Capítulo 5 - Distribuições conjuntas de probabilidade e complementos

Conceição Amado, Ana M. Pires e M. Rosário Oliveira

Em muitas situações, está-se interessado em estudar simultâneamente mais de uma característica numa experiência aleatória. Suponha-se que a experiência é seleccionar aleatoriamente alunos de um certo curso, e o interesse é estudar o perfil "biológico" desses alunos. Pode-se, então considerar que o perfil é composto de:

- peso
- altura
- pressão arterial
- frequência cardíaca
- capacidade respiratória

Ou seja, está-se interessado em cinco variáveis aleatórias que devem ser estudadas simultaneamente. Isto motiva a seguinte definição de um vector aleatório.

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Definição: Considere-se uma experiência aleatória e o seu espaço de resultados Ω. Diz-se que (X, Y) é um vector aleatório, par aleatório ou variável aleatória bidimensional se X e Y forem variáveis aleatórias.

(X, Y) é um vector aleatório:

— **discreto** se X e Y forem variáveis aleatórias discretas;

— **contínuo** se X e Y forem variáveis aleatórias contínuas.

5.1 + 5.2 - Duas v.a. discretas ou contínua

Dadas duas ou mais v.a. o seu comportamento simultâneo é estudado usando as chamadas **distribuições conjuntas**.

Definição: Dadas duas variáveis aleatórias discretas, X e Y, chama-se **função de (massa de) probabilidade conjunta** à função

$$f_{X,Y}(x,y) = P(X = x, Y = y), \quad \forall_{(x,y) \in \mathbb{R}^2},$$

que verifica

i) $f_{X,Y}(x,y) \geq 0, \quad \forall_{(x,y) \in \mathbb{R}^2}$

ii) $\sum_x \sum_y f_{X,Y}(x,y) = 1$

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Esta função em tabela de dupla entrada:

$X \backslash Y$	y_1	y_2	$\cdots$	y_s	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1s}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2s}	$\cdots$
$\ldots$			$\cdots$		$\cdots$
x_r	p_{r1}	p_{r2}	$\cdots$	p_{rs}	$\cdots$
$\ldots$			$\cdots$		$\cdots$

$$p_{ij} = P(X = x_i, Y = y_j)$$

- $p_{ij} \geq 0$
- $\sum_i \sum_j p_{ij} = 1$

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Exemplo 5.1: Considere o lançamento de dois dados perfeitos. Seja

X - v.a. que indica o n.ºde vezes que saiu a face 5

Y - v.a. que indica o n.ºde vezes que saiu a face 6

Valores possíveis: $R_X = \{0, 1, 2\}$ e $R_Y = \{0, 1, 2\}$

$X \sim Bin(2, 1/6)$ e que $Y \sim Bin(2, 1/6)$.

Nota: Isto quer dizer que X e Y têm o mesmo comportamento em termos de valores possíveis e respectivas probabilidades.

Não quer dizer $X = Y$! O que se pode dizer é

X e Y são identicamente distribuídas

Qual é o **comportamento conjunto das duas variáveis**, em termos de probabilidades dos pares de valores (x, y), com $x = 0, 1, 2$ e $y = 0, 1, 2$?

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Verificar que

$$\sum_{x=0}^{2}\sum_{y=0}^{2} P(X = x, Y = y) = 1$$

$$P(X = 0, Y = 0) = \tfrac{4}{6} \times \tfrac{4}{6}$$

$$P(X = 0, Y = 1) = \tfrac{1}{6} \times \tfrac{4}{6} \times 2$$

$$P(X = 0, Y = 2) = \tfrac{1}{6} \times \tfrac{1}{6}$$

$$P(X = 1, Y = 0) = P(X = 0, Y = 1)$$

$$P(X = 1, Y = 1) =) = \tfrac{1}{6} \times \tfrac{1}{6} \times 2$$

$$P(X = 1, Y = 2) = 0$$

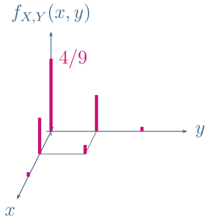
$$P(X = 2, Y = 0) = P(X = 0, Y = 2)$$

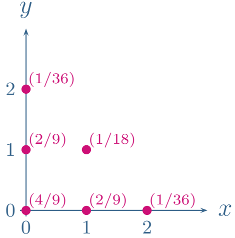
$$P(X = 2, Y = 1) = 0$$

$$P(X = 2, Y = 2) = 0$$

5.1 + 5.2 - Duas v.a. discretas ou contínuas

$X \backslash Y$	0	1	2
0	$\frac{16}{36}$	$\frac{8}{36}$	$\frac{1}{36}$
1	$\frac{8}{36}$	$\frac{2}{36}$	0
2	$\frac{1}{36}$	0	0





5.1 + 5.2 - Duas v.a. discretas ou contínuas

Definição: Seja (X, Y) um v.a. contínuo. Se existir uma função $f_{X,Y}(x,y)$ tal que:

$$F_{(X,Y)}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u,v)\, dvdu, \quad \forall_{(x,y)\in\mathbb{R}^2}$$

então ela diz-se a função de densidade de probabilidade conjunta do v.a. contínuo (X, Y) e satisfaz:

1) $f_{X,Y}(x,y) \geq 0, \quad \forall_{(x,y)\in\mathbb{R}^2}$

2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dxdy = 1$

Nota: Qualquer que seja a região $R \in \mathbb{R}^2$

$$P((X,Y) \in R) = \int_R \int f_{X,Y}(x,y)dxdy.$$

Exemplo 5.2: Seja a f.d.p conjunta do par aleatório (X,Y),

$$f_{(X,Y)}(x,y) = \begin{cases} \frac{1}{2}, & 0 < x < 1, 0 < y < 2 \\ 0, & c.c. \end{cases}$$

É fácil de verificar que é, de facto, uma f.d.p, pois:

(i) $f_{(X,Y)}(x,y) \geq 0, \quad \forall_{(x,y)\in\mathbb{R}^2};$

(ii) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{(X,Y)}(x,y)\, dydx = \int_0^1 \int_0^2 \frac{1}{2} dydx =$
$= \int_0^1 \left[\frac{1}{2}y\right]_0^2 dx = \int_0^1 1dx = [x]_0^1 = 1$

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Definição: Dado o par (X, Y) discreto (contínuo), com função de probabilidade conjunta $P(X = x, Y = y)$ (f.d.p. conjunta $f_{X,Y}(x, y)$), as funções de probabilidade (f.d.p.) marginais de X são:

$$P(X = x) = \sum_{y} P(X = x, Y = y) \ \forall_{x \in \mathbb{R}} \qquad \text{(discreto)}$$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y)\, dy =, \ \forall_{x \in \mathbb{R}} \qquad \text{(contínuo)}$$

e de Y são:

$$P(Y = y) = \textstyle\sum_x P(X = x, Y = y), \ \forall_{y \in \mathbb{R}} \qquad \text{(discreto)}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y)\, dx =, \ \forall_{y \in \mathbb{R}} \qquad \text{(contínuo)}$$

Exemplo 5.1 (cont.) - Funções de probabilidade marginais:

$X \backslash Y$	0	1	2	$P(X=x)$
0	$\frac{16}{36}$	$\frac{8}{36}$	$\frac{1}{36}$	$\frac{25}{36}$
1	$\frac{8}{36}$	$\frac{2}{36}$	0	$\frac{10}{36}$
2	$\frac{1}{36}$	0	0	$\frac{1}{36}$
$P(Y=y)$	$\frac{25}{36}$	$\frac{10}{36}$	$\frac{1}{36}$	1

por exemplo

$P(X=0) = P(X=0, Y=0) + P(X=0, Y=1) + P(X=0, Y=2) =$

$= \sum_{y=0}^{2} P(X=0, Y=y)$ (verificar que X e $Y \sim Bin(2, 1/6)$)

5.1 + 5.2 - Duas v.a. discretas ou contínuas

$$F_{(X,Y)}(x,y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} P(X = x_i, Y = y_j), \quad \forall_{(x,y) \in \mathbb{R}^2} \text{ (discreto)}$$

$$F_{(X,Y)}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v) dv du, \quad \forall_{(x,y) \in \mathbb{R}^2} \text{ (contínuo)}$$

Nota: As propriedades da função de distribuição são análogas às do caso univariado.

Exemplo 5.1 (cont.)

O valor da função distribuição no ponto $(0,1)$:
$F_{(X,Y)}(0,1) = P(X \leq 0, Y \leq 1) = P(X = 0, Y = 0) + P(X = 0, Y = 1) = \frac{24}{36}$

Exemplo 5.2 (cont.)
Cálculo da seguinte probabilidade:

$$P(X \leq \frac{1}{3}, Y \leq 1) = F_{(X,Y)}(\frac{1}{3}, 1) = \int_0^{\frac{1}{3}} \int_0^1 \frac{1}{2} dy dx = \frac{1}{6}$$

5.1 + 5.2 - Duas v.a. discretas ou contínuas

$$E[h(X,Y)] = \sum_x \sum_y h(x,y) P(X=x, Y=y) \qquad \text{(discreto)}$$

$$E[h(X,Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x,y) f_{X,Y}(x,y)\, dy dx \qquad \text{(contínuo)}$$

Diz-se que $E[h(X,Y)]$ existe se $E[|h(X,Y)|] < +\infty$

Casos particulares:

- $E(XY) = \sum_x \sum_y x\, y\, P(X=x, Y=y)$ (discreto)
- $E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x\, y\, f_{X,Y}(x,y)\, dy dx$ (contínuo)
- $E(X) = \sum_x \sum_y x\, P(X=x, Y=y) = \sum_x x \left(\sum_y P(X=x, Y=y) \right) = \sum_x x\, P(X=x)$ (discreto)
- $E[X] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{X,Y}(x,y)\, dy dx = \int_{-\infty}^{+\infty} x f_X(x)\, dx$ (contínuo)

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Exemplo 5.1 (cont.):

$X \backslash Y$	0	1	2	$P(X=x)$
0	16/36	8/36	1/36	25/36
1	8/36	2/36	0	10/36
2	1/36	0	0	1/36
$P(Y=y)$	25/36	10/36	1/36	1

- $E(XY) = \sum_x \sum_y x\, y\, P(X=x, Y=y) =$

$$= 0 \times 0 \times \frac{16}{36} + 0 \times 1 \times \frac{8}{36} + \cdots + 1 \times 1 \times \frac{2}{36} + 2 \times 2 \times 0 = \frac{1}{18}$$

- $E(X) = \sum_x \sum_y x\, P(X=x, Y=y) =$

$$= 0 \times \frac{16}{36} + 0 \times \frac{8}{36} + 0 \times \frac{1}{36} + 1 \times \frac{8}{36} + 1 \times \frac{2}{36} + 2 \times \frac{1}{36} = \frac{1}{3}$$

ou

$$E(X) = \sum_x x\, P(X=x) = 0 \times \frac{25}{36} + 1 \times \frac{10}{36} + 2 \times \frac{1}{36} = \frac{1}{3}$$

ou ainda

$$E(X) = 2 \times \frac{1}{6} = \frac{1}{3}, \text{ dado que } X \sim Bin(2, \tfrac{1}{6})$$

Definição: Dado o par (X, Y) discreto (contínuo), com função de probabilidade conjunta $P(X = x, Y = y)$ (f.d.p conjunta $f_{X,Y}(x,y)$)), chama-se **função de probabilidade (f.d.p.) condicionada** de Y dado x (se $P(X = x) > 0$, $f_X(x) > 0$) à função :
Caso discreto:

$$f_{Y|x}(y) = P(Y = y|X = x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{P(X = x, Y = y)}{P(X = x)}$$

que verifica

1) $f_{Y|x}(y) \geq 0, \quad \forall_y$ e **2)** $\sum_y f_{Y|x}(y) = 1$

Caso contínuo:

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

que verifica

1) $f_{Y|x}(y) \geq 0, \quad \forall_y$ e **2)** $\int_y f_{Y|x}(y)dy = 1$

5.1 + 5.2 - Duas v.a. discretas ou contínuas

- Para cada x fixo, $f_{Y|x}(y)$, tem as propriedades de uma função de probabilidade (é a f.p. da v.a. $Y|X = x$)
- Analogamente define-se $f_{X|y}(x) = P(X = x|Y = y)$.
- E as funções de distribuições condicionadas de $Y|X = x$ e $X|Y = y$ definem-se como usualmente.

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Definição: O valor e esperado condicionado e a variância condicionada de Y dado x (tal que $f_X(x) > 0$) são respectivamente:

Caso discreto:

$$E(Y|X=x) = \sum_y y\, P(Y=y|X=x)$$

$$V(Y|X=x) = E(Y^2|X=x) - E^2(Y|X=x)$$

Caso contínuo:

$$E(Y|X=x) = \int_{-\infty}^{+\infty} y\, f_{Y|x}(y) dy$$

$$V(Y|X=x) = E(Y^2|X=x) - E^2(Y|X=x)$$

5.1 + 5.2 - Duas v.a. discretas ou contínuas

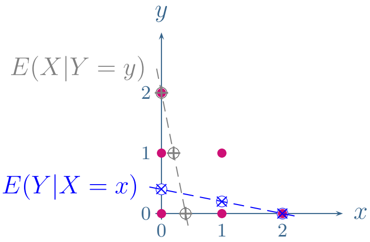
Exemplo 5.1 (cont.):

$X \backslash Y$	0	1	2	$P(X=x)$
0	$16/36$	$8/36$	$1/36$	$25/36$
1	$8/36$	$2/36$	0	$10/36$
2	$1/36$	0	0	$1/36$
$P(Y=y)$	$25/36$	$10/36$	$1/36$	1

$Y \mid X=0$	$Y \mid X=1$
$P(Y=0 \mid X=0) = \frac{16/36}{25/36} = 16/25$	$P(Y=0 \mid X=1) = \frac{8/36}{10/36} = 8/10$
$P(Y=1 \mid X=0) = \frac{8/36}{25/36} = 8/25$	$P(Y=1 \mid X=1) = \frac{2/36}{10/36} = 2/10$
$P(Y=2 \mid X=0) = \frac{1/36}{25/36} = 1/25$	$P(Y=2 \mid X=1) = \frac{0}{10/36} = 0$
$E(Y \mid X=0) = 1 \times \frac{8}{25} + 2 \times \frac{1}{25} = 10/25$	$E(Y \mid X=1) = 1 \times \frac{2}{10} = 2/10$

O que acontece com $Y \mid X=2$?

Representação gráfica de valores esperados condicionados:



Observações:

- Os $E(X|Y)$ e $E(Y|X)$ são v.a.(s) que tomam diferentes valores consoantes os valores fixos para X e Y respectivamente.
- Propriedade destas v.a.(s) :
  - $E(E(X|Y)) = E(X)$ se $E(X)$ existir e $f_Y(y) > 0$.
  - $E(E(Y|X)) = E(Y)$ se $E(Y)$ existir e $f_X(x) > 0$.

5.1 + 5.2 - Duas v.a. discretas ou contínuas

Definição: Seja (X, Y) um vector aleatório. As variáveis aleatórias X e Y (contínuas ou discretas), dizem-se independentes, simbolicamente $X \perp\!\!\!\perp Y$, sse:

$$F_{X,Y}(x, y) = F_X(x)\, F_Y(y), \quad \forall_{(x,y)\in\mathbb{R}^2} \tag{1}$$

A equação (1) é equivalente a:

- $P(X = x, Y = y) = P(X = x)P(Y = y), \quad \forall_{(x,y)\in\mathbb{R}^2}$, se (X, Y) for um vector aleatório discreto;
- $f_{X,Y}(x, y) = f_X(x)\, f_Y(y), \quad \forall_{(x,y)\in\mathbb{R}^2}$, se (X, Y) for um vector aleatório contínuo.

Observação: Podemos dizer, em geral, que as variáveis aleatórias são independentes sse:

$$P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B)$$

para quaisquer acontecimentos A e B definidos no eixo dos xx e no eixo dos yy, respectivamente.

Se $X \perp\!\!\!\perp Y$ então:

- $f_{Y|x}(y) = f_Y(y), \qquad \forall_{(x,y),\ \text{com } f_X(x)>0}$
- $f_{X|y}(x) = f_X(x), \qquad \forall_{(x,y),\ \text{com } f_Y(y)>0}$
- $E(XY) = E(X)E(Y)$.

No **Exemplo 5.1** X e Y serão independentes? E no **Exemplo 5.2**?

5.3 Covariância e correlação: Propriedades

Objectivo: Duas medidas do grau de associação linear de duas variáveis aleatórias:

- covariância;
- correlação.

Definição: A **covariância** de duas variáveis aleatórias X e Y, com valores esperados $E(X)$ e $E(Y)$, respectivamente, representa-se por $\text{cov}(X, Y)$ ou σ_{XY}, é definida por

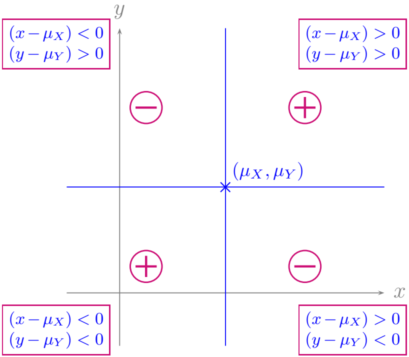
$$\text{cov}(X, Y) = E\left[(X - E(X))(Y - E(Y))\right]$$

e calcula-se por $\displaystyle\sum_x \sum_y (x - E(X))(y - E(Y))P(X = x, Y = y)$

se X e Y forem discretas, ou por

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - E(X))(y - E(Y)) f_{X,Y}(x, y) dx dy$$

se X e Y forem contínuas.


y
$(x-\mu_X)<0$
$(y-\mu_Y)>0$
$(x-\mu_X)>0$
$(y-\mu_Y)>0$
$\ominus$
$\oplus$
(μ_X, μ_Y)
$\oplus$
$\ominus$
x
$(x-\mu_X)<0$
$(y-\mu_Y)<0$
$(x-\mu_X)>0$
$(y-\mu_Y)<0$

Interpretação do valor e sinal covariância:

Mede a variação conjunta de duas variáveis, podendo ser interpretada do modo seguinte:

- se for positiva, as duas variáveis variam em média no mesmo sentido;
- se for negativa, as duas variáveis variam em média em sentidos contrários;
- se for nula, não se verifica nenhuma das tendências anteriores.

5.3 Covariância e correlação: Propriedades

1. $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$
2. $cov(X, Y) \in \mathbb{R}$ e nas (unid$X$)$\times$(unid$Y$)
3. $\text{cov}(X, Y) = \text{cov}(Y, X)$
4. $\text{cov}(X, X) = \text{var}(X)$
5. $\text{cov}(aX + b, cY + d) = a\ c\,\text{cov}(X, Y)$
6. $\text{cov}(X + Z, Y) = \text{cov}(X, Y) + \text{cov}(Z, Y)$
7. Se $X \perp\!\!\!\perp Y$ então $\text{cov}(X, Y) = 0$ ... Muito importante: a proposição inversa não é verdadeira, i.e.

$$\text{cov}(X, Y) = 0 \quad \not\Rightarrow \quad X \text{ e } Y \text{ são independentes}$$

Exercício: Demonstrar estas propriedades

(notar que $E(X + Y) = E(X) + E(Y) \;\; \forall_{X,Y}$)

5.3 Covariância e correlação: Propriedades

Exemplo 5.1 (cont.):

$X \backslash Y$	0	1	2	$P(X=x)$
0	$16/36$	$8/36$	$1/36$	$25/36$
1	$8/36$	$2/36$	0	$10/36$
2	$1/36$	0	0	$1/36$
$P(Y=y)$	$25/36$	$10/36$	$1/36$	1

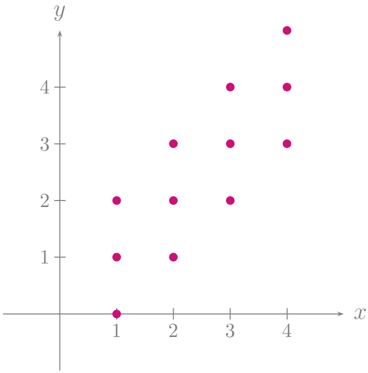
$$E(XY) = \frac{1}{18} \qquad E(X) = E(Y) = \tfrac{1}{3} \qquad \text{(já calculados)}$$

$$\mathrm{cov}(X,Y) = \frac{1}{18} - \frac{1}{9} = -\frac{1}{18}$$

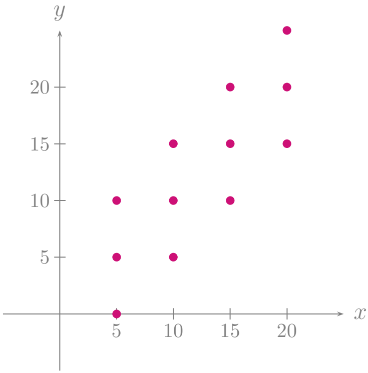
Significa que há uma tendência para Y decrescer quando X cresce e vice-versa.

Podemos saber se essa tendência é “forte” ou “fraca”?

A covariância é uma medida que não permite responder à questão anterior porque é sensível às mudanças de escala:



$\text{cov}(X, Y) = 1.3636$



$\text{cov}(5X, 5Y) = 25(1.3636) = 34.09$

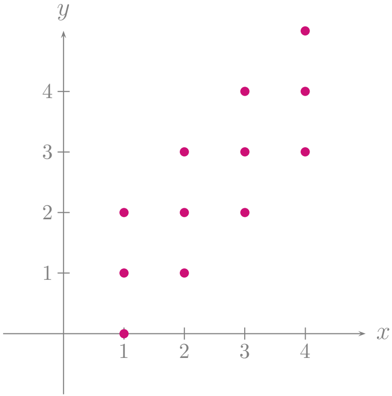
5.3 Covariância e correlação: Propriedades

Definição: A **correlação** ou **coeficiente de correlação** (linear) entre duas variáveis aleatórias, X e Y, representa-se por $\text{corr}(X, Y)$ ou $\rho_{X,Y}$ e é definida por

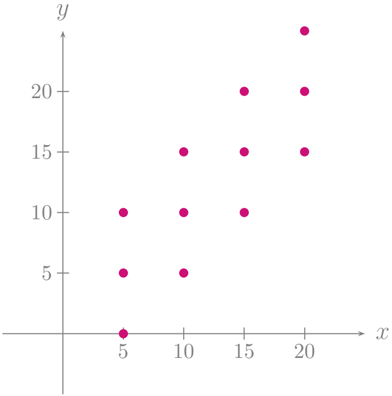
$$\text{corr}(X, Y) = \rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

5.3 Covariância e correlação: Propriedades

O coeficiente de correlação não é alterado quando há mudanças de escala e é adimensional:



$\rho_{X,Y} = 0.81$



$\rho_{5X,5Y} = 0.81$

Propriedades da correlação

1. $-1 \leq \rho_{X,Y} \leq 1$
2. $\rho_{X,Y} = 1$ se e só se $Y = a + bX$, com $b > 0$

   $\rho_{X,Y} = -1$ se e só se $Y = a + bX$, com $b < 0$
3. Se X e Y forem independentes então $\rho_{X,Y} = 0$.
   A proposição inversa não é necessariamente verdadeira, i.e.,

$$\rho_{X,Y} = 0 \quad \not\Rightarrow \quad X \text{ e } Y \text{ são independentes}$$

Demonstrações

1) Considere-se a variável aleatória $\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}$. Usando uma das propriedades da variância vem:

$$V\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) \geq 0 \;\Leftrightarrow\; E\left[\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right)^2\right] - \left[E\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right)\right]^2 \;\Leftrightarrow$$

$$\Leftrightarrow\; \frac{E(X^2)}{\sigma_X^2} + \frac{E(Y^2)}{\sigma_Y^2} + 2\,\frac{E(XY)}{\sigma_X\sigma_Y} - \frac{E^2(X)}{\sigma_X^2} - \frac{E^2(Y)}{\sigma_Y^2} - 2\,\frac{E(X)E(Y)}{\sigma_X\sigma_Y} \geq 0 \;\Leftrightarrow$$

$$\Leftrightarrow\; \frac{E(X^2) - E^2(X)}{\sigma_X^2} + \frac{E(Y^2) - E^2(Y)}{\sigma_Y^2} + 2\,\frac{E(XY) - E(X)E(Y)}{\sigma_X\sigma_Y} \geq 0 \;\Leftrightarrow$$

$$1 + 1 + 2\,\rho_{X,Y} \geq 0 \quad\Leftrightarrow\quad \rho_{X,Y} \geq -1$$

De igual modo $V\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) \geq 0 \;\Leftrightarrow\; \cdots \;\Leftrightarrow\; \rho_{X,Y} \leq 1$

5.3 Covariância e correlação: Propriedades

Demonstrações (cont.)

2) $\rho_{X,Y} = 1 \;\Leftrightarrow\; V\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 0 \;\Leftrightarrow\; \frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = \text{constante}$

$\rho_{X,Y} = -1 \;\Leftrightarrow\; V\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 0 \;\Leftrightarrow\; \frac{X}{\sigma_X} + \frac{Y}{\sigma_Y} = \text{constante}$

3) X e Y independentes $\Leftrightarrow f_{X,Y}(x,y) = f_X(x)\, f_Y(y), \;\Rightarrow$

$$\Rightarrow \quad E(XY) = \int\int x\, y\, f_{X,Y}(x,y) dxdy = \int\int x\, y\, f_X(x)\, f_Y(y) dxdy =$$

$$= \left(\int x\, f_X(x) dx\right)\left(\int y\, f_Y(y) dy\right) = E(X)\, E(Y)$$

isto é:

$$\text{cov}(X,Y) = 0 \quad \Leftrightarrow \quad \rho_{X,Y} = 0$$

5.3 Covariância e correlação: Propriedades

Para mostrar que
$\text{cov}(X, Y) = \text{corr}(X, Y) = 0 \quad \not\Rightarrow \quad X$ e Y são independentes

basta dar um contra-exemplo:

$X \backslash Y$	-1	0	1	$P(X = x)$
-1	0	$1/6$	0	$1/6$
0	$1/12$	$1/12$	$1/12$	$2/3$
1	0	$1/6$	0	$1/6$
$P(Y = y)$	$1/12$	$5/6$	$1/12$	1

I. $E(XY) = 0$, $E(X) = E(Y) = 0$, logo $\text{cov}(X, Y) = 0$

II. No entanto, X e Y não são independentes

(por exemplo, $P(X = -1, Y = -1) \neq P(X = -1) \times P(Y = -1)$)

5.3 Covariância e correlação: Propriedades

Exemplo 5.1 (cont.):

$X \backslash Y$	0	1	2	$P(X = x)$
0	$16/36$	$8/36$	$1/36$	$25/36$
1	$8/36$	$2/36$	0	$10/36$
2	$1/36$	0	0	$1/36$
$P(Y = y)$	$25/36$	$10/36$	$1/36$	1

Vimos antes que $\text{cov}(X, Y) = -1/18$, o que significa que há uma tendência para Y decrescer quando X cresce e vice-versa, e podemos agora responder à questão: essa tendência é "forte" ou "fraca"?

Como X e $Y \sim Bin(2, 1/6)$, tem-se que $V(X) = V(Y) = np(1-p) = 5/18$, logo

$$\text{corr}(X, Y) = \frac{-1/18}{\sqrt{5/18 \times 5/18}} = -\frac{1}{5}$$

o que permite concluir que a sua tendência/relação é "fraca" (por o seu valor estar bastante mais próximo de 0 do que de -1), para além de que as variáveis estão correlacionados linearmente no sentido negativo.

Definição: Dadas p variáveis aleatórias, $X_1, X_2, \ldots, X_p$ e p constantes reais, $c_1, c_2, \ldots, c_p$, diz-se que a variável aleatória Y, definida como

$$Y = c_1 X_1 + c_2 X_2 + \cdots + c_p X_p$$

é uma combinação linear de $X_1, X_2, \ldots, X_p$.

Valor esperado de uma combinação linear

- $E(c_1 X_1 + c_2 X_2) = c_1 E(X_1) + c_2 E(X_2)$

  Demonstração: $\displaystyle\int\int (c_1 x_1 + c_2 x_2) f(x_1, x_2) dx_1 dx_2 =$

  $$= c_1 \int\int x_1 f(x_1, x_2) dx_1 dx_2 + c_2 \int\int x_2 f(x_1, x_2) dx_1 dx_2$$

- $E(c_1 X_1 + c_2 X_2 + \cdots + c_p X_p) = c_1 E(X_1) + c_2 E(X_2) + \cdots + c_p E(X_p)$

Variância de uma combinação linear

- $V(c_1 X_1 + c_2 X_2) = c_1^2 V(X_1) + c_2^2 V(X_2) + 2\, c_1\, c_2 \operatorname{cov}(X_1, X_2)$

Dem.: $V(c_1X_1 + c_2X_2) = E\left[(c_1X_1 + c_2X_2)^2\right] - [E(c_1X_1 + c_2X_2)]^2 =$

$$= E(c_1^2X_1^2 + c_2^2X_2^2 + 2c_1c_2X_1X_2) - (c_1E(X_1) + c_2E(X_2))^2 =$$

$$= c_1^2E(X_1^2) + c_2^2E(X_2^2) + 2c_1c_2E(X_1X_2) -$$
$$c_1^2E^2(X_1) - c_2^2E^2(X_2) - 2c_1c_2E(X_1)E(X_2) =$$

$$= c_1^2\left(E(X_1^2) - E^2(X_1)\right) + c_2^2\left(E(X_2^2) - E^2(X_2)\right) +$$
$$+2c_1c_2\left(E(X_1X_2) - E(X_1)E(X_2)\right)$$

Variância de uma combinação linear (cont.)

- $V(c_1X_1 + \cdots + c_pX_p) = \sum_{i=1}^{p} c_i^2 V(X_i) + 2\sum_{i=1}^{p} \sum_{j=1, j>i}^{p} c_i\, c_j \operatorname{cov}(X_i, X_j)$

- Se $\operatorname{cov}(X_i, X_j) = 0$, para qualquer $i \neq j$, ou seja, se as variáveis aleatórias forem não correlacionadas duas a duas, então

$$V(c_1\, X_1 + \cdots + c_p\, X_p) = \sum_{i=1}^{p} c_i^2 V(X_i)$$

- O mesmo acontece se as variáveis aleatórias forem independentes duas a duas, pois como se viu, independência $\Rightarrow$ covariância (e correlação) nula.

Casos especiais de somas/combinações lineares de variáveis aleatórias

1. Soma de binomiais independentes com a mesma probabilidade de sucesso

$$\left.\begin{array}{l} X_1 \sim Bin(n_1, p) \\ X_2 \sim Bin(n_2, p) \\ X_1 \text{ e } X_2 \text{ independentes} \end{array}\right\} \Rightarrow X_1 + X_2 \sim Bin(n_1 + n_2, p)$$

$$X_i \sim Bin(n_i, p) \text{ independentes } \Rightarrow X_1 + \cdots + X_k \sim Bin\left(\sum_{i=1}^{k} n_i, p\right)$$

Caso especial:

$$X_i \sim Ber(p) \equiv Bin(1, p) \text{ independentes } \Rightarrow X_1 + \cdots + X_n \sim Bin\,(n, p)$$

2. Soma de Poisson's independentes

$$\left.\begin{array}{l} X_1 \sim \textit{Poisson}(\lambda_1) \\ X_2 \sim \textit{Poisson}(\lambda_2) \\ X_1 \text{ e } X_2 \text{ independentes} \end{array}\right\} \Rightarrow\ X_1 + X_2 \sim \textit{Poisson}(\lambda_1 + \lambda_2)$$

$$X_i \sim \textit{Poisson}(\lambda_i) \text{ independentes } \Rightarrow\ X_1 + \cdots + X_k \sim \textit{Poisson}\left(\sum_{i=1}^{k} \lambda_i\right)$$

3. Combinação linear de normais independentes

$$\left.\begin{array}{l} X_1 \sim N(\mu_1, \sigma_1^2) \\ X_2 \sim N(\mu_2, \sigma_2^2) \\ X_1 \text{ e } X_2 \text{ independentes} \end{array}\right\} \Rightarrow \; aX_1 + bX_2 \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$$

$$X_i \sim N(\mu_i, \sigma_i^2) \text{ independentes } \Rightarrow$$

$$\Rightarrow c_1X_1 + \cdots + c_kX_k \sim N\left(\textstyle\sum_{i=1}^k c_i\mu_i, \sum_{i=1}^k c_i^2\sigma_i^2\right)$$

5.6 Teorema do Limite Central

Na maioria das situações é difícil determinar a distribuição da soma de variáveis (mesmo que sejam independentes!). O teorema seguinte justifica a grande utilidade e importância da distribuição normal (quer em probabilidades quer em estatística).

Teorema do Limite Central: Seja $X_1, \ldots, X_n$ uma sucessão de v.a. independentes e identicamente distribuídas com valor esperado $\mu < \infty$ e variância $\sigma^2 < \infty$. Considere-se $S_n = \sum_{i=1}^{n} X_i$, então quando $n \to +\infty$

$$\frac{S_n - E\left(S_n\right)}{\sqrt{V\left(S_n\right)}} = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \overset{a}{\sim} N\left(0,1\right),$$

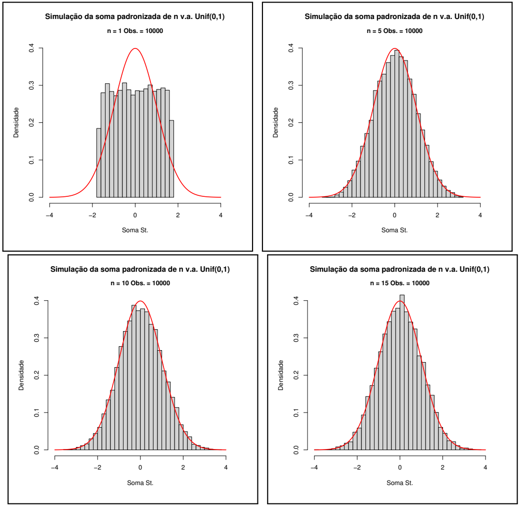
ou seja,

$$\lim_{n \to +\infty} P\left(\frac{S_n - E\left(S_n\right)}{\sqrt{V\left(S_n\right)}} \leq z\right) = \Phi(z)$$

Simulações:

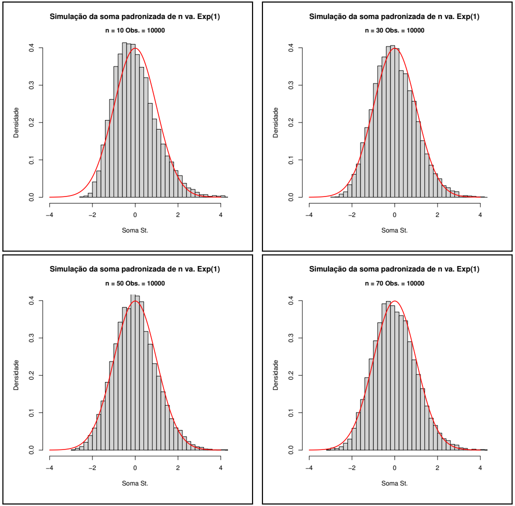
- https://www.youtube.com/watch?v=dlbkaurTAUg(dinâmica)
- Os gráficos das páginas seguintes representam simulações realizadas em R (http://www.r-project.org/).

Somas de Unif[0,1]



5.6 Teorema do Limite Central

Soma de Exponenciais de parâmetro 1



5.6 Teorema do Limite Central

Observações:

- A demonstração do teorema exige algumas ferramentas matemáticas avançadas.
- Observar que,

$$E\left(S_n\right) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = nE(X_1) = n\mu$$

e como $X_1, X_2 \cdots X_n$ são v.a. independentes tem-se

$$V\left(S_n\right) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = nV(X_1) = n\sigma^2$$

- As v.a. $X_1, \ldots, X_n$ podem ser discretas ou contínuas.
- Geralmente considera-se n grande se $n \geq 30$
- As distribuições Binomial e de Poisson podem ser aproximadas pela distribuição normal (na secção anterior vimos que podem ser escritas como somas de variáveis aleatórias).

Aproximações Binomial/Normal e Poisson/Normal:

- $X \sim Bin(n, p)$ pode ser aproximada por $\tilde{X} \sim N(np, np(1-p))$ quando $np > 5$ e $n(1-p) > 5$

- $X \sim Poisson(\lambda)$ pode ser aproximada por $\tilde{X} \sim N(\lambda, \lambda)$ quando $\lambda > 5$

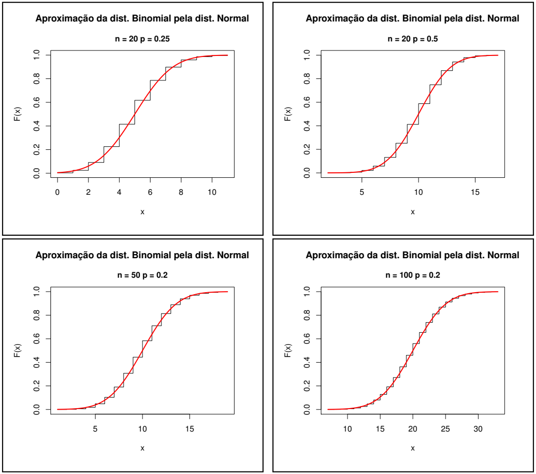
- **Correcção de continuidade:** em geral a aproximação é melhor se fizermos
$$P(X \leq x) \simeq P(\tilde{X} \leq x + 0.5)$$

Exemplo: O número de chamadas de telemóvel registadas a partir de certa "zona" numa hora tem, em condições estacionárias, distribuição de Poisson de parâmetro 1500. Calcule a probabilidade de ocorrerem mais de 1600 chamadas na próxima hora.

$$P(X > 1600) = 1 - P(X \leq 1600) \simeq 1 - P(\tilde{X} \leq 1600.5) = \Phi(-2.59) \simeq 0.0048$$

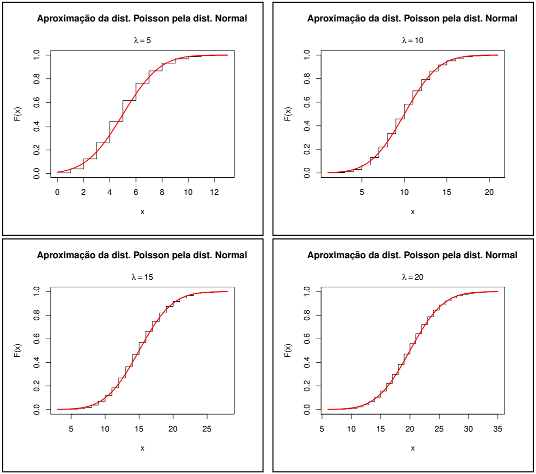
5.6 Teorema do Limite Central

Convergência das funções distribuição



5.6 Teorema do Limite Central

Convergência das funções distribuição



Exercícios

Exercício 5.1: Segundo os cálculos do engenheiro responsável pelo tráfego de uma dada ponte, a carga W (em toneladas) que o tabuleiro dessa ponte pode suportar sem sofrer danos estruturais segue uma distribuição normal, de valor médio 400 toneladas e desvio padrão 40 toneladas. Considere que os pesos dos veículos que nela circulam são variáveis aleatórias normais, independentes, com valor médio 3 toneladas e desvio padrão 0.3 toneladas.

(a) Admita que, em certo momento, estão 100 veículos sobre a ponte. Determine a probabilidade de o peso total desses veículos exceder 400 toneladas.

Exercícios

(b) Prove que o valor esperado e a variância da variável aleatória que representa a diferença entre o peso total de n veículos e a carga W que a ponte pode suportar são, respectivamente, $3n - 400$ e $0.09n + 1600$, admitindo que o peso total e a carga são variáveis aleatórias independentes. Determine ainda o maior valor de n para o qual a probabilidade de ocorrência de danos na estrutura é inferior a 0.1?

Exercício 5.2: Suponha-se que ao adicionar números reais cada número é arredondado previamente para o inteiro mais próximo. Admita-se que os erros de arredondamento são v.a. independentes e identicamente distribuídas com distribuição uniforme contínua no intervalo $[-0.5; 0.5]$ (esta suposição é razoável se desconhecermos à partida tudo sobre os referidos números reais e admitirmos que também eles se distribuem uniformemente e independentemente nalgum intervalo).

a) Qual é a probabilidade de que, ao adicionar 1500 números, o valor absoluto do erro seja superior a 15?

Exercícios

b) Quantos números podem ser somados para que se possa garantir com uma probabilidade de aproximadamente 90% que o erro absoluto não excede 10?