Probabilidades e Estatística

Variáveis Aleatórias

Prof. Caldeira Duarte e Prof.ª Anabela Pereira

Actualizado pelo Prof. António Sardinha (Fevereiro de 2013)
Actualizado pela Prof.ª Paula Pereira (Fevereiro de 2017)

Departamento de Matemática

IPS Instituto Politécnico de Setúbal
Escola Superior de Tecnologia de Setúbal

1 VARIÁVEIS ALEATÓRIAS

Em muitas experiências aleatórias os elementos do espaço de resultados, Ω, são números reais ou conjuntos ordenados de números reais. Assim acontece com o registo de temperaturas, da pluviosidade, no lançamento de dois dados, etc. Mas já o resultado do lançamento de uma moeda ao ar não é um resultado numérico.

Quando Ω não é um conjunto numérico atribui-se muitas vezes a cada elemento ω do espaço de resultados, um número real, atribuição essa que pode ser meramente convencional.

Exemplo 1.1. No lançamento de uma moeda ao ar, o espaço de resultados é o conjunto $\Omega = \{\text{Cara,Coroa}\}$; é usual neste caso fazer a correspondência

ω	X(ω)
Cara	1
Coroa	0

.

Para o mesmo espaço de resultados Ω, é possível estabelecer diferentes correspondências, consoante os objectivos em estudo. ■

Exemplo 1.2. Considere-se uma população de empresas das quais se escolhe uma ao acaso. $\Omega = \{\omega_1,\omega_2,...,\omega_m\}$ será o espaço de resultados e m o número total de empresas.
Podemos definir várias correspondências:
$\omega \to X(\omega)$, sendo $X(\omega)$ o número de empregados da empresa ω,
$\omega \to Y(\omega)$, sendo $Y(\omega)$ o volume de vendas da empresa ω,
ou quaisquer outras, conforme o objectivo em estudo. ■

Supondo agora que só estamos interessados no estudo de uma característica dos elementos de Ω, associemos a cada elemento $\omega \in \Omega$ um número real $X(\omega)$. Estamos assim a definir uma função $X : \Omega \to \mathbb{R}$. Sendo A um acontecimento, chama-se imagem de A por X, e representa-se por $X(A)$, ao conjunto dos valores que X assume para os elementos ω de A, isto é,

$$X(A) = \{X(\omega) : \omega \in A\}.$$

Por outro lado, a cada subconjunto $E \subset \mathbb{R}$, pode fazer-se corresponder o subconjunto $X^{-1}(E)$ formado por todos os elementos $\omega \in \Omega$ tais que $X(\omega) \in E$,

$$X^{-1}(E) = \{\omega : X(\omega) \in E\}.$$

A este conjunto $X^{-1}(E)$ chama-se a imagem inversa de E por X.

Exemplo 1.3. No lançamento de dois dados interessa somente, num dado jogo, a soma dos pontos obtidos. Neste caso, o espaço de resultados é o conjunto

$$\Omega = \{(i, j) : i, j = 1, 2, 3, 4, 5, 6\};$$

defina-se a aplicação

$$X(i, j) = i + j.$$

Sendo $A = \{(1, 1), (1, 2), (2, 1)\}$, a imagem de A por X é $X(A) = \{2, 3\}$; para o acontecimento $B = \{(4, 5), (5, 4), (5, 5), (6, 6)\}$, a imagem de B por X é $X(B) = \{9, 10, 12\}$.

Para o subconjunto real $E_1 = \{2, 3\}$, a imagem inversa de E_1 por X é o acontecimento $X^{-1}(E_1) = \{(1, 1), (1, 2), (2, 1)\}$; se $E_2 = [2, +\infty[$, $X^{-1}(E_2) = \Omega$ e se $E_3 = \left]-\infty, \frac{1}{2}\right]$, $X^{-1}(E_3) = \emptyset$. ■

Estamos agora em condições de perceber a definição de variável aleatória.

Definição 1.4. *Uma variável aleatória (v.a.), X, é uma função com domínio no conjunto Ω dos acontecimentos elementares e contradomínio no conjunto $\mathbb{R}$, cuja imagem inversa de qualquer intervalo $E \subset \mathbb{R}$ é um acontecimento aleatório.*

Observação 1.5. Uma variável aleatória é uma função e não uma variável no sentido em que é habitualmente empregue em Matemática!

1.1 Funções de Distribuição

Considere-se agora uma variável aleatória X, um intervalo real $E =]-\infty, x]$ e a respectiva imagem inversa $X^{-1}(E)$. Pela definição de variável aleatória existe sempre

$$P(X \leq x) = P\left[X^{-1}(E)\right].$$

Como $P(X \leq x)$ depende de x, a igualdade $F(x) = P(X \leq x)$ define uma função real de variável real.

Definição 1.6. *A função $F(x)$ definida por $F(x) = P(X \leq x)$ chama-se a **Função de Distribuição** da variável aleatória X.*

Vamos agora enunciar algumas das propriedades elementares mais importantes das funções de distribuição . As demonstrações destas propriedades podem ser vistas em [3]. Tem-se então que, se $F(x)$ é uma função de distribuição,

Proposição 1.7. $0 \leq F(x) \leq 1$.

Proposição 1.8. *$F(x)$ é uma função não decrescente.*

Proposição 1.9. $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ *e* $F(+\infty) = \lim_{x \to +\infty} F(x) = 1$.

Proposição 1.10. $P(a < X \leq b) = F(b) - F(a)$.

Proposição 1.11. *$F(x)$ é contínua à direita.*

Proposição 1.12. $P(X = a) = F(a) - \lim_{x \to a^-} F(x)$.

Desta proposição pode concluir-se que, se a função de distribuição é contínua para todo o $x \in \mathbb{R}$, tem-se $P(X = x) = 0$, isto é, todos os pontos reais têm probabilidade zero. Mas atenção: isto não significa forçosamente que os acontecimentos $(X = x)$ sejam impossíveis.

Proposição 1.13. *O conjunto de pontos de descontinuidade de qualquer função de distribuição, se não for vazio, é finito ou infinito numerável.*

Iremos agora tratar fundamentalmente de dois tipos de variáveis aleatórias, as do tipo discreto e as do tipo contínuo.

1.2 Variáveis Aleatórias Discretas

Definição 1.14. *Seja X uma variável aleatória e D o conjunto*

$$D = \{a : P(X = a) > 0\}$$

*(conjunto dos pontos de descontinuidade da função de distribuição). A variável aleatória X diz-se do tipo **discreto** quando $P(X \in D) = 1$.*

Quando a variável aleatória é discreta existe um conjunto finito ou infinito numerável,

$$D = \{a_1, a_2, ..., a_n, ...\},$$

tal que,

$$P(X \in D) = \sum_i P(X = a_i) = 1 \quad \text{e} \quad P(X = a_i) > 0, i = 1, 2, ...$$

Definição 1.15. *Seja D o conjunto definido anteriormente. A função,*

$$f(x) = \begin{cases} P(X = x) > 0 & se \quad x \in D \\ 0 & se \quad x \notin D \end{cases}$$

*chama-se **função de probabilidade** da variável aleatória X e verifica as seguintes* propriedades:

$$f(x) \geq 0;$$
$$\sum_x f(x)dx = 1.$$

A função de distribuição de uma variável aleatória discreta pode exprimir-se facilmente em termos da respectiva função de probabilidade:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) = \sum_{x_i \leq x} P(X = x_i).$$

Exemplo 1.16. Seja X a variável aleatória que representa o número de caras saídas no lançamento de 3 moedas equilibradas. O quadro seguinte apresenta a função de probabilidade (ver Figura 1.1) desta variável:

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

Como $F(x) = P(X \leq x)$ então a função de distribuição (ver Figura 1.1) será:

se $x < 0 \Rightarrow F(x) = P(X \leq x) = P(\emptyset) = 0$

se $0 \leq x < 1 \Rightarrow F(x) = P(X = 0) = 1/8 = 0.125$

se $1 \leq x < 2 \Rightarrow F(x) = P(X = 0) + P(X = 1) = 1/8 + 3/8 = 0.5$

se $2 \leq x < 3 \Rightarrow F(x) = P(X = 0) + P(X = 1) + P(X = 2) = 1/8 + 3/8 + 3/8 = 0.875$

se $x \geq 3 \Rightarrow F(x) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = P(\Omega) = 1$

Resumindo,

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 0.125 & , \quad 0 \leq x < 1 \\ 0.5 & , \quad 1 \leq x < 2 \\ 0.875 & , \quad 2 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases} .$$

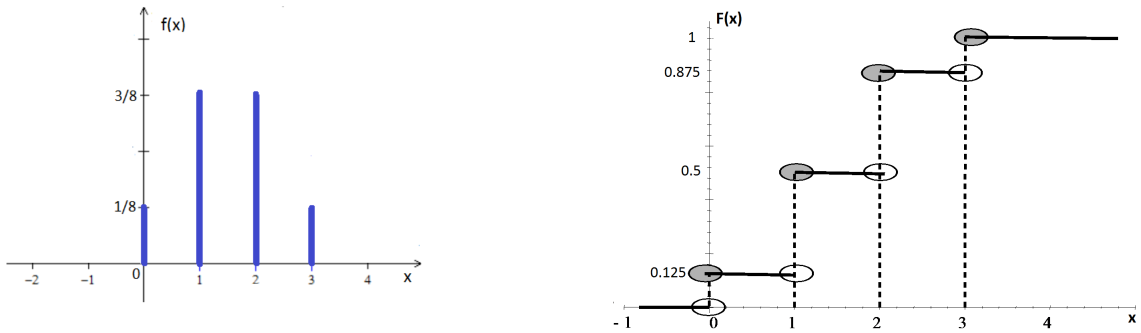

Figura 1.1: Função de probabilidade e Função de distribuição da v.a. discreta do exemplo 1.16

■

1.3 Variáveis Aleatórias Contínuas

Definição 1.17. *Seja X uma variável aleatória e $F(x)$ a respectiva função de distribuição; se,*

$$D = \{a : P(X = a) > 0\} = \emptyset,$$

resulta da proposição 1.12 que $F(x)$ não apresenta descontinuidades. Se, além disso, existe uma função não negativa, $f(x) \geq 0$, tal que para todo o número real x se verifica a relação,

$$F(x) = P\left(X \leq x\right) = \int_{-\infty}^{x} f(u)du,$$

*então a variável aleatória X diz-se **contínua**.*

Definição 1.18. *A função não negativa, $f(x)$, introduzida na definição anterior, chama-se **função de densidade de probabilidade** ou simplesmente função de densidade.*

Exemplo 1.19. Suponha que tem uma variável aleatória contínua, X, com a seguinte função de distribuição (ver Figura 1.2):

$$F\left(x\right) = \begin{cases} 0 & se \quad x \leq 1 \\ \frac{x-1}{2} & se \quad 1 < x < 3 \\ 1 & se \quad x \geq 3 \end{cases} .$$

A derivada $F'(x)$ existe em todos os pontos da recta real excepto em $x = 1$ e $x = 3$. Então $f(x)$, a função de densidade de probabilidade é igual à derivada da função de distribuição em todos os pontos onde exista (ver Figura 1.2), e convenciona-se que é nula nos restantes, isto é

$$f(x) = \begin{cases} F'(x) & , \quad \text{nos pontos onde existe derivada} \\ 0 & , \quad \text{nos outros pontos} \end{cases}$$

então, neste caso, tem-se

$$f(x) = \begin{cases} \frac{1}{2} & se \quad 1 < x < 3 \\ 0 & se \quad x \leq 1 \vee x \geq 3 \end{cases} .$$

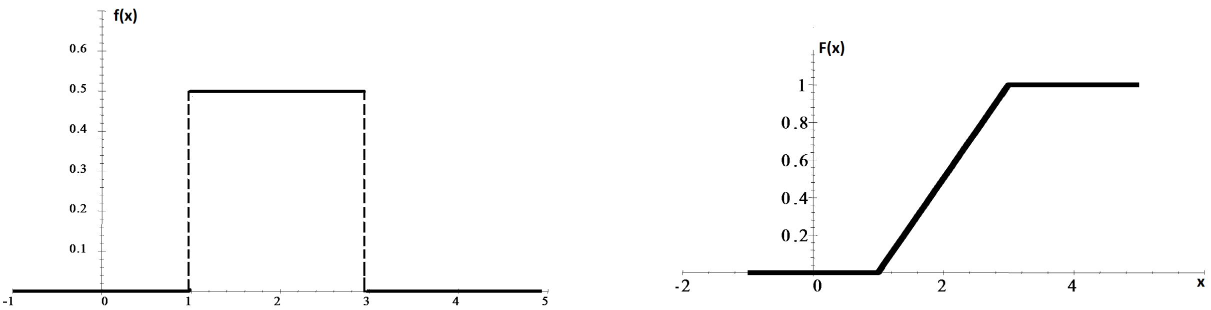

Figura 1.2: Função de densidade de probabilidade e Função de distribuição da v.a. contínua do exemplo 1.19

■

Exemplo 1.20. Suponha que X é uma variável aleatória contínua com a seguinte função densidade de probabilidade

$$f(x) = \begin{cases} 1+x & se \quad -1 \leq x < 0 \\ 1-x & se \quad 0 \leq x \leq 1 \\ 0 & se \quad x < -1 \vee x > 1 \end{cases} .$$

Como $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$ então a função de distribuição (ver Figura 1.3) será:

$$\text{se } x < -1 \quad \Rightarrow \quad F(x) = \int_{-\infty}^{x} 0du = 0$$

$$\text{se } -1 \leq x < 0 \quad \Rightarrow \quad F(x) = \int_{-\infty}^{-1} 0du + \int_{-1}^{x} (1+u)du = \frac{x^2}{2} + x + \frac{1}{2}$$

$$\text{se } 0 \leq x \leq 1 \quad \Rightarrow \quad F(x) = \int_{-\infty}^{-1} 0du + \int_{-1}^{0} (1+u)du + \int_{0}^{x} (1-u)du = -\frac{x^2}{2} + x + \frac{1}{2}$$

$$\text{se } x > 1 \quad \Rightarrow \quad F(x) = \int_{-\infty}^{-1} 0du + \int_{-1}^{0} (1+u)du + \int_{0}^{1} (1-u)du + \int_{1}^{x} 0du = 1$$

Resumindo,

$$F(x) = \begin{cases} 0 & , \quad x < -1 \\ \frac{x^2}{2} + x + \frac{1}{2} & , \quad -1 \le x < 0 \\ -\frac{x^2}{2} + x + \frac{1}{2} & , \quad 0 \le x \le 1 \\ 1 & , \quad x > 1 \end{cases} .$$

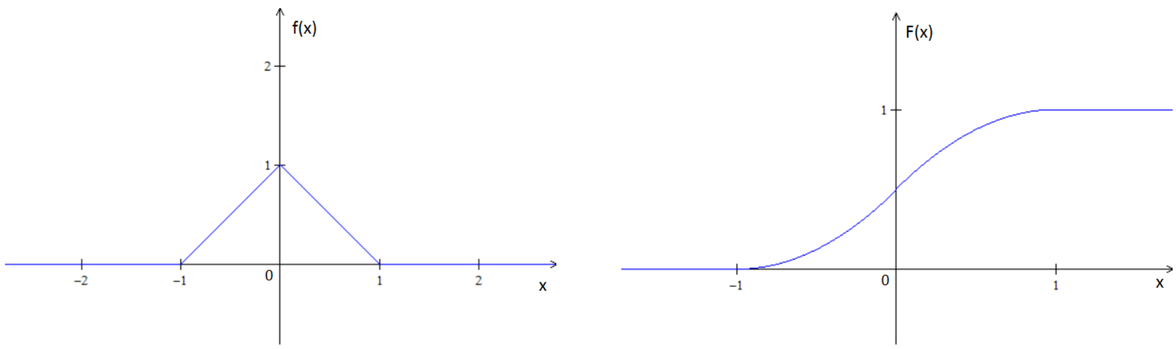

Figura 1.3: Função de densidade de probabilidade e Função de distribuição da v.a. contínua do exemplo 1.20

■

Da definição de função de distribuição e da sua relação com a função de densidade de probabilidade, têm-se as seguintes propriedades:

$$f(x) \ge 0;$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1;$$

$$\int_{a}^{b} f(x)dx = F(b) - F(a) = P(a < X < b).$$

Observação 1.21. Repare-se que $P(a < X < b)$ pode ser interpretada geometricamente como uma área, visto que é calculada através de um integral definido de uma função não negativa (ver Figura 1.4).

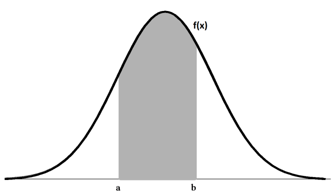

Figura 1.4: $P\,(a < X < b) = \int_a^b f(x)dx$

Observação 1.22. Devido à continuidade da função de distribuição,

$$P(X = a) = P(X = b) = 0,$$

e pode escrever-se

$$P(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b) = P(a < X < b), \quad b > a.$$

1.4 Valores Esperados e Parâmetros

De acordo com [3], o conceito de valor esperado teve a sua origem nos jogos de acaso e foi, segundo se diz, introduzido por Huygens.

Exemplo 1.23. Considere-se um sorteio em que foram vendidos 10000 bilhetes e cujos prémios, em euros, são:
1) um 1^o prémio de 500000;
2) um 2^o prémio de 60000;
3) três 3^{os} prémios de 30000;
4) dez 4^{os} prémios de 10000.

Se fizermos a soma dos produtos dos valores dos prémios que se podem ganhar pelas respectivas probabilidades,

$$500000\left(\frac{1}{10000}\right) + 60000\left(\frac{1}{10000}\right) + 30000\left(\frac{3}{10000}\right) + 10000\left(\frac{10}{10000}\right) = 75 \text{ euros}$$

obtemos aquilo que se chama valor esperado ou esperança matemática do comprador de um bilhete.
Suponhamos que uma pessoa compra sistematicamente um bilhete de uma lotaria deste tipo. Ao fim n repetições, o ganho total é dado por

$$500000.s_n(1) + 60000.s_n(2) + 30000.s_n(3) + 10000.s_n(4),$$

sendo $s_n(i)$ o número de vezes que saiu o i-ésimo prémio, $i = 1, 2, 3, 4$.
O ganho médio foi

$$500000.[s_n(1)/n] + 60000.[s_n(2)/n] + 30000.[s_n(3)/n] + 10000.[s_n(4)/n].$$

A interpretação frequencista do conceito de probabilidade sugere que, para n grande, as frequências relativas são praticamente iguais à probabilidade. Isto significa que ao fim de um grande número de jogadas o ganho médio será aproximadamente igual ao valor da chamada esperança matemática. É por isso que se diz que um jogo é equitativo quando o que se paga para nele participar é igual à esperança matemática do jogador: após um grande número de partidas, o ganho médio por partida não se afastará muito do preço pago para participar em cada partida. ■

Com qualquer distribuição de uma variável aleatória estão sempre associados certos números chamados os **parâmetros da distribuição,** que desempenham um relevante papel na Estatística Matemática. Duas importantes famílias de parâmetros de uma distribuição são: os **momentos** e os **parâmetros de ordem**; iremos, no entanto, apenas abordar os primeiros.

Definição 1.24. *O **momento de ordem** k **em relação à origem** ou momento ordinário de ordem k - k inteiro positivo - de uma variável aleatória é o valor esperado da função $G(X) = X^k$,isto é,*

$$\mu_k = E\left[X^k\right].$$

Se a variável aleatória for discreta,

$$E\left[X^k\right] = \sum_i x_i^k f(x_i);$$

no caso de ser contínua,

$$E\left[X^k\right] = \int_{-\infty}^{+\infty} x^k f(x)dx.$$

Definição 1.25. *O momento de ordem 1 em relação à origem de uma variável aleatória chama-se **esperança matemática**, **média** ou **valor esperado** e representa-se por μ ou $E[X]$.*

Se a variável aleatória for discreta, o valor esperado é

$$\mu = E[X] = \sum_i x_i f(x_i);$$

no caso de ser contínua é

$$\mu = E[X] = \int_{-\infty}^{+\infty} x f(x)dx.$$

Definição 1.26. *O **momento de ordem** k **em relação à média** ou **momento central de ordem** k - k inteiro positivo - de uma variável aleatória é o valor esperado da função $G(X) = (X-\mu)^k$,isto é,*

$$E\left[(X-\mu)^k\right].$$

Se a variável aleatória for discreta,

$$E\left[(X-\mu)^k\right] = \sum_i (x_i - \mu)^k f(x_i);$$

no caso de ser contínua,

$$E\left[(X-\mu)^k\right] = \int_{-\infty}^{+\infty} (x-\mu)^k f(x)dx.$$

Definição 1.27. *O momento central de 2ª ordem de uma variável aleatória X, é chamado **variância** de X, e representa-se habitualmente por $V[X]$ ou σ^2.*

Se a variável aleatória for discreta, a variância é

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_i (x_i - \mu)^2 f(x_i);$$

no caso de ser contínua é

$$\sigma^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx.$$

Definição 1.28. *A raiz quadrada positiva da variância de uma variável aleatória X, chama-se o **desvio padrão** e representa-se habitualmente por σ.*

Os parâmetros σ^2 e σ são medidas de dispersão da variável aleatória em torno do seu valor esperado. Quanto mais concentrada for a distribuição, tanto menor será o valor de σ^2 ou σ. A vantagem de σ em relação a σ^2 é que vem expressa na mesma unidade de medida da variável aleatória.

A seguir apresentam-se algumas propriedades do Valor Esperado e da Variância cuja demonstração se deixa como exercício ao aluno.

Proposição 1.29. *Se X é uma variável aleatória constante, isto é, se $X \equiv a$, então*

$$E\left[X\right] = a.$$

Proposição 1.30. *Se X é uma variável aleatória e a e b são constantes reais*

$$E\left[aX + b\right] = aE\left[X\right] + b.$$

Proposição 1.31. *Seja X uma variável aleatória e $G(X)$ e $H(X)$ funções de X; então*

$$E\left[G(X) + H(X)\right] = E\left[G(X)\right] + E\left[H(X)\right].$$

Proposição 1.32. *Se X é uma variável aleatória,*

$$V\left[X\right] = E\left[X^2\right] - E^2\left[X\right].$$

Demonstração.

$$\begin{aligned} V\left[X\right] &= E\left[(X-\mu)^2\right] = E\left[X^2 - 2\mu X - \mu^2\right] = \\ &= E\left[X^2\right] - 2\mu \underbrace{E\left[X\right]}_{\mu} + \mu^2 = E\left[X^2\right] - 2\mu^2 + \mu^2 = \\ &= E\left[X^2\right] - \mu^2 = E\left[X^2\right] - \left(E\left[X\right]\right)^2 = E\left[X^2\right] - E^2\left[X\right] \end{aligned}$$

□

Proposição 1.33. *Se X é uma variável aleatória,*

$$V\left[X\right] \geq 0.$$

Proposição 1.34. *Se X é uma variável aleatória constante, isto é, se $X \equiv a$, então*

$$V\left[X\right] = 0.$$

Proposição 1.35. *Se X é uma variável aleatória e a e b são constantes reais*

$$V\left[aX + b\right] = a^2 V\left[X\right].$$

1.5 Adenda

Se pretendermos analisar e relacionar duas variáveis aleatórias, X e Y, há parâmetros que caracterizam as ligações existentes entre ambas.

1.6 Covariância

A covariância é uma medida da distribuição conjunta dos valores dos desvios de X e Y em relação às respectivas médias; este parâmetro permite descrever o tipo de relação linear (positiva ou negativa) que existe (ou não) entre as variáveis unidimensionais X e Y.

Definição 1.36. *A **covariância** entre X e Y $[cov\,(X,Y)$ ou $\sigma_{X,Y}]$ define-se como:*

$$cov\,(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right] = \sigma_{X,Y}.$$

Algumas das propriedades da covariância são dadas em seguida.

Proposição 1.37. *Sejam X e Y duas variáveis aleatórias, então a sua covariância pode calcular-se do seguinte modo,*

$$cov\,(X,Y) = E\left[XY\right] - E\left[X\right]E\left[Y\right].$$

Demonstração.

$$\begin{aligned} cov\,(X,Y) &= E\left[(X-\mu_X)(Y-\mu_Y)\right] = \\ &= E\left[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y\right] = \\ &= E\left[XY\right] - \mu_X \underbrace{E\left[Y\right]}_{\mu_Y} - \mu_Y \underbrace{E\left[X\right]}_{\mu_X} + \mu_X\mu_Y = \\ &= E\left[XY\right] - E\left[X\right]E\left[Y\right] \end{aligned}$$

□

Proposição 1.38. *Se X e Y são variável aleatória independentes então,*

$$E\left[XY\right] = E\left[X\right]E\left[Y\right] \quad \text{e} \quad cov\,(X,Y) = 0.$$

O recíproco deste último resultado pode não ser verdadeiro, isto é sendo $cov\,(X,Y) = 0$ não se pode inferir que X e Y sejam independentes, somente que não existe uma relação linear entre as variáveis. Pode contudo existir uma ligação não linear entre X e Y.

Proposição 1.39. *Sejam X e Y duas variáveis aleatórias,*

- $E\left[X \pm Y\right] = E\left[X\right] \pm E\left[Y\right];$
- $V\left[X \pm Y\right] = V\left[X\right] + V\left[Y\right] \pm 2cov\,(X,Y);$
- *Se X e Y são independentes, então* $V\left[X \pm Y\right] = V\left[X\right] + V\left[Y\right].$

Referências

[1] FISZ, M., *Probability Theory and Mathematical Statistics*, Jonh Wiley & Sons, Inc., New York, 1963.

[2] GUIMARÃES, R.C. e CABRAL, J.A.S., *Estatística*, McGraw-Hill de Portugal, Lisboa, 1997.

[3] MURTEIRA, B.J.F., *Probabilidades e Estatística*, McGraw-Hill de Portugal, Lisboa, 1979.

[4] SPIEGEL, M.R., *Probabilidade e Estatística*, Colecção Schaum, McGraw-Hill do Brasil, São Paulo, 1978

[5] OLIVEIRA, J.T., *Probabilidades e Estatística*, vol. I, Escolar Editora, Lisboa,1967.

[6] MELLO, F., *Introdução aos Métodos Estatísticos*, vol. I e II, Cadernos do Instituto de Orientação Profissional, Lisboa, 1973.

[7] MONTGOMERY, D.; RUNGER, G., *Applied Statistics and Probability for Engineers*, Jonh Wiley & Sons, Inc., New York, 2003.