3-Variáveis aleatórias multivariadas

Se o modelo probabilístico envolver mais do que um atributo será necessário generalizar o conceito de variável aleatória passando-se de

$$\omega \in \Omega \to X(\omega) \in \Re$$

para

$$\omega \in \Omega \quad \to \quad \left(X_1(\omega), X_2(\omega), \ldots, X_k(\omega)\right) \in \Re^k$$

isto é, utiliza-se a variável aleatória k-dimensional $(X_1, \ldots, X_k)$.

O estudo centrar-se-á nas variáveis aleatórias bidimensionais: k=2

3.1 Variáveis aleatórias bidimensionais

- **Simplificação:** Em vez de (X_1, X_2) emprega-se (X, Y)

- **Definição 4.1 – Variável aleatória bidimensional**

  Uma v.a. bidimensional, (X, Y), é uma função com domínio Ω e com contradomínio em $\Re^2$, $(X,Y): \omega \in \Omega \quad \rightarrow \quad \big(X(\omega), Y(\omega)\big) \in \Re^2$

3.2 Função de distribuição conjunta

• **Definição – Função de distribuição conjunta**

Seja (X,Y) uma variável aleatória bidimensional. A função real de duas variáveis reais, com domínio $\Re^2$, definida por

$$F_{X,Y}(\mathrm{x},\ \mathrm{y}) = P(X \leq x,\ Y \leq y),$$

é a função de distribuição **conjunta** de (X,Y).

• **Simplificação:** Não havendo confusão, escreve-se $F\ (x,\ y)$ **em vez de** $F_{X,Y}(x,\ y)$

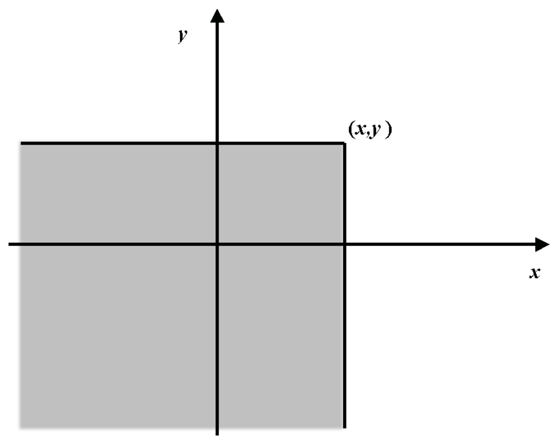


Fig. – Região do plano $\Re^2$ definido pelas desigualdades $X \leq x$ e $Y \leq y$.

- **Propriedades da função de distribuição**

1) $0 \le F(x, y) \le 1.$

2) F é não decrescente, separadamente, em relação a x e em relação a y:

$\Delta x > 0 \Rightarrow F(x, y) \le F(x + \Delta x, y)$ e $\Delta y > 0 \Rightarrow F(x, y) \le F(x, y + \Delta y)$.

3) Para qualquer função de distribuição $F\,(x, y)$,

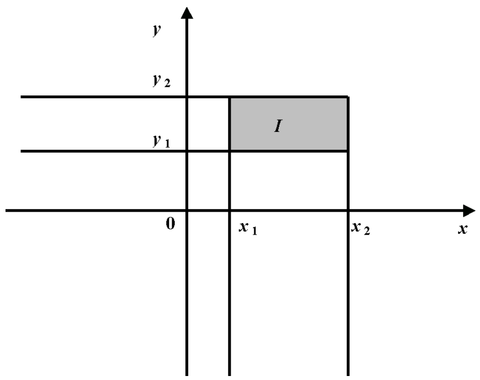
$$F(-\infty, y) = F(x, -\infty) = 0, \; F(+\infty, +\infty) = 1.$$

4) Considere-se o retângulo I de $\Re^2$ com vértices nos pontos,

$$(x_1, y_1), (x_2, y_1), (x_1, y_2), (x_2, y_2), \quad I = \{(x,\ y) : x_1 < x \le x_2,\ y_1 < y \le y_2\}.$$

Então

$$\begin{aligned} P(I) &= P(x_1 < X \le x_2,\ y_1 < Y \le y_2) \\ &= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) \\ &\quad + F(x_1, y_1) \end{aligned}$$



5)Qualquer função de distribuição F é contínua à direita em relação a x e em relação a y,

$$\lim_{s \to 0^+} F\,(x + s, y) = F(x^+, y) = F(x, y) \text{ e}$$
$$\lim_{s \to 0^+} F\,(x, y + s) = F(x, y^+) = F(x, y).$$

.

- **Exemplo -** Cálculo de probabilidades utilizando a função de distribuição

$$F(x, y) = \begin{cases} 0 & \text{se } x < 0 \text{ ou } y < 0 \\ 1 - e^{-x} - e^{-y} + e^{-x-y} & \text{se } x \geq 0 \text{ e } y \geq 0 \end{cases}$$

$P(2 < X \leq 3; 1 < Y \leq 4) =$

3.3 Função de distribuição marginal

- **Funções de distribuição marginais -** cada variável é considerada de forma isolada

$$P(X \leq x) = P(X \leq x,\ Y < +\infty) = \lim_{y \to +\infty} F(x, y) = F(x, +\infty) = F_X(x);$$

$$P(Y \leq y) = P(X < +\infty,\ Y \leq y) = \lim_{x \to +\infty} F(x, y) = F(+\infty, y) = F_Y(y).$$

A distribuição conjunta determina univocamente as distribuições marginais, **mas a inversa não é verdadeira**.

- **Exemplo**:

Calcular as funções de distribuição marginais do exemplo anterior

Calcule também:

$P(X \leq 4) =$ $\qquad\qquad$ $P(Y > 2) =$

3.4 Independência de variáveis aleatórias multivariadas

• **Definição – Variáveis aleatórias independentes**

Sejam B_1 e B_2 dois acontecimentos quaisquer tais que B_1 só envolve X e B_2 apenas se refere a Y. As variáveis aleatórias X e Y são independentes se e só se,

$$P(X \in B_1, Y \in B_2) = P(X \in B_1) \times P(Y \in B_2).$$

De forma equivalente: $F(x, y) = F_X(x) \times F_Y(y)$ para todo o par (x, y)

• **Teorema** – Se X e Y são variáveis aleatórias independentes e se ψ e φ são duas funções então as variáveis aleatórias $U = \psi(X)$, $V = \varphi(Y)$ são também independentes.

5. Variáveis aleatórias discretas multivariadas

- **Definição – Variável aleatória bidimensional discreta**
  (X, Y) é variável aleatória bidimensional discreta se e só se X e Y são variáveis aleatórias discretas.

  Dado $\mathrm{D} = \{(\mathrm{x}, \mathrm{y}) \colon P(X = \mathrm{x},\ Y = \mathrm{y}) > 0\}$, tem-se $P\,[(X,\ Y) \in \mathrm{D}] = 1$.

- **Definição – Função probabilidade conjunta**
  Seja (X, Y) uma variável aleatória bidimensional discreta. A função real de duas variáveis reais, com domínio $\Re^2$, definida por,
  $$f_{X,Y}(x, y) = P(X = x, Y = y)$$
  é a função probabilidade de (X, Y) ou a função probabilidade conjunta das variáveis X e Y.

Propriedades da função probabilidade conjunta:

- $f_{X,Y}(x,y) \geq 0$
- $\sum_{(x,y)\in D} f_{X,Y}(x,y) = 1$
- $P[(X,Y) \in B] = \sum_{(x,y)\in D\cap B} f_{X,Y}(x,y)$
- $F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$

$$= \sum_{s\leq x} \sum_{t\leq y} f_{X,Y}(s,t) \;\; (-\infty < x < +\infty, -\infty < y < +\infty)$$

- **Definição - Função probabilidade marginal**

Considere a variável aleatória bidimensional (X, Y).

Sejam

$F_X \rightarrow$ função de distribuição marginal de X

$D_X = \{x: P(X = x) = F_X(x) - F_X(x^-) > 0\}$, conjunto dos pontos de descontinuidade de F_X

Função probabilidade marginal de X: .

Facilmente se verifica que

$$f_X(x) = \begin{cases} P(X = x) = \sum_{y \in D_Y} f_{X,Y}(x, y) & x \in D_X \\ 0 & x \notin D_X \end{cases}$$

Teorema - As v.a. discretas X e Y são independentes se e só se

$$f_{X,Y}(x, y) = f_X(x) \times f_Y(y) \quad \text{para todo o par } (x, y).$$

Exemplo 3.1 - Lançamento de dois dados. Sejam ***X***: número de pontos obtido com o primeiro dado e ***Y***: o número máximo de pontos obtido no conjunto dos dois dados.

Por exemplo: se sair (1,3) tem-se $X = 1$, $Y = 3$; se sair (4,1) obtém-se $X = 4$, $Y = 4$.

Função probabilidade bidimensional

$X \backslash Y$	1	2	3	4	5	6	$f_X(x)$
1	1/36	1/36	1/36	1/36	1/36	1/36	6/36
2	0	2/36	1/36	1/36	1/36	1/36	6/36
3	0	0	3/36	1/36	1/36	1/36	6/36
4	0	0	0	4/36	1/36	1/36	6/36
5	0	0	0	0	5/36	1/36	6/36
6	0	0	0	0	0	6/36	6/36
$f_Y(y)$	1/36	3/36	5/36	7/36	9/36	11/36	1

- Funções probabilidade marginais. Por exemplo,

$$f_X(4) = P(X = 4) = f(4,4) + f(4,5) + f(4,6) = 6/36 = 1/6\,;$$

$$f_Y(5) = P(Y = 5) = f(1,5) + f(2,5) + f(3,5) + f(4,5) + f(5,5) = 9/36 = 1/4.$$

- As variáveis aleatórias X e Y não são independentes. Por exemplo,

$$f_{X,Y}(4,5) = 1/36 \neq f_X(4)\, f_Y(5) = (6/36)\times(9/36) = 1/24.$$

- As distribuições marginais **não** definem a distribuição conjunta. Poder-se-ia definir outra distribuição conjunta através de $h(x,\, y) = f_X(x)\ f_Y(y)$.
- Cálculo de probabilidades:

$$P(2 < X < 5, Y > 2) =$$

$$P(X < 3) =$$

$$P(Y > 4) =$$

3.6 Variáveis aleatórias contínuas multivariadas

Definição – Variável aleatória bidimensional contínua

A variável aleatória (X,Y), com função de distribuição $F_{X,Y}$, é uma variável aleatória bidimensional contínua se e só se, existe uma função real de duas variáveis reais, não negativa, $f_{X,Y}$, tal que,

$$F_{X,Y}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(u,v)dudv$$

(X,Y) é uma variável aleatória bidimensional contínua se e só se, X e Y são variáveis aleatórias contínuas.

- **Definição – Função densidade conjunta**

  A função $f_{X,Y}$, introduzida na definição anterior, chama-se função densidade (de probabilidade) de (X,Y) ou função densidade conjunta de X e Y.

- **Propriedades da função densidade conjunta**

  - Propriedade 3 das funções de distribuição conjuntas, $F_{X,Y}(+\infty,+\infty)=1$. Logo
  $$F_{X,Y}(+\infty,+\infty)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dxdy=1$$

  - Se $f_{X,Y}(x,y)$ for contínua no ponto (x,y) então $f_{X,Y}(x,y)=\frac{\partial^2 F(x,y)}{\partial x \partial y}$

  - Nos pontos em que não existe segunda derivada, convenciona-se $f_{X,Y}(x,y)=0$.

- Funções densidade marginais:

$f_X(x) = dF_X(x)/dx = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy$ já que
$F_X(x) = F_{X,Y}(x,+\infty) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{X,Y}(u,y)dydu$

$f_Y(y) = dF_Y(y)/dy = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)d$x já que
$F_Y(y) = F_{X,Y}(+\infty,y) = \int_{-\infty}^{y} \int_{-\infty}^{+\infty} f_{X,Y}(x,v)dxdv$

Teorema – As variáveis X e Y dizem-se independentes se e só se,

$f_{X,Y}(x,y) = f_X(x)\, f_Y(y)$ para todo o par (x, y).

3.7 Funções de distribuição condicionais

3.7.1 Variáveis aleatórias discretas

Definição - Função probabilidade condicionada

A função probabilidade de X condicionada pela realização do acontecimento $\{Y = y\}$, com $P(Y = y) > 0$, é dada por,

$$f_{X|Y=y}(x) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad (y \text{ fixo})$$

De modo análogo, se pode definir a função probabilidade de Y condicionada pela realização do acontecimento $\{X = x\}$, com $P(X = x) > 0$, através de,

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

- As funções probabilidade condicionadas gozam de todas as propriedades das f. probabilidade: $\sum_{x \in D_X} f_{X|Y=y}(x) = 1$ e $\sum_{y \in D_Y} f_{Y|X=x}(y) = 1$.

- Tal como para as funções de distribuição marginais também se podem calcular as funções de distribuição condicionadas a partir das funções probabilidade condicionadas:

$F_{X|Y=y}(x) = \sum_{x_i \leq x} f_{X|Y=y}(x_i)$, $\forall x \in \Re$, (y fixo)

e

$F_{Y|X=x}(y) = \sum_{y_j \leq y} f_{Y|X=x}(y_j)$, $\forall y \in \Re$, (x fixo).

Independência de variáveis aleatórias e funções probabilidade condicionadas.
$f_{X|Y=y}(x) = f_X(x)$ se $f_Y(y) > 0$; $f_{Y|X=x}(y) = f_Y(y)$ se $f_X(x) > 0$.

Exemplo 3.2 – Retome-se o exemplo 3.1
Calcule $f_{X|Y=3}(1)$, $f_{X|Y=3}(2)$ e $f_{X|Y=3}(3)$.
Calcule $f_{Y|X=4}(4)$, $f_{Y|X=4}(5)$, $f_{Y|X=4}(6)$

3.7.2 Variáveis aleatórias continuas

Definição – Função densidade condicionada

Seja $f_{XY}(x, y)$ a função densidade conjunta de X e Y. A função densidade de X condicionada por $Y = y$, com $f_Y(y) > 0$, é definida da seguinte maneira:

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad (y \text{ fixo}).$$

De forma análoga $f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$) x fixo).

A função densidade condicionada verifica todas as propriedades de uma função densidade de uma variável aleatória unidimensional.

- As funções de distribuição condicionadas podem ser calculadas a partir das funções densidade condicionadas:

$$F_{X|Y=y}(x) = \int_{-\infty}^{x} f_{X|Y=y}(u)\, du \text{ ,}(\forall x \in \Re) \text{ ,} y \text{ fixo}$$

e

$$F_{Y|X=x}(y) = \int_{-\infty}^{y} f_{Y|X=x}(v)\, dv \text{ ,}(\forall y \in \Re) \text{ ,} x \text{ fixo}$$

- Note-se que $P(X \in A|Y=y) = \int_A f_{X|Y=y}(x)dx$

  Por exemplo $P(a < X < b|Y=y) = \int_a^b f_{X|Y=y}(x)dx$

- Note-se também que $f_{X,Y}(x,y) = f_X(x)f_{Y|X=x}(y) = f_Y(y)f_{X|Y=y}(x)$.

Exemplo 3.3 – Considere-se a função densidade da variável bidimensional (X,Y) dada por

$$f_{X,Y}(x,y) = \begin{cases} (12/7)\ (x^2+xy) & \text{para } 0<x<1 \text{ e } 0<y<1, \\ 0 & \text{para outros } (x,y). \end{cases}$$

Calcule $f_{X|Y=y}(x)$, $f_{Y|X=x}(y)$ e $P(1/3 < X < 2/3|Y=1/3)$.

3.8 Valores esperados de funções de variáveis aleatórias multivariadas

Definição – Valor esperado de função de v.a. bidimensional

Se (X, Y) é uma variável aleatória bidimensional discreta com função probabilidade $f_{X,Y}$, e se ψ é uma função de (X, Y), a expressão

$$E\{\psi(X,Y)\} = \sum_{(x,y)\in D} \psi(x,y) f_{X,Y}(x,y)$$

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

é o valor esperado de $\psi(X, Y)$.

Tal como no caso unidimensional, tem de se verificar a condição de existência do valor esperado.

3.9 Valores esperados marginais

Note-se que a definição anterior aplica-se também ao caso em que $\psi(X,Y)=X$ e $\psi(X,Y)=Y$ e consequentemente

$E\{X\}=\sum_{(x,y)\in D} x f_{X,Y}(x,y)$ e $E\{Y\}=\sum_{(x,y)\in D} y f_{X,Y}(x,y)$

$E\{X\}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f_{X,Y}(x,y)dxdy$ e $E\{Y\}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f_{X,Y}(x,y)dxdy$
se estes valores esperados existirem

Teorema – Dois resultados importantes

Valor esperado de uma soma de v.a. – Se (X,Y) for uma variável aleatória bidimensional e se existirem $E(X)$ e $E(Y)$, então,

$$E(X+Y)=E(X)+E(Y).$$

- **Valor esperado de um produto de v.a. independentes –** Se X e Y forem variáveis aleatórias **independentes** e se existirem $E(X)$ e $E(Y)$, então,

$$E(XY)=E(X)\times E(Y).$$

Exemplo 3.4 – Seja (X, Y) uma v.a. bidimensional discreta com função probabilidade dada por

$X \setminus Y$	3	5	$f_X(x)$
1	0.1	0.3	0.4
2	0.4	0.2	0.6
$f_Y(y)$	0.5	0.5	1

Calcular $E(X) = \mu_X$, $E(Y) = \mu_Y$, $var(X) = \sigma_X^2$, $var(Y) = \sigma_Y^2$, $E(XY)$.

Exemplo 3.5 – Retome-se o exemplo 3.3,

$$f(x, y) = \frac{12}{7}(x^2 + xy) \qquad 0 < x < 1, \quad 0 < y < 1.$$

Calcular $E(X)$, $E(Y)$ e $E(XY)$.

3.10 Momentos em relação à origem

- **Definição – Momentos de ordem *r*+*s* em relação à origem -** Seja (X, Y) uma variável aleatória bidimensional. O valor esperado,

$$\mu'_{rs} = E(X^r Y^s),$$

define, se existir, um momento de ordem $r + s$ em relação à origem da v.a. (X, Y).

- Caso discreto: $\mu'_{rs} = E(X^r Y^s) = \sum_{(x,y)\in D} x^r y^s f_{X,Y}(x, y)$

- Caso contínuo: $\mu'_{rs} = E(X^r Y^s) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^r y^s f_{X,Y}(x, y) dxdy$

- Momentos mais importantes: $r + s = 1 \rightarrow E(X)$ e $E(Y)$

  $r + s = 2 \rightarrow E(X^2)$, $E(Y^2)$ e $E(XY)$

3.11 Momentos em relação à media

Definição – Momentos de ordem ($r + s$) em relação à média - Seja (X, Y) uma variável aleatória bidimensional. O valor esperado,

$$\mu_{rs} = E\{(X-\mu_X)^r (Y-\mu_Y)^s\},$$

define, se existir, um momento de ordem $r + s$ em relação à média da v.a. (X, Y).
Caso discreto:

$$\mu_{rs} = E\{(X-\mu_X)^r (Y-\mu_Y)^s\} = \sum_{(x,y)\in D} (x-\mu_X)^r (y-\mu_Y)^s f_{X,Y}(x,y)$$

- Caso contínuo:

$$\mu_{rs} = E\{(X-\mu_X)^r (Y-\mu_Y)^s\}$$

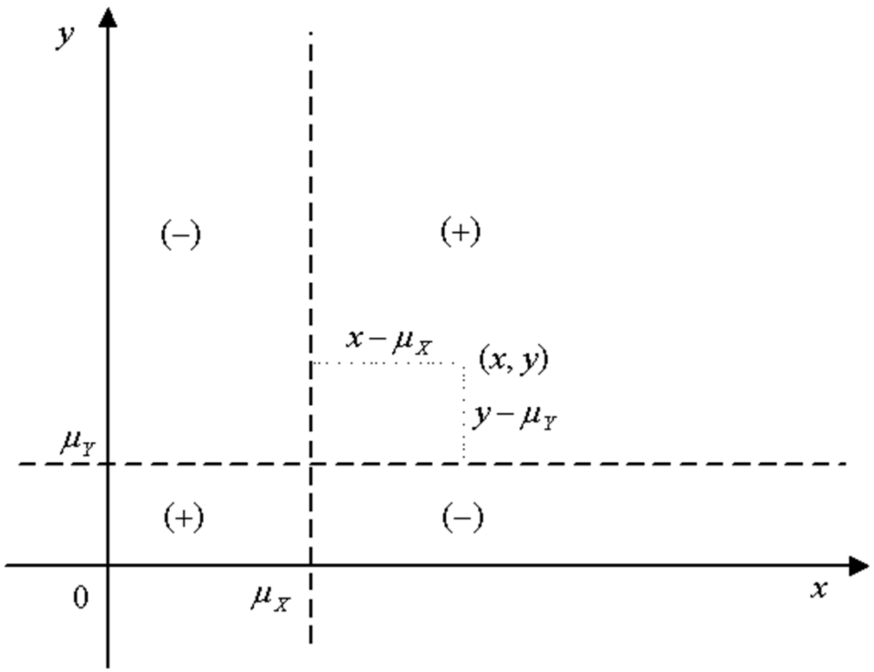
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-\mu_X)^r (y-\mu_Y)^s f_{X,Y}(x,y)\, dxdy$$

- Momentos mais importantes:
  - $r + s = 1 \rightarrow$ sempre nulos
  - $r + s = 2 \rightarrow var(X)$, $var(Y)$ e $cov(X,Y)$

3.12 A covariância

Definição– Covariância - A covariância entre as variáveis aleatórias X e Y é, $\text{cov}(X, Y) = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$, se este valor esperado existir.

Interpretação do conceito de covariância



$\text{cov}(X, Y) = E(XY) - E(X)\, E(Y)$ <u>Resultado importante</u>

Exemplos: retomar os exemplos anteriores e calcular a covariância.

Propriedades da covariância (a , b , c são 3 constantes arbitrárias e assumindo que os valores esperados existem)

- $\text{cov}(c, X) = E(cX) - E(c)\, E(X) = 0$
- $\text{cov}(aX, bY) = ab\,\text{cov}(X, Y)$
- Desigualdade de Cauchy–Schwarz: $\text{cov}(X,Y)^2 \leq \text{var}(X)\ \text{var}(Y)$

- Limitações do conceito de covariância na avaliação da associação entre v.a.:
  - Linearidade
  - Unidades de medida

3.13. O coeficiente de correlação

Definição - Coeficiente de correlação

- O coeficiente de correlação entre as variáveis aleatórias X e Y é dado por,

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}.$$

- Comentários:
  - Resolve o 2º problema
  - $-1 \leq \rho_{X,Y} \leq 1$. $\rho_{X,Y} = \pm 1$ quando e só quando existe relação linear entre as variáveis.
  - Se X e Y são independentes então $cov(X,Y) = 0$. **A recíproca não é verdadeira, isto é, covariância nula não implica independência**.

3.14 Momentos de funções lineares de variáveis aleatórias

- **Teorema –** Se existem segundos momentos para as v.a. X e Y, então

$$\mathrm{var}(aX + bY) = a^2\mathrm{var}(X) + 2ab\,\mathrm{cov}(X, Y) + b^2\mathrm{var}(Y),$$

  para quaisquer constantes a e b.

- Se a=b=1, $\mathrm{var}(X + Y) = \mathrm{var}(X) + 2\mathrm{cov}(X, Y) + \mathrm{var}(Y)$,
- Se a=1 e b=-1, $\mathrm{var}(X - Y) = \mathrm{var}(X) - 2\mathrm{cov}(X, Y) + \mathrm{var}(Y)$,
- Se as variáveis são independentes,

$$\mathrm{var}(aX + bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y),$$

- A covariância é um conceito particularmente importante na construção de carteiras de ativos financeiros

- **Exemplo** – Considerem-se as v.a. X e Y, com função probabilidade conjunta

$X \setminus Y$	–1	0	1	$f_X(x)$
–1	0.00	0.25	0.00	0.25
0	0.25	0.00	0.25	0.50
1	0.00	0.25	0.00	0.25
$f_Y(y)$	0.25	0.50	0.25	1.00

Mostrar que a covariância é nula mas que as variáveis aleatórias não são independentes.

15. Expectativas Condicionais

- A ideia é semelhante ao que se viu anteriormente, mas substituindo a f. probabilidade (densidade) pela função probabilidade (densidade) condicionada.

- **Definição** (expectativas condicionais ou valor esperado condicionado) – Seja a v.a. $Z = \psi(X, Y)$, função das variáveis aleatórias discretas X e Y. O valor esperado de $Z = \psi(X, Y)$ condicionado por $X = x$, é definido por

$$E(Z|X = x) = E\{\psi(X, Y)|X = x\} = \sum_{y \in D_Y} \psi(x, y) f_{Y|X=x}(y),$$

se se verificar a condição habitual de existência do valor esperado.

- De modo análogo, $E(Z|Y = y) = E\{\psi(X, Y)|Y = y\} = \sum_{x \in D_X} \psi(x, y) f_{X|Y=y}(x)$.

- No caso contínuo substitui-se somatórios por integrais (e a função probabilidade condicionada pela função densidade condicionada)

Caso particular mais interessante:

- **Médias condicionais** $E(X|Y=y)=\mu_{X|Y}(y) \rightarrow \psi(X,Y)=X$ e $f_{X|Y=y}(x)$

  ou $E(Y|X=x)=\mu_{Y|X}(x) \rightarrow \psi(X,Y)=Y$ e $f_{Y|X=x}(y)$

Definição - Independência em média

A variável aleatória Y é independente em média da variável aleatória X se e só se,

$E(Y|X=x)=E(Y)$ qualquer que seja x.

A variável aleatória X é independente em média da variável aleatória Y se e só se,

$E(X|Y=y)=E(X)$ qualquer que seja y.

16. A lei das expectativas iteradas

- **A lei das expectativas iteradas (ou lei do valor esperado iterado) –** O valor esperado de $Z = \psi(X, Y)$, se existir, é igual ao valor esperado do seu valor esperado condicionado por X, isto é, $E(Z) = E_X[E_Z(Z|X)]$
- **Exemplo –** (função densidade do Exemplo 3.3)

$$f_{X,Y}(x, y) = \frac{12}{7}(x^2 + xy) \qquad 0 < x < 1, \quad 0 < y < 1.$$

Calcular $E(X)$ diretamente e pela regra do valor esperado iterado.

Definição – Variância condicionada

A variância de Y, condicionada por $X = x$, é definida por,

$$\mathrm{var}(Y|X = x) = E\{\ [Y - E(Y|X = x)]^{\,2}|X = x\}$$

Caso discreto: $\mathrm{var}(Y|X = x) = \sum_{y \in D_Y}[y - E(Y|X = x)]^{\,2} f_{Y|X=x}(y)$

Caso contínuo: $\mathrm{var}(Y|X = x) = \int_{-\infty}^{+\infty}[y - E(Y|X = x)]^{\,2} f_{Y|X=x}(y)dy$

Variância marginal e momentos condicionados

- **Teorema –** Se existe $E(Y^2)$ então, $\mathrm{var}(Y) = \mathrm{var}_X[E(Y|X)] + E_X[\mathrm{Var}(Y|X)]$
- **Teorema** – A covariância de X e Y é igual à covariância de X e o valor esperado de Y condicionado por X, $\mathrm{cov}(X, Y) = \mathrm{cov}[X, E(Y|X)]$.