Material didático preparado pelo professor Maurício Mota para a disciplina CC0311- Probabilidade III ministrado em 2022.2 e agora como bibliografia para a disciplina de Inferência em 2023.1.

# 1 Vetor Aleatório Bidimensional Contínuo .

## 1.1 Definição

A distribuição conjunta do vetor aleatório $(X, Y)$ é caracterizada por uma função, $f(x, y)$, definida $\mathbf{R} \times \mathbf{R}$ com valores reais satisfazendo:

(a) $f(x, y) \geq 0$ para todo par $(x, y) \in \mathbf{R} \times \mathbf{R}$.

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dy\, dx = 1$

(c) Seja $E$ um subconjunto do $\mathbf{R} \times \mathbf{R}$ então:

$$P(E) = \int_E \int f(x, y)\, dy\, dx.$$

(d) O conjunto $A = \{(x, y) \in \mathbf{R} \times \mathbf{R} \mid \mathbf{f(x, y) > 0}\}$ é chamado de suporte da densidade conjunta.

A condição $(b)$ nos diz que o volume sob a superfície representada por $f(x, y)$ é igual a 1. A relação $(c)$ fornece a probabilidade do evento $E$. Se $E = [a, b] \times [c, d]$ então

$$P(E) = \int_a^b \int_c^d f(x, y)\, dy\, dx.$$

## 1.2 Vetor Aleatório Bidimensional com Distribuição Uniforme.

Dizemos que $(X, Y)$ tem um distribuição uniforme em uma região $A$ se sua função de densidade conjunta de probabilidade for constante, isto é,

$$f(x, y) = \frac{1}{c}\ I_A(x, y),$$

com $c =$ área(A).

## 1.3 Distribuições Marginais

As marginais de $X$ e $Y$ são dadas por:

(a)
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ dy.$$

(b)
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\ dx.$$

.

## 1.4 Independência

As variáveis $X$ e $Y$ com densidade conjunta $f(x,y)$ e marginais $f_X(x)$ e $f_Y(y), \textit{respectivamente}$, se

$$f(x,y) = f_X(x)\ f_Y(y),\ \ \forall\ (x,y) \in \mathbf{R} \times \mathbf{R}.$$

Para mostrar a dependência basta encontrar um par $(x_0, y_0)$ tal que:

$$f(x_0, y_0) \neq f_X(x_0)\ f_Y(y_0).$$

## 1.5 Exemplo 1.

Seja $(X,Y)$ um vetor aleatório contínuo com densidade conjunta:

$$f(x,y) = xe^{-x(y+1)}\ \ I_{(0,\infty)}\ (x)\ I_{(0,\infty)}\ (y).$$

a. Qual o suporte da densidade?

b. Mostre que ela é uma legítima função densidade conjunta de $(X,Y)$?

c. Calcule $P(Y > X)$.

d. Identifique a marginal de $X$.

e. Identifique a marginal de $Y$.

f. $X$ e $Y$ são independentes?

**Solução:** O suporte de $(X,Y)$ é dada por:

$$A = (0,\infty) \times (0,\infty).$$

Vamos mostrar a primeira condição
Para $(x,y) \in A$ temos:

$$f(x,y) = xe^{-x(y+1)} > 0 \quad e \quad f(x,y) = 0 \ \text{ para } \ (x,y) \in A^c.$$

logo,

$$f(x,y) \geq 0.$$

Vamos provar a segunda condição:

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\; dy\; dx = \int_{0}^{\infty} \int_{0}^{\infty} \; xe^{-x(y+1)}\; dy\; dx.$$

Logo

$$I = \int_{0}^{\infty} \; e^{-x} \left[ \int_{0}^{\infty} \; xe^{-xy}\; dy \right] \; dx.$$

$$I = \int_{0}^{\infty} \; e^{-x} \; 1 \;\; dx = 1,$$

pois

$$\int_{0}^{\infty} \; xe^{-xy}\; dy = 1$$

que é a integral da exponencial de parâmetro $x$.
Vamos calcular A probabilidade do evento $A$ pedido:

$$P(A) = \int_{0}^{\infty} \int_{x}^{\infty} \; xe^{-x(y+1)}\; dy\; dx$$

$$P(A) = \int_{0}^{\infty} \; e^{-x} \;\; \left[ \int_{x}^{\infty} \; xe^{-xy}\; dy; \;\; \right] ; dx.$$

Note que:

$$\int_{x}^{\infty} \; xe^{-xy}\; dy; = -e^{-xy} \; \Big|_{x}^{\infty} = e^{-x^2}.$$

$$P(A) = \int_{0}^{\infty} \; e^{-x} \;\; e^{-x^2}\; dx = \int_{0}^{\infty} \;\;\; e^{-(x+x^2)}\; dx$$

```
> g=function(x) exp(-(x+x^2))
> I1=integrate(g,0,Inf)$value;I1
[1] 0.5456414
>
```

A marginal de $X$ é dada por:

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\; dy \\
&= \int_{0}^{\infty} xe^{-x(y+1)}\; dy \\
&= \int_{0}^{\infty} xe^{-xy} e^{-x}\; dy \\
&= e^{-x} \int_{0}^{\infty} xe^{-xy}\; dy \\
&= e^{-x}\; I_{(0,\infty)}\; (x),
\end{aligned}$$

visto que $\int_0^{\infty} \; xe^{-xy}\; dy = 1$ que é a integral no suporte da exponencial com parâmetro $\lambda = x$.
Assim

$$X \sim Exp(1).$$

Note que:

$$E(X) = V(X) = 1 \;\; ; \; E(X^2) = 2.$$

Vamos calcular a marginal de $Y$.

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx \\ &= \int_{0}^{\infty} xe^{-(y+1)x}\,dx \\ &= IGG(a=2, b=y+1, c=1) \\ &= \frac{\Gamma(2)}{(1+y)^2} \\ &= \frac{1}{(1+y)^2}\, I_{(0,\infty)}\,(y), \end{aligned}$$

Assim

$$Y \sim F(2,2).$$

Se $Y \sim F(m,n)$
temos que:

$$E(Y) = \frac{m}{n-2}, n > 2$$

$$V(Y) = 2 \times \left[\frac{n}{n-2}\right]^2 \times \frac{m+n-2}{m(n-4}, \; n > 4.$$

Para

$$Y \sim F(2,2)$$

o valor esperado e a variância de $Y$ não existem

$X$ e $Y$ são independentes?
temos que $f(1,1) = e^{-2}$, $f_X(1) = e^{-1}$ e $f_Y(1) = \dfrac{1}{4}$ e $f_X(1)f_Y(1) = \dfrac{e^{-1}}{4}$.
Assim,

$$f(1,1) \neq f_X(1) \times f_Y(1),$$

logo $X$ e $Y$ são dependentes.

## 1.6 Covariância e Correlação

Sejam $X$ e $Y$ variáveis aleatórias com momentos de segunda ordem finitos, isto é,
$E(X^2) < \infty$, $E(Y^2) < \infty$ e $E(XY) < \infty$. Para calcularmos a variância de $S = X + Y$ temos:

$$\begin{aligned} V(S) &= E(S^2) - E^2(S) \\ &= E((X+Y)^2) - [E(X+Y)]^2 \\ &= E(X^2 + Y^2 + 2XY) - [E(X) + E(Y)]^2 \\ &= E(X^2) + E(Y^2) + 2E(XY) - E^2(X) - E^2(Y) - 2E(X)E(Y) \\ &= E(X^2) - E^2(X) + E(Y^2) - E^2(Y) + 2\,[E(XY) - E(X)E(Y)] \\ &= V(X) + V(Y) + 2\,Cov(X,Y). \end{aligned}$$

Vamos definir covariância:

Sejam $X$ e $Y$ duas variáveis aleatórias definidas no mesmo espaço de probabilidade. A covariância entre $X$ e $Y$, denotada por $Cov(X,Y)$ definida por:

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

Dessa definição surge uma fórmula mais operacional:

$$Cov(X,Y) = E(XY) - E(X)E(Y).$$

**Prova:**

$$\begin{aligned}
Cov(X,Y) &= E\left[(X - E(X))(Y - E(Y))\right] \\
&= E\left[XY - XE(Y) - E(X)Y + E(X)E(Y)\right] \\
&= E(XY) - E(XE(Y)) - E(E(X)Y) + E[E(X)E(Y)] \\
&= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\
&= E(XY) - E(X)E(Y).
\end{aligned}$$

**Propriedades da Covariância:**

a. $Cov(X,Y) = Cov(Y,X)$.

**Prova:**

$$Cov(X,Y) = E(X \times Y) - E(X) \times E(Y) = E(Y \times X) - E(Y) \times E(X) = Cov(Y,X).$$

b. $Cov(X,X) = V(X)$.

**Prova:**

$$Cov(X,X) = E(X \times X) - E(X) \times E(X) = E(X^2) - E^2(X) = V(X).$$

c. $Cov(X,a) = 0$, $a$ constante.

**Prova:**

$$Cov(X,a) = E(X \times a) - E(X) \times E(a) = a \times E(X) - a \times E(X) = 0.$$

d. $Cov(aX,bY) = ab\, Cov(X,Y)$, $a$ e $b$ constantes. **Prova:**

$$Cov(aX,bY) = E(aX \times bY) - E(aX) \times E(bY) = ab[E(XY) - E(X)E(Y)] = ab\, Cov(X,Y).$$

e. $Cov(X,Y+Z) = Cov(X,Y) + Cov(X,Z)$.

**Prova:**

$$Cov(X,Y+Z) = E[X \times (Y+Z)] - E(X) \times E(Y+Z),$$

$$Cov(X,Y+Z) = E[XY + XZ] - [E(X)E(Y) + E(X)E(Z)]$$
$$Cov(X,Y+Z) = E[XY] - E(X)E(Y) + E(XZ) - E(X)E(Z) = Cov(X,Y) + Cov(X,Z).$$

f.

$$Cov\left(\sum_{i=1}^{n} a_i\ X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ Cov(X_i, Y_j).$$

**Prova:** Sejam $U = \sum_{i=1}^{n} a_i\ X_i$ e $V = \sum_{j=1}^{m} b_j Y_j$. Logo

$$UV = \sum_{i=1}^{n} a_i\ X_i \sum_{j=1}^{m} b_j Y_j = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ X_i\ Y_j.$$

Logo,

$$E(UV) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ E[X_i\ Y_j].$$

Por outro lado $E[U] = \sum_{i=1}^{n} a_i\ E(X_i)$ e $E[V] = \sum_{j=1}^{m} b_j E[Y_j]$.

Mas,

$$E(U).E(V) = \sum_{i=1}^{n} a_i\ E(X_i) \times \sum_{j=1}^{m} b_j E[Y_j] = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ E(X_i)E(Y_j).$$

Logo,

$$Cov(U, V) = E(UV) - E(U)E(V) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ E[X_i\ Y_j] - \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ E(X_i)E(Y_j).$$

$$Cov(U, V) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ \left(E[X_i\ Y_j] - E(X_i)E(Y_j)\right).$$

Finalmente

$$Cov(U, V) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\ b_j\ Cov(X_i, Y_j).$$

g.

$$V(S_n) = V\left(\sum_{i=1}^{n}\ X_i\right) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i=1}^{n-1}\ \sum_{j=i+1}^{n}\ Cov(X_i, X_j).$$

**Prova:**

$$\begin{aligned}
V(S_n) &= Cov(S_n, S_n) \\
&= Cov\left(\sum_{i=1}^{n}\ X_i, \sum_{j=1}^{n}\ X_j\right) \\
&= \sum_{i=1}^{n}\ \sum_{j=1}^{n}\ Cov(X_i, X_j) \\
&= \sum_{i=1}^{n}\ V(X_i) + \sum_{i=1}^{n}\ \sum_{j=1\ j\neq i}^{n}\ Cov(X_i, X_j) \\
&= \sum_{i=1}^{n}\ V(X_i) + 2\sum_{i=1}^{n-1}\ \sum_{j=i+1}^{n}\ Cov(X_i, X_j),
\end{aligned}$$

pois a covariância é simétrica.

A covariância entre $X$ e $Y$ mede o grau da associação linear das variáveis e é expresso nas unidades medidas das variáveis. Uma medida de associação linear adimensional é o coeficiente de correlação que é definido por:

$$\rho = Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}.$$

**Fato 1:** Se $X$ e $Y$ forem independentes a correlação é nula.
**Prova:**

$$\begin{aligned} Cov(X,Y) &= E(XY) - E(X)E(Y) \\ &= E(X) \times E(Y) - E(X) \times E(Y) \\ &= 0. \end{aligned}$$

Como

$$\rho = Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}} = 0.$$

**Fato 2:** $Cov(X,Y) = 0$ não implica independência. Neste caso as variáveis são ditas não correlacionadas.

**Fato 3:** O coeficiente de correlação varia no intervalo $[-1,1]$.

**Prova:** Sejam $X$ uma variável com média $\mu_1$ e variância $\sigma_1^2$ e $Y$ uma variável com média $\mu_2$ e variância $\sigma_2^2$. Seja $\rho$ o coeficiente de correlação entre $X$ e $Y$.

Considere

$$U = X - E(X) \ \ e \ \ V = Y - E(Y).$$

Logo

$$E(U) = E(V) = 0, \ E(U^2) = Var(U) = Var(X),$$

$$E(V^2) = Var(V) = Var(Y) \ \ e \ \ E(UV) = E\left[(= X - E(X))(Y - E(Y))\right] = Cov(X,Y).$$

Considere a função

$$g(t) = E\left[E(tU + V)^2\right] \geq 0 = E(U^2)t^2 + 2E(UV)t + E(V^2).$$

Para uma função quadrática com coeficiente do segundo grau positivo ser não negativa é preciso que seu discriminante seja menor ou igual a zero. Assim,

$$\begin{aligned} \Delta &= 4E^2(UV) - 4E(U^2)E(V^2) \leq 0 \\ &= E^2(UV) - E(U^2)E(V^2) \leq 0. \end{aligned}$$

Assim

$$E^2(UV) \leq E(U^2)E(V^2).$$

Esta é a famosa desigualdade de Cauchy-Schwarz.

$$[Cov(X,Y)]^2 \leq V(X)V(Y),$$

extraindo a raiz quadrada

$$|Cov(X,Y)| \leq \sqrt{V(X)V(Y)},$$

e finalmente,

$$|\rho| = \frac{|Cov(X,Y)|}{\sqrt{V(X)V(Y)}} \leq 1.$$

## 1.7 Momentos

Seja $(X, Y)$ um vetor aleatório com $E(X) = \mu_1$, $E(Y) = \mu_2$, $Var(X) = \sigma_1^2$, $Var(Y) = \sigma_2^2$ e covariância, $\sigma_{12} = \rho\, \sigma_1 \sigma_2$. O vetor de médias é definido por:

$$\boldsymbol{\mu} = [\mu_1, \mu_2]^\top.$$

A matriz de variâncias-covariâncias, $\Sigma$, é dada por:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix},$$

ou

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

O traço da matriz $\Sigma$ é chamado de Variância Total, isto é,

$$VT = tra(\Sigma) = \sigma_{11} + \sigma_{22} = \sigma_1^2 + \sigma_2^2.$$

O Determinante da matriz $\Sigma$ é chamado de Variância Generalizada , isto é,

$$VG = det(\Sigma) = \sigma_1^2 \times \sigma_2^2 - \rho^2 \times \sigma_1^2 \times \sigma_2^2 = (1 - \rho^2) \times \sigma_1^2 \times \sigma_2^2.$$

Para calcular a covariância entre $X$ e $Y$ precisamos calcular a $E(XY)$. Vamos generalizar para calcular a esperança da função real, $h(X, Y)$, que é definida por:

$$E[h(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)\, f(x, y)\, dy\, dx.$$

Vamos calcular a esperança de $XY$ no exemplo 1:

$$E(XY) = \int_{0}^{\infty} \int_{0y}^{\infty} xy\, xe^{-x(y+1)}\, dy\, dx.$$

$$E(XY) = \int_{0}^{\infty} x\, e^{-x} \left[ \int_{0}^{\infty} y\, xe^{-xy}\, dy \right] dx$$

Mas,

$$\int_{0}^{\infty} y\, xe^{-xy}\, dy = \frac{1}{x},$$

que é a esperança de uma exponencial de parâmetro $\lambda = x$.
Logo,

$$E(XY) = \int_{0}^{\infty} x\, e^{-x}\, \frac{1}{x}\, dx = \int_{0}^{\infty} e^{-x}\, dx = 1.$$

# 2 Distribuição Condicional.

## 2.1 Definição

Seja $(X, Y)$ uma variável aleatória bidimensional contínua (v.a.c) com função densidade de probabilidade conjunta dada por $f(x, y)$ e marginais $f_X(x)$ e $f_Y(y)$. Seja $x$ um ponto do suporte de $X$. A distribuição condicional de $Y|X = x$ é por definição:

$$f_{Y|X=x}(y|x) = \frac{f(x,y)}{f_X(x)}. \tag{1}$$

Verifique que a definição 1 é uma legítima distribuição de probabilidade.

Devemos provar que:

**i)** $f_{Y|X=x}(y|x) \geq 0$

**ii)** $\int_{-\infty}^{\infty} f_{Y|X=x}(y|x)dy = 1$

A propriedade $(i)$ é satisfeita pois no suporte temos $f(x, y) > 0$ e $f_Y(y) > 0$ e portanto temos uma razão positiva. Fora do suporte temos $f(x, y) = 0$ e $f_Y(y) > 0$ logo temos uma razão nula. Assim a condição $(i)$ é satisfeita.

Vamos provar agora a condição $(ii)$

$$\begin{aligned}
\int_{-\infty}^{\infty} f_{Y|X=x}(y|x)dy &= \int_{-\infty}^{\infty} \frac{f(x,y)}{f_X(x)}\, dy \\
&= \frac{1}{f_X(x)} \int_{-\infty}^{\infty} f(x,y)\, dy \\
&= \frac{1}{f_X(x)}\, f_X(x) \\
&= 1.
\end{aligned}$$

## 2.2 Continuação do Exemplo 1

Calcule as distribuições condicionais do Exemplo 1.

A condicional de $Y|X = x$ é dada por:

$$\begin{aligned}
f_{Y|X=x}(y|x) &= \frac{f(x,y)}{f_X(x)} \\
. &= \frac{x\, e^{-x(1+y)}}{e^{-x}} \\
&= x\, e^{-xy}\, I_{(0,\infty)}\, (y).
\end{aligned}$$

Assim,

$$Y|X = x \sim Exp(x).$$

A condicional de $X|Y = y$ é dada por:

$$\begin{aligned}
f_{X|Y=y}(y|x) &= \frac{f(x,y)}{f_Y(y)} \\
. &= \frac{x\, e^{-x(1+y)}}{(1+y)^{-2}} \\
&= (1+y)^2 x\, e^{-(1+y)x}\, I_{(0,\infty)}\, (y).
\end{aligned}$$

Assim,

$$X|Y = y \sim Gama(r = 2, \lambda = (1 + y)).$$

## 2.3 Exemplo 2

Seja $(X, Y)$ com distribuição uniforme na região dada por:

$$A = \{(x, y) \mid 0 \leq x < y \leq 2\}.$$

Um esboço gráfico do suporte é apresentado a seguir:

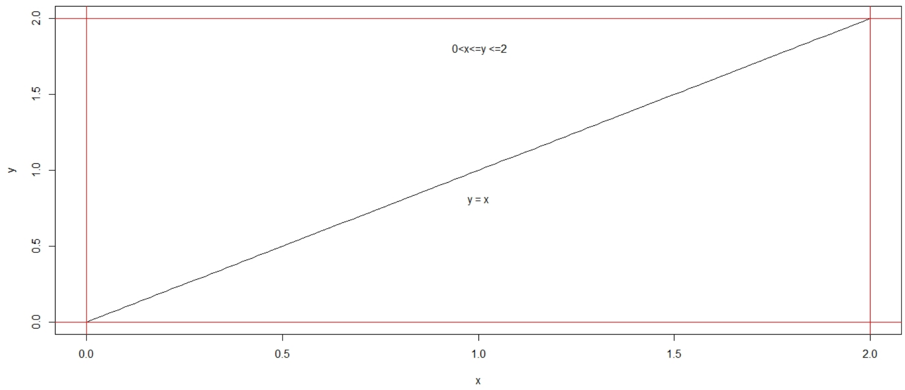


Figura 1:

A região dada é o triângulo com vértices nos pontos $O(0, 0)$, $B(0, 2)$ e $C(2, 2)$ com área igual a 2. Assim a densidade conjunta de $(X, Y)$ é dada por:

$$f(x, y) = \frac{1}{2}\ I_{[0,2]}\ (x)\ I_{[x,2]}\ (y).$$

A marginal de $X$ é dada por:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\ dy = \int_x^2 \frac{1}{2}\ dy = \frac{2-x}{2}\ I_{[0,2]}\ (x),$$

que é a densidade da triangular com $a = 0$, $b = 2$ e $c = 0$.
Assim

$$E(X) = \frac{2}{3}\ ; V(X) = \frac{2}{9} \quad e \quad E(X^2) = \frac{2}{3}.$$

.

A distribuição condicional de $Y|X = x$ é dada por:

$$f_{Y|X=x}(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{\frac{1}{2}}{\frac{2-x}{2}}\ I_{(x,2)}\ (y) = \frac{1}{2-x}\ I_{(x,2)}\ (y). \qquad (2)$$

Dizemos que a condicional de $Y|X = x$ tem distribuição Uniforme com parâmetros $a = x$ e $b = 2$. E no caso geral dizemos

$$Y|X \sim U(X, 2).$$

Vamos calcular a marginal de $Y$. Inicialmente vamos escrever a conjunta de outra maneira:

$$f(x, y) = \frac{1}{2}\ I_{[0,2]}\ (y)\ I_{[0,y]}\ (x).$$

A marginal de $Y$ é dada por:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\ dx = \int_0^y \frac{1}{2}\ dy = \frac{y}{2}\ I_{[0,2]}\ (y),$$

que é a densidade de uma triangular com $a = 0$, $b = 2$, $c = 2$.
Assim,

$$E(Y) = \frac{a+b+c}{3} = \frac{4}{3},$$

$$V(Y) = \frac{a^2+b^2+c^2-ab-ac-bc}{18} = \frac{2}{9},$$

e

$$E(Y^2) = V(Y) + E^2(Y) = \frac{2}{9} + \frac{16}{9} = 2.$$

A distribuição condicional de $X|Y = y$ é dada por:

$$f_{X|Y=y}(y|x) = \frac{f(x,y)}{f_Y(y)} = \frac{\frac{1}{2}}{\frac{y}{2}}\ I_{(0,y)}\ (x) = \frac{1}{y}\ I_{(0,y)}\ (x). \tag{3}$$

Dizemos que a condicional de $X|Y = y$ tem distribuição Uniforme com parâmetros $a = 0$ e $b = y$.
E no caso geral dizemos

$$X|Y \sim U(0, Y).$$

Vamos definir agora os momentos em relação à origem da distribuição condicional.

# 3 Momentos Condicionais em Relação à Origem

$$E(Y^r|X = x) = \int_{-\infty}^{\infty} y^r\ \ f_{Y|X=x}(y|x)\ dy,\ \ r = 1, 2, 3, 4, \ldots. \tag{4}$$

Observe que $E(Y^r|X = x)$ é uma função de $x$.

.
Calcule o primeiro momento em relação à origem da condicional do exemplo 2.

.
Como a distribuição é uniforme temos:

$$E(Y|X = x) = \frac{2+x}{2}.$$

Assim

$$V = E(Y|X) = \frac{X+2}{2} = h(X),$$

é também uma variável aleatória contínua.

.
Vamos achar a lei de $V$, sua média e variância.
Seja $G(v)$ a acumulada de $V$. Logo

$$G(v) = P(V \leq v) = P\left(\frac{X+2}{2} \leq v\right) = P(X \leq 2v - 2) = F_X(2v - 2).$$

A f.d.p. de $V$ é

$$g(v) = 2f_X(2v-2) = 2\frac{(2-2v+2)}{2}\quad I_{(0,2)}\ (2v-2) = 2(2-v)\ I_{(1,2)}\ (v),$$

assim,

$$V \sim triangular(a = 1, b = 2, c = 1)$$

Logo ,

$$E(V) = E[E(Y|X)] = \frac{1+2+1}{3} = \frac{4}{3} = E(Y).$$

A variância de V é dada por:

$$Var(V) = Var\left(E(Y|X)\right) = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18} = \frac{2}{18}.$$

Note que $Y|X \sim U(X, 2)$, assim

$$Var(Y|X) = \frac{(2-X)^2}{12} = \frac{(X-2)^2}{12},$$

e

$$E[Var(Y|X)] = \frac{E[(X-2)^2]}{12} = \frac{E(X^2) - 4E(X) + 4}{12} = \frac{1}{6}.$$

Uma fórmula bastante útil nos diz que:

$$V(Y) = E[Var(Y|X)] + V[E(Y|X)] = \frac{1}{6} + \frac{1}{18} = \frac{2}{9}.$$

O segundo momento em relação à origem da condicional do exemplo 2

$$\begin{aligned}
E(Y^2|X = x) &= \int_x^2 y^2 \; \frac{1}{2-x} \; dy \\
&= \frac{1}{2-x} \int_x^2 y^2 \; dy \\
&= \frac{1}{2-x} \; \frac{y^3}{3} \; \Big|_x^2 \\
&= \frac{1}{2-x} \; \frac{8-x^3}{3} \\
&= \frac{1}{2-x} \; \frac{(2-x)(4+2x+x^2)}{3} \\
&= \frac{4+2x+x^2}{3}.
\end{aligned}$$

Note que

$$E(Y^2|X) = \frac{4 + 2X + X^2}{3}.$$

Assim,

$$E(Y^2) = E\left(E[(Y^2|X)]\right) = \frac{4 + 2E(X) + E(X^2)}{3} = \frac{4 + 2\,\frac{2}{3} + \frac{2}{3}}{3} = 2.$$

## 3.1 Variância Condicional

A variância condicional de $Y|X = x$

$$V(Y|X = x) = \int_{-\infty}^{\infty} (y - E(Y|X = x))^2 f(y|x)\, dy,$$

que também pode calculada como:

$$V(Y|X) = E[Y^2|X] - E^2\, [Y|X].$$

Vamos mostrar agora as propriedades apresentadas:

**Fato 5:** $E(E(Y|X)) = E(Y)$.

**Prova:** Sabemos que $E(Y|X) = h(X)$. Assim,

$$\begin{aligned}
E[E(Y|X)] &= \int_{-\infty}^{\infty} E(Y|X = x)\ \ f_X(x)\, dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\ f(y|x)\ \ f_X(x)\, dydx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\ f(x, y)\, dydx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\ f(x, y)\, dxdy \\
&= \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x, y)\, dx\right] dy \\
&= \int_{-\infty}^{\infty} y\ f_Y(y)\, dy \\
&= E(Y).
\end{aligned}$$

**Fato 6:** $V(Y) = V[E(Y|X)] + E[V(Y|X)]$.

Prova: Vamos mostrar que $V[E(Y|X)] = V(Y) - E[V(Y|X)]$.

$$\begin{aligned}
E\left[V(Y|X)\right] &= E\left[E(Y^2|X) - E^2[Y|X]\right] \\
&= E\left[E(Y^2|X)\right] - E\left[E^2[Y|X]\right] \\
&= E(Y^2) - E\left[E^2[Y|X]\right] \\
&= V(Y) + E^2(Y) - E\left[E^2[Y|X]\right] \\
&= V(Y) - \left[E\left[E^2[Y|X]\right] - [E(Y)]^2\right] \\
&= V(Y) - \left[E\left[E^2[Y|X]\right] - [E(E(Y|X)]^2\right] \\
&= V(Y) - V(E(Y|X)),
\end{aligned}$$

**Fato 7:** $Cov(X, Y) = Cov\left(X, E(Y|X)\right)$.

Prova:

$$\begin{aligned} Cov\left(X, E(Y|X)\right) &= E\left[X\ E(Y|X)\right] - E(X).E[E(Y|X)] \\ &= E\left[E(XY|X)\right] - E(X).E(Y) \\ &= E(XY) - E(X).E(Y) \\ &= Cov(X,Y). \end{aligned}$$

**Fato 8:** Seja $(X, Y)$ um vetor aleatório bidimensional contínuo e $h(X, Y)$ uma função unidimensional das duas variáveis. A esperança condicional de $h(X, Y)$ dado $X = x$ é definida por:

$$E\left[h(X,Y)|X=x\right] = \int_{-\infty}^{\infty} h(x,y) f_{Y|X=x}(y|x)\ \ dy.$$

**Fato 9:** Seja $(X, Y)$ um vetor aleatório bidimensional contínuo. Seja $g(Y)$ uma função real de $Y$ com esperança finita.

Assim

$$E[g(Y)] = E\left[E(g(Y)|X)\right]. \tag{5}$$

Em particular,

$$E[Y] = E\left[E(Y|X)\right]. \tag{6}$$

**Fato 10:** A esperança condicional $\mu_{Y|x} = E(Y|X = x)$ é chamada de curva de regressão de $Y$ em $x$.

**Fato 11:** Se a esperança condicional $\mu_{Y|x} = E(Y|X = x) = a + bx$ então:
então:

$$b = \frac{Cov(X,Y)}{V(X)} = \rho\ \frac{\sigma_Y}{\sigma_X}\ \ e\ \ a = E(Y) - bE(X).$$

Prova: Como $E(Y|X) = a + bX$ temos que:

$$E(Y) = E[E(Y|X)] = E(a + bX) = a + bE(X),$$

e portanto

$$a = E(Y) - bE(X).$$

Multiplicando por $X$ temos

$$XE(Y|X) = E(XY|X) = aX + bX^2.$$

Aplicando o operador esperança temos:

$$E(XY) = E(E(XY|X) = aE(X) + bE(X^2).$$

Logo,

$$E(XY) = [E(Y) - bE(X)]E(X) + bE(X^2) = E(X)E(Y) + b[E(X^2) - E^2(X)],$$

assim,

$$Cov(X,Y) = E(XY) - E(X)E(Y) = bV(X),$$

e portanto

$$b = \frac{Cov(X,Y)}{V(X)} = \frac{\rho\sigma_X\sigma_Y}{\sigma_X^2} = \rho\,\frac{\sigma_Y}{\sigma_X}.$$

**Fato 12:** Se a esperança condicional $\mu_{Y|x} = E(Y|X=x) = a + bx$ e $\mu_{X|y} = E(X|Y=y) = c + dy$ então:

$$bd = \rho^2,$$

e $\rho$ tem o sinal comum de $b$ e $d$.

Prova:

$$\begin{aligned} bd &= \rho\,\frac{\sigma_Y}{\sigma_X} \times \rho\,\frac{\sigma_X}{\sigma_Y} \\ &= \rho^2. \end{aligned}$$

**Fato 13:** Seja $(X,Y)$ um vetor aleatório bidimensional contínuo. Sejam $g_1(.)$ e $g_2(.)$ funções de uma única variável. Então:

a.
$$E[(g_1(Y) + g_1(Y))|X=x] = E[g_1(Y)|X=x] + E[g_2(Y)|X=x].$$

b.
$$E\left[g_1(Y) \times g_2(X)|X=x\right] = g_2(x) \times E[g_1(Y)|X=x].$$

**Fato 14:** Seja $(X,Y)$ um vetor aleatório bidimensional contínuo. Seja $M_Y(t)$ a função geradora de momentos de $Y$. Então:

$$M_Y(t) = E(e^{tY}) = E\left[E(e^{tY}|X)\right] = E\left[M_{Y|X=x}(t)\right].$$

**Fato 15:** Seja $(X,Y)$ um vetor aleatório bidimensional contínuo. Seja $M_{X,Y}(t_1,t_2)$ a função geradora bivariada de momentos de $(X,Y)$. Então:

$$M_{(X,Y)}(t_1,t_2) = E(e^{t_1X + t_2Y}) = \int_{-\infty}^{\infty} e^{t_1x + t_2y} f(x,y)\, dydx.$$

Além disso

$$M_X(t) = M_{(X,Y)}(t,0) \;\; eM_Y(t) = M_{(X,Y)}(0,t) \;\; .$$

# 4 Previsão

Vamos apresentar a seção 7.6 da oitava edição do livro do Sheldon Ross- Probabilidade -Um curso moderno com aplicações.

Às vezes surge uma situação em que o valor de uma variável aleatória $X$ é observado e então, com base no valor observado, tenta-se prever o valor de uma segunda variável aleatória $Y$ . Suponha que $g(X)$ represente o preditor; isto é, se $X$ é observado como sendo igual a $x$, então $g(x)$ nos fornece uma predição para o valor de $Y$. Claramente, queremos escolher $g$ de forma que $g(X)$ se aproxime de $Y$. Um possível critério é escolher $g$ de forma a minimizar

$$E\left[Y - g(X)\right]^2.$$

Mostraremos agora que, de acordo com esse critério, o melhor preditor de $Y$ é

$$g(X) = E(Y|X).$$

Proposição 6.1

$$E\left[Y - g(X)\right]^2 \geq E\left[\left[(Y - E(Y|X))\right]^2\right].$$

Demonstração. Vamos calcular inicialmente a $E\left[(Y - g(X))^2|X\right]$.

$$\begin{aligned}
E\left[(Y - g(X))^2|X\right] &= E\left[(Y - E(Y|X) + E(Y|X) - g(X))^2|X\right] \\
\\
&= E\left[(Y - E(Y|X)|X)^2\right] + \left[(E(Y|X) - g(X))^2|X\right] \\
&\quad +2E\left[(Y - E(Y|X)|X).(E(Y|X) - g(X))|X\right].
\end{aligned}$$

Como dado $X$ , $E(Y|X) - g(X)$ é constante temos que:

$$\begin{aligned}
E\left[(Y - E(Y|X)|X).(E(Y|X) - g(X))|X\right] &= (E(Y|X) - g(X)).E\left[(Y - E(Y|X)|X)\right] \\
&= (E(Y|X) - g(X)).\left[E(Y|X) - E(Y|X)\right] \\
&= 0.
\end{aligned}$$

Assim,

$$E\left[(Y - g(X))^2|X\right] = E\left[(Y - E(Y|X)|X)^2\right] + \left[(E(Y|X) - g(X))^2|X\right],$$

Como $\left[(E(Y|X) - g(X))^2|X\right] \geq 0$ temos que:

$$E\left[(Y - g(X))^2|X\right] \geq E\left[(Y - E(Y|X)|X)^2\right].$$

Vamos calcular a esperança em ambos os lados :

$$E\left[E\left[(Y - g(X))^2|X\right]\right] \geq E\left[E\left[(Y - E(Y|X)|X)^2\right]\right].$$

Logo,

$$E\left[Y - g(X)\right]^2 \geq E\left[\left[(Y - E(Y|X))\right]^2\right].$$

Observação: Um segundo argumento mais intuitivo porém menos rigoroso, que pode ser utilizado para verificar a proposição 6.1 é dado a seguir. É simples verificar que

$$E[(Y - c)^2],$$

é minimizado em $c = E(Y)$. Assim, se queremos predizer o valor de $Y$ quando não dispomos de dados para isso, a melhor predição possível , no sentido de minimizar o erro quadrático médio, é dizer que $Y$ será igual a sua média. Entretanto, se o valor da variável aleatória é observado como sendo $x$, então o problema da predição permanece exatamente igual ao caso anterior com a exceção de que todas as probabilidades e as esperanças estão agora condicionadas ao evento $X = x$. Com isso a melhor predição nesta situação é dizer que $Y$ será igual ao seu valor esperado condicional dado que $X = x$, o que estabelece a Proposição 6.1.

Exemplo 6a. Suponha que o filho de um homem com altura $x$, em cm, atinge uma altura que é normalmente distribuída com média $(x + 2,54)$ e variância 10,16 $cm^2$. Qual é a melhor predição da altura que o filho irá atingir se o seu pai tem 1,83 m de altura?

Solução: Formalmente, este modelo pode ser escrito como: Seja $X$, a altura do pai e $Y$, a altura do filho.

$$Y = X + 2,54 + e,$$

em que $e$ é uma variável aleatória, com média 0 e variância 10,16 e independente de $X$. A melhor predição $E[Y|X = 183]$ é portanto igual a

$$\begin{aligned}
E(Y|X = 183) &= E(X + 2,54 + e|X = 183) \\
&= E(X) + 2,54 + E(e|X) \\
&= E(X) + 2,54 + E(e), \ \ \textit{pela} \text{ independência} \\
&= 183 + 2,54 + 0 \\
&= 185,54\ cm
\end{aligned}$$

Exemplo 6b Suponha que, se um sinal com valor $s$ é enviado do ponto $A$, o valor recebido no ponto $B$ seja normalmente distribuído com parâmetros $(s, 1)$. Se $S$, o valor enviado em $A$, é normalmente distribuído com parâmetros $(\mu, \sigma^2)$, qual é a melhor estimativa para o sinal enviado se $R$, o valor recebido em $B$, é igual a $r$?

Solução.

Pelo enunciado temos que $S \sim N(\mu, \sigma^2)$ e $R|S = s \sim N(s, 1)$.

A f.d.p. de $S$ é dada por:

$$f_S(s) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-\dfrac{(s-\mu)^2}{2\sigma^2}},$$

A f.d.p. de $R|S = s$ é dada por:

$$f_{S|R=r}(s|r) = \frac{1}{\sqrt{2\pi}}\ e^{-\dfrac{(r-s)^2}{2}}.$$

A densidade conjunta de $(S, R)$ é dada por:

$$\begin{aligned}
f_{S,R}(r, s) &= f_S(s) f_{S|R=r}(s|r) \\
&= \frac{1}{2\pi\sigma}\ e^{-\dfrac{1}{2}\left[(r-s)^2 + \dfrac{(s-\mu)^2}{\sigma^2}\right]} \\
&= \frac{1}{2\pi\sigma}\ e^{-\dfrac{1}{2}\left[(s-r)^2 + \dfrac{(s-\mu)^2}{\sigma^2}\right]}
\end{aligned}$$

Vamos analisar a expressão

$$(s - r)^2 + \frac{(s-\mu)^2}{\sigma^2},$$

para completar um quadrado da forma $(s - c)^2$. Vamos utilizar o resultado:

( Livro Estatística Bayesiana, Paulino C.D. et all-pags 147 e 148)

Para completar os quadrados use a identidade:

$$d_1(z - c_1)^2 + d_2(z - c_2)^2 = (d_1 + d_2)(z - c)^2 + \frac{d_1 d_2}{d_1 + d_2}(c_1 - c_2)^2,$$

em que $c = \dfrac{d_1 c_1 + d_2 c_2}{d_1 + d_2}$.

Assim,

$$d_1 = 1,\ d_2 = \frac{1}{\sigma^2},\ \ c_1 = r \ e\ c_2 = \mu,$$

Logo,

$$c = \frac{r + \dfrac{\mu}{\sigma^2}}{1 + \dfrac{1}{\sigma^2}} = \frac{\mu + r\sigma^2}{1 + \sigma^2}.$$

$$d_1 + d_2 = 1 + \frac{1}{\sigma^2} = \frac{1 + \sigma^2}{\sigma^2}.$$

$$\frac{d_1 d_2}{d_1 + d_2} = \frac{1}{1 + \sigma^2},$$

e

$$(c_1 - c_2)^2 = (r - \mu)^2.$$

Note que o núcleo

$$(d_1 + d_2)(z - c)^2 = \frac{(z - c)2}{\dfrac{1}{d_1 + d_2}},$$

Como $1/2$ foi posto em evidência então é o núcleo de uma normal com média $c$ e variância

$$\sigma_*^2 \frac{1}{d_1 + d_2} = \frac{\sigma^2}{1 + \sigma^2}.$$

Mas,

$$\frac{d_1 d_2}{d_1 + d_2}(c_1 - c_2)^2 = \frac{r - \mu}{2(1 + \sigma^2)}.$$

Assim

$$f_{S,R}(r,s) \quad = \quad \frac{1}{2\pi\sigma} \ e^{-\dfrac{(r-\mu)^2}{2(1+\sigma^2)}} \ e^{-\dfrac{(s-c)^2}{2\sigma_*^2}}.$$

A marginal de $R$ é dada por:

$$\begin{aligned}
f_R(s) &= \int_{-\infty}^{\infty} f(r,s)\, ds \\
&= \frac{1}{\sqrt{2\pi(1+\sigma^2)}} \ e^{-\dfrac{(r-\mu)^2}{2(1+\sigma^2)}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1+\sigma^2}}{\sigma} \ e^{-\dfrac{(s-c)^2}{2\sigma_*^2}} \ ds \\
&= \frac{1}{\sqrt{2\pi(1+\sigma^2)}} \ e^{-\dfrac{(r-\mu)^2}{2(1+\sigma^2)}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\ \sigma_*} \ e^{-\dfrac{(s-c)^2}{2\sigma_*^2}} \ ds \\
&= \frac{1}{\sqrt{2\pi(1+\sigma^2)}} \ e^{-\dfrac{(r-\mu)^2}{2(1+\sigma^2)}}.
\end{aligned}$$

Logo, $R \sim N(\mu,\ (1 + \sigma^2))$
A distribuição condicional de $S|R = r$ é dada por:

$$f_{S|R=r}(s|r) = \frac{1}{\sqrt{2\pi}\,\sigma_*}\quad e^{-\dfrac{(s-c)^2}{2\sigma_*^2}},$$

que é normal com:

$$E(S|R=r) = \frac{\mu + r\sigma^2}{1+\sigma^2} = \frac{1}{1+\sigma^2}\,\mu + \frac{\sigma^2}{1+\sigma^2}\,r,$$

e

$$Var(S|R) = \frac{\sigma^2}{1+\sigma^2}.$$

Escrever a média condicional (melhor preditor)como acabamos de fazer é informativo, pois isto mostra que ela é ponderada de $\mu$, o valor do sinal a priori, e r, o valor recebido. Os pesos relativos dados a $\mu$ e $r$ tem entre sia mesma proporção que 1(a variância condicional do sinal recebido quando $s$ é enviado ) tem para $\sigma^2$( a variância do sinal enviado).

Às vezes, acontece da distribuição de probabilidade conjunta de $X$ e $Y$ não ser completamente conhecida, ou se for conhecida, ela ser tal que o cálculo de $E[Y|X=x]$ seja matematicamente intratável. Se, no entanto, a média, a variância e a correlação de $X$ e $Y$ são conhecidos, então podemos pelo menos determinar o melhor preditor de $Y$ com respeito a $X$.

Para obter o melhor preditor de $Y$, $g(X) = a + bX$, precisamos escolher $a$ e $b$ que minimizem

$$h(a,b) = E[(Y - g(X))^2] = E[(Y - a - bX)^2].$$

Vamos desenvolver o quadrado:

$$h(a,b) = E(Y^2 - 2aY - 2bXY + a^2 + b^2X^2 + 2abX),$$

e finalmente,

$$h(a,b) = E(Y^2) - 2aE(Y) - 2bE(XY) + a^2 + b^2E(X^2) + 2abE(X).$$

Vamos calcular as derivadas parciais:

$$\frac{\partial h(a,b)}{\partial a} \quad = \quad -2E(Y) + 2a + 2bE(X) = 0 \tag{7}$$

e

$$\frac{\partial h(a,b)}{\partial b} \quad = \quad -2E(XY) - 2aE(X) + 2bE(X^2) = 0 \tag{8}$$

Assim obtemos o sistema linear:

$$a + E(X)\,b = E(Y)$$

e

$$E(X)\,a + E(X^2)\,b = E(XY)$$

O determinante principal é dado por:

$$\Delta_P = \begin{vmatrix} 1 & E(X) \\ E(X) & E(X^2) \end{vmatrix} = E(X^2) - E^2(X) = V(X).$$

O determinante para a incógnita $b$ é dado por:

$$\Delta_B = \begin{vmatrix} 1 & E(Y) \\ E(X) & E(XY) \end{vmatrix} = E(XY) - E(X)E(Y) = Cov(X,Y).$$

E assim

$$b = \frac{\Delta_b}{\Delta_P} = \frac{Cov(X,Y)}{V(X)} = \frac{\rho\sigma_X\sigma_Y}{V(X)} = \rho\,\frac{\sigma_Y}{\sigma_X}.$$

O valor de $a$ é dado por:

$$a = E(Y) - bE(X).$$

O melhor preditor linear de $Y$ com respeito a $X$
é

$$E(Y) + \rho\,\frac{\sigma_Y}{\sigma_X}(X - E(X)).$$

O erro médio quadrático desse preditor é dado por:

$$EQM = E\left(\left[Y - E(Y) - \rho\,\frac{\sigma_Y}{\sigma_X}[(X - E(X)]\right]^2\right).$$

Assim,

$$EQM = E\left([Y - E(Y)]^2\right) + \left(\rho\,\frac{\sigma_Y}{\sigma_X}\right)^2 E\left([X - E(X)]^2\right) \; - 2\rho\,\frac{\sigma_Y}{\sigma_X}E\left([(X - E(X))(Y - E(Y))\right].$$

Logo,

$$EQM = V(Y) + \rho^2 V(Y) - \rho^2 V(Y) = \left(1 - \rho^2\right) V(Y).$$

Observamos que, se a correlação está próxima de $+1$ ou de $-1$, então o erro quadrático do melhor preditor linear é aproximadamente nulo.

Exemplo 6d Um exemplo no qual a esperança condicional de $Y$ dado $X$ é linear em $X$, e portanto no qual o melhor preditor linear de $Y$ com respeito a $X$ é o melhor preditor possível , é aquele em que $X$ e $Y$ tem uma distribuição normal bivariada. Neste caso

$$E(Y|X = x) = E(Y) + \rho\,\frac{\sigma_Y}{\sigma_X}(X - E(X)).$$

Vamos agora refazer alguns exemplos da seção 7.5.

Um minerador está preso em uma mina contendo 3 portas. A primeira porta leva a um túnel que o levará á saída após 3 horas de viagem. A segunda porta leva a um túnel que fará com que ele retorne à mina após 5 horas de viagem. A terceira porta leva a um túnel que fará com que ele retorne à mina após 7 horas de viagem. Se considerarmos que o minerador pode escolher qualquer uma das portas com igual probabilidade, qual o tempo esperado para que ele chegue à saída?

Solução:

Suponha que $X$ represente o tempo ,em horas, até que o minerador consiga sair e $Y$ o número da porta que ele escolheu (1,2, ou 3). Assim,

A distribuição de $Y$ é dada por:

$$P(Y = y) = \frac{1}{3}\,I_{\{1,2,3\}}(y).$$

Vamos calcular o valor esperado de $X$ usando esperança condicional.

$$\begin{aligned}
E(X) &= E(E(X|Y)) \\
&= E(X|Y = 1)P(Y = 1) + E(X|Y = 2)P(Y = 2) + E(X|Y = 3)P(Y = 3) \\
&= \frac{1}{3}\left[E(X|Y = 1) + E(X|Y = 2) + E(X|Y = 3)\right].
\end{aligned}$$

No entanto,

$$E(X|Y = 1) = 3,$$

com certeza após 3 horas ele conseguirá sair da mina.

Mas,

$$E(X|Y = 2) = E[5 + X^*] = 5 + E(X*) = 5 + E(X),$$

pois se o minerador escolher a segunda porta ele gastará 5 horas e volta ao ponto inicial. Seja $X*$ o tempo que ele levará para deixar a mina e esta variável terá mesma distribuição de $X$.

De maneira similar temos:

$$E(X|Y=3) = E[7+X^*] = 7 + E(X*) = 7 + E(X).$$

Desta maneira

$$E(X) = \frac{1}{3}[3+5+E(X)+7+E(X)] = \frac{1}{3}[15+2E(X)],$$

$$3E(X) = 15 + 2E(X),$$

e

$$E(X) = 15 \;\; horas.$$

Vamos calcular agora $V(X) = E(X^2) - E^2(X.)$

$$\begin{aligned} E(X^2) &= E(E(X^2|Y)) \\ &= E(X^2|Y=1)P(Y=1) + E(X^2|Y=2)P(Y=2) + E(X^3|Y=3)P(Y=3) \\ &= \frac{1}{3}\left[E(X^2|Y=1) + E(X^2|Y=2) + E(X^2|Y=3)\right]. \end{aligned}$$

No entanto,

$$E(X^2|Y=1) = 3^2 = 9,$$

com certeza após 3 horas ele conseguirá sair da mina.

Mas,

$$E(X^2|Y=2) = E[(5+X^*)^2] = 25 + 10E(X*) + E(X*^2) = 25 + 10E(X) + E(X^2) = 175 + E(X^2),$$

pois se o minerador escolher a segunda porta ele gastará 5 horas e volta ao ponto inicial. Seja $X*$ o tempo que ele levará para deixar a mina e esta variável terá mesma distribuição de $X$.

De maneira similar temos:

$$E(X|Y=3) = E[()7+X^*)^2] = 49 + 14E(X*) + E(X*^2) = 49 + 14E(X) + E(X^2) = 259 + E(X^2).$$

Desta maneira

$$E(X^2) = \frac{1}{3}[9+175+E(X^2)+259+E(X^2)] = \frac{1}{3}[334+2E(X^2)],$$

$$3E(X^2) = 334 + 2E(X^2),$$

e

$$E(X^2) = 334 \;\; horas^2.$$

e

$$V(X) = E(X^2) - E^2(X.) = 334 - 225 = 109 \;\; horas^2$$

.

Exemplo 5d

Suponha que o número de pessoas que entram em uma loja de departamentos em determinado dia seja uma variável aleatória com média 50. Suponha ainda que as quantias de dinheiro gastas por esses clientes sejam variáveis independentes com média comum 800 reais. Finalmente, suponha também que a quantia gasta por um cliente seja independente do número total de clientes que entram na loja. Qual é a quantidade esperada de dinheiro gasta na loja em um dado dia?

Solução: Se $N$ representa o número de clientes que entram na loja e $X_i$ a quantidade de dinheiro gasta pelo i-ésimo cliente, então a quantidade total de dinheiro gasta pode ser escrita como

$$S_N = \sum_{i=1}^{N} X_i.$$

Queremos calcular $E(S_N)$ usando argumentos de esperança condicional. Assim,

$$E(S_N) = E\left[S_N|N\right] = E\left[\sum_{i=1}^{N} X_i|N\right].$$

Mas,

$$\begin{aligned}
E\left[\sum_{i=1}^{N} X_i|N=n\right] &= E\left[\sum_{i=1}^{n} X_i|N\right] \\
&= E\left[\sum_{i=1}^{n} X_i\right], \ \textit{pela independencia entre } X_i \textit{ e } N \\
&= n(X), \ \ onde \ \ E(X) = E(X_i).
\end{aligned}$$

O que implica que

$$E\left[\sum_{i=1}^{N} X_i|N\right] = NE(X).$$

Assim,

$$E(S_N) = E\left[S_N|N\right] = E\left[\sum_{i=1}^{N} X_i|N\right] = E[NE(X)] = E(N)E(X),$$

que é a famosa equação de Wald.
Como isso, em nosso exemplo, a quantidade esperada de dinheiro gasta na loja é dada por:

$$E(S_N) = 50 \times 800 = 4000 \ reais.$$

Vamos mostrar agora que:

$$Var(S_N) = E(N)\ Var(X) + [E(X)]^2\ Var(N).$$

Solução:
Sabemos que:

$$\begin{aligned}
Var[S_N|N=n] &= Var(X_1 + X_2 + \ldots + X_n|N=n) \\
&= Var(X_1 + X_2 + \ldots + X_n) \\
&= nVar(X)
\end{aligned}$$

Assim,

$$Var(S_N|N) = Nvar(X).$$

$$\begin{aligned}
Var(S_N) &= E[Var(S_N)|N] + Var[E(S_N|N)] \\
&= E[NVar(X)] + Var[E(X)N] \\
&= E[N]\ Var(X) + [E(X)]^2\ Var(N).
\end{aligned}$$

# 5 Função de Distribuição Acumulada Bidimensional

Seja $(X, Y)$ um vetor aleatório contínuo bidimensional com função densidade conjunta dada por $f_{(X,Y)}(x, y)$. A Função de Distribuição Conjunta Bidimensional Contínua do vetor aleatório $(X, Y)$ é definida por:

$$F : R^2 \to [0, 1]$$

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u, v) dv\, du.$$

1. Propriedades da Função de Distribuição Acumulada Bivariada de Qualquer Tipo.

   a. $F(-\infty\ , y) = \lim_{x \to -\infty} F(x, y) = 0\ \ \forall\, y.$

   b. $F(x,\ -\infty\ ) = \lim_{y \to -\infty} F(x, y) = 0\ \ \forall\, x.$

   c. $\lim_{x \to \infty, y \to \infty} F(x, y) = F(\infty\ , \infty) = 1.$

   d. Se $x_1 < x_2$ e $y_1 < y_2$ , então:

   $$P(x_1 < X \leq x_2,\ y_1 < Y \leq y_2) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0.$$

   e. $F(x, y)$ é contínua à direita em cada argumento, isto é,

   $$\lim_{0<h\to 0} F(x + h, y) = \lim_{0<h\to 0} F(x, y + h) = F(x, y).$$

   Além disso:

   a. $F_X(x) = F_{(X,Y)}(x,\ \infty).$

   b. $F_Y(y) = F_{(X,Y)}(\infty\ , y).$

   c. Se $X$ e $Y$ são independentes então: $F_{(X,Y)}(x, y) = F_X(x)F_Y(y), \forall\ (x, y) \in R^2.$

   d.
   $$max(0, F_X(x) + F_Y(y) - 1) \leq F_{(X,Y)}(x, y) \leq \sqrt{F_X(x)F_Y(y)}, \quad \forall\ (x, y) \in R^2.$$

   Vamos calcular a acumulada bidimensional dos quatro exemplos a seguir:

2. Calcule a função de distribuição acumulada do vetor aleatório bidimensional $(X, Y)$ com f.d.p.c dada por:

$$f(x, y) = e^{-(x+y)} \; I_{(0\,,\,\infty)}\,(x)\, I_{(0\,,\,\infty)}\,(y).$$

3. Calcule a função de distribuição acumulada do vetor aleatório bidimensional $(X, Y)$ com f.d.p.c dada por:

$$f(x, y) = \frac{1}{4} \; I_{(0\,,\,2)}\,(x)\, I_{(0\,,\,2)}\,(y).$$

4. Calcule a função de distribuição acumulada do vetor aleatório bidimensional $(X, Y)$ com f.d.p.c dada por:

$$f(x, y) = \frac{3}{80}\,(x^2 + xy) \; I_{(0\,,\,2)}\,(x)\, I_{(0\,,\,4)}\,(y).$$

5. Calcule a função de distribuição acumulada do vetor aleatório bidimensional $(X, Y)$ com f.d.p.c dada por:

$$f(x, y) = 2 \; I_{(0\,,\,1)}\,(x)\, I_{(x\,,\,1)}\,(y).$$

6. Verifique se $G$ é uma função de distribuição acumulada do vetor aleatório bidimensional $(X, Y)$.

$$G(x, y) = \left[1 - 2\, e^{-2(x+y)}\right] \; I_{(0\,,\,\infty)}\,(x)\, I_{(0\,,\,\infty)}\,(y).$$

# 6 Transformação no Caso Bidimensional

1. Teorema 1: Sejam $X_1$ e $X_2$ variáveis aleatórias com função densidade de probabilidade conjunta $f(x_1, x_2)$. Seja o suporte

$$A = \{(x_1, x_2) \in \mathbf{R}^2; f(x_1, x_2) > 0\}.$$

Suponha que:

i. $y_1 = h_1(x_1, x_2)$ e $y_2 = h_2(x_1, x_2)$ definam uma transformação biunívoca de $A$ em $B$.
ii. As derivadas parciais de primeira ordem de $x_1 = h_1^{-1}(x_1, x_2) = w_1(y_1, y_2)$ e $x_2 = h_2^{-1}(x_1, x_2) = w_2(y_1, y_2)$ sejam funções contínuas em $B$.
iii. O jacobiano da transformação:

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix},$$

seja diferente de zero em $B$.
A densidade de $Y_1 = h_1(X_1, X_2)$ e $Y_2 = h_2(X_1, X_2)$ com suporte $B$ é dada por:

$$g(y_1, y_2) = f(w_1(y_1, y_2), w_2(y_1, y_2)) \;\; |J| \;\; I_B(y_1, y_2).$$

2. Teorema 2: Sejam $X_1$ e $X_2$ variáveis aleatórias com função densidade de probabilidade conjunta $f(x_1, x_2)$. Suponha que o suporte

$$A = \{(x_1, x_2) \in \mathbf{R}^2; f(x_1, x_2) > 0\},$$

possa ser decomposto em uma partição $A_1, A_2, \ldots, A_m$.

Suponha ainda que:

i. $y_1 = h_1(x_1, x_2)$ e $y_2 = h_2(x_1, x_2)$ definam uma transformação biunívoca de $A_i$ em $B$ para $i = 1, 2, \ldots, m$.

ii. As derivadas parciais de primeira ordem de $x_{1i} = h_{1i}^{-1}(x_1, x_2) = w_{1i}(y_1, y_2)$ e $x_{2i} = h_{2i}^{-1}(x_1, x_2) = w_{2i}(y_1, y_2)$,$i = 1, 2, \ldots, m$, sejam funções contínuas em $B$.

iii. O jacobiano da transformação:

$$J_i = \begin{vmatrix} \frac{\partial x_{1i}}{\partial y_1} & \frac{\partial x_{1i}}{\partial y_2} \\ \frac{\partial x_{2i}}{\partial y_1} & \frac{\partial x_{2i}}{\partial y_2} \end{vmatrix},$$

seja diferente de zero em $B$.

A densidade de $Y_1 = h_1(X_1, X_2)$ e $Y_2 = h_2(X_1, X_2)$ com suporte $B$ é dada por:

$$g(y_1, y_2) = \sum_{i=1}^{m} f(w_{1i}(y_1, y_2), w_{2i}(y_1, y_2)) \ |J_i| \ I_B(y_1, y_2).$$

3. Teorema 3: Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias com função densidade de probabilidade conjunta $f(x_1, x_2, \ldots, x_n)$. Seja o suporte

$$A = \{(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n; f(x_1, x_2, \ldots, x_n) > 0\}.$$

Suponha que:

i. $y_1 = h_1(x_1, x_2, \ldots, x_n),\ y_2 = h_2(x_1, x_2, \ldots, x_n), \ldots, y_n = h_n(x_1, x_2, \ldots, x_n)$ definam uma transformação biunívoca de $A$ em $B$.

ii. As derivadas parciais de primeira ordem de $x_1 = h_1^{-1}(x_1, x_2) = w_1(y_1, y_2, \ldots, y_n)$, $x_2 = h_2^{-1}(x_1, x_2, \ldots, x_n) = w_2(y_1, y_2, \ldots, y_n), \ldots, x_n = h_n^{-1}(x_1, x_2, \ldots, x_n) = w_n(y_1, y_2, \ldots, y_n)$ sejam funções contínuas em $B$.

iii. O jacobiano da transformação:

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \ldots & \frac{\partial x_1}{\partial y_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \ldots & \frac{\partial x_2}{\partial y_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \ldots & \frac{\partial x_n}{\partial y_n} \end{vmatrix},$$

seja diferente de zero em $B$.

A densidade de $Y_1 = h_1(X_1, X_2, \ldots, X_n), Y_2 = h_2(X_1, X_2, \ldots, X_n), \ldots, h_n(X_1, X_2), \ldots, X_n)$ com suporte $B$ é dada por:

$$g(y_1, y_2, \ldots, y_n) = f(w_1(y_1, y_2, \ldots, y_n), \ldots, w_n(y_1, y_2, \ldots, y_n)) \ |J| \ I_B(y_1, y_2, \ldots, y_n).$$

4. Teorema 4: Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias com função densidade de probabilidade conjunta $f(x_1, x_2, \ldots, x_n)$. Seja o suporte

$$A = \{(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n; f(x_1, x_2, \ldots, x_n) > 0\},$$

Suponha que ele possa ser decomposto em uma partição $A_1, A_2, \ldots, A_m$

Suponha ainda que:

i. $y_1 = h_1(x_1, x_2, \ldots, x_n),\ y_2 = h_2(x_1, x_2, \ldots, x_n), \ldots, y_n = h_n(x_1, x_2, \ldots, x_n)$ definam uma transformação biunívoca de $A_i$ em $B$, $i = 1, 2, \ldots, m$.

ii. As derivadas parciais de primeira ordem de $x_1 = h_{1i}^{-1}(x_1, x_2) = w_{1i}(y_1, y_2, \ldots, y_n)$, $x_2 = h_{2i}^{-1}(x_1, x_2, \ldots, x_n) = w_{2i}(y_1, y_2, \ldots, y_n), \ldots, x_n = h_{ni}^{-1}(x_1, x_2, \ldots, x_n) = w_{ni}(y_1, y_2, \ldots, y_n)$ sejam funções contínuas em $B$.

iii. O jacobiano da transformação:

$$J_i = \begin{vmatrix} \frac{\partial x_{1i}}{\partial y_1} & \frac{\partial x_{1i}}{\partial y_2} & \ldots & \frac{\partial x_{1i}}{\partial y_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial x_{2i}}{\partial y_1} & \frac{\partial x_{2i}}{\partial y_2} & \ldots & \frac{\partial x_{2i}}{\partial y_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial x_{ni}}{\partial y_1} & \frac{\partial x_{ni}}{\partial y_2} & \ldots & \frac{\partial x_{ni}}{\partial y_n} \end{vmatrix},$$

seja diferente de zero em $B$, $i = 1, 2, \ldots, m$.

A densidade de $Y_1 = h_1(X_1, X_2, \ldots, X_n), \ldots, X_n), Y_2 = h_2(X_1, X_2, \ldots, X_n), \ldots, h_n(X_1, X_2), \ldots, X_n)$ com suporte $B$ é dada por:

$$g(y_1, y_2, \ldots, y_n) = \sum_{i=1}^{m} f(w_{1i}(y_1, y_2, \ldots, y_n), \ldots, w_{ni}(y_1, y_2, \ldots, y_n)) \;\; |J_i| \;\; I_B(y_1, y_2, \ldots, y_n).$$

# 7 Transformações do Vetor Aleatório Contínuo Bidimensional (X,Y).

Seja $(X, Y)$ um vetor aleatório contínuo bidimensional com função densidade de probabilidade conjunta dada por $f(x, y)$ com suporte $A$. Sejam $f_X(x)$ e $f_Y(y)$ as marginais.

## 7.1 Função Densidade de Probabilidade da Soma S=X+Y.

$$f_S(s) = \int_{-\infty}^{\infty} f(x, s - x)dx,$$

se $X$ e $Y$ forem Independentes

$$f_S(s) = \int_{-\infty}^{\infty} f_X(x) f_Y(s - x)dx.$$

Prova: Considere a variável auxiliar $A = X$.
Temos que:
$x = a = w_1(s, a)$ e $y = s - a = w_2(s, a)$.
O Jacobiano da transformação é dado por:

$$J = \begin{vmatrix} \frac{\partial x}{\partial s} = 0 & \frac{\partial x}{\partial a} = 1 \\ \\ \frac{\partial y}{\partial s} = 1 & \frac{\partial y}{\partial a} = -1 \end{vmatrix} = -1,$$

A única condição é que o Jacobiano seja diferente de zero em $B$.

Assim

$$|J| = 1.$$

A densidade de $S = h_1(X,Y) = X + Y$ e $A = h_2(X,Y) = X$ com suporte $B$ é dada por:

$$g(s,a) = f(w_1(s,a), w_2(s,a)) \ \ |J| \ \ I_B(s,a).$$

Assim,

$$g(s,a) = f(a, s-a) \quad I_B(s,a).$$

A densidade de $S$ é dada por:

$$f_S(s) = \int_{-\infty}^{\infty} \ f(a, s-a)da,$$

que pode ser posto na forma $x = a$

$$f_S(s) = \int_{-\infty}^{\infty} \ f(x, s-x)dx.$$

Se $X$ e $Y$ forem independentes forem

$$f_S(s) = \int_{-\infty}^{\infty} \ f_X(x)F_y(s-x)dx.$$

Exemplos: Calcule a distribuição da soma $S = X + Y$.

1.
$$f(x,y) = (x+y) \ I_A(x) \ I_A(y), A = [0,1].$$

Solução: Como $0 \leq x \leq 1$ e $0 \leq y \leq 1$

temos $0 \leq x + y \leq 2$, isto é, $0 \leq s \leq 2$.

A densidade de $S$ é dada por:

$$\begin{aligned} f_S(s) &= \int_{-\infty}^{\infty} f(x, s-x)dx \\ &= \int_{-\infty}^{\infty} (x + s - x) \ I_{[0,1]}(x) \ I_{[0,1]}(s-x)dx \\ &= s \int_{-\infty}^{\infty} \ I_{[0,1]}(x) \ I_{[0,1]}(s-x)dx \\ &= s \int_a^b \ 1 \ dx \\ &= s(b-a), \end{aligned}$$

A variação de $x$ entre $a$ e $b$ sai das inequações:

$$I_{[0,1]}(x) = 1,$$

o que acarreta

$$0 \leq x \leq 1 \tag{9}$$

.

Por outro lado
$$I_{[0,1]}(s-x)=1,$$
nos leva a:
$0 \leq s-x \leq 1$ ou

$$s-1 \leq x \leq s \tag{10}$$

.
De (1) e (2) temos:

$$max(0,s-1) \leq x \leq min(1,s) \tag{11}$$

.
logo $a = max(0,s-1)$ e $b = min(1,s)$
A densidade de $S$ é dada por:

$$f_S(s) = s[(min(1,s) - max(0,s-1)]\ I_{[0,2]}\ (s).$$

Que também pode ser posta na forma:

$$f_S(s) = s^2\ I_{[0,1]}\ (s) + s(2-s)\ I_{[1,2]}.$$

2. $X \sim Exp(1)$, $Y \sim Exp(1)$, independentes. A conjunta de $(X,Y)$ é dada por:

$$f(x,y) = exp[-(x+y)]\ I_{(0,\infty)}(x)\ I_{(0,\infty)}(y).$$

Como $x>0$ e $y>0$ temos que $s=x+y>0$.
A densidade de $S$ é dada por:

$$\begin{aligned}
f_S(s) &= \int_{-\infty}^{\infty} f(x,s-x)dx \\
&= \int_{-\infty}^{\infty} exp[-(x+s-x)]\ I_{(0,\infty)}(x)\ I_{(0,\infty)}(s-x)dx \\
&= exp(-s)\ \int_{-\infty}^{\infty}\ I_{(0,\infty)}(x)\ I_{(0,\infty)}(s-x)dx \\
&= exp(-s)\ \int_a^b\ 1\ dx \\
&= exp(-s)(b-a),
\end{aligned}$$

A variação de $x$ entre $a$ e $b$ sai das inequações:

$$I_{(0,\infty)}(x)=1,$$

o que acarreta

$$x>0 \tag{12}$$

.
Por outro lado
$$I_{(0,\infty)}(s-x)=1,$$
nos leva a:

$$s - x > 0$$

$$x \leq s \tag{13}$$

.

De (4) e (5) temos:

$$0 \leq x \leq s \tag{14}$$

.

logo $a = 0, \quad b = s$ e $b - a = s$

A densidade de $S$ é dada por:

$$f_S(s) = s\, e^{-s}\, I_{(0,\infty}\,(s),$$

que é a densidade da gama (2,1).

3. $X \sim U[0,2] \quad Y \sim U[0,2]$, $X$ e $Y$ independentes.

   A conjunta de $(X, Y)$ é dada por:

$$f(x,y) = \frac{1}{4}\; I_A(x)\; I_A(y), A = [0,2]).$$

   Solução: Como $0 \leq x \leq 2$ e $0 \leq y \leq 2$

   temos $0 \leq x + y \leq 4$, isto é, $0 \leq s \leq 4$.

   A densidade de $S$ é dada por:

$$\begin{aligned}
f_S(s) &= \int_{-\infty}^{\infty} f(x, s-x)dx \\
&= \int_{-\infty}^{\infty} \frac{1}{4}\; I_{[0,2]}(x)\; I_{[0,2]}(s-x)dx \\
&= \frac{1}{4} \int_{-\infty}^{\infty} I_{[0,2]}(x)\; I_{[0,2]}(s-x)dx \\
&= \frac{1}{4}s \int_{a}^{b} 1\; dx \\
&= \frac{b-a}{4},
\end{aligned}$$

   A variação de $x$ entre $a$ e $b$ sai das inequações:

$$I_{[0,2]}(x) = 1,$$

   o que acarreta

$$0 \leq x \leq 2 \tag{15}$$

   .

   Por outro lado

$$I_{[0,2]}(s-x) = 1,$$

   nos leva a:

   $0 \leq s - x \leq 2$ ou

$$s - 2 \leq x \leq s \tag{16}$$

.

De (1) e (2) temos:

$$max(0, s-2) \leq x \leq min(2, s) \tag{17}$$

.

logo $a = max(0, s-2)$ e $b = min(2, s)$

A densidade de $S$ é dada por:

$$f_S(s) = \frac{(min(2, s) - max(0, s-2)}{4} \ \ I_{[0,4]} \ (s).$$

Que também pode ser posta na forma:

$$f_S(s) = \frac{s}{4} \ I_{[0,2]} \ (s) + \frac{4-s}{4} \ I_{[2,4]},$$

que é a densidade da triangular com $a = 0, \ \ b = 4$ e $c = 2$.

4. Seja $(X, Y)$ com distribuição uniforme na região:

$$A = \{(x, y) \in R^2 \mid x > 0, \ y > 0 \ e \ x + y \leq 1\}.$$

A área de $A$ vale 1/2. Assim a f.d.p. c. de $(X, Y)$ é dada por:

$$f(x, y) = 2I_{(0,1)} \ (x) \ \ I_{(0,1-x)} \ (y).$$

Como $0 < x + y \leq 1$ temos que $0 < s \leq 1$

A densidade de $S$ é dada por:

$$\begin{aligned} f_S(s) &= \int_{-\infty}^{\infty} f(x, s-x) dx \\ &= \int_{-\infty}^{\infty} 2 \ I_{(0,1]}(x) \ I_{(0,1-x]}(s-x) dx \\ &= 2 \int_a^b \ 1 \ dx \\ &= 2(b-a). \end{aligned}$$

A variação de $x$ entre $a$ e $b$ sai das inequações:

$$I_{[0,1]}(x) = 1,$$

o que acarreta

$$0 \leq x \leq 1 \tag{18}$$

.

Por outro lado

$$I_{[0,1-x]}(s-x) = 1,$$

nos leva a:

$$0 \leq s - x \leq 1 - x$$

ou
Da inequação

$$0 \leq s - x$$

temos que

$$x \leq s \tag{19}$$

.
Da inequação

$$s - x \leq 1 - x$$

temos que

$$s \leq 1.$$

De (10) e (11) temos:

$$0 \leq x \leq min(1, s) = s \tag{20}$$

.
logo $a = 0$ e $b = s$.
A densidade de $S$ é dada por:

$$f_S(s) = 2s \;\; I_{(0,1]} \; (s),$$

que é a densidade da beta(2,1).

## 7.2 Função Densidade de Probabilidade da Diferença D=X-Y.

$$f_D(d) = \int_{-\infty}^{\infty} f(d + y, y) dy,$$

se $X$ e $Y$ forem Independentes

$$f_D(d) = \int_{-\infty}^{\infty} f_X(d + y) f_Y(y) dx.$$

Prova: Considere a variável auxiliar $A = Y$. Temos que: $d = x - y$ e $a = y$ .Assim
$x = d + a = w_1(d, a)$ e $y = d = w_2(d, a)$.
O Jacobiano da transformação é dado por:

$$J = \begin{vmatrix} \frac{\partial x}{\partial d} = 1 & \frac{\partial x}{\partial a} = 1 \\ \frac{\partial y}{\partial d} = 0 & \frac{\partial y}{\partial a} = 1 \end{vmatrix} = 1,$$

seja diferente de zero em $B$.
Assim $|J| = 1$.
A densidade de $D = h_1(X, Y) = X - Y$ e $A = h_2(X, Y) = Y$ com suporte $B$ é dada por:

$$g(d, a) = f(w_1(d, a), w_2(d, a)) \;\; |J| \;\; I_B(d, a).$$

Assim,

$$g(d, a) = f(d + a, a) \;\; ; I_B(d, a).$$

A densidade de $D$ é dada por:

$$f_D(d) = \int_{-\infty}^{\infty} f(d+a,a)da,$$

que pode ser posto na forma $y = a$

$$f_D(d) = \int_{-\infty}^{\infty} f(d+y,y)dy.$$

Se $X$ e $Y$ forem independentes forem

$$f_D(d) = \int_{-\infty}^{\infty} f_X(d+y)f_Y(y)dy.$$

Exemplos: Calcule a distribuição da diferença $D = X - Y$.

5. $X \sim U[0,1] \quad Y \sim U[0,1]$, $X$ e $Y$ independentes.
   A conjunta de $(X,Y)$ é dada por:

$$f(x,y) = \ I_A(x)\ I_A(y), A = [0,1]).$$

Solução: Como $0 \le x \le 1$ e $0 \le y \le 1$
temos $-1 \le x - y \le 1$, isto é, $-1 \le d \le 1$.
A densidade de $D$ é dada por:

$$\begin{aligned}
f_D(d) &= \int_{-\infty}^{\infty} f(d+y,y)dy \\
&= \int_{-\infty}^{\infty} I_{(0,1)}(d+y)I_{(0,1)}(y) \\
&= \int_a^b dy \\
&= (b-a),
\end{aligned}$$

A variação de $x$ entre $a$ e $b$ sai das inequações:

$$I_{[0,1]}(y) = 1,$$

o que acarreta

$$0 \le y \le 1 \tag{21}$$

.

Por outro lado

$$I_{[0,1]}(d+y) = 1,$$

nos leva a:

$$0 \le d + y \le 1$$

ou

$$-d \le y \le 1 - d$$

temos que

$$-d \le y \le 1 - d \tag{22}$$

.

De (10) e (11) temos:

$$max(0, -d) \leq y \leq min(1, 1-d) \tag{23}$$

.

logo $a = max(0, -d)$ e $b = min(1, 1-d)$ e

$$b - a = min(1, 1-d) - max(0, -d).$$

Sabemos que

$$2min(a, b) = a + b - |a - b|.$$

Logo,

$$2min(1, 1-d) = 1 + 1 - d - |d| = 2 - d - |d|.$$

Logo,

$$2min(1, 1-d) = 1 + 1 - d - |d| = 2 - d - |d|.$$

Sabemos que

$$2max(a, b) = a + b + |a - b|.$$

Logo,

$$2max(0, -d) = 0 - d + |d| = -d + |d|.$$

Assim,

$$2(b - a) = 2[min(1, 1-d) - max(0, -d)] = [2 - d - |d| + d - |d|] = 2[1 - |d|]$$

Assim,

$$f_D(d) = 2[1 - |d|] \;\; I_{[-1,1]} \; (d),$$

que é a densidade da triangular com parâmetros $a = -1$, $b = 1$ e $c = 0$.

6. $X \sim Exp(1)$, $Y \sim Exp(1)$, independentes.
   A conjunta de $(X, Y)$ é dada por:

$$f(x, y) = exp[-(x + y)] \; I_{(0,\infty)}(x) \; I_{(0,\infty)}(y).$$

Como $x > 0$ e $y > 0$ temos que $-\infty < d = x - y < \infty$.

A densidade de $D$ é dada por:

$$\begin{aligned}
f_D(d) &= \int_{-\infty}^{\infty} f(d+y, y)dy \\
&= \int_{-\infty}^{\infty} exp[-(d+y+y)]\ I_{(0,\infty)}(d+y)\ I_{(0,\infty)}(y)dy \\
&= exp(-d)\ \int_{-\infty}^{\infty} exp[-2y]\ I_{(0,\infty)}(d+y)\ I_{(0,\infty)}(y)dy \\
&= exp(-d)\ \int_{max(0,-d)}^{\infty} \quad exp[-2y]\ dy \\
&= \frac{1}{2}\ exp(-d)\ exp[-2max(0,-d)] \\
&= \frac{1}{2}\ \ exp[-(d+2max(0,-d)] \\
&= \frac{1}{2}\ \ exp[-|d|].
\end{aligned}$$

Vamos explicar com detalhes a mágica utilizada.
A variação de $y$ sai das inequações:

$$I_{(0,\infty)}(y) = 1,$$

o que acarreta

$$y > 0 \tag{24}$$

.
Por outro lado

$$I_{(0,\infty)}(d+y) = 1,$$

nos leva a:

$$d + y > 0$$

$$y \leq -d \tag{25}$$

.
De (13) e (14) temos:

$$y \leq max(0, -d) \tag{26}$$

.
Por outro lado sabemos que

$$max(a, b) = \frac{a + b + |a - b|}{2},$$

assim,

$$2max(a, b) = a + b + |a - b|.$$

Fazendo $a = 0$ e $b = -d$ temos:

$$2max(0, -d) = 0 - d + |0 + d| = -d + |d|,$$

que nos leva a:

$$d + 2max(0, -d) = |d|.$$

A densidade de $D$ é dada por:

$$f_D(d) = \frac{1}{2}\ exp[-|d|]\ I_{(-\infty,\infty)}\ (d),$$

que é a densidade da Laplace padrão (0,1).

## 7.3 Função Densidade de Probabilidade do Produto U=XY.

$$f_U(u) = \int_{-\infty}^{\infty} \frac{1}{|x|} f(x, \frac{u}{x}) dx,$$

se $X$ e $Y$ forem Independentes

$$f_U(u) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_X(x) f_Y(\frac{u}{x})) dx.$$

Prova: Considere a variável auxiliar $A = X$. Temos que: $u = xy$ e $a = x$ .Assim
$x = a = w_1(u, a)$ e $y = \frac{u}{a} = w_2(u, a)$.
O Jacobiano da transformação é dado por:

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} = 0 & \frac{\partial x}{\partial a} = 1 \\ \frac{\partial y}{\partial u} = \frac{1}{a} & \frac{\partial y}{\partial a} = -\frac{u}{a^2} \end{vmatrix} = -\frac{1}{a},$$

seja diferente de zero em $B$, isto é, que $a$ não seja nulo..
Assim

$$|J| = \frac{1}{|a|}.$$

A densidade de $U = h_1(X, Y) = XY$ e $A = h_2(X, Y) = X$ com suporte $B$ é dada por:

$$g(u, a) = f(w_1(u, a), w_2(u, a))\ \ |J|\ \ I_B(u, a).$$

Assim,

$$g(u, a) = ;\frac{1}{|a|}; f(a, \frac{u}{a})\ \ ; I_B(u, a).$$

A densidade de $U$ é dada por:

$$f_U(u) = \int_{-\infty}^{\infty}\ \frac{1}{|a|} f(a, \frac{u}{a}) da,$$

que pode ser posto na forma $x = a$

$$f_U(u) = \int_{-\infty}^{\infty}\ \frac{1}{|a|}; f(a, \frac{u}{x}) dx.$$

Se $X$ e $Y$ forem independentes forem

$$f_U(u) = \int_{-\infty}^{\infty}\ \frac{1}{|x|}; f_X(x) f_Y(\frac{u}{x}) dx.$$

Exemplos: Calcule a distribuição do produto $U = XY$

7. $X \sim U[0,1] \quad Y \sim U[0,1]$, $X$ e $Y$ independentes.
A conjunta de $(X,Y)$ é dada por:

$$f(x,y) = \; I_A(x)\; I_A(y), A = [0,1]).$$

Solução: Como $0 \le x \le 1$ e $0 \le y \le 1$
temos $0 \le xy \le 1$, isto é, $0 \le u \le 1$.
A densidade de $U$ é dada por:

$$\begin{aligned} f_U(u) &= \int_{-\infty}^{\infty} \frac{1}{|x|}\; f_X(x) f_Y(\frac{u}{x})dx \\ &= \int_{-\infty}^{\infty} \frac{1}{|x|}\; I_{(0,1)}(\frac{u}{x})\; I_{(0,1)}(x)dx \\ &= \int_u^1 \frac{1}{x}dx \\ &= -ln(u)\; I_{(0,1)}\;\; (u) \end{aligned}$$

A f.d.p. de $U = XY$ é dada por:

$$f_U(u) = -ln(u)\; I_{(0,1)}\;\; (u).$$

8. $X \sim Exp(1)$, $Y \sim Exp(1)$, independentes.
A conjunta de $(X,Y)$ é dada por:

$$f(x,y) = exp[-(x+y)]\; I_{(0,\infty)}(x)\; I_{(0,\infty)}(y).$$

Como $x > 0$ e $y > 0$ temos que $u = xy > 0$.
A densidade de $U$ é dada por:

$$\begin{aligned} f_U(u) &= \int_{-\infty}^{\infty} \frac{1}{|x|}\; f_X(x) f_Y(\frac{u}{x})dx \\ &= \int_{-\infty}^{\infty} \frac{1}{|x|} exp[-(\frac{u}{x} + x)]\; I_{(0,\infty)}(\frac{u}{x})\; I_{(0,\infty)}(x)dx \\ &= \int_0^{\infty} \frac{1}{x} exp[-(\frac{u}{x} + x)]dx, \end{aligned}$$

que não tem uma forma fechada.
A f.d.p de $U = XY$ é dada por:

$$f_U(u) = \int_0^{\infty} \frac{1}{x} exp[-(\frac{u}{x} + x)]dx \;\; I_{(0,\infty)}\; (u).$$

## 7.4 Função Densidade de Probabilidade do Quociente $V = \frac{X}{Y}$.

$$f_V(v) = \int_{-\infty}^{\infty} |y| f(vy, y)dy,$$

se $X$ e $Y$ forem Independentes

$$f_V(v) = \int_{-\infty}^{\infty} |y| f_X(vy) f_Y(y)dy.$$

Prova: Considere a variável auxiliar $A = Y$. Temos que: $v = \dfrac{x}{y}$ e $a = y$. Assim
$x = va = w_1(v, a)$ e $y = a = w_2(v, a)$.
O Jacobiano da transformação é dado por:

$$J = \begin{vmatrix} \frac{\partial x}{\partial v} = a & \frac{\partial x}{\partial a} = v \\ \frac{\partial y}{\partial v} = 0 & \frac{\partial y}{\partial a} = 1 \end{vmatrix} = a,$$

que é diferente de zero em $B$ desde que $a$ não seja nulo.
Assim $|J| = |a|$.
A densidade de $V = h_1(X, Y) = \dfrac{X}{Y}$ e $A = h_2(X, Y) = Y$ com suporte $B$ é dada por:

$$g(v, a) = f(w_1(v, a), w_2(v, a)) \;\; |J| \;\; I_B(v, a).$$

Assim,

$$g(v, a) = |a| \; f(va, a) \;\; ; I_B(v, a).$$

A densidade de $V$ é dada por:

$$f_V(v) = \int_{-\infty}^{\infty} \;\; |a| \; f(va, a) da,$$

que pode ser posto na forma $y = a$

$$f_V(v) = \int_{-\infty}^{\infty} \;\; |y| \; f(vy, y) dy.$$

Se $X$ e $Y$ forem independentes forem

$$f_D(d) = \int_{-\infty}^{\infty} \;\; f_X(d + y) f_Y(y) dy.$$

Exemplos: Calcule a distribuição do quociente $V = \dfrac{X}{Y}$.

9. $X \sim Exp(1)$, $Y \sim Exp(1)$, independentes.
A conjunta de $(X, Y)$ é dada por:

$$f(x, y) = exp[-(x + y)] \; I_{(0,\infty)}(x) \; I_{(0,\infty)}(y).$$

Como $x > 0$ e $y > 0$ temos que $v = \dfrac{x}{y} > 0$.
A densidade de $V$ é dada por:

$$\begin{aligned} f_V(v) &= \int_{-\infty}^{\infty} |x| \; f_X(vy) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} |x| exp[-(vy + y)] \; I_{(0,\infty)}(vy) \; I_{(0,\infty)}(y) dy \\ &= \int_{0}^{\infty} y exp[-(1 + v)y] dy \\ &= \frac{1}{1 + v} \int_{0}^{\infty} y(1 + v) exp[-(1 + v)y] dy \\ &= \frac{1}{1 + v} \; E(Y*), \;\; Y* \sim exp(1 + v) \\ &= \frac{1}{(1 + v)^2} \;\; I_{(0,\infty)} \; (v), \end{aligned}$$

que é a densidade da $F(2,2)$.

10. $X \sim U[0,1] \quad Y \sim U[0,1]$, $X$ e $Y$ independentes.
A conjunta de $(X,Y)$ é dada por:

$$f(x,y) = \; I_A(x)\; I_A(y), A = [0,1]).$$

Solução: Como $0 \leq x \leq 1$ e $0 \leq y \leq 1$
temos $0 \leq \dfrac{x}{y} \leq \infty$, isto é, $0 \leq v \leq \infty$.
A densidade de $V$ é dada por:

$$\begin{aligned} f_V(v) &= \int_{-\infty}^{\infty} \frac{|}{y}|\; f_X(vy) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} |y|\; I_{(0,1)}(vy)\; I_{(0,1)}(y) dy \\ &= \int_a^b y dy \\ &= \frac{b^2 - a^2}{2} \end{aligned}$$

A variação de $y$ sai das inequações:

$$I_{(0,1)}(y) = 1,$$

o que acarreta

$$0 < y < 1 \tag{27}$$

.
Por outro lado

$$I_{(0,\infty)}(vy) = 1,$$

nos leva a:

$$0 < vy < 10$$

$$0 < y < 1/v \tag{28}$$

.
De (19) e (20) temos:

$$0 < y \leq min(1, 1/v) \tag{29}$$

.
A f.d.p. de $V = \dfrac{X}{Y}$ é dada por:

$$f_V(v) = \frac{[min(1,1/v)]^2}{2}\; I_{(0,\infty)}\;\;(v),$$

que pode ser posta na forma:

$$f_V(v) = \frac{1}{2}\; I_{(0,1)}\;\;(v) + \frac{1}{2v^2}\; I_{(1,\infty)}\;\;(v)$$

11. $X \sim N(0,1) \quad Y \sim N(0,1)$, $X$ e $Y$ independentes.
A conjunta de $(X,Y)$ é dada por:

$$f(x,y) = \frac{1}{2\pi}\, exp[-\frac{x^2+y^2}{2}]\ \ I_A(x)\ I_A(y), \ \ A = (-\infty, \infty).$$

Solução: Como $-\infty < x < \infty \quad \text{e} \quad -\infty < y < \infty$
temos $-\infty < \dfrac{x}{y} < \infty$, isto é, $-\infty < v \ < \infty$.
A densidade de $V$ é dada por:

$$\begin{aligned}
f_V(v) &= \int_{-\infty}^{\infty} |y|\ f_X(vy) f_Y(y) dy \\
&= \int_{-\infty}^{\infty} |y|\ \frac{1}{2\pi}\ exp[-\frac{v^2y^2+y^2}{2}]\ I_{(-\infty,\infty)}(vy)\ I_{(-\infty,\infty)}(y) dy \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} |y|\ exp[-\frac{(1+v^2)y^2}{2}]\ dy \\
&= \frac{2}{2\pi}\int_{0}^{\infty} y\ exp[-\frac{(1+v^2)y^2}{2}] \\
&= \frac{1}{\pi}\ IGG(a=2, b=\frac{1+v^2}{2}, c=2) \\
&= \frac{1}{\pi}\ \frac{1}{2\ \frac{1+v^2}{2}} \\
&= \frac{1}{\pi(1+v^2)}\ I_{(-\infty,\infty)}\ (v),
\end{aligned}$$

que é a densidade da Cauchy padrão.

## 7.5 Função Densidade de Probabilidade do Máximo V=Max(X,Y).

$$f_V(v) = \int_{-\infty}^{v} [f(u,v) + f(v,u)] du \ \ I_B(v),$$

se $X$ e $Y$ forem Independentes

$$f_V(v) = [f_X(v)F_Y(v) + F_X(v)f_Y(v)]\ \ I_B(v).$$

se $X$ e $Y$ forem Independentes e Identicamente Distribuídas(i.i.d)

$$f_V(v) = 2F(v)f(v)I_A\ (v),$$

em que $F$ é a função de distribuição acumulada com de $X$ e de $Y$ e $f$ a f.d.p. com suporte $A$.

1. $X \sim Exp(1)$, $Y \sim Exp(1)$, independentes. Qual a densidade de $V = Max(X,Y)$?

A densidade comum de $X$ e de $Y$ é dada por:

$$f(x) = exp(-x)\ I_{(0,\infty)}(x).$$

A acumulada comum de $X$ e de $Y$ é dada por:

$$F(x) = (1 - exp(-x))\ \ I_{(0,\infty)}(x).$$

Como $x > 0$ e $y > 0$ temos que $v = \max(x, y) > 0$.
a densidade de $V = Max(X, Y)$ é dada por:

$$f_V(v) = 2F(v)f(v)I_A\ (v),$$

$$f_V(v) = 2(1 - e^{-v})\ e^{-v}I_A\ (v),\ \ A = (0, \infty).$$

## 7.6 Função Densidade de Probabilidade do Mínimo U=Min(X,Y).

$$f_U(u) = \int_u^{\infty} [f(u, v) + f(v, u)]dv\ \ I_B(u),$$

se $X$ e $Y$ forem Independentes

$$f_U(u) = \{f_X(v)\,[1 - F_Y(u)] + [1 - F_X(u)]\,f_Y(u)\}\ \ I_B(u).$$

se $X$ e $Y$ forem Independentes e Identicamente Distribuídas(i.i.d)

$$f_U(u) = 2[1 - F(u)]f(u)I_A(u),$$

em que $F$ é a função de distribuição acumulada com de $X$ e de $Y$ e $f$ a f.d.p. com suporte $A$.

1. $X \sim Exp(1)$, $Y \sim Exp(1)$, independentes. Qual a densidade de $U = Min(X, Y)$?
   A densidade comum de $X$ e de $Y$ é dada por:

   $$f(x) = exp(-x)\ I_{(0,\infty)}(x).$$

   A acumulada comum de $X$ e de $Y$ é dada por:

   $$F(x) = (1 - exp[-x])\ \ I_{(0,\infty)}(x).$$

   A sobrevivência comum de $X$ e de $Y$ é dada por:

   $$S(x) = exp[-x]\ I_{(0,\infty)}(x).$$

   Como $x > 0$ e $y > 0$ temos que $v = \min(x, y) > 0$.
   a densidade de $U = Min(X, Y)$ é dada por:

$$f_U(u) = 2S(u)f(u)I_A\ (u),$$

$$f_U(u) = 2\ e^{-u}\ e^{-u}\ I_A\ (u) = 2\ e^{-2u}\ I_A\ (u), \quad A = (0, \infty),$$

que é a densidade de uma exponencial com parâmetro $\lambda = 2$.

## 7.7 Função Densidade de Probabilidade Conjunta de (U,V)=(Min(X,Y) , Max(X,Y)).

$$f_{(U,V)}(u, v) = [f_{(X,Y)}(u, v) + f_{(X,Y)}(v, u)]I_{\{-\infty<u<v<\infty\}},$$

e $X$ e $Y$ forem Independentes

$$f_{(U,V)}(u, v) = [f_X(u)f_Y(v) + f_X(v)f_Y(u)]I_{\{-\infty<u<v<\infty\}},$$

se $X$ e $Y$ forem Independentes e Identicamente Distribuídas(i.i.d)

$$f_{(U,V)}(u, v) = 2f(u)f(v)I_{\{-\infty<u<v<\infty\}}.$$

1. $X \sim Exp(1), Y \sim Exp(1)$, independentes. Qual a densidade conjunta de $(U, V) = (Min(X, Y)\ ,\ \ Max(X, Y))$.?

   A densidade comum de $X$ e de $Y$ é dada por:

   $$f(x) = exp(-x)\ I_{(0,\infty)}(x).$$

   A densidade conjunta de $(U, V)$ é dada por:

   $$f_{(U,V)}(u, v) = 2f(u)f(v)I_{\{-\infty<u<v<\infty\}}.$$

   Assim,

   $$f_{(U,V)}(u, v) = 2\ e^{-u}\ e^{-v} =\ e^{-u-v}\ \ I_{\{0<u<v<\infty\}}.$$

   A marginal de $U$ é dada por:

   $$f_U(u) = e^{-u}\ \int_u^\infty e^{-v}\ dv = 2\ e^{-u}\ e^{-u},$$

   $$f_U(u) = 2\ e^{-2u}\ I_{(0,\infty)}(u).$$

   A marginal de $V$ é dada por:

   $$f_V(v) = e^{-v}\ \int_0^v e^{-u}\ du = 2\ e^{-v}\ (1 - e^{-v}),$$

   $$f_V(v) = 2\ e^{-v}\ (1 - e^{-v})\ I_{(0,\infty)}(v).$$