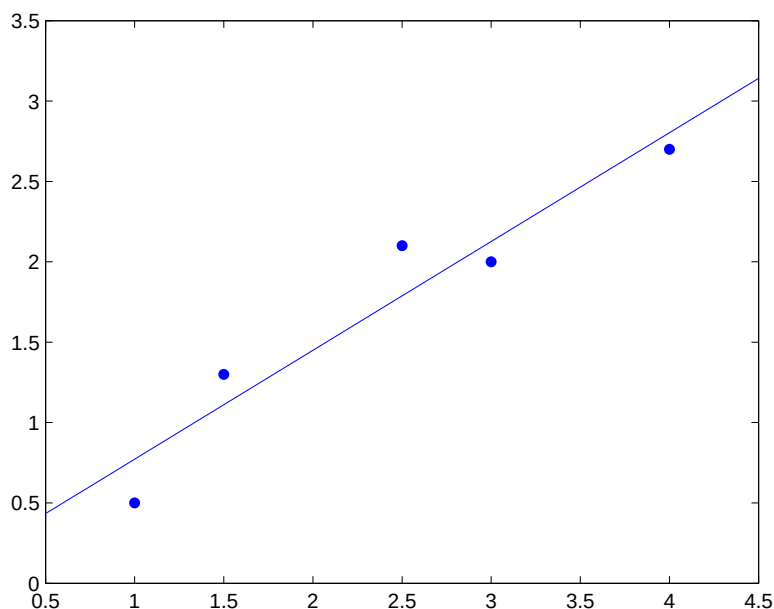
Així doncs, la recta de regressió és, aproximadament,

$$y = 0.0947 + 0.6772x.$$

D'altra banda, el coeficient de correlació mostral ve donat per

$$r = \frac{s_{xy}}{s_x s_y} = \frac{0.772}{\sqrt{1.14} \sqrt{0.5696}} \approx 0.958.$$

Per tant, com que $|r| = 0.958$ és un valor proper a 1, l'ajustament és bo. El corresponent diagrama de dispersió amb la recta de regressió és el següent:



On es pot apreciar que, efectivament, l'ajustament és raonablement bo, com ja indica el coeficient de correlació.

2.1.1 Intervals de confiança

En aquesta secció, treballarem sota la hipòtesi que les variables aleatòries d'on procedeixen les dades mostrals són normals amb condicions d'independència, la qual cosa ens permet aprofundir en l'anàlisi de regressió lineal.

Suposem que les V.A. $\{Y_i\}_{i=1}^n$ que generen els elements y de la mostra, obtingudes a partir dels valors $\{x_i\}_{i=1}^n$, són independents entre si i segueixen una distribució normal de mitjana $\mu = \alpha + \beta x_i$ i variància σ^2 (independent de x_i), és a dir, $Y_i \sim N(\alpha + \beta x_i, \sigma^2)$, $1 \leq i \leq n$. Aleshores podem expressar cada V.A. Y_i de la següent manera:

$$Y_i = \alpha + \beta x_i + \varepsilon_i, \quad 1 \leq i \leq n,$$

on $\{\varepsilon_i\}_{i=1}^n$ són V.A. independents entre si, distribuïdes normalment, amb mitjana

$$E(\varepsilon_i) = E(Y_i - (\alpha + \beta x_i)) = E(Y_i) - (\alpha + \beta x_i) = \mu - (\alpha + \beta x_i) = 0$$

i variància

$$\text{var}(\varepsilon_i) = \text{var}(Y_i - (\alpha + \beta x_i)) = \text{var}(Y_i) = \sigma^2.$$

Segons hem vist anteriorment, tenim un procediment d'obtenció d'una **estimació puntual** dels paràmetres α i β , que són, respectivament, a i b , sota un procediment de mínims quadrats. El que veurem tot seguit,

és un procediment de càlcul d'**interval·ls de confiança** per a aquests paràmetres, és a dir, interval·ls que contenen α i β amb una certa probabilitat, els quals serà possible calcular gràcies a les hipòtesis prèvies.

Per a procedir, també necessitem introduir un altre paràmetre mostral, que anomenem **variància**, i que ve donat per

$$s^2 = \frac{1}{n} \sum_{i=1}^n (y_i - (a + bx_i))^2 = \frac{1}{n} \sum_{i=1}^n (y_i - a - bx_i)^2,$$

el qual constitueix una mesura de l'error de l'ajustament (note's que és precisament la quantitat que es minimitza al procediment de mínims quadrats). Anomenem **desviació típica** a l'arrel quadrada positiva de la variància, és a dir, $s = +\sqrt{s^2}$.

Els estimadors (no esbiaixats) de cada paràmetre, útils per argumentar l'obtenció de l'expressió dels interval·ls de confiança corresponents, són els següents:

- Estimador de β :

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

- Estimador d' α :

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}.$$

- Estimador de σ^2 :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}x_i)^2.$$

Emprant la teoria de probabilitat vista, es pot comprovar que els estimadors compleixen les propietats que es presenten tot seguit.

Propietats

1. $\hat{\beta} \sim N\left(\beta, \frac{\sigma}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}\right).$
2. $\hat{\alpha} \sim N\left(\alpha, \frac{\sigma \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}{\sqrt{n \sum_{i=1}^n (x_i - \bar{x})^2}}\right).$
3. $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2.$
4. $\hat{\alpha}$ i $\hat{\beta}$ són independents de $\hat{\sigma}^2.$

Emprant aquestes propietats convenientment, es pot comprovar que els interval·ls de confiança per a cadascun dels paràmetres de dependència de Y sobre X , amb un nivell de significació ε , són els següents:

- **Interval de confiança per a α :**

$$a \pm t_{n-2, 1-\frac{\varepsilon}{2}} \sqrt{\frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n (y_i - a - bx_i)^2}{n(n-2) \sum_{i=1}^n (\bar{x} - x_i)^2}} = a \pm t_{n-2, 1-\frac{\varepsilon}{2}} \sqrt{\frac{(s_x^2 + \bar{x}^2)s^2}{(n-2)s_x^2}}.$$

- Interval de confiança per a β :

$$b \pm t_{n-2, 1-\frac{\varepsilon}{2}} \sqrt{\frac{\sum_{i=1}^n (y_i - a - bx_i)^2}{(n-2) \sum_{i=1}^n (x_i - \bar{x})^2}} = b \pm t_{n-2, 1-\frac{\varepsilon}{2}} \sqrt{\frac{s^2}{(n-2)s_x^2}}.$$

- Interval de confiança per a σ :

$$\left[\frac{\sum_{i=1}^n (y_i - a - bx_i)^2}{x_{n-2, 1-\frac{\varepsilon}{2}}}, \frac{\sum_{i=1}^n (y_i - a - bx_i)^2}{x_{n-2, \frac{\varepsilon}{2}}} \right] = \left[\frac{ns^2}{x_{n-2, 1-\frac{\varepsilon}{2}}}, \frac{ns^2}{x_{n-2, \frac{\varepsilon}{2}}} \right].$$

Els corresponents intervals de confiança per a la dependència de X sobre Y són els mateixos, intercanviant els papers de X i de Y i amb a i b els paràmetres de la recta de regressió de X sobre Y .

Exemple Calcular els corresponents intervals de confiança per a α , β i σ^2 amb una significació d'un 5% per a les dades de l'exemple anterior:

$$\{(1, 0.5), (1.5, 1.3), (2.5, 2.1), (3, 2), (4, 2.7)\}.$$

Solució Farem ús d'alguns paràmetres calculats anteriorment, com ara, la mitjana i variància de x , $\bar{x} = 2.4$ i $s_x^2 = 1.14$, respectivament, a més dels paràmetres de recta de regressió de Y sobre X , $a \approx 0.0947$, $b \approx 0.6772$.

D'altra banda, es necessita calcular la variància

$$\begin{aligned} s^2 &= \frac{1}{n} \sum_{i=1}^n (y_i - a - bx_i)^2 \\ &\approx \frac{(0.5 - 0.0947 - 0.6772 \cdot 1)^2}{5} + \frac{(1.3 - 0.0947 - 0.6772 \cdot 1.5)^2}{5} + \frac{(2.1 - 0.0947 - 0.6772 \cdot 2.5)^2}{5} \\ &\quad + \frac{(2 - 0.0947 - 0.6772 \cdot 3)^2}{5} + \frac{(2.7 - 0.0947 - 0.6772 \cdot 4)^2}{5} \approx 0.0468. \end{aligned}$$

Amb açò, tenim que l'interval de confiança per a α amb una significació $\varepsilon = 0.05$ és, tenint en compte que $t_{n-2, 1-\frac{\varepsilon}{2}} = t_{3, 0.975} \approx 3.18245$,

$$\begin{aligned} a \pm t_{n-2, 1-\frac{\varepsilon}{2}} \sqrt{\frac{(s_x^2 + \bar{x}^2)s^2}{(n-2)s_x^2}} &\approx 0.0947 \pm 3.18245 \sqrt{\frac{(1.14 + 2.4^2) \cdot 0.0468}{3 \cdot 1.14}} \approx 0.0947 \pm 0.9779 \\ &= [-0.8832, 1.0726]. \end{aligned}$$

D'altra banda, l'interval de confiança per a β resulta

$$b \pm t_{n-2, 1-\frac{\varepsilon}{2}} \sqrt{\frac{s^2}{(n-2)s_x^2}} \approx 0.6772 \pm 0.0947 \pm 3.18245 \sqrt{\frac{0.0468}{3 \cdot 1.14}} \approx 0.6772 \pm 0.3723 = [0.3049, 1.0495].$$

Finalment, l'interval de confiança per a σ^2 , tenint en compte que $x_{n-2, 1-\frac{\varepsilon}{2}} = x_{3, 0.975} \approx 9.34840$ i $x_{n-2, \frac{\varepsilon}{2}} = x_{3, 0.025} \approx 0.21579$, és

$$\left[\frac{ns^2}{x_{n-2, 1-\frac{\varepsilon}{2}}}, \frac{ns^2}{x_{n-2, \frac{\varepsilon}{2}}} \right] \approx \left[\frac{5 \cdot 0.0468}{9.34840}, \frac{5 \cdot 0.0468}{0.21579} \right] \approx [0.025, 1.0846].$$

2.1.2 Prediccions

Continuant amb el model simple normal anterior, suposem que a partir d'una mostra amb n observacions s'ha estimat $E(Y|X = x) = \alpha + \beta x$ mitjançant la recta de regressió $y = a + bx$ i, amb aquestes condicions, es desitja predir el valor y_0 d'una nova observació quan $x = x_0$.

La manera més senzilla de procedir és mitjançant una **predicció puntual**, que consisteix simplement a substituir el valor x_0 a l'expressió de la recta de regressió, $y_0 = a + bx_0$. El valor y_0 es pot interpretar com una estimació *puntual* de $E(Y|X = x_0) = \alpha + \beta x_0$.

No obstant això, si emprem les tècniques anteriors, també podem parlar del que es coneix com **interval de predicció**, que contenen el valor $\alpha + \beta x_0$ amb una certa probabilitat. En el cas lineal, l'interval de predicció per a un cert valor x_0 , amb la corresponent predicció puntual $y_0 = a + bx_0$, és

$$y_0 \pm t_{n-2, 1-\frac{\alpha}{2}} \sqrt{\frac{ns^2}{n-2} \left(\frac{n+1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)} = y_0 \pm t_{n-2, 1-\frac{\alpha}{2}} \sqrt{\frac{s^2}{n-2} \left(n+1 + \frac{(x_0 - \bar{x})^2}{s_x^2} \right)}.$$

Exemple Considerem les dades dels exemples anterior,

$$\{(1, 0.5), (1.5, 1.3), (2.5, 2.1), (3, 2), (4, 2.7)\}.$$

Es demana calcular a partir d'elles un interval de predicció per a $x_0 = 5$ amb un nivell de significació d'un 5%.

Solució La predicció puntual resulta en aquest cas

$$y_0 = a + bx_0 \approx 0.0947 + 0.6772 \cdot 5 = 3.4807.$$

Així doncs, l'interval de predicció resulta

$$\begin{aligned} y_0 \pm t_{n-2, 1-\frac{\alpha}{2}} \sqrt{\frac{s^2}{n-2} \left(n+1 + \frac{(x_0 - \bar{x})^2}{s_x^2} \right)} &\approx 3.4807 \pm 3.18245 \sqrt{\frac{0.0468}{3} \left(6 + \frac{(5 - 2.4)^2}{1.14} \right)} \\ &\approx 3.4807 \pm 1.373 = [2.1077, 4.8537]. \end{aligned}$$

Exercici 18 Amb les dades dels exemples anteriors, calcula la recta de regressió de X sobre Y , els intervals de confiança per a α , β i σ^2 (on, en aquest cas, $x = \alpha + \beta y$) i l'interval de predicció per a $y_0 = 3.5$.

Indicació: És suficient amb intercanviar el paper de x i de y a totes les fórmules.

2.2 Altres ajustaments

A banda de l'ajustament lineal, existeixen una infinitat de possibilitats d'ajustar les dades a unes altres classes de funcions. Aquesta consideració és especialment útil si es detecta que la naturalesa de les dades no és lineal, sinó que respon a una altra classe de funció. Es presenten a continuació un seguit d'exemples dels ajustaments de dades més representatius.

2.2.1 Ajustament parabòlic

Suposem que se sap que el tipus de funció de la qual procedeix un cert conjunt de dades al pla, $\{(x_i, y_i)\}_{i=1}^n$, no és lineal, sinó quadràtica, és a dir, de la forma $f(x) = \alpha + \beta x + \gamma x^2$. El plantejament que farem ara, doncs, és ajustar els paràmetres de la funció $g(x) = a + bx + cx^2$ resolent el Sistema d'Equacions (2.1).

Operant sobre aquest, es pot comprovar que l'expressió del sistema resultant és la següent:

$$\left. \begin{aligned} na + \left(\sum_{i=1}^n x_i\right)b + \left(\sum_{i=1}^n x_i^2\right)c &= \sum_{i=1}^n y_i \\ \left(\sum_{i=1}^n x_i\right)a + \left(\sum_{i=1}^n x_i^2\right)b + \left(\sum_{i=1}^n x_i^3\right)c &= \sum_{i=1}^n x_i y_i \\ \left(\sum_{i=1}^n x_i^2\right)a + \left(\sum_{i=1}^n x_i^3\right)b + \left(\sum_{i=1}^n x_i^4\right)c &= \sum_{i=1}^n x_i^2 y_i \end{aligned} \right\}$$

Exemple Ajustar parabòlicament les dades següents:

$$\{(-2, 8.39), (-0.7, 4.66), (1.3, 8.89), (2.1, 29.42), (4.3, 42.87), (5.2, 73.12)\}.$$

Solució En primer lloc, calculem tots els coeficients que apareixen al sistema d'equacions.

$$\begin{aligned} n &= 6 \\ \sum_{i=1}^n x_i &= -2 - 0.7 + 1.3 + 2.1 + 4.3 + 5.2 = 10.2 \\ \sum_{i=1}^n x_i^2 &= (-2)^2 + (-0.7)^2 + 1.3^2 + 2.1^2 + 4.3^2 + 5.2^2 = 56.12 \\ \sum_{i=1}^n x_i^3 &= (-2)^3 + (-0.7)^3 + 1.3^3 + 2.1^3 + 4.3^3 + 5.2^3 = 223.23 \\ \sum_{i=1}^n x_i^4 &= (-2)^4 + (-0.7)^4 + 1.3^4 + 2.1^4 + 4.3^4 + 5.2^4 = 1111.586 \\ \sum_{i=1}^n y_i &= 8.39 + 4.66 + 8.89 + 29.42 + 42.87 + 73.12 = 167.35 \\ \sum_{i=1}^n x_i y_i &= (-2) \cdot 8.39 + (-0.7) \cdot 4.66 + 1.3 \cdot 8.89 + 2.1 \cdot 29.42 + 4.3 \cdot 42.87 + 5.2 \cdot 73.12 = 617.862 \\ \sum_{i=1}^n x_i^2 y_i &= (-2)^2 \cdot 8.39 + (-0.7)^2 \cdot 4.66 + 1.3^2 \cdot 8.89 + 2.1^2 \cdot 29.42 + 4.3^2 \cdot 42.87 + 5.2^2 \cdot 73.12 = 2950.4408 \end{aligned}$$

Per tant, el sistema queda de la següent manera:

$$\left. \begin{aligned} 6a + 10.2b + 56.12c &= 167.35 \\ 10.2a + 56.12b + 223.23c &= 617.862 \\ 56.12a + 223.23b + 1111.586c &= 2950.4408 \end{aligned} \right\}$$

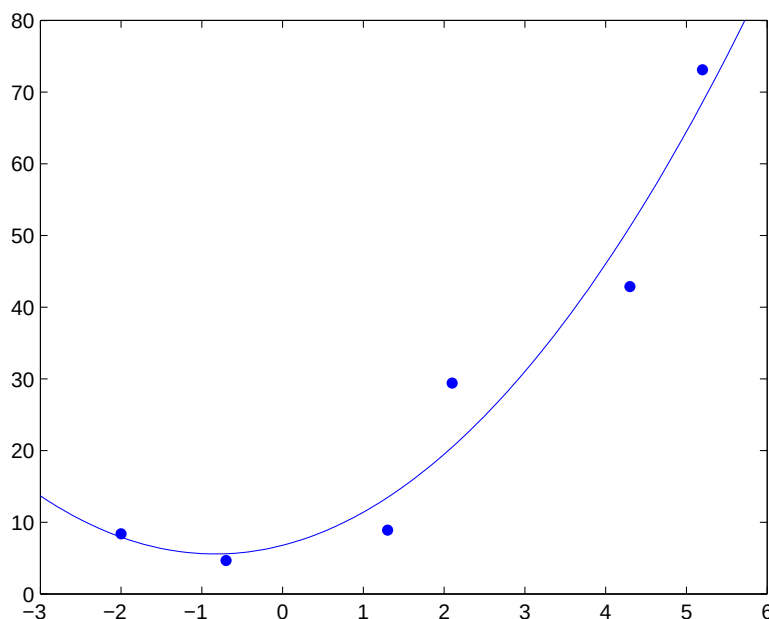
Resolent, s'obté:

$$\begin{cases} a \approx 6.78 \\ b \approx 2.89 \\ c \approx 1.73 \end{cases}$$

Així doncs, la paràbola que resol el problema de mínims quadrats és, de manera aproximada,

$$6.78 + 2.89x + 1.73x^2.$$

El corresponent diagrama de dispersió, juntament amb la gràfica de la paràbola, té l'aspecte següent:



Exercici 19 Repeteix aquest exemple realitzant ara un ajustament lineal. Compara els resultats obtinguts.

2.2.2 Ajustament hiperbòlic

En aquest cas ajustem una funció hiperbòlica de la forma $f(x) = \alpha + \frac{\beta}{x}$, mitjançant $y = g(x) = a + \frac{b}{x}$ a partir d'unes dades en el pla $\{(x_i, y_i)\}_{i=1}^n$. Per tal d'obtenir a i b per mínims quadrats, podem fer el canvi de variable $z = \frac{1}{x}$, de manera que l'expressió anterior es converteix en la recta $y = a + bz$.

A la pràctica, la manera de procedir és convertir els elements de la submostra X mitjançant el canvi $z_i = \frac{1}{x_i}$, $1 \leq i \leq n$ i obtenir a i b mitjançant el procediment habitual per a la recta de regressió que ajusta, en aquest cas, els valors de la mostra $\{(z_i, y_i)\}_{i=1}^n$ fent

$$b = \frac{s_{zy}}{s_z^2}, \quad a = \bar{y} - b\bar{z},$$

on

$$s_{zy} = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n z_i y_i - \bar{z}\bar{y}, \quad s_z^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 = \frac{1}{n} \sum_{i=1}^n z_i^2 - \bar{z}^2.$$

D'igual manera que en cas lineal, es pot calcular també el coeficient de correlació per

$$r = \frac{s_{zy}}{s_z s_y}.$$

Exemple Es demana efectuar un ajustament hiperbòlic de les dades següents:

$$\{(0.5, 0.09), (1, 0.83), (1.7, 1.34), (2.8, 1.75), (3.2, 1.9)\}.$$

Solució Separant per components, obtenim les submostres

$$X = \{0.5, 1, 1.7, 2.8, 3.2\},$$

$$Y = \{0.09, 0.83, 1.34, 1.75, 1.9\}.$$

Fent el canvi $z_i = \frac{1}{x_i}$, obtenim la nova submostra transformada:

$$Z = \frac{1}{X} = \left\{ \frac{1}{0.5}, \frac{1}{1}, \frac{1}{1.7}, \frac{1}{2.8}, \frac{1}{3.2} \right\} \approx \{2, 1, 0.5882, 0.3571, 0.3125\}.$$

Tenim per tant

$$\bar{z} \approx \frac{2 + 1 + 0.5882 + 0.3571 + 0.3125}{5} \approx 0.8516,$$

$$\bar{y} = \frac{0.09 + 0.83 + 1.34 + 1.75 + 1.9}{5} = 1.182,$$

$$s_z^2 \approx \frac{2^2 + 1^2 + 0.5882^2 + 0.3571^2 + 0.3125^2}{5} - 0.8516^2 \approx 0.3891,$$

$$s_y^2 = \frac{0.09^2 + 0.83^2 + 1.34^2 + 1.75^2 + 1.9^2}{5} - 1.182^2 \approx 0.4359,$$

$$s_{zy} \approx \frac{2 \cdot 0.09 + 1 \cdot 0.83 + 0.5882 \cdot 1.34 + 0.3571 \cdot 1.75 + 0.3125 \cdot 1.9}{5} - 0.8516 \cdot 1.182 \approx -0.4032.$$

Així doncs, s'obté

$$b = \frac{s_{zy}}{s_z^2} \approx -1.0362, \quad a = \bar{y} - b\bar{z} \approx 2.0644.$$

Per tant, la hipèrbola que ajusta millor les dades en termes de mínims quadrats és aproximadament

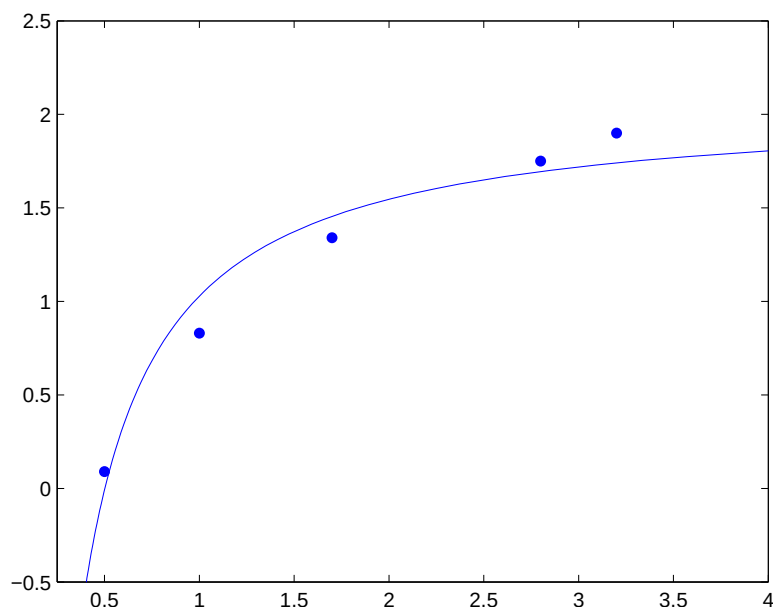
$$y = 2.0644 - \frac{1.0362}{x},$$

amb coeficient de correlació

$$r = \frac{s_{zy}}{s_z s_y} \approx -0.979.$$

Com que $|r|$ està prop de la unitat, es pot considerar que l'ajustament és bo. El signe negatiu indica que la dependència entre les dades transformades és negativa, és a dir que la tendència és una disminució de y en augmentar z i viceversa. Com que $z = \frac{1}{x}$ (funció decreixent), en termes de les variables originals es dona la relació oposada, és a dir, una tendència de augment de y en augmentar x i viceversa.

A la gràfica següent es presenta el diagrama de dispersió de les variables originals amb la funció de l'ajustament, on es poden comprovar visualment totes les afirmacions prèvies.



Exercici 20 Ajusta hiperbòlicament les dades

$$\{(0, 3), (0.5, 5), (1, 6), (2, 7)\}.$$

Calcula el coeficient de correlació i interpreta els resultats.

2.2.3 Ajustament potencial

Suposem que la mostra $\{(x_i, y_i)\}_{i=1}^n$ procedeix a partir d'una funció potencial de la forma $f(x) = \alpha x^\beta$. Per tal d'ajustar α, β , considerem la funció $y = g(x) = ax^b$. Podem obtenir a i b per mínims quadrats reduint de nou el problema al cas lineal. Per a fer-ho, prenem logaritmes a ambdós costats de la igualtat i apliquem les seues propietats, tot resultant

$$\log(y) = \log(ax^b) = \log(a) + \log(x^b) = \log(a) + b \log(x).$$

Si fem els següents canvis de variable i paràmetres

$$\left. \begin{aligned} z &= \log(x) \\ u &= \log(y) \\ c &= \log(a) \end{aligned} \right\}$$

aleshores tenim l'equació d'una recta amb les noves variables i paràmetres:

$$u = c + bz.$$

c i b s'obtenen de la mateixa manera que en el cas de la recta de regressió, amb la precaució que ara es treballa amb les noves variables, és a dir,

$$b = \frac{s_{zu}}{s_z^2}, \quad c = \bar{u} - b\bar{z}.$$

Per tal d'obtenir a , cal desfer el canvi $c = \log(a)$, tot obtenint-se

$$a = e^c.$$

De la mateixa manera que abans, es pot calcular també el coeficient de correlació per

$$r = \frac{s_{zu}}{s_z s_u}.$$

Exemple *Ajustar potencialment les dades següents:*

$$\{(0.25, 1.39), (0.8, 2.22), (1.27, 2.89), (1.98, 2.89), (2.1, 3.59), (2.55, 4.63), (3, 4.37)\}.$$

Solució *Separant per submostres, es té*

$$\begin{aligned} X &= \{0.25, 0.8, 1.27, 1.98, 2.1, 2.55, 3\}, \\ Y &= \{1.39, 2.22, 2.89, 2.89, 3.59, 4.63, 4.37\}. \end{aligned}$$

Fent els corresponents canvis de variable $z_i = \log(x_i)$ i $u_i = \log(y_i)$, s'obtenen les següents submostres transformades:

$$\begin{aligned} Z &= \log(X) = \{\log(0.25), \log(0.8), \log(1.27), \log(1.98), \log(2.1), \log(2.55), \log(3)\}, \\ &\approx \{-1.39, -0.22, 0.24, 0.68, 0.74, 0.94, 1.10\}, \\ U &= \log(Y) = \{\log(1.39), \log(2.22), \log(2.89), \log(2.89), \log(3.59), \log(4.63), \log(4.37)\} \\ &\approx \{0.33, 0.80, 1.06, 1.06, 1.28, 1.53, 1.47\}. \end{aligned}$$

D'altra banda, es té

$$\begin{aligned} \bar{z} &\approx \frac{-1.39 - 0.22 + 0.24 + 0.68 + 0.74 + 0.94 + 1.10}{7} \approx 0.3, \\ \bar{u} &\approx \frac{0.33 + 0.80 + 1.06 + 1.06 + 1.28 + 1.53 + 1.47}{7} \approx 1.08, \\ s_z^2 &\approx \frac{(-1.39)^2 + (-0.22)^2 + 0.24^2 + 0.68^2 + 0.74^2 + 0.94^2 + 1.10^2}{7} - 0.3^2 \approx 0.64, \\ s_u^2 &\approx \frac{0.33^2 + 0.80^2 + 1.06^2 + 1.06^2 + 1.28^2 + 1.53^2 + 1.47^2}{7} - 1.08^2 \approx 0.15, \\ s_{zu} &\approx \frac{(-1.39) \cdot 0.33 + (-0.22) \cdot 0.80 + 0.24 \cdot 1.06 + 0.68 \cdot 1.06 + 0.74 \cdot 1.28 + 0.94 \cdot 1.53 + 1.10 \cdot 1.47}{7} \\ &\quad - 0.3 \cdot 1.08 \approx 0.3. \end{aligned}$$

Obtenim per tant

$$b = \frac{s_{zu}}{s_z^2} \approx 0.47, \quad c = \bar{u} - b\bar{z} \approx 0.94.$$

Desfent el canvi, resulta

$$a = e^c \approx 2.55$$

i, per tant, la funció potencial aproximada que ajusta les dades per mínims quadrats és

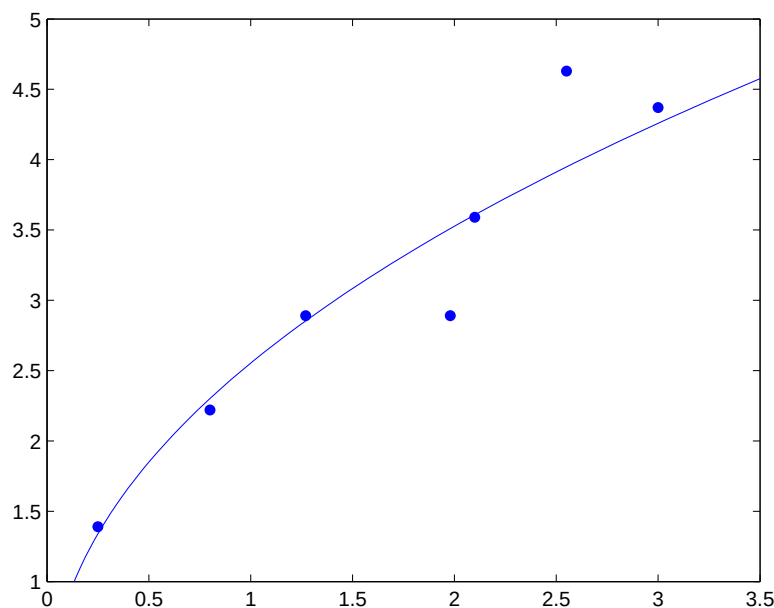
$$2.55x^{0.47},$$

amb coeficient de correlació

$$r = \frac{s_{zu}}{s_z s_u} \approx 0.97,$$

per tant, ens tornem a trobar en un cas d'ajustament raonablement bo.

El diagrama de dispersió amb la corresponent funció de l'ajustament és el següent:



2.2.4 Ajustament exponencial

En aquest cas, treballem sota la hipòtesi que la mostra $\{(x_i, y_i)\}_{i=1}^n$ ha estat generada a partir d'una funció exponencial de la forma $f(x) = \alpha\beta^x$. Llavors, es procedeix a ajustar-les a la funció $y = g(x) = ab^x$. Novament, podem simplificar aquest problema al cas lineal prenent logaritmes:

$$\log(y) = \log(ab^x) = \log(a) + \log(b^x) = \log(a) + \log(b)x.$$

Fent els canvis de variable i paràmetres

$$\left. \begin{aligned} u &= \log(y) \\ c &= \log(a) \\ d &= \log(b) \end{aligned} \right\}$$

l'expressió anterior es converteix en la següent recta

$$u = c + dx,$$

on c i d s'obtenen mitjançant el procediment d'ajustament lineal amb les noves variables,

$$d = \frac{s_{xu}}{s_x^2}, \quad c = \bar{u} - d\bar{x}.$$

Per tal d'obtenir a i b , cal desfer els canvis de variable $c = \log(a)$ i $d = \log(b)$, és a dir, es calculen per $a = e^c$ i $b = e^d$.

D'altra banda, el coeficient de correlació es pot calcular per

$$r = \frac{s_{xu}}{s_x s_u}.$$

Exemple *Ajusta exponencialment les dades següents*

$$\{(0.4, 1.88), (0.9, 1.42), (1.2, 1.06), (1.6, 0.84), (2.1, 0.62)\}.$$

Solució *Per components, tenim*

$$X = \{0.4, 0.9, 1.2, 1.6, 2.1\},$$

$$Y = \{1.88, 1.42, 1.06, 0.84, 0.62\}.$$

Fem el canvi $u_i = \log(y_i)$, tot obtenint

$$U = \log(Y) = \{\log(1.88), \log(1.42), \log(1.06), \log(0.84), \log(0.62)\} \approx \{0.63, 0.35, 0.06, -0.17, -0.48\}.$$

D'altra banda,

$$\bar{x} = \frac{0.4 + 0.9 + 1.2 + 1.6 + 2.1}{5} = 1.24,$$

$$\bar{u} \approx \frac{0.63 + 0.35 + 0.06 - 0.17 - 0.48}{5} \approx 0.08,$$

$$s_x^2 \approx \frac{0.4^2 + 0.9^2 + 1.2^2 + 1.6^2 + 2.1^2}{5} - 1.24^2 \approx 0.34,$$

$$s_u^2 \approx \frac{0.63^2 + 0.35^2 + 0.06^2 - 0.17^2 - 0.48^2}{5} - 0.08^2 \approx 0.15,$$

$$s_{xu} \approx \frac{0.4 \cdot 0.63 + 0.9 \cdot 0.35 + 1.2 \cdot 0.06 + 1.6 \cdot (-0.17) + 2.1 \cdot (-0.48)}{5} - 1.24 \cdot 0.08 \approx -0.23.$$

Amb açò, obtenim els termes de la recta de regressió per a les variables i paràmetres resultants de la transformació.

$$d = \frac{s_{su}}{s_x^2} \approx -0.67, \quad c = \bar{u} - d\bar{x} \approx 0.9.$$

Podem desfer el canvi mitjançant

$$a = e^c \approx 2.47, \quad b = e^d \approx 0.51.$$

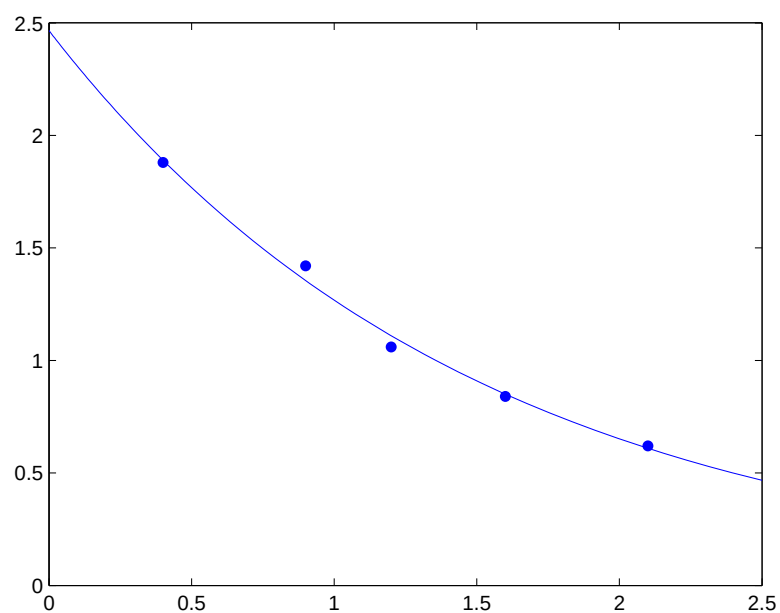
Així doncs, la funció exponencial de l'ajustament resulta, de manera aproximada,

$$y = 2.47 \cdot 0.51^x.$$

Finalment, obtenim el coeficient de correlació:

$$r = \frac{s_{xu}}{s_x s_u} \approx -0.9968,$$

la qual cosa indica que l'ajustament és excel·lent per estar $|r|$ molt a prop de 1. Podem comprovar en el diagrama de dispersió amb la funció de l'ajustament, que es presenta a continuació, que efectivament aquesta funció ajusta molt bé les dades.

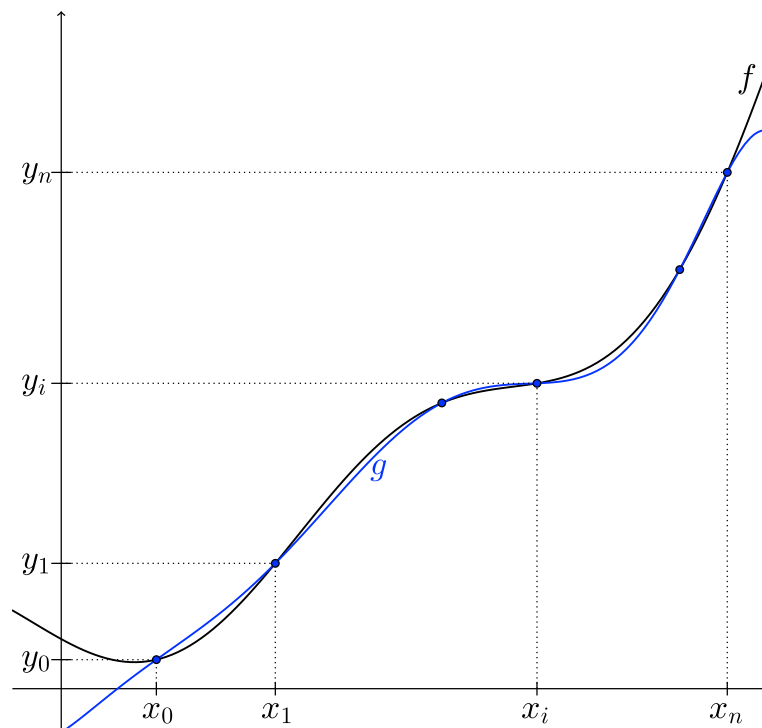


Capítol 3

Interpolació polinòmica

La **interpolació** és una tècnica molt útil que s'empra per a reconstruir valors d'una funció f en punts on aquesta no es coneix, a partir d'un conjunt de $n + 1$ valors o **nodes** $\{x_i\}_{i=0}^n$ on sí que es coneix, és a dir, disposem també d'altres $n + 1$ valors $\{y_i\}_{i=0}^n$ tals que $y_i = f(x_i)$, també anomenats **valors nodals**.

Amb la notació anterior, suposem que els nodes estan ordenats i són distints entre si, és a dir, $x_0 < x_1 < \dots < x_n$. Aleshores, el problema d'interpolació consisteix a trobar una funció g a partir de la informació nodal disponible, $\{(x_i, y_i)\}_{i=0}^n$, complint que $g(x) \approx f(x)$. Com que se sap que $y_i = f(x_i)$, $0 \leq i \leq n$, la condició que imposem per tal que g siga una funció interpoladora de f és que també verifiqui la condició $y_i = g(x_i)$, $0 \leq i \leq n$.



A pesar de tot, el plantejament anterior, amb les hipòtesis sobre g , és massa genèric. Hi un nombre infinit de funcions que compleixen aquesta condició i a més, com és natural, no totes són una bona aproximació de f fora dels nodes on aquesta es coneix. Per tant, el que anem fer és treballar amb la família de funcions més simples possible que poden complir aquestes condicions, que són les **funcions polinòmiques**. Així doncs, el problema es redueix al de trobar un polinomi p complint

$$p(x_i) = y_i, \quad 0 \leq i \leq n.$$

Donat que açò indueix la verificació de $n + 1$ condicions simultàneament, aquest polinomi genèric hauria de tenir almenys $n + 1$ graus de llibertat per tal d'aspirar a l'existència de solució; de fet, per tal d'aspirar també a la unicitat, caldria que en tinguera exactament $n + 1$. Els polinomis que verifiquen aquesta condició són els de grau menor o igual que n , ja que tenen $n + 1$ coeficients:

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n = \sum_{j=0}^n a_jx^j.$$

Així doncs, la imposició de les $n + 1$ condicions de la forma $p(x_i) = y_i$, $0 \leq i \leq n$, genera un sistema de

$n + 1$ equacions amb $n + 1$ incògnites, $(a_0, a_1, \dots, a_n)$, que presenta el següent aspecte:

$$\left. \begin{aligned} a_0 + x_0 a_1 + x_0^2 a_2 + \dots + x_0^{n-1} a_{n-1} + x_0^n a_n &= y_0 \\ a_0 + x_1 a_1 + x_1^2 a_2 + \dots + x_1^{n-1} a_{n-1} + x_1^n a_n &= y_1 \\ &\vdots \\ a_0 + x_n a_1 + x_n^2 a_2 + \dots + x_n^{n-1} a_{n-1} + x_n^n a_n &= y_n \end{aligned} \right\}$$

El resultat següent assegura que el sistema anterior és sempre compatible determinat o, el que és el mateix, l'existència i unicitat del polinomi interpolador.

Teorema Donat un conjunt de $n + 1$ punts al pla, $\{(x_i, y_i)\}_{i=0}^n$, existeix un únic polinomi p de grau menor o igual que n (és a dir, $p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{n-1} x^{n-1} + a_n x^n$) tal que

$$p(x_i) = y_i, \quad 0 \leq i \leq n.$$

Definició El polinomi anterior rep el nom de **polinomi interpolador** de f als nodes $\{x_i\}_{i=0}^n$.

Exemple Calcular el polinomi interpolador que passa pels punts $\{(1, 2), (3, 0)\}$.

Solució Hem de trobar un polinomi p complint $p(1) = 2$ i $p(3) = 0$. Com que tenim 2 condicions, el grau del polinomi és $n \leq 2 - 1 = 1$, és a dir, una recta de la forma $p(x) = a_0 + a_1 x$. Imposant les dues condicions anteriors, obtenim el següent sistema 2×2 .

$$\left. \begin{aligned} a_0 + a_1 &= 2 \\ a_0 + 3a_1 &= 0 \end{aligned} \right\}$$

Resolent, s'obté

$$\left\{ \begin{aligned} a_0 &= 3 \\ a_1 &= -1 \end{aligned} \right.$$

Per tant, el polinomi interpolador és $p(x) = 3 - x$.

Exemple Siga $f : \mathbb{R} \rightarrow \mathbb{R}$ donada per $f(x) = \sin(\pi x) \quad \forall x \in \mathbb{R}$. Es demana calcular el polinomi interpolador de f als nodes $\{0, \frac{1}{6}, \frac{1}{2}\}$. Calcula el valor de dit polinomi interpolador a $x = \frac{1}{3}$ i compara el resultat amb $f(\frac{1}{3})$.

Solució Cal obtenir un polinomi p complint les condicions següents

$$\left. \begin{aligned} p(0) &= f(0) = \sin(0) = 0 \\ p\left(\frac{1}{6}\right) &= f\left(\frac{1}{6}\right) = \sin\left(\frac{\pi}{6}\right) = \frac{1}{2} \\ p\left(\frac{1}{2}\right) &= f\left(\frac{1}{2}\right) = \sin\left(\frac{\pi}{2}\right) = 1 \end{aligned} \right\}$$

Com que ara tenim tres condicions el grau de polinomi és $n \leq 3 - 1 = 2$, és a dir, una paràbola de la forma $p(x) = a_0 + a_1x + a_2x^2$, on els seus coeficients, aplicant les condicions anteriors, són solució del següent sistema d'equacions 3×3

$$\left. \begin{array}{rcl} a_0 & & = 0 \\ a_0 + \frac{1}{6}a_1 + \frac{1}{36}a_2 & = & \frac{1}{2} \\ a_0 + \frac{1}{2}a_1 + \frac{1}{4}a_2 & = & 1 \end{array} \right\}$$

la solució del qual és

$$\left\{ \begin{array}{l} a_0 = 0 \\ a_1 = \frac{7}{2} \\ a_2 = -3 \end{array} \right.$$

Per tant, el polinomi interpolador és

$$p(x) = \frac{7}{2}x - 3x^2.$$

D'altra banda,

$$f\left(\frac{1}{3}\right) = \sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2},$$

mentre que

$$p\left(\frac{1}{3}\right) = \frac{7}{2} \cdot \frac{1}{3} - 3\left(\frac{1}{3}\right)^2 = \frac{5}{6}.$$

Així doncs,

$$f\left(\frac{5}{6}\right) - p\left(\frac{5}{6}\right) = \frac{\sqrt{3}}{2} - \frac{5}{6} \approx 0.0327.$$

La darrera part de l'exemple anterior motiva la introducció i estudi del concepte de l'**error d'interpolació**, que veurem tot seguit.

3.1 Error d'interpolació

Sota certes condicions de diferenciabilitat, i en disposició d'informació sobre la derivada d'un cert grau de f , es pot fer un estudi teòric sobre l'error comés a l'interpolació a un cert punt. El resultat següent especifica exactament quina és l'expressió de l'error per als polinomis interpoladors.

Teorema Siguen $a, b \in \mathbb{R}$ tals que $a < b$, siga f tal que $f \in C^{(n+1)}((a, b))$ (és a dir, f és diferenciable $n + 1$ voltes amb derivada $(n + 1)$ -èssima, $f^{(n+1)}$, contínua), siga $\{x_i\}_{i=0}^n$ un conjunt de $n + 1$ nodes complint $a < x_0 < x_1 < \dots < x_n < b$ i $\{y_i\}_{i=0}^n$ un conjunt de $n + 1$ valors donats per $y_i = f(x_i)$, $0 \leq i \leq n$. Aleshores, $\forall x \in (a, b)$ el polinomi interpolador p als punts $\{(x_i, y_i)\}_{i=0}^n$ compleix que $\exists \xi = \xi(x) \in (x_0, x_n) \subseteq (a, b)$ (ξ depèn de x) tal que

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)(x - x_1) \cdots (x - x_n).$$

En particular, si denotem per $M = \max_{x_0 \leq t \leq x_n} |f^{(n+1)}(t)|$, es té la següent fita de l'error d'interpolació en el punt x :

$$|f(x) - p(x)| \leq \frac{M}{(n+1)!} |(x-x_0)(x-x_1)\cdots(x-x_n)| = \frac{M}{(n+1)!} \left| \prod_{i=0}^n (x-x_i) \right|.$$

Exemple Calculem la fita teòrica de l'error d'interpolació per a l'exemple anterior al punt on s'ha obtingut l'aproximació, $x = \frac{1}{3}$. Com que el polinomi interpolador s'ha calculat a partir de $n+1 = 3$ punts, hem de calcular, en primer lloc, $f^{(3)} = f'''$. D'altra banda, es té que $f(x) = \sin(\pi x)$, i per tant es pot comprovar fàcilment que

$$f'''(x) = -\pi^3 \cos(\pi x).$$

Ara hem de fitar el mòdul d'aquesta derivada des de $x_0 = 0$ fins a $x_n = x_2 = \frac{1}{2}$, tot resultant:

$$M = \max_{0 \leq t \leq \frac{1}{2}} |f'''(t)| = \pi^3 \cdot \max_{0 \leq t \leq \frac{1}{2}} |\cos(\pi t)| = \pi^3,$$

ja que el màxim per a $t \in (0, \frac{1}{2})$ de $|\cos(\pi t)|$ s'assoleix a $t = 0$, en què $\cos(\pi \cdot 0) = 1$. La fita de l'error per a un $x \in \mathbb{R}$ qualsevol és, en aquest cas,

$$|f(x) - p(x)| \leq \frac{M}{(n+1)!} |(x-x_0)(x-x_1)(x-x_2)| = \frac{\pi^3}{3!} \left| (x-0) \left(x - \frac{1}{6}\right) \left(x - \frac{1}{2}\right) \right|.$$

En particular, per a $x = \frac{1}{3}$ es té

$$\left| f\left(\frac{5}{6}\right) - p\left(\frac{5}{6}\right) \right| \leq \frac{\pi^3}{3!} \left| \left(\frac{1}{3} - 0\right) \left(\frac{1}{3} - \frac{1}{6}\right) \left(\frac{1}{3} - \frac{1}{2}\right) \right| = \frac{\pi^3}{648}.$$

Notem que açò és coherent amb l'error obtingut de manera exacta a l'exemple anterior, ja que, efectivament, es compleix

$$0.0327 \approx \left| f\left(\frac{5}{6}\right) - p\left(\frac{5}{6}\right) \right| \leq \frac{\pi^3}{648} \approx 0.0478.$$

3.2 Fórmules d'interpolació

Fins al moment, l'única tècnica que hem vist per tal d'obtenir l'expressió d'un polinomi interpolador és a partir de l'obtenció dels seus coeficients, resolent directament el sistema induït per les condicions que es deriven dels punts per on ha de passar. Açò és una tècnica viable si el nombre de nodes d'interpolació és petit, com als exemples que hem vist fins ara; no obstant això, no és una tècnica viable (ni eficient) en el cas que el nombre de punts augmente significativament.

És per això que en aquesta secció anem a veure un seguit de tècniques que permeten calcular d'una manera més ràpida i senzilla les expressions dels polinomis interpoladors, que són especialment útils quan hi ha un nombre elevat de nodes, que dificulta excessivament la resolució manual del sistema d'equacions per a l'obtenció dels coeficients.

3.2.1 Fórmula de Lagrange

El mètode de Lagrange consisteix en la descomposició d'un polinomi interpolador, complint $p(x_i) = y_i$, $0 \leq i \leq n$, i que es desitja calcular, en una combinació lineal dels valors nodals $\{y_i\}_{i=0}^n$:

$$p(x) = p_0(x)y_0 + p_1(x)y_1 + \dots + p_n(x)y_n = \sum_{i=0}^n p_i(x)y_i.$$

El que es fa ara és imposar una sèrie de condicions sobre els polinomis $\{p_i\}_{i=0}^n$ per tal que es complisca la condició de polinomi interpolador, $p(x_i) = y_i$, $0 \leq i \leq n$.

En primer lloc, volem que $p(x_0) = y_0$. Per tal d'aconseguir-ho, i tenint en compte la condició anterior, serà suficient amb fer que $p_0(x_0) = 1$ (terme que multiplica a y_0) i, d'altra banda, els termes que multipliquen a la resta de y_i , $i \neq 0$, verifiquen $p_0(x_i) = 0$. D'eixa manera, aconseguim garantir la primera condició:

$$p(x_0) = p_0(x_0)y_0 + p_1(x_0)y_1 + \dots + p_n(x_0)y_n = 1 \cdot y_0 + 0 \cdot y_1 + \dots + 0 \cdot y_n = y_0.$$

La idea ara és aplicar el mateix criteri corresponent a la resta de casos. En general, per tal que es garantisca la condició $p(x_i) = y_i$, és suficient amb imposar $p_i(x_i) = 1$ i $p_i(x_j) = 0$ per a $j \neq i$, ja que d'eixa manera s'aconsegueix el que es vol obtenir:

$$\begin{aligned} p(x_i) &= p_0(x_i)y_0 + \dots + p_{i-1}(x_i)y_{i-1} + p_i(x_i)y_i + p_{i+1}(x_i)y_{i+1} + \dots + p_n(x_i)y_n \\ &= 0 \cdot y_0 + \dots + 0 \cdot y_{i-1} + 1 \cdot y_i + 0 \cdot y_{i+1} + \dots + 0 \cdot y_n = y_i. \end{aligned}$$

La imposició d'aquestes $n + 1$ condicions sobre cadascun dels p_i es tradueix com se segueix:

$$p_i(x_j) = \delta_i^j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Es pot comprovar que els polinomis que verifiquen cadascuna de les condicions anteriors estan unívocament determinats i venen donats per:

$$p_i(x) = \frac{(x - x_0) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_0) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)} = \frac{\prod_{j=0, j \neq i}^n (x - x_j)}{\prod_{j=0, j \neq i}^n (x_i - x_j)}, \quad 0 \leq i \leq n.$$

En efecte, ja que si $j \neq i$ aleshores, com que el numerador de $p(x)$ conté un producte de monomis, on un d'ells és de la forma $(x - x_j)$, llavors avaluant a $x = x_j$ aquest monomi s'anul·la, i per tant $p_i(x_j) = 0$. D'altra banda, s'observa que si $x = x_i$ el numerador coincideix amb el denominador, i per tant el resultat del quocient és la unitat, és a dir, $p_i(x_i) = 1$.

El conjunt de polinomis $\{p_i\}_{i=0}^n$ rep el nom de **base de Lagrange**.

Exemple Considerem l'exemple amb què hem treballat amb anterioritat i anem a calcular el polinomi interpolador pel mètode de Lagrange, tot comprovant que el resultat és el mateix que l'obtingut amb el procediment anterior.

Cerquem un polinomi p verificant

$$\left. \begin{aligned} p(0) &= 0 \\ p\left(\frac{1}{6}\right) &= \frac{1}{2} \\ p\left(\frac{1}{2}\right) &= 1 \end{aligned} \right\}$$

Per a fer-ho, calculem primer els polinomis de la base de Lagrange:

$$p_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - \frac{1}{6})(x - \frac{1}{2})}{(0 - \frac{1}{6})(0 - \frac{1}{2})} = 12x^2 - 8x + 1.$$

$$p_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 0)(x - \frac{1}{2})}{(\frac{1}{6} - 0)(\frac{1}{6} - \frac{1}{2})} = -18x^2 + 9x.$$

$$p_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 0)(x - \frac{1}{6})}{(\frac{1}{2} - 0)(\frac{1}{2} - \frac{1}{6})} = 6x^2 - x.$$

Per tant, l'expressió del polinomi interpolador resulta

$$p(x) = p_0(x)y_0 + p_1(x)y_1 + p_2(x)y_2 = (12x^2 - 8x + 1) \cdot 0 + (-18x^2 + 9x) \cdot \frac{1}{2} + (6x^2 - x) \cdot 1 = -3x^2 + \frac{7}{2}x,$$

que, efectivament, coincideix amb l'obtingut resolent el sistema.

3.2.2 Fórmula de Newton

El mètode que es presenta tot seguit pot ser aplicat **únicament** quan tots els nodes de les abscisses $\{x_i\}_{i=0}^n$ estan equiespaiats, és a dir, a distància constant $x_{i+1} - x_i = h$, $0 \leq i < n$, ($h > 0$ independent de i); o el que és mateix, es poden escriure com $x_i = x_0 + ih$, $0 \leq i \leq n$.

Per exemple, per a $n = 2$, els nodes $x_0 = 0$, $x_1 = \frac{1}{2}$, $x_2 = 1$ estan equiespaiats, ja que $x_1 - x_0 = \frac{1}{2} - 0 = \frac{1}{2}$ i $x_2 - x_1 = 1 - \frac{1}{2} = \frac{1}{2}$, amb $h = x_1 - x_0 = x_2 - x_1 = \frac{1}{2}$; per contra, els nodes $x_0 = 0$, $x_1 = \frac{1}{6}$ i $x_2 = \frac{1}{2}$ no estan equiespaiats, ja que $x_1 - x_0 = \frac{1}{6} - 0 = \frac{1}{6}$ i $x_2 - x_1 = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$, i per tant $x_1 - x_0 \neq x_2 - x_1$.

Si els nodes verifiquen estar equiespaiats, la manera de procedir és la que es presenta a continuació.

En primer lloc, es calcula, columna per columna, la següent taula de **diferències no dividides**:

y	Δy	$\Delta^2 y$	$\dots$	$\Delta^n y$
y_0	$\Delta y_0 = y_1 - y_0$	$\Delta^2 y_0 = \Delta y_1 - \Delta y_0$	$\dots$	$\Delta^n y_0 = \Delta^{n-1} y_1 - \Delta^{n-1} y_0$
y_1	$\Delta y_1 = y_2 - y_1$	$\Delta^2 y_1 = \Delta y_2 - \Delta y_1$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\ddots$	
y_{n-2}	$\Delta y_{n-2} = y_{n-1} - y_{n-2}$	$\Delta^2 y_{n-2} = \Delta y_{n-1} - \Delta y_{n-2}$		
y_{n-1}	$\Delta y_{n-1} = y_n - y_{n-1}$			
y_n				

Finalment, a partir dels valors obtinguts a la taula, podem obtenir el polinomi interpolador de dues maneres diferents:

- **Primera fórmula de Newton:**

$$p(x) = y_0 + q\Delta y_0 + \frac{q(q-1)}{2!}\Delta^2 y_0 + \dots + \frac{q(q-1)\dots(q-n+1)}{n!}\Delta^n y_0,$$

$$\text{on } q = \frac{x - x_0}{h}.$$

- **Segona fórmula de Newton:**

$$p(x) = y_n + q\Delta y_{n-1} + \frac{q(q+1)}{2!}\Delta^2 y_{n-2} + \dots + \frac{q(q+1)\cdots(q+n-1)}{n!}\Delta^n y_0,$$

$$\text{on } q = \frac{x - x_n}{h}.$$

Errors d'interpolació de la fórmula de Newton

En cas d'estar en condicions d'aplicar de fórmula de Newton (nodes equiespaiats), la fórmula de l'error d'interpolació se simplifica en cada cas com se segueix:

- **Error de la primera fórmula de Newton:**

$$f(x) - p(x) = f^{(n+1)}(\xi) \frac{q(q-1)\cdots(q-n)}{(n+1)!} h^{n+1},$$

$$\text{on } q = \frac{x - x_0}{h} \text{ i } \xi \in (x_0, x_n). \text{ Per tant, l'error es pot fitar per}$$

$$|f(x) - p(x)| \leq M \frac{|q(q-1)\cdots(q-n)|}{(n+1)!} h^{n+1},$$

$$\text{on } M = \max_{x_0 \leq t \leq x_n} |f^{(n+1)}(t)|.$$

- **Error de la segona fórmula de Newton:**

$$f(x) - p(x) = f^{(n+1)}(\xi) \frac{q(q+1)\cdots(q+n)}{(n+1)!} h^{n+1},$$

$$\text{on } q = \frac{x - x_n}{h} \text{ i } \xi \in (x_0, x_n). \text{ Per tant, l'error en aquest cas es pot fitar per}$$

$$|f(x) - p(x)| \leq M \frac{|q(q+1)\cdots(q+n)|}{(n+1)!} h^{n+1},$$

$$\text{on } M = \max_{x_0 \leq t \leq x_n} |f^{(n+1)}(t)|.$$

Exemple Siga $f : (-1, +\infty) \rightarrow \mathbb{R}$ donada per $f(x) = \frac{1}{1+x}$. Calcula el polinomi interpolador de f per als nodes $\{0, \frac{1}{3}, \frac{2}{3}, 1\}$ al valor $x = \frac{1}{2}$. Calcula una fita de l'error per a eixe punt i compara'l amb l'error exacte.

Solució Avaluant als nodes corresponents, obtenim:

$$\left. \begin{array}{lclcl} p(0) & = & f(0) & = & \frac{1}{1+0} = 1 \\ p\left(\frac{1}{3}\right) & = & f\left(\frac{1}{3}\right) & = & \frac{1}{1+\frac{1}{3}} = 0.75 \\ p\left(\frac{2}{3}\right) & = & f\left(\frac{2}{3}\right) & = & \frac{1}{1+\frac{2}{3}} = 0.6 \\ p(1) & = & f(1) & = & \frac{1}{1+1} = 0.5 \end{array} \right\}$$

La taula de diferències dividides resulta per tant

y	Δy	$\Delta^2 y$	$\Delta^3 y$
1	$0.75 - 1 = -0.25$	$-0.15 - (-0.25) = 0.1$	$0.05 - 0.1 = -0.05$
0.75	$0.6 - 0.75 = -0.15$	$-0.1 - (-0.15) = 0.05$	
0.6	$0.5 - 0.6 = -0.1$		
0.5			

D'altra banda, $h = \frac{1}{3}$ i, com que es demana únicament el valor del polinomi interpolador a $x = \frac{1}{2} = 0.5$, podem calcular directament el valor sense necessitat d'obtenir el polinomi interpolador en funció d'un valor x genèric. Així doncs, emprant per exemple la primer fórmula de Newton, calculem

$$q = \frac{x - x_0}{h} = \frac{0.5 - 0}{\frac{1}{3}} = 1.5$$

i per tant es té

$$\begin{aligned} p(0.5) &= y_0 + q\Delta y_0 + \frac{q(q-1)}{2!}\Delta^2 y_0 + \frac{q(q-1)(q-2)}{3!}\Delta^3 y_0 \\ &= 1 + 1.5 \cdot (-0.25) + \frac{1.5 \cdot 0.5}{2} \cdot 0.1 + \frac{1.5 \cdot 0.5 \cdot (-0.5)}{6} \cdot (-0.05) = \frac{213}{320} = 0.665625. \end{aligned}$$

L'error exacte és per tant

$$f(0.5) - p(0.5) = \frac{1}{1 + \frac{1}{2}} - \frac{213}{320} = \frac{2}{3} - \frac{213}{320} = \frac{1}{960} \approx 0.001042.$$

Per altra banda, la fita de l'error per a un $x \in (1, +\infty)$ genèric resulta ser en aquest cas

$$|f(x) - p(x)| \leq M \left| \frac{q(q-1)(q-2)(q-3)}{4!} \right| h^4,$$

amb

$$M = \max_{0 \leq t \leq 1} |f^{IV}(t)| = \max_{0 \leq t \leq 1} \left| \frac{24}{(x+1)^5} \right| = 24.$$

Substituint en cada cas per a $x = 0.5$,

$$|f(0.5) - p(0.5)| \leq 24 \left| \frac{1.5(1.5-1)(1.5-2)(1.5-3)}{4!} \right| \left(\frac{1}{3} \right)^4 = \frac{1}{144} \approx 0.00694.$$

Comprovem novament que la fita de l'error està en concordància amb els resultats teòrics, ja que

$$\frac{1}{960} = |f(0.5) - p(0.5)| \leq \frac{1}{144}.$$

Exercici 21 Repeteix l'exemple anterior emprant la segona fórmula de Newton i la corresponent fórmula de la fita de l'error. Comprova que els resultats obtinguts són els mateixos.

3.2.3 Fórmula de les diferències dividides

En el cas que els nodes de les abscisses $\{x_i\}_{i=0}^n$ no estiguen equiespaiats, les fórmules de Newton descrites anteriorment no són aplicables i cal cercar altres procediments. Per tal d'abordar aquest cas, primer necessitem introduir el concepte de **diferències dividides**.

Donada la informació nodal en el pla, $\{(x_i, y_i)\}_{i=0}^n$, amb $y_i = f(x_i)$, es defineixen les diferències dividides de primer ordre com

$$f[x_i, x_{i+1}] = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad 0 \leq i \leq n-1.$$

Les de segon ordre, es defineixen a partir de les de primer ordre per

$$f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i}, \quad 0 \leq i \leq n-2.$$

En general, definim les diferències dividides d'ordre j a partir de les diferències dividides d'ordre $j-1$ com

$$f[x_i, x_{i+1}, \dots, x_{i+j}] = \frac{f[x_{i+1}, x_{i+2}, \dots, x_{i+j}] - f[x_i, x_{i+1}, \dots, x_{i+j-1}]}{x_{i+j} - x_i}, \quad 0 \leq i \leq n-j.$$

Per simplicitat de notació, denotarem d'ara endavant per D_i^j a la corresponent diferència dividida d'ordre j ,

$$D_i^j = f[x_i, x_{i+1}, \dots, x_{i+j}], \quad 0 \leq i \leq n-j.$$

Els valors obtinguts a les diferències dividides es corresponen amb aproximacions a un terme proporcional a la derivada del mateix ordre que la diferència dividida. Concretament,

$$D_i^j \approx \frac{f^{(j)}(x)}{j!}, \quad x \in (x_0, x_n).$$

Per tal de calcular el polinomi interpolador mitjançant diferències dividides, es construeix primer una taula columna per columna de la següent manera:

x	y	D	D^2	$\dots$	D^n
x_0	y_0	$D_0 = \frac{y_1 - y_0}{x_1 - x_0}$	$D_0^2 = \frac{D_1 - D_0}{x_2 - x_0}$	$\dots$	$D_0^n = \frac{D_1^{n-1} - D_0^{n-1}}{x_n - x_0}$
x_1	y_1	$D_1 = \frac{y_2 - y_1}{x_2 - x_1}$	$D_1^2 = \frac{D_2 - D_1}{x_3 - x_1}$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	
x_{n-2}	y_{n-2}	$D_{n-2} = \frac{y_{n-1} - y_{n-2}}{x_{n-1} - x_{n-2}}$	$D_{n-2}^2 = \frac{D_{n-1} - D_{n-2}}{x_n - x_{n-2}}$		
x_{n-1}	y_{n-1}	$D_{n-1} = \frac{y_n - y_{n-1}}{x_n - x_{n-1}}$			
x_n	y_n				

I el polinomi interpolador es calcula per

$$p(x) = y_0 + D_0(x - x_0) + D_0^2(x - x_0)(x - x_1) + \dots + D_0^n(x - x_0)(x - x_1) \cdots (x - x_{n-1}).$$

Exemple Calculem el polinomi interpolador de la funció $f(x) = \sin(\pi x)$ als nodes $\{0, \frac{1}{6}, \frac{1}{2}, 1\}$ mitjançant diferències dividides.

El primer que fem, és calcular la taula de valors corresponent

x	y	D	D^2	D^3
0	0	$\frac{\frac{1}{2} - 0}{\frac{1}{6} - 0} = 3$	$\frac{\frac{3}{2} - 3}{\frac{1}{2} - 0} = \frac{-\frac{3}{2}}{\frac{1}{2}} = -3$	$\frac{-\frac{21}{5} - (-3)}{1 - 0} = -\frac{6}{5} = -1.2$
$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1 - \frac{1}{2}}{\frac{1}{2} - \frac{1}{6}} = \frac{3}{2} = 1.5$	$\frac{-2 - \frac{3}{2}}{1 - \frac{1}{6}} = -\frac{21}{5} = -4.2$	
$\frac{1}{2}$	1	$\frac{0 - 1}{1 - \frac{1}{2}} = -2$		
1	0			

Per tant, l'expressió del polinomi interpolador és

$$\begin{aligned} p(x) &= y_0 + D_0(x - x_0) + D_0^2(x - x_0)(x - x_1) + D_0^3(x - x_0)(x - x_1)(x - x_2) \\ &= 0 + 3(x - 0) - 3(x - 0) \left(x - \frac{1}{6}\right) - 1.2(x - 0) \left(x - \frac{1}{6}\right) \left(x - \frac{1}{2}\right) = \frac{17x - 11x^2 - 6x^3}{5} \\ &= -1.2x^3 - 2.2x^2 + 3.4x. \end{aligned}$$

Podem emprar ara el polinomi per a aproximar valors de $f(x)$. Per exemple, per a $x = \frac{5}{6}$, es té $f(\frac{5}{6}) = \sin(\frac{5\pi}{6}) = \frac{1}{2}$. Una aproximació d'aquest valor és per tant

$$p\left(\frac{5}{6}\right) = -1.2 \left(\frac{5}{6}\right)^3 - 2.2 \left(\frac{5}{6}\right)^2 + 3.4 \cdot \frac{5}{6} = \frac{11}{18}.$$

Aleshores l'error exacte és

$$f\left(\frac{5}{6}\right) - p\left(\frac{5}{6}\right) = \frac{1}{2} - \frac{11}{18} = -\frac{1}{9}.$$

D'altra banda, una fita de l'error en aquest cas ve donada per

$$|f(x) - p(x)| \leq \frac{M}{4!} |(x - x_0)(x - x_1)(x - x_2)(x - x_3)|,$$

on

$$M = \max_{t \in [x_0, x_3]} |f^{IV}(t)| = \max_{t \in [0, 1]} |\pi^4 \sin(\pi x)| = \pi^4.$$

En particular, per a $x = \frac{5}{6}$ resulta

$$0.1111 \approx \frac{1}{9} = \left| f\left(\frac{5}{6}\right) - p\left(\frac{5}{6}\right) \right| \leq \frac{\pi^4}{4!} \left| \left(\frac{5}{6} - 0\right) \left(\frac{5}{6} - \frac{1}{6}\right) \left(\frac{5}{6} - \frac{1}{2}\right) \left(\frac{5}{6} - 1\right) \right| = \frac{5\pi^4}{3888} \approx 0.1253.$$

Exercici 22 Calcula mitjançant el mètode de les diferències dividides el polinomi que interpola la funció prèvia als nodes $\{0, \frac{1}{6}, \frac{1}{2}\}$ i comprova que el resultat és el mateix que l'obtingut amb les tècniques anteriors.

3.2.4 Mètodes d'Aitken i Neville

Els mètodes que tot seguit presentem són especialment útils (i ràpids) en el cas que solament es necessite obtenir el valor del polinomi interpolador en un punt específic, sense necessitat de calcular la seua expressió general.

Donat un conjunt de nodes $\{(x_i, y_i)\}_{i=0}^n$, es descriu a continuació el procediment a seguir en cada cas per tal de calcular el valor $p(x)$, per a un valor x específic.

Mètode d'Aitken

Es calcula la següent taula de valors:

x	$P_0 = y$	P_1
x_0	$P_{0,0} = y_0$	
x_1	$P_{0,1} = y_1$	$P_{1,1} = \frac{(x_1 - x)P_{0,0} - (x_0 - x)P_{0,1}}{x_1 - x_0}$
x_2	$P_{0,2} = y_2$	$P_{1,2} = \frac{(x_2 - x)P_{0,0} - (x_0 - x)P_{0,2}}{x_2 - x_0}$
$\vdots$	$\vdots$	$\vdots$
x_{n-1}	$P_{0,n-1} = y_{n-1}$	$P_{1,n-1} = \frac{(x_{n-1} - x)P_{0,0} - (x_0 - x)P_{0,n-1}}{x_{n-1} - x_0}$
x_n	$P_{0,n} = y_n$	$P_{1,n} = \frac{(x_n - x)P_{0,0} - (x_0 - x)P_{0,n}}{x_n - x_0}$

P_2	$\dots$	P_n
$P_{2,2} = \frac{(x_2 - x)P_{1,1} - (x_1 - x)P_{1,2}}{x_2 - x_1}$		
$\vdots$	$\ddots$	
$P_{2,n-1} = \frac{(x_{n-1} - x)P_{1,1} - (x_1 - x)P_{1,n-1}}{x_{n-1} - x_1}$	$\dots$	
$P_{2,n} = \frac{(x_n - x)P_{1,1} - (x_1 - x)P_{1,n}}{x_n - x_1}$	$\dots$	$P_{n,n} = \frac{(x_n - x)P_{n-1,n-1} - (x_{n-1} - x)P_{n-1,n}}{x_n - x_{n-1}}$

I el valor del polinomi interpolador en x es calcula per $p(x) = P_{n,n}$.

Exemple Donada $f(x) = \sin(\pi x)$, es demana calcular el valor del polinomi interpolador f en els nodes $\{0, \frac{1}{6}, \frac{1}{2}, 1\}$ al punt $x = \frac{5}{6}$ mitjançant el mètode d'Aitken.

Solució Calculem la taula de valors especificada anteriorment per a aquest cas particular:

x	$P_0 = y$	P_1	P_2	P_3
0	0			
$\frac{1}{6}$	$\frac{1}{2}$	$\frac{(\frac{1}{6} - \frac{5}{6}) \cdot 0 - (0 - \frac{5}{6}) \cdot \frac{1}{2}}{\frac{1}{6} - 0} = \frac{5}{2}$		
$\frac{1}{2}$	1	$\frac{(\frac{1}{2} - \frac{5}{6}) \cdot 0 - (0 - \frac{5}{6}) \cdot 1}{\frac{1}{2} - 0} = \frac{5}{3}$	$\frac{(\frac{1}{2} - \frac{5}{6}) \cdot \frac{5}{2} - (\frac{1}{6} - \frac{5}{6}) \cdot \frac{5}{3}}{\frac{1}{2} - \frac{1}{6}} = \frac{5}{6}$	
1	0	$\frac{(1 - \frac{5}{6}) \cdot 0 - (0 - \frac{5}{6}) \cdot 0}{1 - 0} = 0$	$\frac{(1 - \frac{5}{6}) \cdot \frac{5}{2} - (\frac{1}{6} - \frac{5}{6}) \cdot 0}{1 - \frac{1}{6}} = \frac{1}{2}$	$\frac{(1 - \frac{5}{6}) \cdot \frac{5}{6} - (\frac{1}{2} - \frac{5}{6}) \cdot \frac{1}{2}}{1 - \frac{1}{2}} = \frac{11}{18}$

Per tant, $p(\frac{5}{6}) = \frac{11}{18}$, que efectivament coincideix amb el resultat obtingut emprant diferències dividides.

Mètode de Neville

Es calcula la taula de valors següent:

x	$P_0 = y$	P_1
x_0	$P_{0,0} = y_0$	$P_{1,0} = \frac{(x_1 - x)P_{0,0} - (x_0 - x)P_{0,1}}{x_1 - x_0}$
x_1	$P_{0,1} = y_1$	$P_{1,1} = \frac{(x_2 - x)P_{0,1} - (x_1 - x)P_{0,2}}{x_2 - x_1}$
$\vdots$	$\vdots$	$\vdots$
x_{n-2}	$P_{0,n-2} = y_{n-2}$	$P_{1,n-2} = \frac{(x_{n-1} - x)P_{0,n-2} - (x_{n-2} - x)P_{0,n-1}}{x_{n-1} - x_{n-2}}$
x_{n-1}	$P_{0,n-1} = y_{n-1}$	$P_{1,n-1} = \frac{(x_n - x)P_{0,n-1} - (x_{n-1} - x)P_{0,n}}{x_n - x_{n-1}}$
x_n	$P_{0,n} = y_n$	

P_2	$\dots$	P_n
$P_{2,0} = \frac{(x_2 - x)P_{1,0} - (x_0 - x)P_{1,1}}{x_2 - x_0}$	$\dots$	$P_{n,0} = \frac{(x_n - x)P_{n-1,0} - (x_0 - x)P_{n-1,1}}{x_n - x_0}$
$P_{2,1} = \frac{(x_3 - x)P_{1,1} - (x_1 - x)P_{1,2}}{x_3 - x_1}$	$\dots$	
$\vdots$	$\ddots$	
$P_{2,n-2} = \frac{(x_{n-1} - x)P_{1,n-2} - (x_{n-3} - x)P_{1,n-1}}{x_{n-1} - x_{n-3}}$		

El valor del polinomi interpolador en x es calcula per $p(x) = P_{n,0}$.

Exercici 23 Repeteix l'exemple anterior emprant el mètode de Neville i comprova que els resultats són els mateixos.

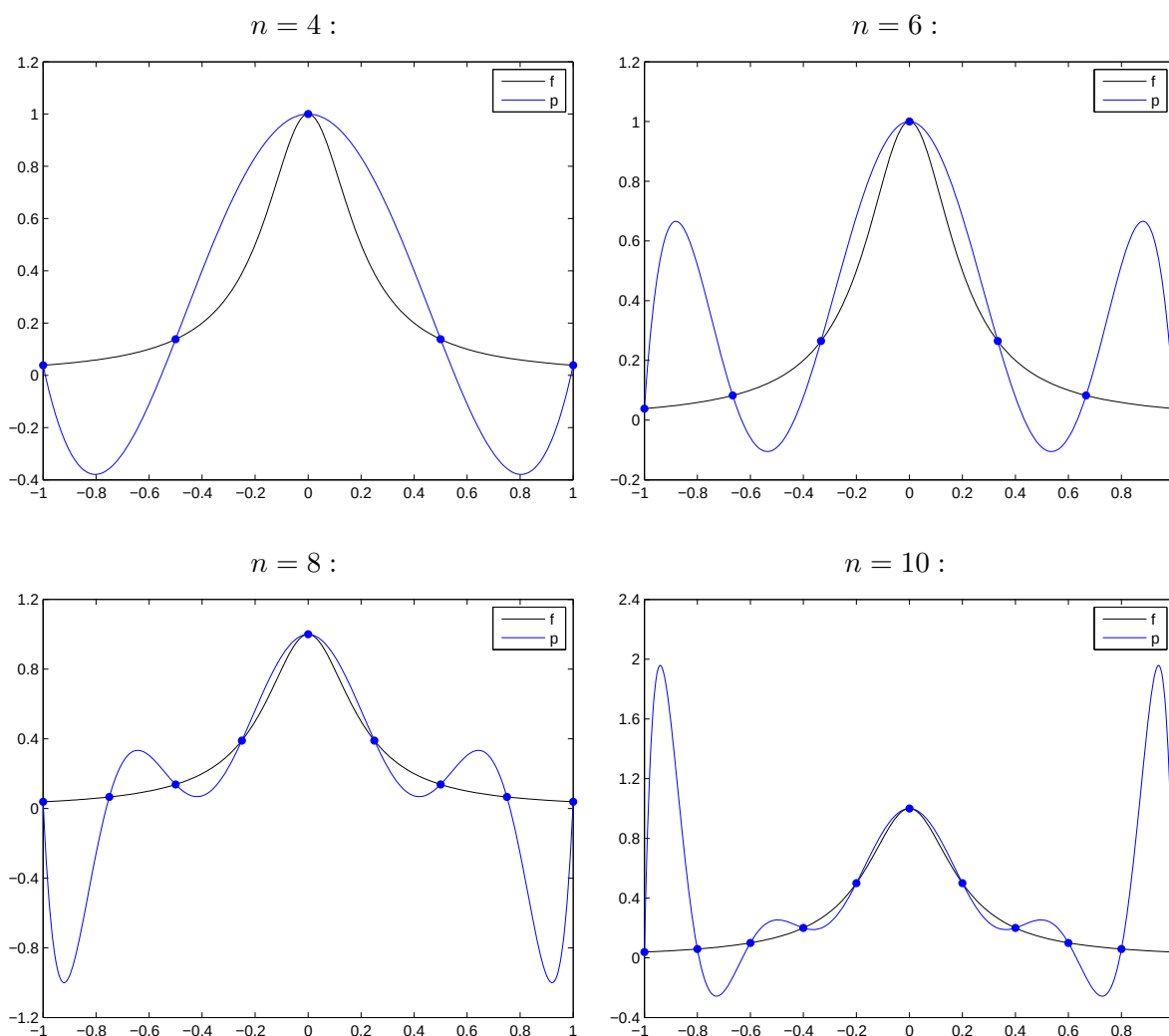
3.3 Interpolació a trossos

Hi ha casos en què la interpolació polinòmica amb un nombre elevat de punts no és adequada per a certs tipus de funcions, especialment quan el valor de les derivades d'aquestes augmenta sense control segons augmenta la grau de derivació.

Un exemple de funció que verifica la propietat anterior és $f : \mathbb{R} \rightarrow \mathbb{R}$ donada per

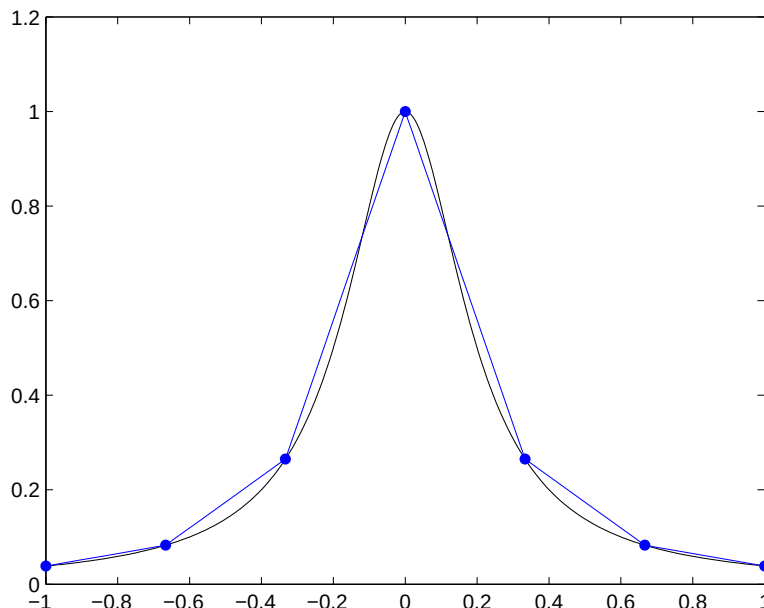
$$f(x) = \frac{1}{1 + 25x^2}, \quad x \in \mathbb{R}.$$

Si calculem els corresponents polinomis interpoladors amb $n+1$ nodes equidistants sobre l'interval $[-1, 1]$, $x_j = -1 + \frac{2j}{n}$, $0 \leq j \leq n$, podem veure a les gràfiques que es presenten a continuació que els resultats no són satisfactoris; de fet, les oscil·lacions s'agregen en augmentar el nombre de nodes.



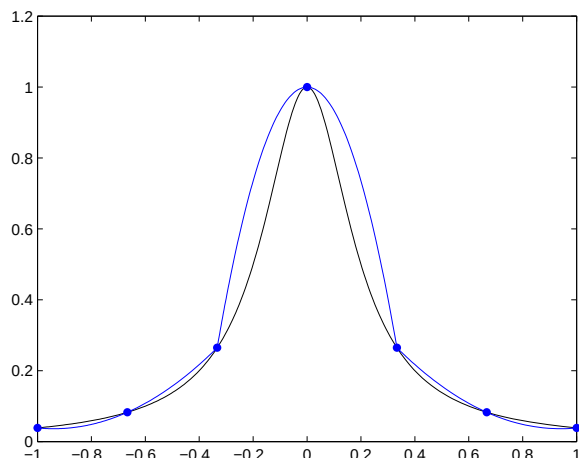
El fenomen anterior, conegut com **fenomen de Runge**, pot suposar un problema si, per motius de necessitat de precisió, es vol treballar amb un nombre elevat de nodes. Una manera de pal·liar-lo és mitjançant el que es coneix com **interpolació a trossos** o **segmentària**. La idea consisteix en el següent: en lloc de calcular un polinomi interpolador p als $n + 1$ nodes, calculem en el seu lloc diversos polinomis interpoladors en conjunts més petits de nodes.

Per exemple, en el cas anterior $n = 6$ (7 nodes), una interpolació segmentària amb polinomis de grau 1 consisteix simplement en connectar els nodes consecutius per segments:

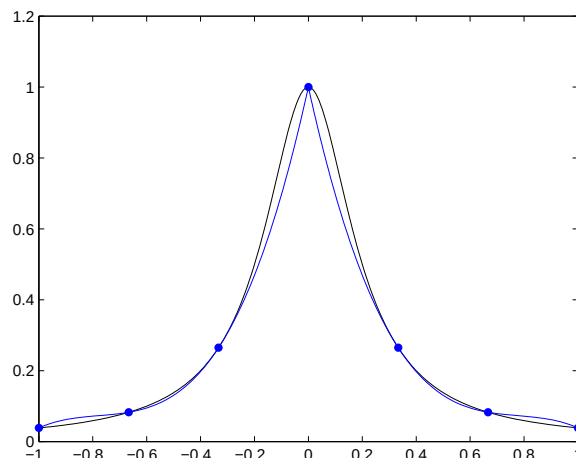


Si es desitja refinar més, en lloc de segments, podem, per exemple, connectar paràboles generades a partir de 3 punts consecutius o cúbiques a partir de 4 punts consecutius. El primer cas correspon a calcular tres polinomis interpoladors (paràboles): p_0 , associat als nodes x_0, x_1, x_2 ; p_1 , associat a x_2, x_3, x_4 ; i p_2 associat a x_4, x_5, x_6 . Quant al segon cas, es calculen dos polinomis interpoladors (cúbiques): p_0 , associat als nodes x_0, x_1, x_2, x_3 ; i p_1 , associat a x_3, x_4, x_5, x_6 .

Paràboles:



Cúbiques:



3.3.1 Splines cúbics

La funció resultant de la interpolació segmentària a tot el conjunt de nodes, tot i ser contínua, en general no és diferenciable als punts de connexió entre polinomis interpoladors. Existeix una alternativa més raonable que permet solucionar el problema anterior i que a més millora, en general, la precisió dels resultats. Aquesta ferramenta es coneix com **interpolació per splines**, encara que en aquest cas ens limitarem a treballar, a mode il·lustratiu, amb **splines cúbics**.

Siga $g(x)$ la funció polinòmica que volem obtenir per a interpolar als punts $\{(x_i, y_i)\}_{i=0}^n$, amb $y_i = f(x_i)$, $0 \leq i \leq n$ per a una certa funció f . Aquesta funció g és de la forma

$$g(x) = g_i(x), \quad x \in [x_i, x_{i+1}], \quad 0 \leq i < n,$$

amb $\{g_i\}_{i=0}^{n-1}$ n funcions polinòmiques complint

1. $g_i(x_i) = y_i$, $g_i(x_{i+1}) = y_{i+1}$, $0 \leq i < n$.
2. $g'_i(x_{i+1}) = g'_{i+1}(x_{i+1})$, $0 \leq i < n - 1$.
3. $g''_i(x_{i+1}) = g''_{i+1}(x_{i+1})$, $0 \leq i < n - 1$.
4. $g''_0(x_0) = 0$, $g''_{n-1}(x_n) = 0$.

La primera condició es correspon amb la d'interpolació en si mateixa (igualtat als nodes corresponents), mentre que la segona i la tercera es corresponen amb condicions de compatibilitat per tal de fer coincidir la primera i la segona derivada als punts de connexió. La darrera condició és tècnica i s'imposa perquè el nombre d'equacions coincidís amb el de les incògnites en el cas que g_i siguin polinomis de grau 3 (sistema $4n \times 4n$). Resolent, s'obté

$$g_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, \quad 0 \leq i < n,$$

on

$$\begin{cases} a_i = y_i \\ b_i = \frac{y_{i+1} - y_i}{h_i} - \frac{M_i h_i}{3} - \frac{M_{i+1} h_i}{6} \\ c_i = \frac{M_i}{2} \\ d_i = \frac{M_{i+1} - M_i}{6h_i} \end{cases}$$

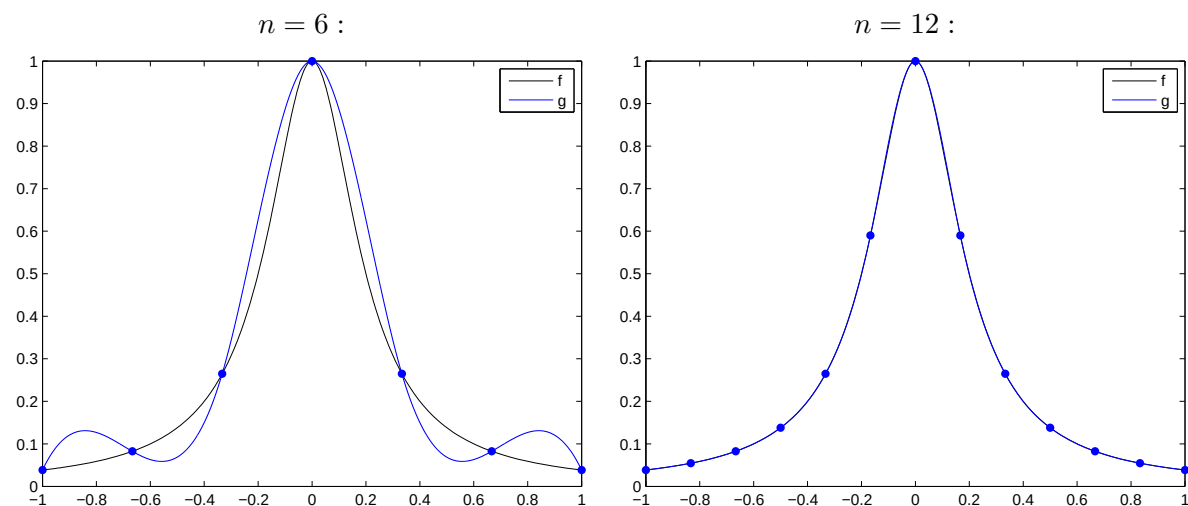
essent

$$\begin{aligned} h_i &= x_{i+1} - x_i, \quad 0 \leq i < n, \\ M_i &= g''_i(x_i), \quad 0 \leq i < n, \\ M_n &= g''_{n-1}(x_n). \end{aligned}$$

Per la condició de la quarta propietat, es té $M_0 = M_n = 0$ i, d'altra banda, la resta d' M_i , $1 \leq i \leq n - 1$, es determinen a partir del següent sistema d'equacions $(n - 1) \times (n - 1)$:

$$h_{i-1}M_{i-1} + 2(h_{i-1} + h_i)M_i + h_iM_{i+1} = 6 \left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \right), \quad 1 \leq i \leq n - 1.$$

Amb aquesta tècnica podem comprovar que si interpolem la funció anterior, el comportament oscil·latori de la interpolació minva segons augmenta n , com mostren les figures següents:



On es comprova la reducció del comportament oscil·latori en augmentar n . De fet, per a $n = 12$ la interpolació per splines cúbics ja dóna un resultat excel·lent.

Capítol 4

Integració numèrica

En aquest darrer capítol veurem algunes tècniques per a aproximar solucions de processos d'integració, com són el càlcul d'**integrals definides** i **equacions diferencials**. Començarem estudiant els mètodes per a aproximar numèricament integrals definides mitjançant les **fórmules de Newton-Cotes**.

4.1 Fórmules de Newton-Cotes

Recordem que la definició geomètrica d'integral definida d'una funció f sobre un interval $[a, b]$, $\int_a^b f(x)dx$, es correspon amb l'àrea que tanca la funció amb respecte de l'eix X al llarg de dit interval (tenint en compte que l'àrea pot ser negativa, corresponent a regions on f ho és).

Tenint en compte que l'àrea que tanca f a intervals petits, $[x, x+h]$ és aproximadament l'àrea del rectangle de base h (la longitud de l'interval) i altura $f(x)$ (el valor de la funció en x), tenim que si f és contínua en $[a, b]$, dita àrea es pot obtenir com a límit d'una suma infinita d'aproximacions d'aquesta per àrees de rectangles amb base infinitesimal:

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} h_n f(x_{i,n}),$$

on $x_{i,n} = a + ih_n$, $0 \leq i \leq n$, amb $h_n = \frac{b-a}{n}$.

Per a resoldre analíticament (és a dir, de manera exacta) una integral definida, disposem de ferramentes com ara la **regla de Barrow**:

$$\int_a^b f(x)dx = F(b) - F(a),$$

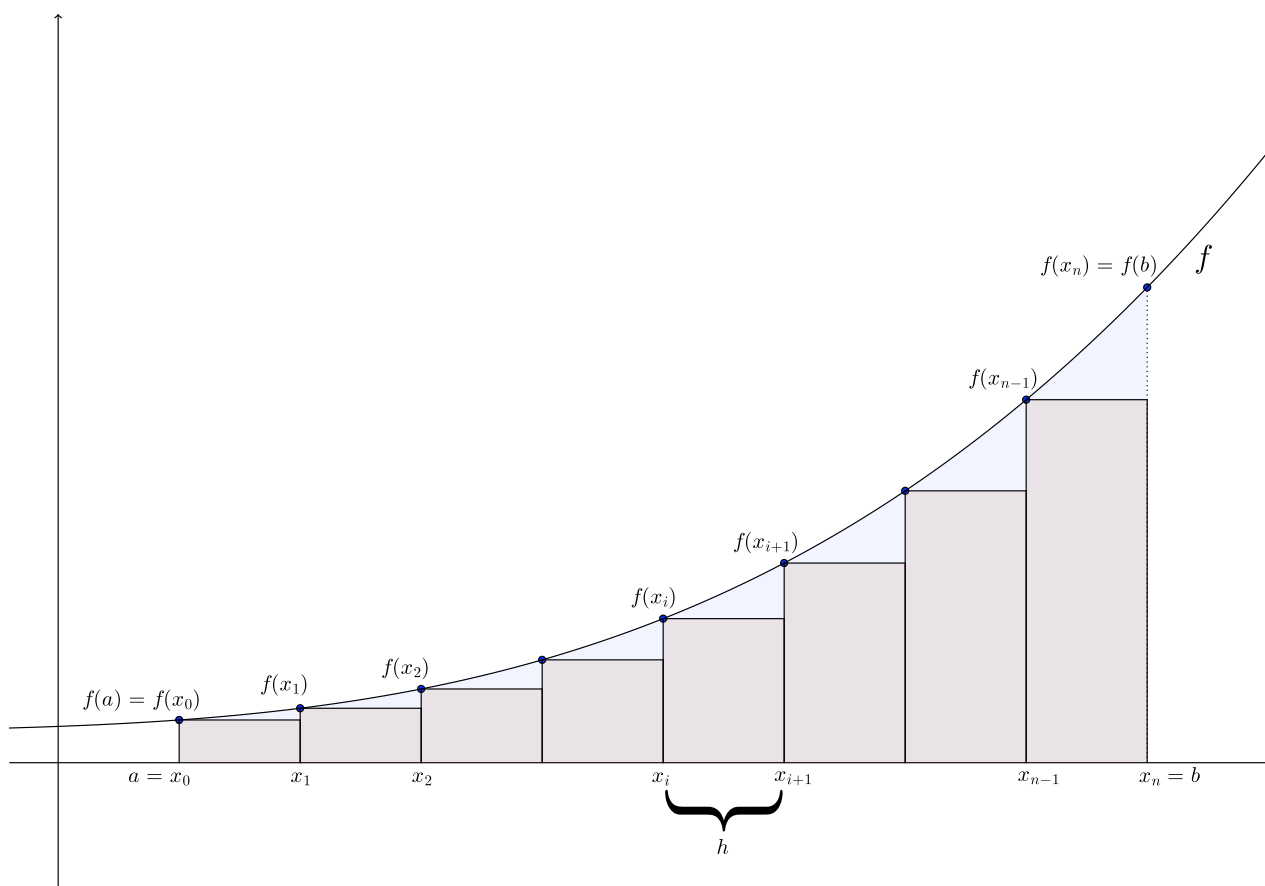
on F és una primitiva de f , $F'(x) = f(x)$.

No obstant això, aquest procés pot ser molt complicat o directament impossible depenent de la complexitat de la funció f . De fet, hi ha molts casos en què això ocorre. Per exemple, funcions ja vistes, com ara la funció de densitat de la distribució normal i la de les seues distribucions deduïdes, que no tenen primitiva elemental coneguda (i de fet està demostrat que la seua primitiva no admet una expressió analítica).

Açò motiva l'estudi de mètodes que, tot i manca la possibilitat de calcular el valor exacte d'una certa integral, proporcionen una manera d'aproximar-la tant com es vulga, fins assolir el nivell de precisió desitjat. La manera més simple de fer-ho és substituir el límit anterior per una expressió amb un valor de n fix, tot resultant una aproximació de la integral.

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n-1} h f(x_i),$$

on $x_i = a + ih$, $0 \leq i \leq n$, amb $h = \frac{b-a}{n}$.



Aquesta manera d'aproximar la integral es correspon específicament a la situació mostrada a la figura anterior, on s'aproxima per l'àrea dels rectangles de color roig el valor exacte de la integral, és a dir, l'àrea que deixa la funció f sobre l'interval $[a, b]$, en blau. Aquest mètode d'aproximació rep el nom **regla del rectangle** i és el mètode conceptualment més senzill que hi ha per a aproximar integrals, però lluny de ser eficient, ja que convergeix molt lentament a l'àrea exacta.

Veurem a continuació una manera de generar mètodes més precisos emprant ferramentes d'interpolació, ja vistes en el tema anterior. La idea consisteix en aproximar per trossos la funció per un polinomi interpolador; aleshores, com que la primitiva d'un polinomi sempre té expressió analítica i, de fet, es pot calcular amb facilitat, cal esperar aproximacions raonables de la integral definida sense que això implique un cost excessiu.

Així doncs, donada una franja entre uns nodes determinats, $x_i \leq x \leq x_{i+q}$ aproximarem f pel corresponent polinomi interpolador p_i als nodes $\{(x_j, y_j)\}_{j=i}^{i+q}$, on $y_j = f(x_j)$, $0 \leq j \leq q$. D'aquesta manera, com que $f(x) \approx p_i(x)$ per a $x \in [x_i, x_{i+q}]$, aleshores també es compleix

$$\int_{x_i}^{x_{i+q}} f(x) dx \approx \int_{x_i}^{x_{i+q}} p_i(x) dx.$$

Si apliquem aquest procediment a totes les seccions d'aquests tipus que componen l'interval $[a, b]$, llavors

$$\int_a^b f(x) dx \approx \sum_{i=0}^{\frac{n}{q}-1} \int_{x_{iq}}^{x_{(i+1)q}} p_{iq}(x) dx,$$

amb n múltiple de q .

La clau del procediment és l'obtenció d'una expressió coneguda per a cada secció $[x_i, x_{i+r}]$, és a dir, el valor de la integral en funció dels valors nodals $y_i = f(x_i)$. Per tal de simplificar la notació, denotarem per $[a, b]$ a una secció específica.

Prenem p_q el polinomi interpolador de grau q complint $p_q(x_j) = y_j$, on $x_j = a + jh$, essent $h = \frac{b-a}{q}$, $0 \leq j \leq q$. Aleshores podem escriure

$$f(x) = p_q(x) + R_q(x),$$

on $R_q = f - p_q$ és l'error d'interpolació. Prenent integrals, es té

$$\int_a^b f(x)dx = \int_a^b p_q(x)dx + \underbrace{\int_a^b R_q(x)dx}_{=R_q[f]},$$

on $R_q[f]$ rep el nom d'**error de quadratura** o **residu d'integració** per a la funció f .

Com que p_q es pot expressar segons la fórmula de Lagrange com

$$p_q(x) = \sum_{i=0}^q = \frac{\prod_{j=0}^q (x - x_j)}{\prod_{j=0}^q (x_i - x_j)} y_i,$$

llavors, l'aproximació de la integral es pot escriure com

$$\int_a^b p_q(x)dx = \sum_{i=0}^q A_i y_i, \quad (4.1)$$

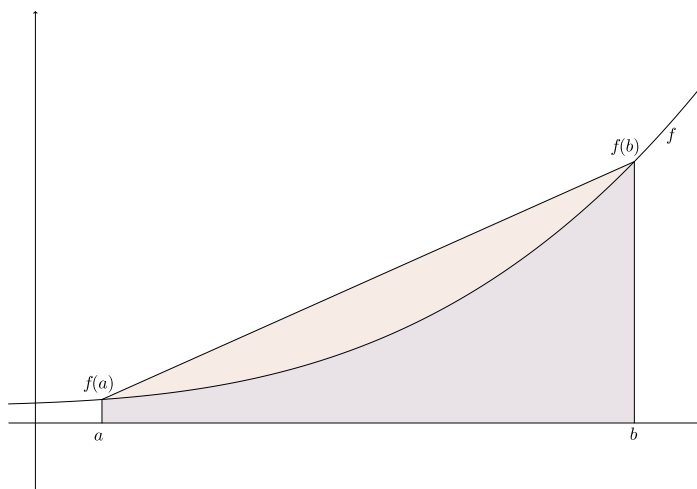
on

$$A_i = \frac{\int_a^b \prod_{j=0}^q (x - x_j) dx}{\prod_{j=0}^q (x_i - x_j)} y_i.$$

Per a cada q , és a dir, cada grau d'interpolació, els coeficients A_i són uns nombres específics, que determinen unívocament el mètode d'aproximació de la integral d'acord amb la **Fórmula general de Newton-Cotes (4.1)**. En el nostre cas, ens limitarem a estudiar els casos $q \in \{1, 2, 3\}$.

4.1.1 Fórmula dels trapezis

Correspon a fer $q = 1$ en la fórmula general de Newton-Cotes. En termes geomètrics, consisteix en fer interpolació lineal, o el que és el mateix, unir els punts de la funció corresponents als nodes amb un segment i calcular l'àrea del trapezi resultant, com mostra la següent figura.



En aquest cas resulta $A_0 = A_1 = \frac{b-a}{2}$, i per tant

$$\int_a^b f(x)dx \approx \int_a^b p_1(x)dx = \frac{b-a}{2}(f(a) + f(b)).$$

En aquest cas, si $f \in C^2(a, b)$, l'error es pot escriure com

$$R_1[f] = \frac{h^3}{12}f''(\xi), \quad \xi \in (a, b).$$

Fórmula dels trapezis composta

És el resultat d'aplicar la regla anterior a cada segment $[x_i, x_{i+1}]$, on $x_i = a + ih$, amb $h = \frac{b-a}{n}$, $0 \leq i \leq n$. En altres paraules, per tal d'aproximar la integral de f en $[a, b]$, dividim aquest en n subintervalls $[x_i, x_{i+1}]$, $0 \leq i < n$ i a cadascun d'aquests se li aplica una fórmula trapezoidal. Emprant el cas anterior, l'aproximació de la integral queda per tant de la següent manera:

$$\begin{aligned} \int_a^b f(x)dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx \approx \sum_{i=0}^{n-1} \frac{h}{2}(y_i + y_{i+1}) \\ &= \frac{h}{2}(y_0 + y_1) + \frac{h}{2}(y_1 + y_2) + \frac{h}{2}(y_2 + y_3) + \dots + \frac{h}{2}(y_{n-2} + y_{n-1}) + \frac{h}{2}(y_{n-1} + y_n) \\ &= \frac{h}{2}(y_0 + y_1 + y_1 + y_2 + y_2 + \dots + y_{n-1} + y_{n-1} + y_n) \\ &= \frac{h}{2}(y_0 + 2(y_1 + y_2 + \dots + y_{n-1}) + y_n). \end{aligned}$$

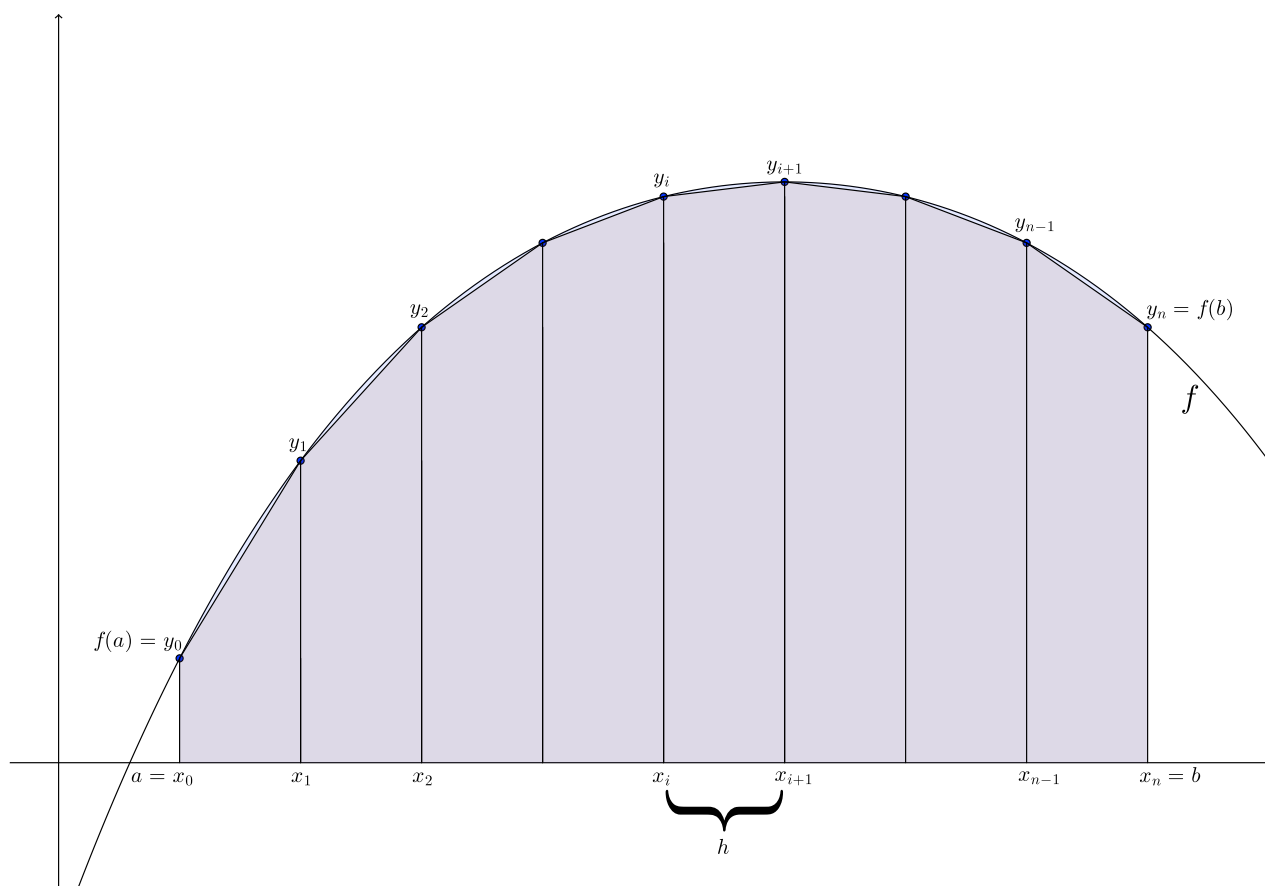
Si $f \in C^2(a, b)$, l'error per a la regla dels trapezis composta es pot escriure com

$$R_1 = \frac{nh^3}{12}f''(\xi) = \frac{(b-a)h^2}{12}f''(\xi), \quad \xi \in [a, b].$$

Una fita de l'error és per tant

$$|R_1| \leq \frac{nh^3}{12} \max_{x \in [a, b]} |f''(x)| = \frac{(b-a)h^2}{12} \max_{x \in [a, b]} |f''(x)|.$$

La següent figura il·lustra el procediment geomètric que representa la regla dels trapezis composta.



Exemple Siga $f : \mathbb{R} \rightarrow \mathbb{R}$ donada per $f(x) = \frac{2}{1+x^2}$. Es demana aproximar $\int_{-1}^1 f(x)dx$ mitjançant la regla del trapezi amb 7 punts ($n = 6$). Calcula una fita de l'error i compara amb l'error exacte.

Solució En aquest cas, tenim $a = -1$ i $b = 1$. Per tant, $h = \frac{b-a}{n} = \frac{1-(-1)}{6} = \frac{1}{3}$. Així doncs, $x_i = a + ih = -1 + \frac{i}{3}$, $0 \leq i \leq 6$. Aleshores, la taula de valors associada als nodes $\{(x_i, y_i)\}_{i=0}^6$ resulta:

x	-1	$-\frac{2}{3}$	$-\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$	1
$y = f(x)$	$f(-1) = 1$	$f(-\frac{2}{3}) = \frac{18}{13}$	$f(-\frac{1}{3}) = \frac{9}{5}$	$f(0) = 2$	$f(\frac{1}{3}) = \frac{9}{5}$	$f(\frac{2}{3}) = \frac{18}{13}$	$f(1) = 1$

Aplicant la fórmula dels trapezis, resulta per tant

$$\begin{aligned} \int_{-1}^1 f(x)dx &\approx \frac{h}{2}(y_0 + 2(y_1 + y_2 + y_3 + y_4 + y_5) + y_6) = \frac{\frac{1}{3}}{2} \left(1 + 2 \left(\frac{18}{13} + \frac{9}{5} + 1 + \frac{9}{5} + \frac{18}{13} \right) + 1 \right) \\ &= \frac{203}{65} \approx 3.1231. \end{aligned}$$

D'altra banda, la solució exacta de la integral es pot calcular com

$$\begin{aligned} \int_{-1}^1 f(x)dx &= \int_{-1}^1 \frac{2}{1+x^2} dx = [2 \arctan(x)]_{x=-1}^{x=1} = 2 \arctan(1) - 2 \arctan(-1) = \frac{\pi}{2} - \left(-\frac{\pi}{2} \right) = \pi \\ &\approx 3.1416. \end{aligned}$$

Notem que les regles de quadratura aplicades a aquest exemple particular ens permeten aproximar numèricament el valor de π .

L'error exacte comés és

$$|R_1| = \left| \pi - \frac{203}{65} \right| \approx 1.8516 \cdot 10^{-2}.$$

D'altra banda, per obtenir una fita de l'error hem de calcular f'' , l'expressió de la qual és

$$f''(x) = \frac{12x^2 - 4}{(1 + x^2)^3},$$

d'on pot comprovar-se que

$$\max_{x \in [-1, 1]} |f''(x)| = |f''(0)| = |-4| = 4.$$

Així doncs, la fita de l'error es calcula per

$$|R_1| \leq \frac{(b-a)h^2}{12} \max_{x \in [a, b]} |f''(x)| = \frac{(1 - (-1))^2 \left(\frac{1}{3}\right)^2}{12} \cdot 4 = \frac{4}{27} \approx 1.1482 \cdot 10^{-1},$$

que és coherent amb el resultat obtingut per a l'error exacte.

L'exemple anterior motiva algunes preguntes. Una d'elles, per exemple, és si és possible fitar fins la precisió desitjada el valor d'una integral definida, o el que és el mateix, saber quin és el nombre de punts mínim en què s'ha de dividir l'interval d'integració a partir del qual l'error és menor que una certa tolerància. Això ens pot permetre, entre moltes altres coses, calcular aproximacions tan bones com es desitge del número π , que ha aparegut a l'exemple anterior. Afortunadament, la resposta a aquesta qüestió és afirmativa.

En efecte, suposem que $\varepsilon > 0$ és l'error màxim o **tolerància** desitjada. Dit d'una altra manera, volem que $|R_1| < \varepsilon$. En particular, si imposem que la fita de l'error també verifiqui la condició anterior, aleshores l'error exacte també ho complirà. Per tant, imposem

$$\frac{(b-a)h^2}{12} M_2 < \varepsilon,$$

on $M_2 = \max_{x \in [a, b]} |f''(x)|$.

Com que $h = \frac{b-a}{n}$, aleshores, substituint en l'expressió anterior, es té

$$\frac{(b-a)^3}{12n^2} M_2 < \varepsilon.$$

Resolent la inequació per a n , s'obté

$$n > \sqrt{\frac{M_2(b-a)^3}{12\varepsilon}}.$$

Així doncs, el primer n que garanteix l'acompliment de la fita és el nombre natural que queda justament a la dreta del valor corresponent al membre de la dreta de la desigualtat anterior.

Exemple La integral definida $\int_0^{0.5} e^{x^2}$ no pot calcular-se de manera exacta, ja que la funció $f(x) = e^{x^2}$ no té primitiva elemental coneguda. Tot i així, podem aproximar-la tant com vulguem amb una precisió o tolerància predeterminada.

Per exemple, podem aproximar-la garantint un error menor que $3 \cdot 10^{-3}$ emprant la fórmula dels trapezis. Per a fer-ho, calculem primer

$$M_2 = \max_{x \in [0, 0.5]} |f''(x)| = \max_{x \in [0, 0.5]} |(4x^2 + 2)e^{x^2}| = |f''(0.5)| = 3\sqrt[4]{e}.$$

Aleshores, n ha d'acomplir

$$n > \sqrt{\frac{M_2(b-a)^3}{12\varepsilon}} = \sqrt{\frac{3\sqrt[4]{e}(0.5-0)^3}{12 \cdot 3 \cdot 10^{-3}}} \approx 3.6573.$$

Així doncs, és suficient prendre $n = 4$ (el primer natural que verifica ser superior a la quantitat obtinguda), és a dir, un total de 5 punts. Els nodes a triar, tenint en compte que $h = \frac{b-a}{n} = \frac{0.5-0}{4} = \frac{1}{8} = 0.125$, són per tant

$$x_i = a + ih = \frac{i}{8} = 0.125i, \quad 0 \leq i \leq 4.$$

La taula de valors resulta en aquest cas

x	$y = f(x)$
0	$f(0) = e^0 = 1$
0.125	$f(0.125) = e^{0.125^2} = e^{0.015625}$
0.25	$f(0.25) = e^{0.25^2} = e^{0.0625}$
0.375	$f(0.375) = e^{0.375^2} = e^{0.140625}$
0.5	$f(0.5) = e^{0.5^2} = e^{0.25}$

I aplicant la fórmula dels trapezis es té

$$\int_0^{0.5} e^{x^2} \approx \frac{h}{2}(y_0 + 2(y_1 + y_2 + y_3) + y_4) = 0.0625(1 + 2(e^{0.015625} + e^{0.0625} + e^{0.140625}) + e^{0.25}) \approx 0.5467.$$

4.1.2 Fórmula de Simpson 1/3

Correspon a fer $q = 2$ en la fórmula de Newton-Cotes. En aquest cas, s'obté $A_0 = \frac{b-a}{3}$, $A_1 = \frac{4(b-a)}{3}$, $A_2 = \frac{b-a}{3}$. L'expressió queda per tant

$$\int_a^b f(x)dx \approx \frac{b-a}{3} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right).$$

Si $f \in C^4(a, b)$, l'expressió de l'error exacte es pot escriure com

$$R_2[f] = \frac{(b-a)^5}{90} f^{IV}(\xi), \quad \xi \in (a, b).$$

Fórmula de Simpson 1/3 composta

Dividint l'interval $[a, b]$ en $n = 2m$ (múltiple de 2) subintervalls iguals i aplicant la regla anterior a cadascun d'ells, $[x_i, x_{i+1}]$, on $x_i = a + ih$, amb $h = \frac{b-a}{n}$, $0 \leq i \leq n$, obtenim

$$\begin{aligned}\int_a^b f(x)dx &= \sum_{i=0}^{m-1} \int_{x_i}^{x_{i+1}} f(x)dx \approx \sum_{i=0}^{m-1} \frac{h}{3}(y_{2i} + 4y_{2i+1} + y_{2i+2}) \\ &= \frac{h}{3}(y_0 + 4y_1 + y_2) + \frac{h}{3}(y_2 + 4y_3 + y_4) + \dots + \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n) \\ &= \frac{h}{3}(y_0 + 4(y_1 + y_3 + \dots + y_{n-3} + y_{n-1}) + 2(y_2 + y_4 + \dots + y_{n-4} + y_{n-2}) + y_n).\end{aligned}$$

Si $f \in C^4(a, b)$, l'error de la regla de Simpson 1/3 es pot expressar com

$$R_2 = \frac{mh^5}{90} f^{IV}(\xi) = \frac{(b-a)h^4}{180} f^{IV}(\xi), \quad \xi \in (a, b),$$

que es pot fitar com

$$|R_2| \leq \frac{(b-a)h^4}{180} \max_{x \in [a, b]} |f^{IV}(x)|.$$

Exemple Siga $f : \mathbb{R} \rightarrow \mathbb{R}$ donada per $f(x) = \frac{2}{1+x^2}$. Es demana aproximar $\int_{-1}^1 f(x)dx$ mitjançant la regla de Simpson 1/3 amb 7 punts ($n = 6$), calcular una fita de l'error i comparar amb l'error exacte.

Solució Aprofitant les dades ja calculades per a aquesta funció per al cas dels trapezis, tenim, emprant ara la fórmula de Simpson 1/3, i tenint en compte que $h = \frac{b-a}{n} = \frac{1-(-1)}{6} = \frac{1}{3}$,

$$\begin{aligned}\int_a^b f(x)dx &\approx \frac{h}{3}(y_0 + 4(y_1 + y_3 + y_5) + 2(y_2 + y_4) + y_6) \\ &= \frac{1}{3} \left(1 + 4 \left(\frac{18}{13} + 2 + \frac{18}{13} \right) + 2 \left(\frac{9}{5} + \frac{9}{5} \right) + 1 \right) = \frac{1838}{585} \approx 3.14188.\end{aligned}$$

L'error exacte resulta per tant

$$|R_2| = \left| \pi - \frac{1838}{585} \right| \approx 2.8769 \cdot 10^{-4}.$$

De l'error obtingut es pot comprovar que l'aproximació de π és considerablement millor que l'obtinguda amb la regla dels trapezis per al mateix nombre de punts.

D'altra banda, tenint en compte que

$$f^{IV}(x) = \frac{240x^4 - 480x^2 + 48}{(1+x^2)^3},$$

amb

$$\max_{x \in [-1, 1]} |f^{IV}(x)| = |f^{IV}(0)| = |48| = 48,$$

la fita de l'error s'obté per

$$|R_2| \leq \frac{(b-a)h^4}{180} \max_{x \in [a, b]} |f^{IV}(x)| = \frac{(1-(-1))(\frac{1}{3})^4}{180} \cdot 48 = \frac{8}{1215} \approx 6.5844 \cdot 10^{-3},$$

que efectivament és un valor superior a l'error exacte obtingut.

De la mateixa manera que en el cas dels trapezis, podem plantejar-nos un problema consistent en l'obtenció d'una aproximació d'una certa integral definida assegurant que l'error és menor que un cert valor ε , ço és, $|R_2| < \varepsilon$. Per a fer-ho, és suficient garantir que la fita de l'error siga menor que ε ; en aquest cas,

$$\frac{(b-a)h^4}{180}M_4 < \varepsilon,$$

on $M_4 = \max_{x \in [a,b]} |f^{IV}(x)|$.

Tenint en compte que $h = \frac{b-a}{n}$, aïllant n a la inequació anterior s'obté la condició

$$n > \sqrt[4]{\frac{M_4(b-a)^5}{180\varepsilon}}.$$

Com que en aquest cas n ha de ser un múltiple de 2, el primer valor de n que verifica la condició desitjada és el valor corresponent a arrodonir al primer nombre parell que queda per sobre de la quantitat obtinguda al segon membre de la desigualtat.

Exemple Calculem una aproximació de $\int_0^{0.5} e^{x^2} dx$ mitjançant la regla de Simpson 1/3 garantint un error menor que 10^{-4} .

Donat que $f(x) = e^{x^2}$, es té

$$f^{IV}(x) = (16x^4 + 48x^2 + 12)e^{x^2},$$

i per tant

$$M_4 = \max_{x \in [0,0.5]} |f^{IV}(x)| = |f^{IV}(0.5)| = 25\sqrt[4]{e}.$$

Aleshores, n ha de verificar

$$n > \sqrt[4]{\frac{M_4(b-a)^5}{180\varepsilon}} = \sqrt[4]{\frac{25\sqrt[4]{e}(0.5-0)^5}{180 \cdot 10^{-5}}} \approx 2.7322.$$

Per tant, el nombre parell més proper que queda per sobre d'eixa quantitat és $n = 4$.

Aplicant la regla de Simpson 1/3, aprofitant les dades calculades per a l'exemple corresponent per a la regla dels trapezis, s'obté

$$\int_0^{0.5} e^{x^2} dx \approx \frac{h}{3}(y_0 + 4(y_1 + y_3) + 2y_2 + y_4) = \frac{1}{3}(1 + 4(e^{0.015625} + e^{0.140625}) + 2e^{0.0625} + e^{0.25}) \approx 0.544999.$$

Exercici 24 Repeteix l'exemple anterior garantint ara un error menor que $2.5 \cdot 10^{-6}$.

4.1.3 Fórmula de Simpson 3/8

Es correspon amb el cas $q = 3$ en la fórmula de Newton-Cotes, on $A_0 = A_3 = \frac{3(b-a)}{8}$, $A_1 = A_2 = \frac{9(b-a)}{8}$. L'expressió queda per tant com se segueix

$$\int_a^b f(x) dx \approx \frac{3(b-a)}{8} \left(f(a) + 3f\left(\frac{a+b}{3}\right) + 3f\left(\frac{2(a+b)}{3}\right) + f(b) \right).$$

Si $f \in C^4(a, b)$, l'expressió de l'error exacte es pot escriure com

$$R_2[f] = \frac{3(b-a)^5}{80} f^{IV}(\xi), \quad \xi \in (a, b).$$

Fórmula de Simpson 3/8 composta

Dividint l'interval $[a, b]$ en $n = 3m$ (múltiple de 3) subintervalls iguals i aplicant la regla anterior a cadascun d'ells, $[x_i, x_{i+1}]$, on $x_i = a + ih$, amb $h = \frac{b-a}{n}$, $0 \leq i \leq n$, s'obté

$$\begin{aligned} \int_a^b f(x)dx &= \sum_{i=0}^{m-1} \int_{x_i}^{x_{i+1}} f(x)dx \approx \sum_{i=0}^{m-1} \frac{h}{3} (y_{3i} + 3y_{3i+1} + 3y_{3i+2} + y_{3i+3}) \\ &= \frac{h}{3} (y_0 + 3y_1 + 3y_2 + y_3) + \frac{h}{3} (y_3 + 3y_4 + 3y_5 + y_6) + \dots + \frac{h}{3} (y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n) \\ &= \frac{h}{3} (y_0 + 3(y_1 + y_2 + y_4 + y_5 + \dots + y_{n-2} + y_{n-1}) + 2(y_3 + y_6 + \dots + y_{n-3}) + y_n). \end{aligned}$$

Si $f \in C^4(a, b)$, l'error de la regla de Simpson 3/8 es pot expressar com

$$R_2 = \frac{3mh^5}{80} f^{IV}(\xi) = \frac{(b-a)h^4}{80} f^{IV}(\xi), \quad \xi \in (a, b),$$

que es pot fitar com

$$|R_2| \leq \frac{(b-a)h^4}{80} \max_{x \in [a, b]} |f^{IV}(x)|.$$

Exercici 25 Siga $f : \mathbb{R} \rightarrow \mathbb{R}$ donada per $f(x) = \frac{2}{1+x^2}$. Aproxima $\int_{-1}^1 f(x)dx$ mitjançant la regla de Simpson 3/8 amb 7 punts ($n = 6$), calcula una fita de l'error i compara amb l'error exacte. Compara amb els resultats obtinguts amb les tècniques anteriors.

D'igual manera com s'ha fet en els dos casos anteriors, donada una tolerància $\varepsilon > 0$, podem trobar la condició que ha de verificar n per tal que l'error de la aproximació siga menor que aquesta. Calcant els càlculs realitzats per a la regla de Simpson 1/3, la condició resulta

$$n > \sqrt[4]{\frac{M_4(b-a)^5}{80\varepsilon}},$$

on $M_4 = \max_{x \in [a, b]} |f^{IV}(x)|$.

Exercici 26 Aproxima $\int_0^{0.5} e^{x^2} dx$ mitjançant la regla de Simpson 3/8 garantint un error menor que $1.5 \cdot 10^{-6}$. Garanteix l'aproximació obtinguda un error menor que $6.5 \cdot 10^{-7}$? I de $6 \cdot 10^{-7}$? Per què?

Exercici 27 Aproxima el valor de $\log(2)$:

- (a) Amb un error menor que 10^{-2} emprant la fórmula dels trapezis, Simpson 1/3 i Simpson 3/8.
- (b) Amb un error menor que 10^{-4} emprant les fórmules de Simpson.

En cada cas, compara amb l'error exacte i comprova que es verifiquen les fites.

Indicació: Escribeu $\log(2)$ com la integral definida d'una certa funció $f(x)$.

4.2 Integració numèrica d'equacions diferencials

Una **equació diferencial** és una relació entre una funció y amb una o diverses de les seues derivades. Per exemple, la funció $y(x) = e^x$ compleix $y'(x) = e^x = y(x)$, i per tant es compleix la relació

$$y' = y,$$

o el que és el mateix,

$$y - y' = 0.$$

Un altre exemple és emprant $y(x) = e^{x^2}$, que compleix $y'(x) = 2xe^{x^2} = 2xy(x)$, i per tant en aquest cas es verifica la relació

$$y' = 2xy,$$

que es pot escriure com

$$y' - 2xy = 0.$$

Un darrer exemple on apareixen derivades d'ordre superior a 1 és $y(x) = \sin(x)$, complint-se $y''(x) = -\sin(x) = -y(x)$, i per tant verificant-se l'equació

$$y'' = -y \longrightarrow y'' + y = 0.$$

En general, una equació diferencial pot escriure's de la següent manera:

$$F(x, y, y', y'', \dots, y^{(n)}) = 0.$$

L'ordre de la major derivada que apareix en l'equació determina l'**ordre** de l'equació diferencial i la màxima potència el seu **grau**.

En el nostre cas, ens limitarem a treballar amb equacions diferencial d'ordre 1 (és a dir, que només involucren derivades de primer ordre) de la forma

$$y' = f(x, y)$$

amb una condició inicial $y(x_0) = y_0$. f rep el nom de **funció pendent**.

Per exemple, l'equació

$$y' - xy^2 + \cos(x) = 0$$

es pot escriure com

$$y' = xy^2 - \cos(x)$$

i per tant pot expressar-se de la forma anterior com $y' = f(x, y)$, on $f(x, y) = xy^2 - \cos(x)$.

El procediment de resolució d'una equació diferencial consisteix en l'obtenció d'una funció y complint la igualtat especificada per l'expressió de l'equació al mateix temps que la condició inicial. No obstant això, aquest procediment és en general complicat i els casos en què es pot calcular la solució analíticament són molt reduïts. Aquest fet motiva el disseny d'estratègies per a aproximar les solucions de manera numèrica.

En general, un mètode numèric d'aquest tipus consisteix en calcular una aproximació de $y(x_n)$ per a un cert nombre de passos, n , des del pas 0, corresponent a passar de la condició inicial $y(x_0) = y_0$, on es coneix el valor exacte de y , a una aproximació de $y(x_1)$, $y_1 \approx y(x_1)$, fins al pas $n - 1$, on es passa d'una aproximació de $y(x_{n-1})$, $y_{n-1} \approx y(x_{n-1})$ a l'aproximació desitjada, $y_n \approx y(x_n)$. Així doncs, a cada pas s'avancen $h = \frac{x_n - x_0}{n}$ unitats a les abscisses, tot obtenint-se en general al pas i una aproximació de $y(x_i)$, on $x_i = x_0 + ih$, $0 \leq i \leq n$.

Dins dels possibles mètodes numèrics associats a aquest tipus d'equacions, començarem estudiant el més senzill, que és el **mètode d'Euler**.

4.2.1 Mètode d'Euler

El mètode d'Euler es basa en l'aproximació de la funció pendent mitjançant diferències dividides. Suposem que estem en el pas i , on hem obtingut a partir de passos previs una aproximació de $y(x_i)$, $y_i \approx y(x_i)$. Llavors:

$$f(x_i, y_i) \approx f(x_i, y(x_i)) = y'(x_i) \approx \frac{y(x_{i+1}) - y(x_i)}{h} \approx \frac{y(x_{i+1}) - y_i}{h}.$$

Per tant, aïllant s'obté

$$y(x_{i+1}) \approx y_i + hf(x_i, y_i).$$

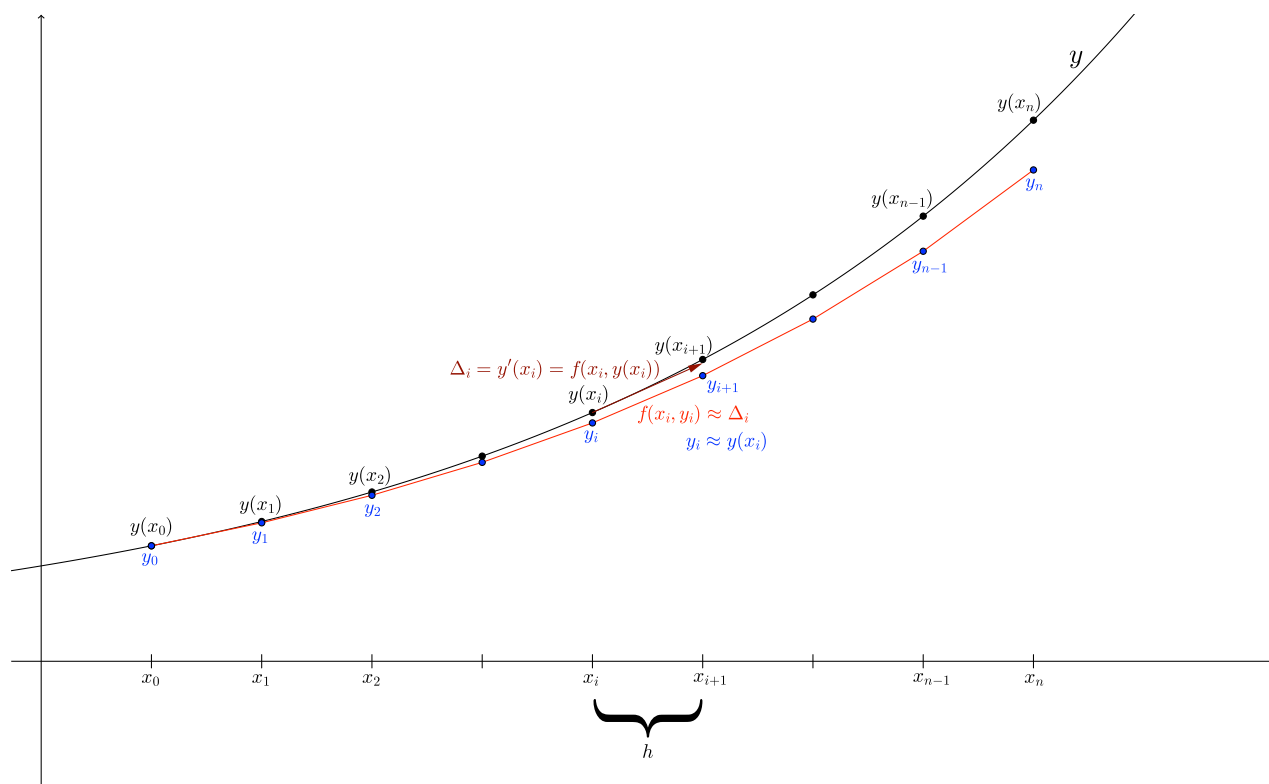
Per tant, si definim

$$y_{i+1} = y_i + hf(x_i, y_i),$$

llavors es té que $y_{i+1} \approx y(x_{i+1})$.

Així doncs, el mètode d'Euler es pot escriure com es segueix

$$\begin{aligned} y_0 &= y(x_0), \\ y_{i+1} &= y_i + hf(x_i, y_i), \quad 0 \leq i < n. \end{aligned}$$



D'altra banda, l'expressió de l'error corresponent a aplicar el pas i suposant que partim del valor exacte, és a dir, $y_i = y(x_i)$, també conegut com **error local de truncament**, té la forma

$$R = \frac{h^2}{2} y''(\xi), \quad \xi \in (x_i, x_{i+1}).$$

Exemple Considerem l'equació

$$y' - y + x - 3 = 0$$

amb condició inicial $y(1) = 0$. Es demana obtenir una aproximació de $y(2)$ mitjançant el mètode d'Euler amb 4 passos.

Solució Aillem y' de l'equació:

$$y' = y - x + 3,$$

per tant, la funció pendent en aquest cas és

$$f(x, y) = y - x + 3.$$

D'altra banda, hem de fer $n = 4$ passos, amb $x_0 = 1$ (punt on es coneix la condició inicial) i $x_n = x_4 = 2$ (punt on volem obtenir l'aproximació de la solució). Així doncs,

$$h = \frac{x_n - x_0}{n} = \frac{x_4 - x_0}{4} = \frac{2 - 1}{4} = \frac{1}{4} = 0.25.$$

Per tant, $x_i = x_0 + ih = 1 + 0.25i$, obtenint-se per tant els nodes

$$\{x_0 = 1, x_1 = 1.25, x_2 = 1.5, x_3 = 1.75, x_4 = 2\}.$$

Emprant la condició inicial, tenim que

$$y_0 = y(x_0) = y(1) = 0.$$

Aplicant ara els 4 passos del mètode d'Euler demanats, tenim, fent $i = 0$ en la fórmula del mètode:

$$y_1 = y_0 + hf(x_0, y_0) = y_0 + h(y_0 - x_0 + 3) = 0 + 0.25 \cdot (0 - 1 + 3) = 0.5 \approx y(x_1) = y(1.25).$$

Continuant amb el pas $i = 1$:

$$y_2 = y_1 + hf(x_1, y_1) = y_1 + h(y_1 - x_1 + 3) = 0.5 + 0.25 \cdot (0.5 - 1.25 + 3) = 1.0625 \approx y(x_2) = y(1.5).$$

Per a $i = 2$, s'obté:

$$y_3 = y_2 + hf(x_2, y_2) = y_2 + h(y_2 - x_2 + 3) = 1.0625 + 0.25 \cdot (1.0625 - 1.5 + 3) = 1.703125 \approx y(x_3) = y(1.75).$$

Finalment, per al darrer pas, corresponent a $i = 3$, es té:

$$y_4 = y_3 + hf(x_3, y_3) = y_3 + h(y_3 - x_3 + 3) = 1.703125 + 0.25 \cdot (1.703125 - 1.75 + 3) = 2.44140625.$$

Per tant, l'aproximació de $y_4 \approx y(x_4) = y(2)$ és el darrer valor calculat, és a dir, $y_4 = 2.44140625$.

En aquest problema concret, podem calcular la solució exacta, que és

$$y(x) = e^{x-1} + x - 2.$$

Per tant, el valor exacte de $y(2)$ és

$$y(2) = e^{2-1} + 2 - 2 = e \approx 2.71828.$$

L'error exacte resulta per tant

$$|y(x_4) - y_4| = |y(2) - y_4| = |e - 2.44140625| \approx 0.2769.$$

Exercici 28 Repeteix l'exemple anterior fent 8 passos.

4.2.2 Mètode d'Euler millorat

El mètode d'Euler és el més simple i per tant el més senzill d'aplicar; no obstant això, és massa imprecís, com es pot comprovar a l'exemple anterior.

Tot seguit presentarem una versió millorada d'aquest mètode, anomenat **mètode d'Euler millorat** o **mètode de Runge-Kutta de segon ordre** que consisteix en refinar l'aproximació de la funció pendent en cada pas. En aquest cas, l'avançament temporal s'efectua aproximant la pendent per la mitja aritmètica de dues pendents calculades en punts diferents. Concretament, aquest mètode es formula com

$$y_0 = y(x_0),$$

$$y_{i+1} = y_i + \frac{h}{2}(K_{i,1} + K_{i,2}), \quad 0 \leq i < n,$$

on

$$K_{i,1} = f(x_i, y_i), \quad 0 \leq i < n,$$

$$K_{i,2} = f(x_i + h, y_i + hK_{i,1}) = f(x_{i+1}, y_i + hK_{i,1}), \quad 0 \leq i < n.$$

Exemple Repetim l'exemple anterior aplicant el mètode d'Euler millorat amb 2 passos. Per tant, en aquest cas es té que $x_n = x_2 = 2$,

$$h = \frac{x_n - x_0}{n} = \frac{x_2 - x_0}{2} = \frac{2 - 1}{2} = 0.5$$

i els nodes resultants són $x_i = x_0 + ih$, $0 \leq i \leq 2$:

$$\{x_0 = 1, x_1 = 1.5, x_2 = 2\}.$$

D'una banda, tenim

$$y_0 = y(x_0) = y(1) = 0.$$

Per a efectuar el pas corresponent a $i = 0$, calculem primer les pendents K :

$$K_{0,1} = f(x_0, y_0) = f(1, 0) = 0 - 1 + 3 = 2,$$

$$K_{0,2} = f(x_1, y_0 + hK_{0,1}) = f(1.5, 0 + 0.5 \cdot 2) = f(1.5, 1) = 1 - 1.5 + 3 = 2.5.$$

Per tant,

$$y_1 = y_0 + \frac{h}{2}(K_{0,1} + K_{0,2}) = 0 + 0.25 \cdot (2 + 2.5) = 1.125 \approx y(x_1) = y(1.5).$$

Finalment, per al darrer pas associat a $i = 1$ s'obtenen les pendents

$$K_{1,1} = f(x_1, y_1) = f(1.5, 1.125) = 1.125 - 1.5 + 3 = 2.625,$$

$$K_{1,2} = f(x_2, y_1 + hK_{1,1}) = f(2, 1.125 + 0.5 \cdot 2.625) = f(2, 2.4375) = 2.4375 - 2 + 3 = 3.4375.$$

Per tant,

$$y_2 = y_1 + \frac{h}{2}(K_{1,1} + K_{1,2}) = 1.125 + 0.25 \cdot (2.625 + 3.4375) = 2.640625.$$

L'error en aquest cas resulta per tant

$$|e - 2.640625| \approx 0.0777,$$

un resultat millorat que en el cas el mètode d'Euler simple, fins i tot havent fet en aquest cas la meitat de passos.

Exercici 29 Repeteix l'exemple anterior fent 4 passos i compara l'error obtingut amb el de l'exercici anterior (mètode d'Euler simple amb 8 passos).

4.2.3 Mètode de Runge-Kutta de quart ordre

Per acabar, es presenta a continuació un mètode encara més acurat que els anteriors, consistent en fer una mitjana ponderada de quatre pendents obtingudes a cada pas, en lloc d'una mitjana aritmètica de 2 pendents com en el cas del mètode d'Euler millorat.

Aquest mètode s'escriu com es segueix

$$y_0 = y(x_0),$$

$$y_{i+1} = y_i + \frac{h}{6}(K_{i,1} + 2K_{i,2} + 2K_{i,3} + K_{i,4}), \quad 0 \leq i < n,$$

on

$$K_{i,1} = f(x_i, y_i), \quad 0 \leq i < n,$$

$$K_{i,2} = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}K_{i,1}\right), \quad 0 \leq i < n,$$

$$K_{i,3} = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}K_{i,2}\right), \quad 0 \leq i < n,$$

$$K_{i,4} = f(x_i + h, y_i + hK_{i,3}) = f(x_{i+1}, y_i + hK_{i,3}), \quad 0 \leq i < n.$$

Exemple Apliquem el mètode de Runge-Kutta a l'equació dels exemples anteriors amb un únic pas. Per tant, en aquest cas tenim $n = 1$, amb $x_n = x_1 = 2$ i

$$h = \frac{x_n - x_0}{n} = \frac{x_1 - x_0}{1} = \frac{2 - 1}{1} = 1,$$

amb els nodes

$$\{x_0 = 1, x_1 = 2\}.$$

Aplicant la condició inicial:

$$y_0 = y(x_0) = y(1) = 0.$$

Calculem en primer lloc les pendents per al primer (i últim) pas $i = 0$:

$$K_{0,1} = f(x_0, y_0) = f(1, 0) = 0 - 1 + 3 = 2,$$

$$K_{0,2} = f\left(x_0 + \frac{h}{2}, y_0 + \frac{h}{2}K_{0,1}\right) = f\left(1 + \frac{1}{2}, 0 + \frac{1}{2} \cdot 2\right) = f(1.5, 1) = 1 - 1.5 + 3 = 2.5,$$

$$K_{0,3} = f\left(x_0 + \frac{h}{2}, y_0 + \frac{h}{2}K_{0,2}\right) = f\left(1 + \frac{1}{2}, 0 + \frac{1}{2} \cdot 2.5\right) = f(1.5, 1.25) = 1.25 - 1.5 + 3 = 2.75,$$

$$K_{0,4} = f(x_1, y_0 + hK_{0,3}) = f(2, 0 + 1 \cdot 2.75) = f(2, 2.75) = 2.75 - 2 + 3 = 3.75.$$

S'obté per tant

$$y_1 = y_0 + \frac{h}{6}(K_{0,1} + 2K_{0,2} + 2K_{0,3} + K_{0,4}) = 0 + \frac{1}{6}(2 + 2 \cdot 2.5 + 2 \cdot 2.75 + 3.75) = \frac{65}{24} = 2.708\bar{3} \approx y(x_1) = y(2).$$

L'error és per tant

$$\left|e - \frac{65}{24}\right| \approx 0.00995,$$

que és menor que l'obtingut en els casos anteriors, havent efectuat únicament un pas.

Exercici 30 Repeteix l'exemple anterior amb 2 passos i compara els resultats obtinguts amb els dels exercicis anteriors.