Capítulo 3

Vetores Aleatórios

3.1 Introdução

Em geral, um mesmo fenômeno aleatório pode envolver várias variáveis de interesse. O estudo conjunto dessas variáveis é o objeto deste capítulo. As idéias desenvolvidas no capítulo anterior são, agora, estendidas para o que chamaremos de *vetor aleatório*. De fato, muitos dos conceitos serão apresentados para duas variáveis, mas serão, na sua maioria, válidos para um número finito de variáveis. Na Seção 3.2 a generalização da idéia de variável aleatória é apresentada com a correspondente descrição de sua estrutura probabilística. No estudo do comportamento conjunto das variáveis, definimos o conceito de independência. A independência tem um papel importante em probabilidade, amostragem e em muitos métodos estatísticos. Na Seção 3.3, funções de variáveis aleatórias são discutidas. Ainda, nessa seção, descrevemos o método do Jacobiano aplicado para transformação de vetores aleatórios.

3.2 Vetores Aleatórios

A realização de um fenômeno aleatório pode ensejar muitas perguntas de interesse. Em termos gerais, quando uma coleta de dados é realizada, quase sempre, ela envolve várias variáveis. Trabalhar conjuntamente com duas ou mais variáveis é, assim, um caminho natural para responder muitas perguntas práticas. Mesmo em fenômenos aleatórios relativamente simples podemos definir mais de uma variável de interesse. Nesta seção, formalizamos o tratamento probabilístico de conjuntos de variáveis aleatórias. Para distinguir do caso unidimensional, usaremos negrito quando tivermos mais de uma variável.

Definição 3.1: Vetor Aleatório

Seja $(\Omega, \mathcal{F}, \mathcal{P})$ um espaço de probabilidade. Definimos como *vetor aleatório, variável aleatória multidimensional* ou *variável aleatória multivariada* a uma função $\mathbf{X}$ de Ω em $\mathbb{R}^m$ que seja $\mathcal{F}$-mensurável. Em outras palavras, a função representada por $\mathbf{X}(w) = (X_1(w), X_2(w), \ldots, X_m(w))$ é tal que para todo $i = 1, 2, \ldots m$ e todo $I_i \subset \mathbb{R}$, temos $X_i^{-1}(I_i) \in \mathcal{F}$. □

Na definição de vetor aleatório, está implícito que $X_1, X_2, \ldots, X_m$ são variáveis aleatórias no mesmo espaço de probabilidade $(\Omega, \mathcal{F}, \mathcal{P})$. Isto segue do requisito de que a imagem inversa de cada um dos X_i's esteja em $\mathcal{F}$, ou seja, para todo $I_i \subset \mathbb{R}$ o conjunto $X_i^{-1}(I_i)$ é um evento. Assim,

$$\{w : X_1(w) \leq x_1, \ldots, X_m(w) \leq x_m\} = \bigcap_{i=1}^{m} \{w : X_i(w) \leq x_i\},$$

também é um evento, pois é a intersecção de elementos da σ-álgebra $\mathcal{F}$.

Da mesma forma que no caso univariado, definimos uma função que caracteriza, de modo único, o comportamento probabilístico do vetor aleatório.

Definição 3.2: Função de Distribuição Conjunta

A *função de distribuição conjunta* de $\mathbf{X}$ é definida por

$$F(\mathbf{x}) = F(x_1, x_2, \ldots, x_m) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_m \leq x_m),$$

para qualquer $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$. □

Exemplo 3.1: Uma máquina automática faz furos em uma chapa de aço e trabalha diariamente até quebrar. Existem dois tipos de chapas com diferentes durezas. Essas chapas alimentam o funcionamento da máquina de uma forma aleatória. Assim, ao quebrar, a máquina poderia estar furando qualquer uma das chapas. No momento de quebra da máquina, duas variáveis são de interesse: X_1 com os valores 0 e 1, indicando o tipo da chapa em operação e X_2 o número de dias desde a última quebra. Admitimos a seguinte função de distribuição conjunta para as variáveis:

$$F(x_1, x_2) = \begin{cases} 0, & x_1 < 0 \text{ ou } x_2 < 0; \\ \frac{2}{3} \sum\limits_{j=0}^{[x_2]} (\frac{1}{2})^{j+1}, & 0 \leq x_1 < 1, x_2 \geq 0; \\ \sum\limits_{j=0}^{[x_2]} (\frac{1}{2})^{j+1}, & x_1 \geq 1, x_2 \geq 0; \end{cases}$$

em que $[x_2]$ indica o maior inteiro menor ou igual a x_2.

A partir dessa estrutura de probabilidade, poderíamos avaliar o efeito do tipo de chapa no funcionamento da máquina, bem como estabelecer programas de manutenção preventiva. □

Exemplo 3.2: Considere uma central de reservas de uma companhia aérea e, para uma chamada ao acaso, estamos interessados em duas quantidades aleatórias. A variável X_1 é o tempo de espera, em minutos, para o cliente iniciar seu atendimento. A segunda variável de interesse é o tempo de atendimento em minutos, representado por X_2. Vamos supor que o comportamento conjunto dessas variáveis se dá segundo a função de distribuição conjunta apresentada a seguir:

$$F(x_1, x_2) = \begin{cases} 0, & x_1 < 0 \text{ ou } x_2 < 0; \\ 1 - e^{-x_1} - e^{-2x_2} + e^{-(x_1+2x_2)}, & x_1 \geq 0, x_2 \geq 0. \end{cases}$$

A qualidade do serviço de reservas dessa companhia poderia, então, ser avaliado levando em conta as probabilidades dos tempos de demora e atendimento. □

Na próxima proposição, apresentamos algumas propriedades da função de distribuição conjunta.

Proposição 3.1: Propriedades da Função de distribuição Conjunta

Seja $\mathbf{X}$ um vetor aleatório em $(\Omega, \mathcal{F}, \mathcal{P})$ então, para qualquer $\mathbf{x} \in \mathbb{R}^m$, $F(\mathbf{x})$ satisfaz as seguintes propriedades:

(FC1) $F(\mathbf{x})$ é não decrescente em cada uma de suas coordenadas;
(FC2) $F(\mathbf{x})$ é contínua à direita em cada uma de suas coordenadas;
(FC3) Se para algum $j, x_j \rightarrow -\infty$, então $F(\mathbf{x}) \rightarrow 0$ e, ainda, se para todo $j, x_j \rightarrow \infty$, então $F(\mathbf{x}) \rightarrow 1$;
(FC4) $F(\mathbf{x})$ é tal que, para $\forall\, a_i, b_i \in \mathbb{R}, a_i < b_i, 1 \leq i \leq m$, temos $P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2, \ldots, a_m < X_m \leq b_m) \geq 0$.

Demonstração:

As propriedades são verificadas, observando-se o comportamento dos respectivos eventos.

Para (FC1), considere um j qualquer fixado e $a_j \leq b_j$. Observe que o conjunto $\bigcap_{\substack{i=1 \\ i \neq j}}^{m} \{w : X_i(w) \leq x_i\} \cap \{w : X_j(w) \leq a_j\}$ está contido na seguinte intersecção de conjuntos: $\bigcap_{\substack{i=1 \\ i \neq j}}^{m} \{w : X_i(w) \leq x_i\} \cap \{w : X_j(w) \leq b_j\}$. Logo, a aplicação de probabilidade nesses conjuntos, verifica a propriedade.

A demonstração de (FC2) é análoga ao caso univariado.

Para (FC3), seja x_j tal que $x_j \rightarrow -\infty$ então $\{w : X_j(w) \leq x_j\} \downarrow \emptyset$ e, portanto, $F(\mathbf{x}) \rightarrow 0$. Se todos os $x_{j's}$ vão para ∞, então cada um dos eventos

$\{w : X_j(w) \leq x_j\}$ cresce para Ω. Assim, a intersecção de todos eles cresce também para Ω. Dessa forma, a propriedade (FC3) está demonstrada.

A propriedade (FC4) é válida pela definição de probabilidade, desde que estejamos calculando a probabilidade de um conjunto de $\mathcal{F}$. Mas esse é o caso, pois para cada $i, i = 1, 2, \cdots, m$ temos

$$\{w : a_i < X_i(w) \leq b_i\} \in \mathcal{F} \text{ e, portanto, } \bigcap_{i=1}^{m} \{w : a_i < X_i(w) \leq b_i\} \in \mathcal{F};$$

e, assim, a demonstração da proposição está completa. □

Essas propriedades também podem servir como definição. Isto é, qualquer função de $\mathbb{R}^m$ em $[0, 1]$, satisfazendo (FC1)-(FC4), é função de distribuição conjunta de algum vetor aleatório. A demonstração não será feita e o leitor interessado pode consultar Tucker [67].

A propriedade (FC4) parece tão óbvia que poderíamos questionar a necessidade de mencioná-la. Cabe notar que, no caso unidimensional, não foi necessário incluirmos (FC4) para caracterizar a função de distribuição, mas no caso multidimensional ela é essencial. Como veremos no exemplo abaixo, existem funções multivariadas que satisfazem todas as propriedades exceto (FC4), não sendo, portanto, função de distribuição de nenhum vetor aleatório.

Uma forma conveniente de apresentar a propriedade (FC4) é através da sucessiva aplicação de operadores-diferença. Sendo I_i o intervalo $(a_i, b_i]$, o operador resulta na diferença entre os valores da função quando somente a coordenada i é alterada de a_i para b_i. Denotando por $\triangle_{i,I_i}$ temos, por exemplo,

$$\triangle_{2,I_2} F(\mathbf{x}) = F(x_1, b_2, \ldots, x_m) - F(x_1, a_2, \ldots, x_m).$$

Assim, a propriedade (FC4) pode ser escrita como $\triangle_{1,I_1} \ldots \triangle_{m,I_m} F(\mathbf{x}) \geq 0$.

Exemplo 3.3: (James [81])

Considere a seguinte função (ver região S em figura a seguir):

$$G(x, y) = \begin{cases} 1, & \text{em } S = \{(x, y) : x \geq 0, y \geq 0 \text{ e } x + y \geq 1\}; \\ 0, & \text{caso contrário.} \end{cases}$$

A região S é apresentada em figura a seguir

As propriedades (FC1)-(FC3) são verificadas sem dificuldade. Por exemplo, para (FC3), tomando um ponto $(x, y) \in S$, observamos que a continuidade à direita se verifica nas duas coordenadas, pois a fronteira faz parte de S. Também, para pontos $(x, y) \notin S$, a continuidade à direita é verificada, uma vez que G tem sempre o valor 0 nessa região.

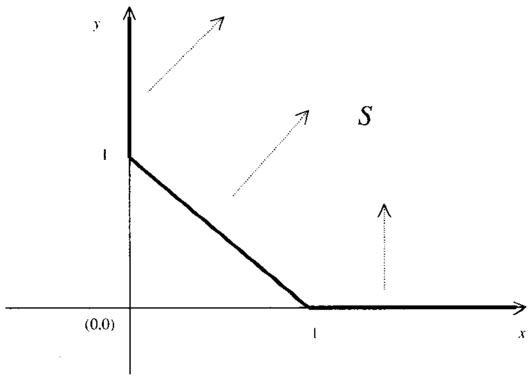

Figura 3.1: Região S para a função $G(x, y)$.

A função G não satisfaz (FC4) e, portanto, não pode ser função de distribuição conjunta. Observe que

$$P(0 < X \leq 1, 0 < Y \leq 1) = G(1,1) - G(1,0) - G(0,1) + G(0,0) = -1,$$

que contradiz (FC4). □

No cálculo de probabilidades com vetores aleatórios, levamos em conta a classificação de cada uma das variáveis envolvidas. As definições, apresentadas a seguir, auxiliam os cálculos nos vários casos de interesse. Inicialmente, obtemos a distribuição de cada variável isoladamente, dada a função de distribuição conjunta do vetor aleatório. Note que, exceto em situações especiais, não se pode obter a conjunta do vetor partindo das funções de distribuição de cada componente.

Definição 3.3: Função de Distribuição Marginal

Seja $F(\mathbf{x})$ a função de distribuição de $(X_1, X_2, \ldots, X_m)$. Para cada k, $k = 1, 2, \cdots, m$, definimos a *função de distribuição marginal* de X_k por:

$$F_{X_k}(x_k) = \lim_{\substack{x_i \to \infty \\ \forall i \neq k}} F(\mathbf{x}),$$

em que o limite é aplicado em todas as coordenadas, exceto k. □

Definição 3.4: Vetor Discreto, Probabilidade Conjunta e Marginal

Se as variáveis do vetor aleatório são discretas, temos um *vetor aleatório discreto*. Sua *função de probabilidade conjunta* é definida da seguinte forma:

$$p(\mathbf{x}) = p(x_1, x_2, \cdots, x_m) = P(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m).$$

A *função de probabilidade marginal* de X_k, $k = 1, 2, \cdots, m$ é dada por:

$$p_{X_k}(x_k) = P(X_k = x_k) = \sum_{\substack{x_i \\ \forall i \neq k}} p(\mathbf{x}) = \sum_{\substack{x_i \\ \forall i \neq k}} P(X_1 = x_1, \ldots, X_m = x_m);$$

ou seja, a marginal vem da soma em todas as coordenadas, exceto k. □

Note que a idéia da marginal se aplica para mais de uma variável. Se desejamos a função de probabilidade conjunta (marginal) de algumas das m variáveis, somamos em todas as coordenadas, exceto as das variáveis de interesse.

Proposição 3.2: Propriedades da Função de Probabilidade Conjunta

Seja $\mathbf{X}$ um vetor aleatório discreto em $(\Omega, \mathcal{F}, \mathcal{P})$. Então sua função de probabilidade conjunta (fpc) satisfaz as propriedades:

(fpc1) $p(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \mathbb{R}^m$;

(fpc2) $\sum_{\mathbf{x}} p(\mathbf{x}) = 1$.

Demonstração:

As provas são análogas ao caso univariado e serão omitidas. □

A função de probabilidade e a função de distribuição podem ser obtidas, uma a partir da outra, de forma similar ao que é feito no caso unidimensional.

No caso especial de duas variáveis discretas, é conveniente a representação da função de probabilidade através da *tabela de dupla entrada*:

$X_1 \backslash X_2$	b_1	$\cdots$	b_s	$P(X_1 = x_1)$
a_1	$p(a_1, b_1)$	$\cdots$	$p(a_1, b_s)$	$p_{X_1}(a_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_r	$p(a_r, b_1)$	$\cdots$	$p(a_r, b_s)$	$p_{X_1}(a_r)$
$P(X_2 = x_2)$	$p_{X_2}(b_1)$	$\cdots$	$p_{X_2}(b_s)$	1

Nessa tabela, o corpo representa a distribuição conjunta e as últimas linha e coluna fornecem as distribuições marginais. Na tabela apresentada X_1 assume os valores $a_1, a_2, \cdots, a_r$ e X_2 os valores $b_1, b_2, \cdots, b_s$. Observe que as distribuições marginais das duas variáveis são obtidas através da soma de linhas e colunas correspondentes.

Exemplo 3.4: Duas moedas equilibradas são lançadas de forma independente e definimos as variáveis aleatórias X e Y da seguinte forma:

X : número de caras nos dois lançamentos;

Y : função indicadora de faces iguais nos dois lançamentos.

O espaço de probabilidade $(\Omega, \mathcal{F}, \mathcal{P})$ pode ser construído como usual. Consideramos Ω todos os 4 pares de resultados possíves e $\mathcal{F}$ a σ-álgebra das partes de Ω. A função de probabilidade $\mathcal{P}$ atribui igual probabilidade aos elementos de Ω.

Para calcular as probabilidades conjuntas sempre nos reportamos aos pontos de Ω, pois as variáveis aleatórias são funções das ocorrências do espaço amostral. Representando por C e K, os eventos cara e coroa, respectivamente, temos:

Eventos	prob.	X	Y
(C, C)	$1/4$	2	1
(C, K)	$1/4$	1	0
(K, C)	$1/4$	1	0
(K, K)	$1/4$	0	1

Coletando os valores comuns dos pares (X, Y), obtemos a probabilidade conjunta apresentada, a seguir, na forma de uma tabela de dupla entrada.

$X\backslash Y$	0	1	$P(X = x)$
0	0	$1/4$	$1/4$
1	$1/2$	0	$1/2$
2	0	$1/4$	$1/4$
$P(Y = y)$	$1/2$	$1/2$	1

Para calcular a função de distribuição conjunta de X e Y é necessário separar em vários casos, percorrendo intervalos de valores em cada variável. O gráfico a seguir apresenta as várias regiões, correspondentes aos diversos casos. As fronteiras de cada região não foram assinaladas, para evitar congestionar o gráfico.

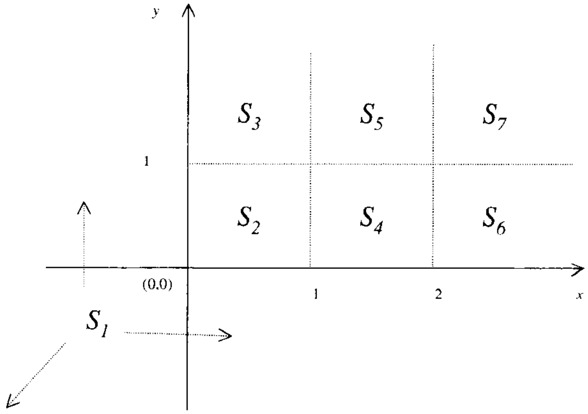

Figura 3.2: Partição dos valores de (X, Y).

Dessa forma, a função de distribuição conjunta pode ser expressa da seguinte forma:

$$F_{X,Y}(x,y) = \begin{cases} 0, & \text{em } S_1 = \{(x,y) : x < 0 \text{ ou } y < 0\}; \\ 0, & \text{em } S_2 = \{(x,y) : 0 \leq x < 1, 0 \leq y < 1\}; \\ 1/4, & \text{em } S_3 = \{(x,y) : 0 \leq x < 1, y \geq 1\}; \\ 1/2, & \text{em } S_4 = \{(x,y) : 1 \leq x < 2, 0 \leq y < 1\}; \\ 3/4, & \text{em } S_5 = \{(x,y) : 1 \leq x < 2, y \geq 1\}; \\ 1/2, & \text{em } S_6 = \{(x,y) : x \geq 2, 0 \leq y < 1\}; \\ 1, & \text{em } S_7 = \{(x,y) : x \geq 2, y \geq 1\}. \end{cases}$$

Observe que podemos resumir a apresentação da função se agregarmos os valores comuns. Entretanto, optamos por essa forma inicial de apresentação para evidenciar o cuidado que é preciso ter no estabelecimento das regiões convenientes para o cálculo de $F_{X,Y}$. Fica a cargo do leitor reescrever a função numa forma mais resumida e também verificar que estão satisfeitas as propriedades de função de distribuição. □

Exemplo 3.5: A tabela abaixo apresenta valores de probabilidade conjunta de duas variáveis X e Y em função de constantes reais a e b. Vamos determinar condições para seus valores, de modo que as propriedades da função de probabilidade conjunta estejam satisfeitas.

$X \setminus Y$	0	1	2	3
0	a	$a+b-\frac{1}{5}$	b	$\frac{1}{5}$
1	$\frac{1}{5}$	$2(a+b-\frac{1}{5})$	$\frac{1}{5}-a$	$\frac{1}{5}-b$

A primeira condição requer que as probabilidades conjuntas sejam não negativas. Assim, temos que atender simultaneamente às desigualdades:

$$a \geq 0; a+b-\frac{1}{5} \geq 0; b \geq 0; \frac{1}{5}-a \geq 0; \frac{1}{5}-b \geq 0.$$

Logo ficamos restritos à $R_1 = \left\{(a,b) : 0 \leq a \leq \frac{1}{5}, 0 \leq b \leq \frac{1}{5}, a+b \geq \frac{1}{5}\right\}$.

Impondo que a soma de todas as probabilidades conjuntas seja igual a 1, vem

$$a + (a+b-\frac{1}{5}) + b + \cdots + (\frac{1}{5} - b) = 1.$$

Daí, resulta em $R_2 = \left\{(a,b) : a+b = \frac{4}{15}\right\}$. Os valores de a e b precisam estar na intersecção dos conjuntos R_1e R_2. Assim,

$$R = R_1 \cap R_2 = \left\{(a,b) : \frac{1}{15} \leq a \leq \frac{1}{5}, \frac{1}{15} \leq b \leq \frac{1}{5}, a+b = \frac{4}{15}\right\}.$$

Tomando as constantes no conjunto R, garantimos que a tabela acima representa uma função de probabilidade conjunta de duas variáveis. □

Definição 3.5: Vetor Contínuo, Densidade Conjunta e Marginal

Denominamos *vetor aleatório contínuo* o vetor aleatório cujas componentes são variáveis aleatórias contínuas. Em outras palavras, um vetor aleatório é contínuo se, dada a função de distribuição F, existe uma função $f : \mathbb{R}^m \longrightarrow \mathbb{R}^+$, denominada *função densidade conjunta* (fdc), tal que

$$F(\mathbf{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} f(\mathbf{y})\, dy_1\, dy_2 \cdots dy_m.$$

A *função densidade marginal* é dada pela expressão:

$$f_{X_k}(x_k) = \underbrace{\int_{x_1} \cdots \int_{x_m}}_{(i \neq k)} f(\mathbf{x})\, \underbrace{dx_1\, dx_2 \cdots dx_m}_{(i \neq k)}.$$

□

Note que a função densidade $f(\mathbf{x})$ é obtida a partir de F por sucessivas derivadas parciais, generalizando, assim, o caso univariado. Se desejamos a marginal conjunta de l variáveis, procedemos de modo similar ao caso discreto, integrando todas as variáveis exceto as de interesse.

Proposição 3.3: Propriedades da Função Densidade Conjunta

Seja $\mathbf{X}$ um vetor aleatório contínuo em $(\Omega, \mathcal{F}, \mathcal{P})$. Então sua função densidade conjunta (fdc) satisfaz as propriedades:

(fdc1) $f(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \mathbb{R}^m$;

(fdc2) $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\mathbf{x})\, dx_1\, dx_2 \cdots dx_m = 1.$

Demonstração:

As provas serão omitidas, mas são similares ao caso univariado. □

Exemplo 3.6: A função de distribuição conjunta de X e Y é dada por:

$$F_{X,Y}(x,y) = \begin{cases} 0, & \text{se } x < 0 \text{ ou } y < 0; \\ \frac{x}{5}(1 - e^{-y}), & \text{se } 0 \leq x < 5, y \geq 0; \\ 1 - e^{-y}, & \text{se } x \geq 5, y \geq 0. \end{cases}$$

Vamos verificar as propriedades da função de distribuição na região $0 \leq x < 5$ e $y \geq 0$. Nas demais regiões, as provas são similares.

Para provar (FC1) escolha $x_1 \leq x_2$ em $[0, 5)$ e fixe $y \geq 0$. Então,

$$F_{X,Y}(x_1, y) = \frac{x_1}{5}(1 - e^{-y}) \leq \frac{x_2}{5}(1 - e^{-y}) = F_{X,Y}(x_2, y).$$

Considere, agora, $y_1 \leq y_2$ em $[0, \infty)$ mantendo x fixo ($x \in [0, 5)$). Temos,

$$F_{X,Y}(x, y_1) = \frac{x}{5}(1 - e^{-y_1}) \leq \frac{x}{5}(1 - e^{-y_2}) = F_{X,Y}(x, y_2),$$

uma vez que $e^{-y_1} \geq e^{-y_2}$. Logo, $F_{X,Y}$ é não decrescente em cada uma das coordenadas e (FC1) está satisfeita nessa região (complete para as outras!).

As propriedades (FC2) e (FC3) são verificadas sem dificuldade.

Para (FC4) devemos provar que $F_{X,Y}(x,y)$ é tal que, para $\forall a_i, b_i \in \mathbb{R}$, $a_i < b_i$ e intervalo $I_i = (a_i, b_i]$, $1 \leq i \leq 2$, temos

$$P(a_1 < X \leq b_1, a_2 < Y \leq b_2) = \triangle_{1,I_1}\triangle_{2,I_2}F_{X,Y}(x,y) \geq 0.$$

Podemos escrever essa condição (ver exercício), da seguinte forma:

$$F_{X,Y}(b_1,b_2) + F_{X,Y}(a_1,a_2) - F_{X,Y}(a_1,b_2) - F_{X,Y}(b_1,a_2) \geq 0.$$

Considere que I_i, $i = 1,2$ estão na região $0 \leq x < 5$ e $y \geq 0$. Temos

$$\frac{b_1}{5}(1 - e^{-b_2}) + \frac{a_1}{5}(1 - e^{-a_2}) - \frac{a_1}{5}(1 - e^{-b_2}) - \frac{b_1}{5}(1 - e^{-a_2}),$$

que, após simplificação, torna-se $\frac{(b_1-a_1)}{5}(e^{-a_2} - e^{-b_2})$. Essa é uma quantidade não negativa, pois $a_2 < b_2$. Para as outras escolhas de região para os intervalos I_i 's, a verificação de (FC4) é análoga.

Assim, valem as propriedades de função de distribuição conjunta.

Para obter as funções de distribuição marginais, passamos ao limite, lembrando que $F_{X,Y}(x,y)$ é nula se $x < 0$ ou $y < 0$:

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x,y) = \begin{cases} \lim\limits_{y\to\infty} \frac{x}{5}(1 - e^{-y}) = x/5, & \text{se } 0 \leq x < 5; \\ \lim\limits_{y\to\infty} (1 - e^{-y}) = 1, & \text{se } x \geq 5. \end{cases}$$

$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x,y) = \lim_{x\to\infty} (1 - e^{-y}) = 1 - e^{-y}, \; y \geq 0.$$

As funções densidades podem ser obtidas, a partir da derivação das respectivas funções de distribuição.

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x\,\partial y}F_{X,Y}(x,y) = \begin{cases} \frac{1}{5}e^{-y}, & \text{se } 0 \leq x \leq 5, y \geq 0; \\ 0, & \text{caso contrário.} \end{cases}$$

$$f_X(x) = \frac{\partial}{\partial x}F_X(x) = \begin{cases} \frac{1}{5}, & \text{se } 0 \leq x \leq 5; \\ 0, & \text{caso contrário.} \end{cases}$$

$$f_Y(y) = \frac{\partial}{\partial y}F_Y(y) = \begin{cases} e^{-y}, & \text{se } y \geq 0; \\ 0, & \text{caso contrário.} \end{cases}$$

Note que as densidades de X e Y correspondem aos modelos $U_c(0,5)$ e *Exp*(1), respectivamente. Elas poderiam, também, ser obtidas através das marginais da densidade conjunta. □

Exemplo 3.7: (Dudewicz e Mishra [88])

Vamos obter o valor de k, de modo que a função abaixo seja a densidade conjunta de três variáveis aleatórias contínuas X, Y e Z.

$$f(x,y,z) = \begin{cases} kxy^2z & \text{se } 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq \sqrt{2}; \\ 0 & \text{caso contrário.} \end{cases}$$

A função será positiva, se k for positivo. Para valer a segunda propriedade da função densidade conjunta, impomos que a integral tripla de menos a mais infinito seja 1. Isto é

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

Temos

$$\int_0^{\sqrt{2}}\int_0^1\int_0^1 kxy^2z\, dx\, dy\, dz = \int_0^{\sqrt{2}}\int_0^1 ky^2z\,\frac{1}{2}\, dy\, dz = \int_0^{\sqrt{2}} kz\,\frac{1}{6}\, dz = \frac{k}{6}.$$

Segue então que $k = 6$.

A marginal de X pode ser obtida, integrando-se em y e z para todos os valores possíveis. Temos, para $0 \leq x < 1$,

$$f_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\, dy\, dz = \int_0^{\sqrt{2}}\int_0^1 6xy^2z\, dy\, dz = 2x.$$

Pela conjunta, sabemos que X tem valores apenas em $[0, 1]$ e, portanto,

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

Logo, $X \sim Beta(2, 1)$. As outras marginais são calculadas de modo similar. □

Se as variáveis não são do mesmo tipo, ainda é possível a obtenção da função de distribuição conjunta. Nesses casos não existe nomenclatura padrão. Para a situação em que temos mistura da densidade com a função de probabilidade, usaremos a notação das variáveis contínuas, indicando por f essas funções. As propriedades são equivalentes aos casos sem mistura.

Exemplo 3.8: Função Mista (DeGroot e Schervich [02])

Considere duas variáveis aleatórias X e Y, sendo X discreta e Y contínua. Nós podemos usar a função mista de probabilidade conjunta para caracterizar o comportamento probabilístico conjunto dessas variáveis. Para calcular a probabilidade de (X, Y) pertencer a uma certa região do plano xy, somamos os valores para a variável X e integramos para os de Y.

Vamos supor que a função mista de probabilidade conjunta de X e Y é dada por:

$$f(x,y) = \begin{cases} \frac{xy^{x-1}}{3}, & \text{se } x = 1, 2, 3 \text{ e } 0 \leq y \leq 1; \\ 0, & \text{caso contrário.} \end{cases}$$

Para verificar se essa função poderá gerar probabilidades, observamos que é positiva e vamos mostrar que ao percorrer todos os valores possíveis para X e Y, via soma ou integral conforme o caso, obtemos o valor 1. Existindo o resultado dessa dupla operação, podemos fazê-la em qualquer ordem. Entretanto, é mais fácil fazer a integral primeiro. Temos

$$\sum_{x=1}^{3} \int_0^1 f(x,y)dy = \sum_{x=1}^{3} \int_0^1 \frac{xy^{x-1}}{3} dy = \sum_{x=1}^{3} \frac{1}{3} = 1;$$

logo $f(x, y)$ pode ser considerada uma probabilidade. Por exemplo,

$$P(X \geq 2, Y \geq 1/2) = \sum_{x=2}^{3} \int_{1/2}^{1} \frac{xy^{x-1}}{3} dy = \sum_{x=2}^{3} \left(\frac{1-(1/2)^x}{3}\right) = 13/24.$$

Para obter a função mista de distribuição conjunta, teremos que separar várias regiões de valores do par (x, y). Após algum trabalho, obtemos:

$$F(x,y) = \begin{cases} 0, & \text{se } x < 1 \text{ ou } y < 0; \\ \frac{y}{3}, & \text{se } 1 \leq x < 2, 0 \leq y < 1; \\ \frac{y+y^2}{3}, & \text{se } 2 \leq x < 3, 0 \leq y < 1; \\ \frac{y+y^2+y^3}{3}, & \text{se } x \geq 3, 0 \leq y < 1; \\ \frac{1}{3}, & \text{se } 1 \leq x < 2, y \geq 1; \\ \frac{2}{3}, & \text{se } 2 \leq x < 3, y \geq 1; \\ 1, & \text{se } x \geq 3, y \geq 1. \end{cases}$$

□

Ao tratar com várias variáveis é comum haver interesse em obter o comportamento condicional de uma delas em função de outras. Os valores das variáveis induzem eventos e, portanto, ao condicionar na ocorrência de certos valores de uma variável, estamos, de fato, condicionando em eventos. Dessa forma, a definição de probabilidade condicional pode também ser aplicada para o condicionamento entre variáveis aleatórias.

Observe que, na definição de uma variável aleatória X em $(\Omega, \mathcal{F}, \mathcal{P})$, requeremos que o conjunto $\{w : X(w) \leq x\}$ estivesse em $\mathcal{F}$, para todo x real. Assim, a ocorrência de valores de uma variável pode ser descrita de duas formas: pelo próprio conjunto de valores assumidos, isto é, por um elemento de $\mathcal{B}(\mathbb{R})$ (σ-álgebra de Borel em $\mathbb{R}$) ou por um evento da σ-álgebra $\mathcal{F}$ que foi induzido por esse boreliano através da variável aleatória. Num estudo mais formal do condicionamento em variáveis aleatórias, essas duas formas podem ser usadas alternativamente conforme a conveniência da situação.

Vamos restringir nossa apresentação ao caso bivariado, mas as idéias podem ser estendidas para o caso multivariado.

Definição 3.6: Distribuição Condicional

Sejam X e Y duas variáveis em $(\Omega, \mathcal{F}, \mathcal{P})$ e $B_1, B_2 \in \mathcal{B}(\mathbb{R})$ com $P(Y \in B_2) > 0$. Definimos a *distribuição condicional de X dado $Y \in B_2$*, pela expressão

$$P(X \in B_1 | Y \in B_2) = \frac{P([X \in B_1] \cap [Y \in B_2])}{P(Y \in B_2)}.$$

Se $P(Y \in B_2) = 0$, definimos $P(X \in B_1 | Y \in B_2) = P(X \in B_1)$. □

Apresentamos, a seguir, algumas fórmulas úteis para o cálculo com probabilidades condicionais. Iniciamos com o condicionamento em uma variável discreta.

Seja Y uma variável discreta e $y \in \mathbb{R}$ tal que $P(Y = y) > 0$. Sendo X uma variável qualquer e $B \in \mathcal{B}(\mathbb{R})$, temos

$$P(X \in B | Y = y) = \frac{P([X \in B] \cap [Y = y])}{P(Y = y)},$$

que, com alguma manipulação, resulta na expressão:

$$P(X \in B) = \sum_y P(Y = y) P(X \in B | Y = y).$$

Essa é uma versão, para variáveis aleatórias, da Lei da Probabilidade Total.

Tomando $B = (-\infty, x]$, $x \in \mathbb{R}$, podemos obter a *função de distribuição condicional de X, dado* $Y = y$:

$$F_{X|Y}(x|\, Y = y) = \frac{P([X \leq x] \cap [Y = y])}{P(Y = y)}.$$

Como conseqüência, temos:

$$F_X(x) = \sum_y P(Y = y) F_{X|Y}(x|\, Y = y).$$

Em algumas situações, é útil escrever a função de distribuição conjunta a partir da condicional:

$$F_{X,Y}(x, y) = \sum_{k: k \leq y} P(Y = k) F_{X|Y}(x|\, Y = k).$$

Essa expressão é similar à regra do produto de probabilidades, aplicada aqui para variáveis aleatórias e função de distribuição.

Se X também for discreta, a probabilidade condicional pode ser reescrita na notação de função de probabilidade da seguinte forma:

$$p_{X|Y}(x|\, y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}.$$

As outras fórmulas se adaptam convenientemente.

Vamos considerar, agora, o condicionamento em variáveis contínuas. Sendo Y contínua, é freqüente um abuso de notação indicando um condicionamento em eventos do tipo $[Y = y]$ que tem probabilidade zero. Nesse caso, a probabilidade condicional deve ser entendida como o limite de probabilidades condicionadas em $[y - \epsilon \leq Y \leq y + \epsilon]$ quando $\epsilon \to 0$. Isto é, para X uma variável qualquer e $B \in \mathcal{B}(\mathbb{R})$,

$$P(X \in B|\, Y = y) = \lim_{\epsilon \to 0} P(X \in B|\, y - \epsilon \leq Y \leq y + \epsilon).$$

Se X e Y são ambas variáveis contínuas, podemos usar a interpretação acima para definir a *densidade condicional de X dado Y*:

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

cuja expressão nada mais é do que a conjunta dividida pela marginal. A função de distribuição condicional é definida de modo similar. Observe que o relacionamento entre função densidade e função de distribuição também se preserva com o condicionamento. Isto é,

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(z|y)\, dz.$$

Como conseqüência, as seguintes fórmulas são também válidas:

$$F_{X,Y}(x,y) = \int_{-\infty}^{y} f_Y(z)\, F_{X|Y}(x|z)\, dz;$$
$$F_X(x) = \int_{-\infty}^{\infty} f_Y(y)\, F_{X|Y}(x|y)\, dy.$$

Ressaltamos que, de um modo mais geral, a função de distribuição condicional pode ser caracterizada como a solução das seguintes equações (já apresentadas):

$$F_{X,Y}(x,y) = \sum_{k:k\leq y} P(Y=k) F_{X|Y}(x|\,Y=k), \text{ se } Y \text{ é discreta;}$$
$$F_{X,Y}(x,y) = \int_{-\infty}^{y} f_Y(z)\, F_{X|Y}(x|z)\, dz, \text{ se } Y \text{ é contínua.}$$

Assim, conhecendo-se a conjunta e a marginal, a função de distribuição condicional será aquela que satisfaz a conveniente equação, dentre as duas acima.

Exemplo 3.9: (Continuação do Exemplo 3.8)

No Exemplo 3.8, a variável X é discreta, Y é contínua e sua função mista de probabilidade conjunta é a seguinte:

$$f(x,y) = \begin{cases} \frac{xy^{x-1}}{3}, & \text{se } x = 1,2,3 \text{ e } 0 \leq y \leq 1; \\ 0, & \text{caso contrário.} \end{cases}$$

Vamos obter algumas probabilidades condicionais calculando suas expressões, apenas, nos intervalos em que não se anulam. Temos

$$f(y|\,X=2)=\frac{f(2,y)}{P(X=2)}=\frac{2y/3}{1/3}=2y,\ 0\le y\le 1;$$
$$f(x|\,Y=1/2)=\frac{f(x,1/2)}{f_Y(1/2)}=\frac{\frac{1}{3}[x(1/2)^{x-1}]}{11/12}=\frac{4}{11}[x(1/2)^{x-1}],x=1,2,3;$$

os cálculos das marginais foram omitidos, mas podem ser feitos sem dificuldade.

Para as funções de distribuição condicionais teríamos:

$$\text{Para } 0\le y\le 1\text{: } F_{Y|X}(y|\,X=2)=\int_{-\infty}^{y}f(z|\,X=2)\,dz=\int_{-\infty}^{y}2z\,dz=y^2,$$

e, portanto,

$$F_{Y|X}(y|\,X=2)=\begin{cases}0, & \text{se } y<0;\\ y^2, & \text{se } 0\le y\le 1;\\ 1, & \text{se } y>1.\end{cases}$$

Por outro lado, para $F_{X|Y}(x|\,Y=1/2)$, $x\in\mathbb{R}$ temos

$$F_{X|Y}(x|\,Y=1/2)=\sum_{k=1}^{\min(3,[x])}f_{X|Y}(k|\,Y=1/2);$$

que resulta em

$$F_{X|Y}(x|\,Y=1/2)=\begin{cases}0, & \text{se } x<1;\\ \frac{4}{11}, & \text{se } 1\le x<2;\\ \frac{8}{11}, & \text{se } 2\le x<3;\\ 1, & \text{se } x\ge 3.\end{cases}$$

Outras quantidades podem ser obtidas de modo similar. □

Exemplo 3.10: As variáveis X e Y têm densidade dada por

$$f(x,y)=(x+y),0\le x\le 1,0\le y\le 1.$$

Vamos obter a densidade condicional $f_{X|Y}$. Inicialmente, calculamos a marginal de Y. Temos

$$f_Y(y)=\int_0^1(x+y)\,dx=y+\frac{1}{2},\ 0\le y\le 1.$$

Então, para qualquer y fixado, com $0 \leq y \leq 1$:

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{x+y}{y+1/2} = \frac{2(x+y)}{2y+1} \text{ ; para } 0 \leq x \leq 1.$$

Fora desse intervalo, $f_{X|Y}(x|y)$ é zero.

A função de distribuição condicional pode ser obtida a partir da integração da densidade condicional. Assim, para qualquer y fixado com $0 \leq y \leq 1$ e $0 \leq x \leq 1$, vem:

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{2(z+y)}{2y+1}\, dz = \frac{x(2y+x)}{2y+1},$$

e, portanto,

$$F_{X|Y}(x|y) = \begin{cases} 0, & \text{se } x < 0; \\ \frac{x(2y+x)}{2y+1}, & \text{se } 0 \leq x < 1; \\ 1, & \text{se } x \geq 1. \end{cases}$$

A função de distribuição conjunta pode ser obtida a partir da condicional, da seguinte forma:

$$F_{X,Y}(x,y) = \int_{-\infty}^{y} f_Y(z) F_{X|Y}(x|z)\, dz = \begin{cases} 0, & \text{se } x < 0 \text{ ou } y < 0; \\ \frac{1}{2}(x^2y + xy^2), & \text{se } 0 \leq x < 1, 0 \leq y < 1; \\ \frac{1}{2}(x + x^2), & \text{se } 0 \leq x < 1, y \geq 1; \\ \frac{1}{2}(y + y^2), & \text{se } x \geq 1, 0 \leq y < 1; \\ 1, & \text{se } x \geq 1, y \geq 1. \end{cases}$$

Deixamos ao leitor a tarefa de verificar se essa expressão está correta, o que pode ser feito calculando seu valor diretamente pela densidade conjunta. □

Um dos conceitos mais importantes em Probabilidade e Estatística é o da *independência entre variáveis*. Muitas das técnicas estatísticas só podem ser aplicadas em amostras cujos elementos foram supostamente coletados de forma independente. Nesses casos, as variáveis observadas em uma unidade amostral são independentes daquelas de uma outra unidade. De fato, a independência é um requisito importante que permite resolver, com rigor matemático e sem aproximações, muitos problemas de interesse prático. Como em muitas situações, a independência não pode ser assumida, sua imposição serve como uma aproximação da realidade e origina, nesses casos, uma solução também aproximada.

Definição 3.7: Independência entre Variáveis

Duas variáveis aleatórias, X e Y em $(\Omega, \mathcal{F}, \mathcal{P})$, são *independentes* se a informação sobre uma delas não altera a probabilidade de ocorrência da outra. Isto é, para qualquer $B_1, B_2 \in \mathcal{B}(\mathbb{R})$,

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

Uma definição alternativa e equivalente é:

$$X, Y \text{ independentes se, e somente se,}$$

$$\forall B_1, B_2 \in \mathcal{B}(\mathbb{R}), P(X \in B_1, Y \in B_2) = P(X \in B_1)\, P(Y \in B_2).$$

□

Note que não basta valer a igualdade para algumas escolhas dos borelianos B_1 e B_2 na σ-álgebra $\mathcal{B}(\mathbb{R})$. É necessário valer para todas.

A definição alternativa é muito útil e freqüentemente mais prática de ser usada. Em termos da função de distribuição, a definição alternativa pode ser escrita como:

$$X, Y \text{ independentes } \Leftrightarrow F_{X,Y}(x, y) = F_X(x)\, F_Y(y), \forall (x, y) \in \mathbb{R}^2.$$

Para as discretas, podemos escrever uma definição equivalente com o uso de funções de probabilidade:

$$X, Y \text{ independentes } \Leftrightarrow p_{X,Y}(x, y) = p_X(x)\, p_Y(y), \forall (x, y) \in \mathbb{R}^2.$$

Para as contínuas, a condição de independência usa as densidades:

$$X, Y \text{ independentes } \Leftrightarrow f_{X,Y}(x, y) = f_X(x)\, f_Y(y), \forall (x, y) \in \mathbb{R}^2,$$

em que ressaltamos que esta última igualdade deve valer exceto, eventualmente, em conjuntos com probabilidade zero.

Em geral, dizemos que as variáveis $X_1, X_2, \cdots, X_n$ definidas em um mesmo espaço de probabilidade $(\Omega, \mathcal{F}, \mathcal{P})$ são independentes, se $\forall B_i \in \mathcal{B}(\mathbb{R}), i = 1, 2, \cdots, n$, tivermos:

$$P(X_1 \in B_1, \cdots, X_n \in B_n) = \prod_{i=1}^{n} P(X_i \in B_i).$$

Segue dessa expressão que qualquer sub-família dessas variáveis também é formada por variáveis independentes. A verificação é simples, bastando tomar os convenientes B_i's $= \mathbb{R}$.

Exemplo 3.11: Com base em resultados do Posto de Saúde do bairro, estabeleceu-se a função de probabilidade conjunta entre os números diários de crianças atendidas com alergia (X) e com pneumonia (Y). Na tabela abaixo, apresentamos a conjunta e as marginais para essas variáveis.

$X \backslash Y$	0	1	2	$P(X = x)$
0	1/16	1/16	1/8	1/4
1	1/8	1/8	0	1/4
2	1/16	1/8	1/8	5/16
3	0	1/8	1/16	3/16
$P(Y = y)$	1/4	7/16	5/16	1

Se as variáveis fossem independentes, o produto das marginais daria a função de probabilidade conjunta. Entretanto, isso não ocorre pois

$$P(X = 1, Y = 2) = 0 \neq P(X = 1) \times P(Y = 2) = 5/64.$$

Pela definição de independência, basta haver um par em que isto não aconteça para quebrar a propriedade de independência. Portanto, os números diários de crianças atendidas com alergia e com pneumonia são variáveis dependentes.

Condicionado à ocorrência ou não de casos de pneumonia, como se comporta o número de crianças alérgicas? Pela dependência, esperamos comportamentos diferentes. Vamos calcular a função de probabilidade condicional de X dado o evento $[Y > 0]$ e comparar seu resultado com o condicionamento em $[Y = 0]$. Temos

$$\begin{aligned} P(X = 0|Y > 0) &= \frac{P(X = 0, Y > 0)}{P(Y > 0)} \\ &= \frac{P(X = 0, Y = 1) + P(X = 0, Y = 2)}{P(Y = 1) + P(Y = 2)} \\ &= 1/4; \end{aligned}$$

enquanto

$$P(X = 0|Y = 0) = \frac{P(X = 0, Y = 0)}{P(Y = 0)} = \frac{1/16}{1/4} = 1/4.$$

Assim, a probabilidade de não haver casos de crianças alérgicas $(X = 0)$, não sofre influência da ocorrência ou não de casos de pneumonia. Entretanto, para outros valores de X a situação é diferente, conforme apresentamos na tabela abaixo:

x	$P(X = x \mid Y > 0)$	$P(X = x \mid Y = 0)$
0	1/4	1/4
1	1/6	1/2
2	1/3	1/4
3	1/4	0

Observe que as probabilidades condicionais somente são iguais para $[X = 0]$. Para os outros valores, a probabilidade condicional de X é crescente ou decrescente, se condicionada ao evento $[Y > 0]$ ou $[Y = 0]$, respectivamente. Dessa forma, enfatizamos que as variáveis são dependentes. □

Exemplo 3.12: A função densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x, y) = \frac{1}{2x} I_{(0,2]}(x) I_{(0,x]}(y).$$

As marginais são calculadas através de integração:

$$f_X(x) = \int_0^x \frac{1}{2x}\, dy = \frac{1}{2},\ 0 < x \leq 2;$$
$$f_Y(y) = \int_y^2 \frac{1}{2x}\, dx = \frac{1}{2}(\log 2 - \log y),\ 0 < y \leq 2.$$

É imediato verificar que $f_X(x)$ satisfaz as propriedades de uma densidade. Observe que para calcular a densidade de Y, os limites de integração obedeceram à condição de que os valores da variável X precisam ser sempre superiores ao de Y. A expressão obtida é efetivamente uma densidade, pois é não negativa (a função logarítmica é crescente) e também

$$\int_0^2 f_Y(y)\, dy = \int_0^2 \frac{1}{2}(\log 2 - \log y) dy = \Big[\frac{1}{2}(y \log 2 - y(\log y - 1))\Big]_0^2 = 1.$$

As densidades condicionais são calculadas pelo quociente entre a conjunta e a marginal. Vamos indicar apenas os valores não nulos, temos:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1}{x(\log 2 - \log y)}, y \leq x \leq 2, y \text{ fixado}, y \in (0,2);$$
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1}{x}, 0 < y \leq x, x \text{ fixado}, x \in (0,2).$$

É importante ressaltar que, no cálculo de densidades condicionais, a expresssão resultante é sempre função do valor usado no condicionamento. Dessa forma, $f_{X|Y}(x|y)$ depende de y. Se estamos integrando essa função em x, o valor y pode ser considerado constante. Por exemplo, para $y \in (0,2)$ vem

$$P(X < 1|Y = y) = \int_y^1 f_{X|Y}(x|y)\,dx = \begin{cases} 0, & \text{se } 1 \leq y < 2; \\ \int_y^1 \frac{1}{x(\log 2 - \log y)} dx, & \text{se } 0 < y < 1. \end{cases}$$

Então,

$$P(X < 1|Y = y) = \begin{cases} 0, & \text{se } 1 \leq y < 2; \\ \frac{\log y}{\log y - \log 2}, & \text{se } 0 < y < 1. \end{cases}$$

Temos também que, para $x \in (0,2)$, temos

$$\begin{aligned} P(Y < 1/2|X = x) &= \int_0^{\max(1/2,x)} f_{Y|X}(y|x)\,dy \\ &= \begin{cases} \int_0^x \frac{1}{x}\,dy, & \text{se } 0 < x \leq 1/2; \\ \int_0^{1/2} \frac{1}{x}\,dy, & \text{se } \frac{1}{2} < x \leq 2; \end{cases} \\ &= \begin{cases} 1, & \text{se } 0 < x \leq 1/2; \\ \frac{1}{2x}, & \text{se } \frac{1}{2} < x \leq 2. \end{cases} \end{aligned}$$

Outras probabilidades podem ser calculadas de modo similar. Deixamos ao leitor a tarefa de verificar que $f_{X|Y}(x|y)$ e $f_{Y|X}(y|x)$ satisfazem as propriedades de uma densidade.

Finalmente, observamos que as variáveis não são independentes, pois a densidade conjunta não é o produto das marginais. Ou seja, temos

$$f_{X,Y}(x,y) = \frac{1}{2x} I_{(0,2]}(x) I_{(0,x]}(y),$$

enquanto que

$$f_X(x)f_Y(y) = \frac{1}{4x}(\log 2 - \log y\,)I_{(0,2]}(x)I_{(0,2]}(y);$$

e elas não são iguais para quase todo $(x, y) \in \mathbb{R}^2$. Em outras palavras, existe um conjunto, com probabilidade positiva, em que a condição de independência não se verifica. □

Encerramos essa seção com a apresentação de dois modelos multivariados. O modelo apresentado a seguir generaliza a distribuição Binomial ao tratar com experimentos que têm mais de duas alternativas de resultado.

Definição 3.8: Modelo Multinomial

Considere um experimento com m possíveis resultados, cada um com probabilidade $p_i \geq 0$, $i = 1, 2, \cdots, m$ e $\sum\limits_{i=1}^{m} p_i = 1$. Esse experimento é repetido n vezes de forma independente e observamos as variáveis $X_1, X_2, \cdots, X_m$, que correspondem ao número de ocorrências de cada um dos possíveis resultados dessas repetições. O vetor aleatório $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ segue o modelo *Multinomial* com função de probabilidade conjunta dada por:

$$p_{\mathbf{X}}(k_1, k_2, \ldots, k_m) = \frac{n!}{k_1!\,k_2! \ldots k_m!}\, p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m},$$

com $\sum\limits_{i=1}^{m} p_i = 1$ e $\sum\limits_{i=1}^{m} k_i = n$, $k_i \in \mathbb{N}$, $0 \leq k_i \leq n$. □

Exemplo 3.13: Para 10 lançamentos independentes de um dado equilibrado seja X_i o número de ocorrências da face i, $i = 1, 2, \ldots, 6$. O vetor aleatório $\mathbf{X} = (X_1, X_2, \ldots, X_6)$ segue o modelo Multinomial com parâmetros $p_i = 1/6$, $i = 1, 2, \cdots, 6$. Assim,

$$p_{\mathbf{X}}(k_1, k_2, \ldots, k_6) = \frac{10!}{k_1!\,k_2! \ldots k_m!}\,(\frac{1}{6})^{10}, \text{ com } k_1 + k_2 + \ldots + k_6 = 10.$$

□

Um dos modelos multivariados contínuos mais importante é aquele que generaliza a distribuição Normal. Vamos apresentar o caso bi-variado e recomendamos a consulta às referências para o modelo geral.

Definição 3.9: Modelo Normal Bivariado

Dizemos que o vetor aleatório $\mathbf{X} = (X_1, X_2)$ segue o modelo *Normal Bivariado* se sua função densidade é dada por

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$
$$e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]},$$

com $(x_1, x_2) \in \mathbb{R}^2$ e os parâmetros $\mu_1, \mu_2, \sigma_1, \sigma_2$ e ρ satisfazem às seguintes condições: μ_1 e μ_2 estão em $\mathbb{R}$, $\sigma_1 > 0, \sigma_2 > 0$ e $-1 \leq \rho \leq 1$. □

A interpretação dos parâmetros do modelo Normal Bivariado depende de conceitos apresentados no Capítulo 5. Entretanto, adiantamos que μ_1 e μ_2 são as médias e σ_1 e σ_2 são os desvios-padrão de X_1 e X_2, respectivamente. O parâmetro ρ é o coeficiente de correlação entre as variáveis. Usualmente, os parâmetros μ_1 e μ_2 são denominados parâmetros de localização, enquanto que σ_1 e σ_2 são parâmetros de escala.

Exemplo 3.14: O vetor (X, Y) segue o modelo Normal Bivariado com parâmetros $\mu_X = \mu_Y = 1, \sigma_X = \sigma_Y = 1$ e $\rho = 1/2$. Então sua densidade conjunta é dada por:

$$f_{X,Y}(x, y) = \frac{1}{\pi\sqrt{3}} e^{-\frac{2}{3}(x^2+y^2-xy)}, (x, y) \in \mathbb{R}^2.$$

Para a marginal de X temos

$$\begin{aligned}
f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy &= \int_{-\infty}^{\infty} \frac{1}{\pi\sqrt{3}} e^{-\frac{2}{3}(x^2+y^2-xy)}dy \\
&= e^{-\frac{2}{3}x^2} \int_{-\infty}^{\infty} \frac{1}{\pi\sqrt{3}} e^{-\frac{2}{3}(y^2-xy)}dy \\
&= e^{-\frac{2}{3}x^2} e^{\frac{2}{3}\frac{x^2}{4}} \int_{-\infty}^{\infty} \frac{1}{\pi\sqrt{3}} e^{-\frac{(y-\frac{x}{2})^2}{\frac{3}{2}}} dy \\
&= \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{\frac{3}{4}}} e^{-\frac{(y-\frac{x}{2})^2}{2(\sqrt{\frac{3}{4}})^2}} dy,
\end{aligned}$$

em que, na última passagem, ajeitamos o integrando para obter a densidade correspondente a uma Normal com $\mu = x/2$ e $\sigma = \sqrt{\frac{3}{4}}$. Logo, concluímos que

$$f_X(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}, -\infty < x < \infty;$$

que corresponde à densidade da $N(0,1)$. De modo análogo, teríamos $Y \sim N(0,1)$.

As variáveis não são independentes pois verificamos, facilmente, que o produto das marginais não resulta na densidade conjunta. □

Exercícios da Seção 3.2:

1. Sejam X e Y variáveis aleatórias em $(\Omega, \mathcal{F}, \mathcal{P})$ com função de distribuição conjunta dada por $F_{X,Y}$. Mostre que $P(a_1 < X \leq b_1, a_2 < Y \leq b_2)$, com a_1, b_1, a_2 e $b_2 \in \mathbb{R}$, pode ser escrita como:

$$F_{X,Y}(b_1, b_2) + F_{X,Y}(a_1, a_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(b_1, a_2).$$

2. Mostre que H não é função de distribuição conjunta.

$$H(x,y) = \begin{cases} 1, & \text{se } \max(x,y) \geq 0 \text{ ou } x^2 + y^2 \leq 1; \\ 0, & \text{caso contrário.} \end{cases}$$

3. A densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x,y) = \frac{1}{2} I_A(x,y) \text{ com } A = \{(x,y) \in \mathbb{R}^2; -1 \leq x \leq 1, 0 \leq y \leq 1\}.$$

a. Calcule as marginais e verifique se X e Y são independentes.
b. Obtenha a densidade condicional de Y dado que $X > 0$.
c. Determine a função de distribuição conjunta entre X e Y.

4. Sejam X e Y com densidade conjunta:

$$f_{X,Y}(x,y) = \frac{1}{\pi} I_A(x,y) \text{ com } A = \{(x,y) \in \mathbb{R}; x^2 + y^2 \leq 1\}.$$

a. Calcule as marginais e verifique se X e Y são independentes.
b. Determine a densidade condicional de X dado que $Y = 1/2$.

5. A densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{4}, & \text{se } -2 \leq x < 0, 0 \leq y \leq 1; \\ \frac{1}{4}, & \text{se } 0 \leq x \leq 1, -2 \leq y < 0; \\ 0, & \text{caso contrário.} \end{cases}$$

a. Calcule as marginais e verifique se X e Y são independentes.
b. Determine a função de distribuição conjunta entre X e Y.
c. Obtenha a função de distribuição condicional de Y dado $X = x$.

6. Seja $Y \sim Poisson(\lambda)$ e $X|(Y = n) \sim B(n, p)$.
a. Calcule a função de probabilidade de X.
b. Determine a função de probabilidade condicional de Y dado $X = x$.

7. Um milionário excêntrico, uma vez por semana, deixa seu escritório com X milhares de reais no bolso. Ao caminhar para sua casa vai distribuindo esse dinheiro aos eventuais pedintes que encontra. Admita que X tem densidade de probabilidade $f_X(x) = \frac{x}{8} I_{(0,4)}(x)$ e, também que o dinheiro que lhe resta ao chegar em casa, denotado por Y, tem probabilidade uniforme entre zero e o dinheiro com que deixou o escritório. Isto é, $Y|(X = x) \sim U_c[0, x]$.
a. Calcule a densidade conjunta entre X e Y.
b. Determine a densidade marginal de Y.

8. A função de distribuição conjunta de (X, Y) é dada por:

$$F_{X,Y}(x,y) = \begin{cases} 0, & \text{se } x < 0 \text{ ou } y < 0 \text{ ou } x \geq y; \\ \frac{2x^2y^2 - x^4}{16}, & \text{se } 0 \leq x < y, 0 \leq y < 2; \\ \frac{8x^2 - x^4}{16}, & \text{se } 0 \leq x < 2, y \geq 2; \\ 1, & \text{se } x \geq 2, y \geq 2, x < y. \end{cases}$$

a. Obtenha as funções de distribuição marginais de X e Y.
b. Calcule a densidade conjunta entre X e Y.
c. Calcule as densidades marginais de X e Y de duas maneiras diferentes.

9. Considere a função:

$$f_{X|Y}(x|y) = \begin{cases} \frac{y^x e^{-y}}{x!}, & \text{se } x = 0, 1, \cdots \text{ e } y > 0; \\ 0, & \text{caso contrário.} \end{cases}$$

a. Mostre que, para cada y fixado, $f(.|y)$ é uma função de probabilidade.
b. Determine a conjunta de X e Y se $Y \sim Exp(1)$.

c. Nas condições de (b), obtenha a marginal de X.

10. Seja $F(x, y)$ a função mista de distribuição conjunta de (X, Y). Suponha que F corresponde à densidade Uniforme Contínua sobre o intervalo $(0, 1)$ do eixo y de $\mathbb{R}^2$.
a. Obtenha F.
b. Mostre que $F_X(x)$ é contínua exceto se $x = 0$.
c. Prove que $F_Y(y)$ é contínua.
d. Você diria que (X, Y) é contínua se, e só se, sua função de distribuição conjunta induz probabilidade zero a cada ponto $(x, y) \in \mathbb{R}^2$?

3.3 Funções de Variáveis Aleatórias

Sendo X uma variável aleatória em $(\Omega, \mathcal{F}, \mathcal{P})$, a função ou transformação $h : X \to \mathbb{R}$ também será uma variável aleatória no mesmo espaço de probabilidade. Dado o conhecimento de X através de, por exemplo, sua função de distribuição, desejamos obter o comportamento de $h(X)$. De modo geral, podemos ter uma família de funções atuando sobre um vetor aleatório $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ para produzir um novo vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_r)$ de dimensão r, eventualmente igual a m. Nesse caso, $\mathbf{X}$ é um vetor com função de distribuição conjunta conhecida e desejamos o comportamento de $\mathbf{Y}$. As dificuldades envolvidas nessa tarefa dependem das variáveis iniciais e da função de interesse. Em termos matemáticos gerais, dizemos que $\mathbf{Y}$ é uma transformação de $\mathbf{X}$.

Para obter o comportamento probabilístico de transformações uma técnica muito conveniente, principalmente no caso discreto, é o chamado *método direto*. Ele consiste em realizar operações algébricas simples, aplicando a definição da transformação diretamente na expressão da função de distribuição (ou função densidade ou de probabilidade conforme o caso).

Sejam X_1 e X_2 duas variáveis discretas com função de probabilidade conjunta $p(x_1, x_2)$ e defina $Y = h(X_1, X_2)$. A variável Y também será discreta com valores no contra-domínio da função h. Sua função de probabilidade é dada por:

$$p_Y(y) = P(Y = y) = P(h(X_1, X_2) = y) = \sum_{(x_1, x_2) \in A_y} p(x_1, x_2),$$

em que $A_y = \{(x_1, x_2) : h(x_1, x_2) = y\}$. Isto é, para cada y fixado, a soma se dá em todos os pares (x_1, x_2) cuja aplicação da função h resulta no valor y. A função de distribuição de Y pode ser obtida de maneira análoga.

No caso de funções de variáveis contínuas, é mais conveniente obter primeiro a função de distribuição. Por exemplo, considere duas variáveis contínuas X_1 e X_2 com função densidade conjunta $f(x_1, x_2)$. Seja $Y = h(X_1, X_2)$, temos

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(h(X_1, X_2) \leq y) = \\ &= P((X_1, X_2) \in B_y) = \iint\limits_{B_y} f(x_1, x_2)\, dx_1\, dx_2, \end{aligned}$$

sendo $B_y = \{(x_1, x_2) : h(x_1, x_2) \leq y\}$. A densidade de Y pode ser calculada por derivação de $F_Y(y)$. Entretanto, às vezes, a densidade pode ser obtida diretamente se for possível reescrever a integral dupla acima na forma:

$$F_Y(y) = \int_{-\infty}^{y} g(w)\, dw.$$

Nesse caso, g é a densidade da variável Y.

Exemplo 3.15: A função de probabilidade conjunta de X e Y é dada na tabela abaixo:

$X\backslash Y$	0	1	2
0	$1/4$	$1/8$	$1/8$
1	$1/4$	0	$1/4$

Vamos obter o comportamento de $X + Y$ e $X - Y$. Uma forma prática de obter as funções de probabilidade dessas variáveis é construir, inicialmente, uma tabela com seus valores a partir da conjunta de X e Y.

(X, Y)	$p(x, y)$	$X + Y$	$X - Y$
$(0, 0)$	$1/4$	0	0
$(0, 1)$	$1/8$	1	-1
$(0, 2)$	$1/8$	2	-2
$(1, 0)$	$1/4$	1	1
$(1, 2)$	$1/4$	3	-1

Segue, então, que as funções de probabilidade de interesse são

$X + Y$	0	1	2	3
prob.	$1/4$	$3/8$	$1/8$	$1/4$

e

$X - Y$	-2	-1	0	1
prob.	$1/8$	$3/8$	$1/4$	$1/4$

Para a conjunta entre $X+Y$ e $X-Y$, utilizamos novamente a tabela acima coletando os eventuais pares comuns (que nesse caso não existem):

$X+Y \backslash X-Y$	-2	-1	0	1
0	0	0	1/4	0
1	0	1/8	0	1/4
2	1/8	0	0	0
3	0	1/4	0	0

A partir dessa conjunta, outros cálculos de probabilidade podem ser feitos. □

Exemplo 3.16: Seja X uma variável aleatória contínua com densidade $f(x) = \frac{1}{2} I_{[1,3]}(x)$. Vamos obter a função de distribuição e a função densidade de $Y = e^X$.

Observe que Y será contínua e com valores no intervalo $[e, e^3]$. Sendo F_Y sua função de distribuição, temos de imediato que $F_Y(y) = 0$ se $y < e$ e $F_Y(y) = 1$ se $y \geq e^3$. Para $e \leq y < e^3$,

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y) =$$
$$= P(X \leq \log y) = \int_{-\infty}^{\log y} \frac{1}{2} I_{[1,3]} dx = \frac{\log y - 1}{2}.$$

Dessa forma,

$$F_Y(y) = \begin{cases} 0, & \text{se } y < e; \\ \frac{\log y - 1}{2}, & \text{se } e \leq y < e^3; \\ 1, & \text{se } y \geq e^3. \end{cases}$$

O leitor pode verificar que a densidade de Y é dada por $f_Y(y) = \frac{1}{2y} I_{[e,e^3]}$. □

Exemplo 3.17: Considere a variável aleatória contínua X com valores entre 0 e 3 e densidade $f_X(x) = \frac{1}{3} I_{[0,3]}(x)$. A partir dessa variável, uma nova variável é definida da seguinte forma:

$$Y = \begin{cases} 3, & \text{se } 2 < X \leq 3; \\ X, & \text{caso contrário.} \end{cases}$$

A variável Y será do tipo misto. Sempre que X estiver em $(2,3]$, o valor de Y será 3. A função de distribuição de Y será dada por:

$$F_Y(y) = \begin{cases} 0, & \text{se } y < 0; \\ y/3, & \text{se } 0 \leq y < 2; \\ 2/3, & \text{se } 2 \leq y < 3; \\ 1, & \text{se } y \geq 3. \end{cases}$$

O gráfico de $F_Y(y)$ é representado a seguir e pode-se observar que no ponto $y = 3$ a função é descontínua.

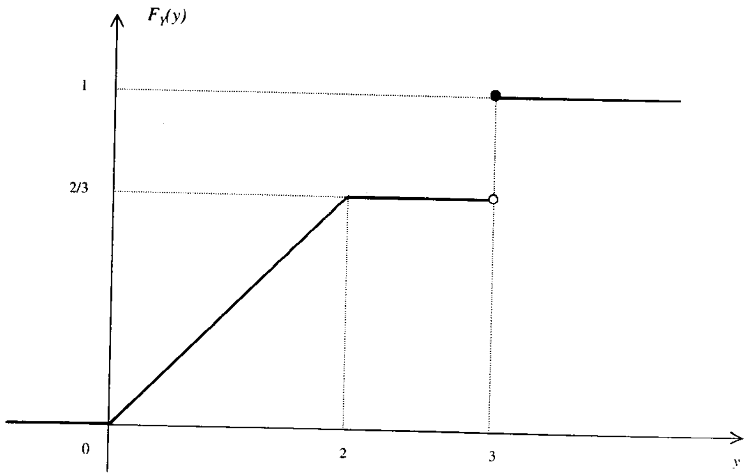

Figura 3.3: Função de Distribuição Mista.

Suponha, agora, que queremos determinar $P(Y < 2 | Y \geq 1)$. Temos:

$$P(Y < 2 | Y \geq 1) = \frac{P([Y < 2] \cap [Y \geq 1])}{P(Y \geq 1)}$$

$$= \frac{P(1 \leq Y < 2)}{P(Y \geq 1)}$$

$$= \frac{F_Y(2^-) - F_Y(1^-)}{1 - F_Y(1^-)} = 1/2.$$

□

Exemplo 3.18: A função densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x,y) = e^{-(x+y)} I_{(0,\infty)}(x) I_{(0,\infty)}(y).$$

Vamos determinar a função de distribuição de $Z = 2X + Y$. Observe que a variável Z é sempre positiva. Dessa forma, para $z < 0$, temos $F_Z(z) = 0$. Para $z \geq 0$,

$$F_Z(z) = P(2X + Y \leq z) = P((X,Y) \in B_z) = \iint_{B_z} f_{X,Y}(x,y)dx\,dy,$$

com $B_z = \{(x,y) : 2x + y \leq z\}$. Note que, no cálculo da integral, os limites precisam levar em conta, além de B_z, a definição de $f_{X,Y}(x,y)$. Assim,

$$F_Z(z) = \iint_{B_z} e^{-(x+y)} I_{(0,\infty)}(x) I_{(0,\infty)}(y)\,dx\,dy = \int_0^z \int_0^{\frac{z-y}{2}} e^{-(x+y)}\,dx\,dy,$$

então,

$$F_Z(z) = \begin{cases} 0, & \text{se } z < 0; \\ 1 + e^{-z} - 2e^{-\frac{z}{2}}, & \text{se } z \geq 0. \end{cases}$$

Pode-se verificar que F_Z satisfaz as propriedades de função de distribuição e, por derivação, podemos obter a densidade de Z. □

O próximo exemplo apresenta uma situação que pode surpreender vários leitores. Partindo de variáveis discretas relativamente simples, tomamos uma função dessas variáveis, também relativamente simples, para produzir uma nova variável bem mais complicada e que não tem nada de simples! O exemplo retoma o conceito de função singular apresentado no capítulo anterior.

Exemplo 3.19: Função Singular-tipo Cantor (Dudewicz e Mishra [88])

Considere sucessivos lançamentos independentes de uma moeda equilibrada. Definindo como sucesso a ocorrência de *cara*, temos que $X_1, X_2, \cdots$, formam um sequência de variáveis de Bernoulli independentes. Assim,

$$P(X_i = 1) = 1/2\,, i = 1, 2, \cdots.$$

Sendo a ocorrência de sucesso no lançamento i recompensada por $3/4^i$ (em alguma unidade monetária), estamos interessados na função de distribuição do ganho total. Conforme veremos essa função será do tipo singular, isto é, será contínua e com derivada igual a zero em quase toda parte (a massa de probabilidade onde isto não ocorre é igual a 0).

Denotando por W a variável ganho vem

$$W = \sum_{i=1}^{\infty} \frac{3}{4^i} X_i \,.$$

Observe que $0 \le W \le 1$, pois para *cara* em todos os lançamentos (isto é, $X_1 = X_2 = \cdots = 1$), temos $\sum_{i=1}^{\infty} \frac{3}{4^i} = 1$. Daí decorre que já podemos saber alguns valores parciais da função de distribuição. Para $w < 0$, ela tem valor 0 e para $w \ge 1$, vale 1. Lembrando que as funções de distribuição são contínuas à direita, o ponto $w = 0$ poderia ser um candidato a uma descontinuidade à esquerda e, assim, F não seria contínua. Entretanto, isso não ocorre uma vez que:

$$\begin{aligned} F(0) = P(W \le 0) &= F(0^-) + P(W = 0) \\ &= 0 + P(X_1 = X_2 = \cdots = 0) \\ &= \lim_{k\to\infty} \left(\frac{1}{2}\right)^k = 0. \end{aligned}$$

Então, temos os seguintes valores parciais para F:

$$F(w) = P(W \le w) = \begin{cases} 0, & \text{se } w \le 0; \\ 1, & \text{se } w \ge 1, \end{cases}$$

restando o cálculo de F para os outros intervalos.

Observando o resultado de X_1 (o 1o. lançamento), podemos estabelecer outras condições sobre os intervalos de valores para W:

$$\begin{aligned} &\text{Se } X_1 = 1 \ \Rightarrow W \ge 3/4; \\ &\text{Se } X_1 = 0 \ \Rightarrow W \le \frac{3}{4^2} + \frac{3}{4^3} + \cdots = \frac{1}{4}. \end{aligned}$$

Logo $P(\frac{1}{4} < W < \frac{3}{4}) = 0$. Vamos ver o que acontece nos pontos extremos desse intervalo. Temos, pela independência dos X_i's,

$$P(W = 1/4) = P(X_1 = 0, X_2 = X_3 \cdots = 1) = \frac{1}{2} \times \frac{1}{2} \times \cdots = \lim_{k\to\infty} \left(\frac{1}{2}\right)^k = 0;$$

de modo análogo, concluímos $P(W = 3/4) = 0$. De fato, podemos repetir esse argumento de modo que, para qualquer valor k, teremos $P(W = k) = 0$. Então, para o intervalo $\frac{1}{4} \le w \le \frac{3}{4}$, $P(W \le w) = P(X_1 = 0) = 1/2$.

Resumindo, o que temos até agora para F, é o seguinte:

$$F(w) = P(W \le w) = \begin{cases} 0, & \text{se } w \le 0; \\ 1/2, & \text{se } \frac{1}{4} \le w \le \frac{3}{4}; \\ 1, & \text{se } w \ge 1. \end{cases}$$

Temos 2 intervalos em que F ainda não está definida, a saber, $(0, 1/4)$ e $(3/4, 1)$. O gráfico correspondente ao que já calculamos é apresentado na figura a seguir.

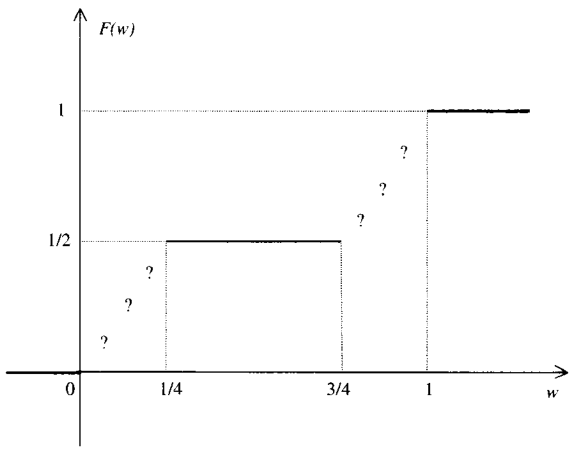

Figura 3.4: Construção da Função de Distribuição- tipo Cantor.

Antes de prosseguir, observamos que o ganho, a partir do 2o. lançamento, tem a mesma distribuição que $W/4$. Note que

$$\sum_{i=2}^{\infty} \frac{3}{4^i} X_i = \frac{3}{4} \sum_{i=2}^{\infty} \frac{1}{4^{i-1}} X_i = \frac{1}{4} \Big(\sum_{j=1}^{\infty} \frac{3}{4^j} X_{j+1} \Big).$$

Tendo em vista que as variáveis $X_1, X_2, \cdots$, são identicamente distribuídas e estamos fazendo uma soma infinita, o termo entre parêntesis tem a mesma distribuição de W. Usaremos esse resultado no que segue.

Para $0 < w < 1/4$ temos

$$\begin{aligned}
F(w) = P(W \leq w) &= P(W \leq w, X_1 = 0) + P(W \leq w, X_1 = 1) \\
&= P\Big(\sum_{i=1}^{\infty} \frac{3}{4^i} X_i \leq w, X_1 = 0\Big) + 0 \\
&= P\Big(\sum_{i=2}^{\infty} \frac{3}{4^i} X_i \leq w, X_1 = 0\Big) \\
&= P\Big(\sum_{i=2}^{\infty} \frac{3}{4^i} X_i \leq w\Big) P(X_1 = 0) \\
&= \frac{1}{2} P(\frac{W}{4} \leq w) \\
&= \frac{1}{2} F(4w).
\end{aligned}$$

De acordo com o que já sabemos de F, temos para $1/4 < 4w < 3/4$, $F(4w) = 1/2$. Assim, para $1/16 \leq w \leq 3/16$ vem $F(w) = 1/4$, uma vez que os extremos não contribuem com probabilidade positiva. Restam sem definição os intervalos $(0, 1/16)$ e $(3/16, 1/4)$.

Para $3/4 < w < 1$, procedemos de modo análogo.

$$\begin{aligned}
F(w) = P(W \leq w) &= P(W \leq w, X_1 = 0) + P(W \leq w, X_1 = 1) \\
&= \frac{1}{2} + P\Big(\sum_{i=1}^{\infty} \frac{3}{4^i} X_i \leq w, X_1 = 1\Big) \\
&= \frac{1}{2} + P\Big((\frac{3}{4} + \sum_{i=2}^{\infty} \frac{3}{4^i} X_i) \leq w, X_1 = 1\Big) \\
&= \frac{1}{2} + P\Big(\sum_{i=2}^{\infty} \frac{3}{4^i} X_i \leq w - \frac{3}{4}\Big) P(X_1 = 1) \\
&= \frac{1}{2} + \frac{1}{2} P(\frac{W}{4} \leq w - \frac{3}{4}) \\
&= \frac{1}{2} + \frac{1}{2} F(4(w - \frac{3}{4})).
\end{aligned}$$

Então, $F(w) = 3/4$ para $13/16 \leq w \leq 15/16$ e resta definir o valor de F em $(3/4, 13/16)$ e $(15/16, 1)$. O gráfico de F com os valores já conhecidos até o momento é representado na figura a seguir.

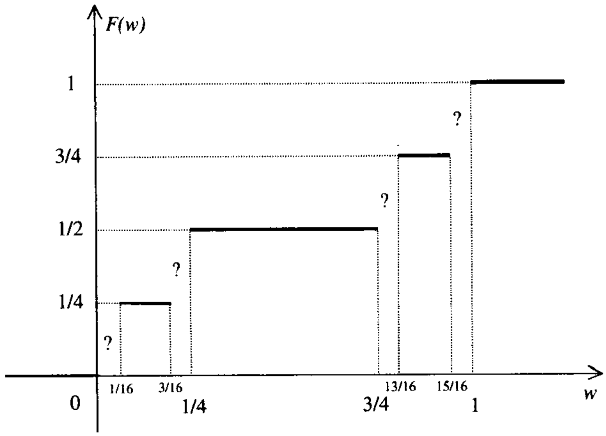

Figura 3.5: Construção da Função de Distribuição- tipo Cantor (continuação).

Observe que, nos três sub-intervalos de $(0, 1)$ já analisados, a função de distribuição é constante. O maior salto, que indicaria descontinuidade de F, não poderia ultrapassar 1/4 pois trata-se de uma função monótona não decrescente. Os intervalos já definidos têm comprimento total de $3/4$ e cada um dos quatro restantes tem comprimento $1/16$.

Usando o mesmo raciocínio desenvolvido acima, conseguimos definir F em um sub-intervalo de cada um dos intervalos ainda sem definição. Por exemplo, no intervalo $(0, 1/16)$, definimos $F(w) = 1/8$ para $1/64 \leq w \leq 3/64$ e restarão, sem definição, os intervalos $(0, 1/64)$ e $(3/64, 3/16)$.

De modo geral, cada intervalo restante é dividido em 4 partes e, para os dois sub-intervalos centrais, F será constante e igual à média dos valores já atribuídos aos intervalos mais próximos, que já têm probabilidade determinada. O valor máximo de um eventual salto de F está limitado à metade da diferença entre seus valores vizinhos. Usando indução matemática, após n etapas, teremos já definido o valor de F em $1 + 2 + 2^2 + 2^3 + \cdots 2^{n-1}$ intervalos. Eles terão comprimento total de $2^{-1} + 2^{-2} + 2^{-3} + \cdots 2^{-n} = 1 - 1/2^n$. Tomando o limite, quando o número de etapas vai a infinito, concluímos: F será contínua em todo $\mathbb{R}$ e com $F(0) = 0$ e $F(1) = 1$. Os sub-intervalos de $[0, 1]$, em que F é constante, têm comprimento total igual a 1. Além disso, como a derivada de uma constante é zero, temos $F'(w) = 0$ para quase todo $w \in \mathbb{R}$. Logo, F é uma função de distribuição do tipo singular. A construção exibida neste exemplo é uma variante do que é conhecido como *Função de Cantor* (ver Chae [80]). $\square$

Em pesquisa ou aplicações, o uso de recursos computacionais têm se tornado freqüente nas áreas de Probabilidade e Estatística. Dentre as diversas técnicas utilizadas, a simulação se destaca. Uma das peças chave em simulações, nas áreas mencionadas, é poder gerar valores de uma variável com certa distribuição de probabilidade. Os métodos computacionais desenvolvidos para esse fim se utilizam de uma transformação especial, definida a seguir.

Definição 3.10: Transformação Integral de Probabilidade

Seja X uma variável aleatória com função de distribuição F. A transformação de X tal que $Y = F(X)$ é denominada *transformação integral de probabilidade*. □

Em simulações, o uso da transformação acima depende da possibilidade de inverter F. A inversa tem domínio em $[0, 1]$, mas se F tiver saltos ou for em escada, ela não existirá. Na próxima definição, uma função que faz as vezes de inversa é apresentada e, por abuso de notação, também será representada por F^{-1}.

Definição 3.11: Inversa Generalizada de F

Seja F uma função de distribuição qualquer. A *inversa generalizada*, denotada por F^{-1} é definida da seguinte forma:

$$F^{-1}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\}. \qquad \square$$

Note que, se a inversa de F existe no sentido usual, ela coincide com a inversa generalizada. Por exemplo, se F for estritamente crescente e contínua, então $\forall x \in \mathbb{R}$, existirá apenas um y tal que $F(x) = y$. Nesse caso, pela definição acima teríamos $F^{-1}(y) = x$, mesmo resultado de uma inversa usual. Na figura abaixo, apresentamos alguns casos de F^{-1}. Na próxima proposição, demonstramos uma relação que será a base para resultados úteis em simulação.

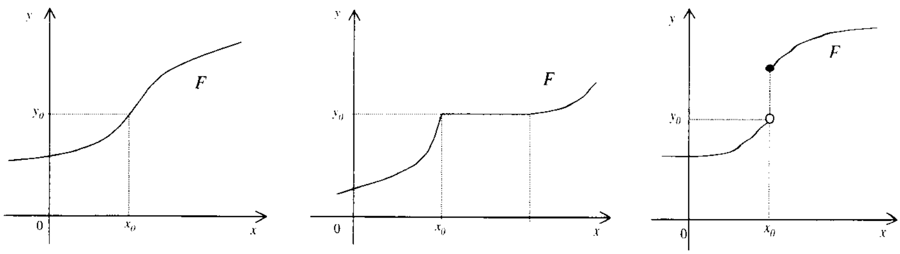

Figura 3.6: Formas alternativas para F com $F^{-1}(y_0) = x_0$.

Proposição 3.4:

Seja F uma função de distribuição qualquer e F^{-1} sua inversa generalizada. Temos:

$$F^{-1}(y) \leq t \text{ se, e somente se, } y \leq F(t).$$

Demonstração:

Inicialmente, observamos que, pela definição de F^{-1}, existe uma seqüência x_n, tal que $x_n \in \{x \in \mathbb{R} : F(x) \geq y\}$ e $x_n \downarrow F^{-1}(y)$. Da continuidade à direita de F, temos $F(x_n) \to F(F^{-1}(y))$ e ainda $F(F^{-1}(y)) \geq y$.

Supondo que $F^{-1}(y) \leq t$, vamos verificar que $y \leq F(t)$. Como F é não decrescente, segue $F(F^{-1}(y)) \leq F(t)$. Da conclusão do parágrafo anterior, vem $y \leq F(F^{-1}(y))$ e, portanto, vale $y \leq F(t)$.

Para a outra implicação, supomos $y \leq F(t)$. Dessa forma, t pertence ao conjunto $\{x \in \mathbb{R} : F(x) \geq y\}$ e, pelo argumento sobre a seqüência x_n no ínicio da demonstração, segue que $F^{-1}(y) \leq t$. □

Com a aplicação da proposição anterior, vamos verificar um importante resultado para variáveis contínuas.

Proposição 3.5:

Seja X uma variável aleatória contínua com função de distribuição F. Então $Y = F(X)$ terá distribuição Uniforme Contínua em $[0, 1]$. Vale a recíproca, sendo $U \sim U_c[0, 1]$, então $X = F^{-1}(U)$ terá função de distribuição F.

Demonstração:

Se X é contínua, F é uma função contínua e o mesmo ocorre com $Y = F(X)$. A função de distribuição F será um número entre 0 e 1 e, portanto, $P(Y < 0) = P(F(X) < 0) = 0$ e $P(Y \leq 1) = P(F(X) \leq 1) = 1$.

Para $0 < y < 1$,

$$P(Y \geq y) = P(F(X) \geq y) = P(X \geq F^{-1}(y)) = 1 - P(X \leq F^{-1}(y));$$

em que usamos a proposição anterior e, na última igualdade, a continuidade de X. Essa continuidade ainda implica que $F(F^{-1}(y)) = y$. Então

$$P(Y \geq y) = 1 - P(X \leq F^{-1}(y)) = 1 - F(F^{-1}(y)) = 1 - y.$$

Como Y é contínua, segue $P(Y > y) = P(Y \geq y)$ e, dessa forma, Y terá distribuição Uniforme Contínua em $[0, 1]$.

Considere, agora, a recíproca. Suponha que $U \sim U_c[0,1]$ e $X = F^{-1}(U)$. Temos

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

Logo a variável aleatória X tem distribuição F. □

Em simulação é comum a necessidade de gerar valores de uma certa variável com uma distribuição determinada F. Pela proposição acima, isto pode ser feito gerando um valor aleatório da Uniforme Continua em $[0,1]$, digamos u, e tomando $x = F^{-1}(u)$. Valores do tipo u são conhecidos como *números pseudo-aleatórios*. Esses números imitam a escolha, feita ao acaso, de valores do intervalo $[0,1]$.

Os números pseudo-aleatórios têm essa denominação pois são gerados por algoritimos determinísticos e devem aparentar terem sido sorteados de forma independente e aleatória. Isto significa que, tomando uma grande quantidade deles, esperaríamos que estivessem espalhados uniformemente em $[0,1]$. Muitos programas de computador já incluem rotinas de geração de números aleatórios em suas bibliotecas. O próximo exemplo ilustra essas idéias.

Exemplo 3.20: Geração de Valores Aleatórios (DeGroot e Schervich [02])

Suponha que temos a disposição algum *software* com um gerador de números pseudo-aleatórios e independentes da $U_c[0,1]$. Desejamos gerar valores de uma variável X que tem densidade $f(x) = \frac{1}{2}(2-x)I_{(0,2)}(x)$.

A correspondente função de distribuição é dada por:

$$F(x) = \begin{cases} 0, & x \leq 0; \\ x - \frac{x^2}{4}, & 0 < x \leq 2; \\ 1, & x > 2. \end{cases}$$

Para obter a inversa resolvemos a equação $y = F(x)$ em x, sendo que $y \in (0,1)$ é o domínio da função inversa. Basta considerar a expressão de $F(x)$ para o intervalo $(0,2]$. A solução, com uma das raízes sendo desprezada, fornece: $x = F^{-1}(y) = 2(1 - \sqrt{1-y})$ para $y \in (0,1)$.

Suponha que geramos, aleatoriamente e de forma independente, três valores da distribuição $U_c[0,1]$: $y_1 = 0,4125$; $y_2 = 0,0894$ e $y_3 = 0,8302$.

Aplicando esses valores na expressão da função inversa, obtemos: $x_1 = 0,4670$; $x_2 = 0,0914$ e $x_3 = 1,1759$.

Dessa forma, obtemos três observações independentes da variável X. □

Um importante resultado referente a um vetor de variáveis aleatórias independentes é aquele que garante que funções disjuntas de variáveis independentes sejam independentes. Um caso particular é demonstrado no próximo exemplo.

Exemplo 3.21: Famílias disjuntas de variáveis independentes

Sejam $X_1, X_2, \cdots, X_n$ variáveis aleatórias independentes num mesmo espaço de probabilidade. Com o uso de duas funções g_1 e g_2 defina as variáveis Y_1 e Y_2 tais que:

$$Y_1 = g_1(X_1, X_2, \cdots, X_m) \text{ e } Y_2 = g_2(X_{m+1}, X_{m+2}, \cdots, X_n); 1 \leq m < n.$$

O que podemos dizer sobre a independência entre Y_1 e Y_2?

Inicialmente, observamos que as variáveis são funções disjuntas de $X_1, X_2, \cdots, X_n$. Sendo que Y_1 é função das primeiras m variáveis, enquanto Y_2 das restantes $n - m$. Assim, temos que $g_1 : \mathbb{R}^m \to \mathbb{R}$ e $g_2 : \mathbb{R}^{n-m} \to \mathbb{R}$.

Sejam B_1 e B_2 dois borelianos quaisquer dos reais, lembramos que

$$g_1^{-1}(B_1) = \{(x_1, x_2, \cdots, x_m) \in \mathbb{R}^m : g_1(x_1, x_2, \cdots, x_m) \in B_1\},$$

e, então, $[Y_1 \in B_1]$ é equivalente a $[(X_1, X_2, \cdots, X_m) \in g_1^{-1}(B_1)]$. De modo análogo, $[Y_2 \in B_2]$ e $[(X_{m+1}, X_{m+2}, \cdots, X_n) \in g_2^{-1}(B_2)]$ são equivalentes.

Para verificar a independência entre Y_1 e Y_2, note que $P([Y_1 \in B_1], [Y_2 \in B_2])$ é igual a

$$P\Big([(X_1, X_2, \cdots, X_m) \in g_1^{-1}(B_1)], [(X_{m+1}, X_{m+2}, \cdots, X_n) \in g_2^{-1}(B_2)]\Big),$$

o qual, por sua vez, é igual ao produto

$$P\Big([(X_1, X_2, \cdots, X_m) \in g_1^{-1}(B_1)]\Big)\; P\Big([(X_{m+1}, X_{m+2}, \cdots, X_n) \in g_2^{-1}(B_2)]\Big),$$

pela definição de independência de vetores aleatórios. Mas, essa última probabilidade pode ser escrita, utilizando novamente o argumento de equivalência mencionado acima, como $P([Y_1 \in B_1])\, P([Y_2 \in B_2])$. Fica, assim, verificado a independência entre Y_1 e Y_2. □

Algumas funções de variáveis aleatórias merecem destaque, pois aparecem com grande freqüência em aplicações. As expressões referentes ao comportamento probabilístico da soma, diferença, produto, quociente, mínimo e máximo entre variáveis aleatórias são apresentadas nas próximas proposições.

Proposição 3.6: Densidade da Soma e da Diferença de Variáveis

Dadas duas variáveis aleatórias contínuas X e Y com densidade conjunta $f_{X,Y}$, então a função densidade da soma e da diferença entre essas variáveis são dadas por, respectivamente:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)\, dx;$$
$$f_{X-Y}(w) = \int_{-\infty}^{\infty} f_{X,Y}(w+y, y)\, dy.$$

Demonstração:

Para a soma, com $B_z = \{(x, y) : x + y \leq z\}$, temos:

$$F_{X+Y}(z) = P(X+Y \leq z) = P((X,Y) \in B_z) = \int\int_{B_z} f_{X,Y}(x,y)dx\,dy;$$

Então,

$$F_{X+Y}(z) = \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{z-x} f_{X,Y}(x,y)dy\Big]dx = \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{z} f_{X,Y}(x,u-x)du\Big]dx,$$

em que fizemos a substituição $y = u - x$.

Para obter a densidade tomamos a derivada:

$$\begin{aligned}\frac{\partial F_{X+Y}(z)}{\partial z} &= \frac{\partial}{\partial z}\Big\{\int_{-\infty}^{\infty}\Big[\int_{-\infty}^{z} f_{X,Y}(x,u-x)du\Big]dx\Big\} \\ &= \frac{\partial}{\partial z}\Big\{\int_{-\infty}^{z}\Big[\int_{-\infty}^{\infty} f_{X,Y}(x,u-x)dx\Big]du\Big\},\end{aligned}$$

assim, concluímos que $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)\, dx$.

Para a diferença $X - Y$, com $B_w = \{(x, y) : x - y \leq w\}$, temos

$$F_{X-Y}(w) = P(X-Y \leq w) = P((X,Y) \in B_w) = \int\int_{B_w} f_{X,Y}(x,y)dx\,dy =$$
$$= \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{w+y} f_{X,Y}(x,y)dx\Big]dy = \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{w} f_{X,Y}(u+y,y)du\Big]dy,$$

em que fizemos a substituição $x = u + y$. Invertendo a ordem de integração e derivando, obtemos a densidade $f_{X-Y}(w) = \int_{-\infty}^{\infty} f_{X,Y}(w + y, y)\, dy$. □

Observe que se X e Y forem independentes a densidade conjunta pode ser escrita como o produto das densidades marginais, simplificando as expressões. No caso independente, a densidade da soma recebe o nome de *convolução* das densidades de X e Y. A proposição anterior tem, também, uma versão para o caso discreto que não será apresentada.

Exemplo 3.22: A densidade conjunta das variáveis X e Y é dada pela expressão:

$$f_{X,Y}(x, y) = \frac{1}{3}\,(x + y) I_{(0,2]}(x) I_{(0,1]}(y).$$

Vamos obter a densidade de $X + Y$. Temos,

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)\, dx = \int_{-\infty}^{\infty} \frac{1}{3}\, z\, I_{(0,2]}(x) I_{(0,1]}(z - x)\, dx.$$

Avaliando o comportamento dos indicadores, teremos (ver figura a seguir):

$z < 0$ ou $z > 3$: $I_{(0,1]}(z - x) = 0$;

$0 \leq z < 1$: $I_{(0,1]}(z - x) = 1$ somente se $0 < x \leq z$;

$1 \leq z < 2$: $I_{(0,1]}(z - x) = 1$ somente se $z - 1 < x \leq z$;

$2 \leq z \leq 3$: $I_{(0,1]}(z - x) = 1$ somente se $z - 1 < x \leq 2$;

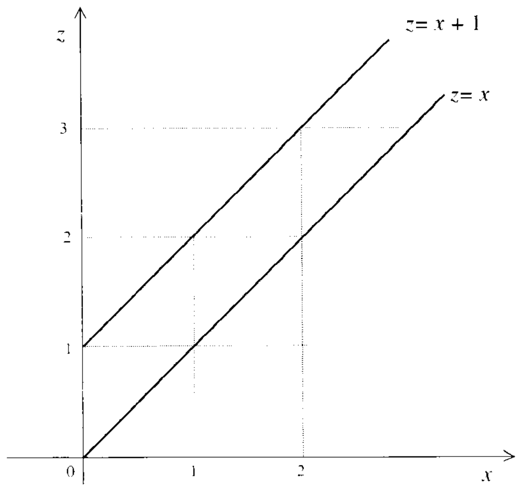

Figura 3.7: Região de Integração- função soma.

Então,
$$0 \le z < 1 : f_{X+Y}(z) = \int_0^z \frac{1}{3}\, z\, dx = \frac{z^2}{3};$$
$$1 \le z < 2 : f_{X+Y}(z) = \int_{z-1}^z \frac{1}{3}\, z\, dx = \frac{z}{3};$$
$$2 \le z \le 3 : f_{X+Y}(z) = \int_{z-1}^2 \frac{1}{3}\, z\, dx = \frac{z(3-z)}{3}.$$

A densidade procurada satisfaz as propriedades de densidade e será dada por:

$$f_{X+Y}(z) = \begin{cases} \frac{z^2}{3}, & \text{se } 0 \le z < 1; \\ \frac{z}{3}, & \text{se } 1 \le z < 2; \\ \frac{z(3-z)}{3}, & \text{se } 2 \le z \le 3; \\ 0, & \text{caso contrário.} \end{cases}$$

Recomendamos ao leitor tentar obter essa densidade via função de distribuição. □

Apresentamos, a seguir, alguns exemplos de relação entre modelos que envolvem soma de variáveis.

Exemplo 3.23: Relação entre Bernoulli e Binomial

Sejam $X_1, X_2, \cdots, X_n$ variáveis independentes com distribuição Bernoulli de parâmetro p. Defina $Y = X_1 + X_2 + \cdots + X_n$, então Y tem distribuição Binomial de parâmetros n e p.

Para verificar esse resultado, observe que o evento $[Y = k]$ significa a ocorrência de k sucessos e $n - k$ fracassos, dentre os n experimentos. Como existem várias alternativas na ordem de ocorrência dos sucessos, precisamos multiplicar pelo número delas, o que é dado pela combinação de n tomados k a k. A relação também poderia ser verificada por indução a partir da fórmula de convolução para funções de probabilidade. □

Exemplo 3.24: Relação entre Exponencial e Gama

Sejam X_1 e X_2 variáveis independentes com mesma distribuição Exponencial de parâmetro λ. Vamos verificar que $X_1 + X_2$ têm distribuição Gama com parâmetros $\alpha = 2$ e $\beta = \lambda$. Pela independência de X_1 e X_2 temos:

$$f_{X_1,X_2}(x,y) = \lambda^2 e^{-\lambda(x+y)},\ 0 < x < \infty, 0 < y < \infty.$$

Então

$$
\begin{aligned}
F_{X_1+X_2}(z) = P(X_1 + X_2 \leq z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_{X_1,X_2}(x,y)\,dx\,dy \\
&= \int_0^z \int_0^{z-x} \lambda^2 e^{-\lambda(x+y)} dy\,dx \\
&= \int_0^z \lambda\, e^{-\lambda x}\,(1 - e^{-\lambda(z-x)})dx \\
&= 1 - e^{-\lambda z} - \lambda z\, e^{-\lambda z}.
\end{aligned}
$$

Para obter a densidade, calculamos a derivada:

$$
f_{X_1+X_2}(z) = \frac{\partial F_{X_1+X_2}(z)}{\partial z} = \lambda^2 z\, e^{-\lambda z}.
$$

Essa última expressão corresponde à densidade Gama de parâmetros $\alpha = 2$ e $\beta = \lambda$. De modo geral, é possível verificar que a soma de n Exponenciais independentes de mesmo parâmetro λ terá distribuição *Gama*(n, λ). □

Exemplo 3.25: Relação entre Poisson e Gama

Sejam Y_1 e Y_2 intervalos entre ocorrências de um certo evento, isto é, Y_1 e Y_2 indicam o tempo de espera para a 1a. e o tempo entre a 1a. e 2a. ocorrências, respectivamente. Definimos X_t como o número de ocorrências até o instante t. Temos a seguinte equivalência de eventos:

$$
[\text{2a. ocorrência antes de } t] = [\text{Pelo menos 2 ocorrências antes de } t];
$$

$$
[Y_1 + Y_2 < t] = [X_t \geq 2].
$$

Suponha agora que Y_1 e Y_2 são variáveis independentes com distribuição Exponencial de parâmetro λ. Sabemos que $Y_1 + Y_2$ têm distribuição Gama com parâmetros $\alpha = 2$ e $\beta = \lambda$. Dessa forma, pela igualdade dos eventos acima, temos

$$
P(X_t \geq 2) = P(Y_1 + Y_2 < t) = \int_0^t \frac{\lambda^2}{\Gamma(2)}\, y\, e^{-\lambda y}\, dy;
$$

que, aplicando a técnica de integral por partes, resulta em

$$
P(X_t \geq 2) = P(Y_1 + Y_2 < t) = 1 - e^{-\lambda t} - \lambda t\, e^{-\lambda t}.
$$

Note que a mesma expressão resultaria se o modelo Poisson (λt) fosse usado para X_t. De fato, para qualquer n, a variável X_t terá distribuição de Poisson com parâmetro λt. A prova pode ser feita por indução finita e integração por partes. □

Exemplo 3.26: Reprodução de Normais

Dizemos que a distribuição Normal se reproduz, porque a soma de Normais independentes também tem distribuição Normal. Essa propriedade vale também para outros modelos, entre eles, Binomial, Poisson e Gama.

Vamos verificar que a soma de duas Normais Padrão independentes é Normal com $\mu = 0$ e $\sigma^2 = 2$. O caso geral pode ser demonstrado, mais facilmente, com o uso da função geradora de momentos a ser definida no Capítulo 5.

Sejam X_1 e X_2 duas variáveis aleatórias independentes com distribuição $N(0, 1)$. Para a distribuição da soma temos

$$F_{X_1+X_2}(z) = P(X_1 + X_2 \leq z) = = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_{X_1,X_2}(x, y)\, dx\, dy;$$

que resulta em

$$\begin{aligned}
F_{X_1+X_2}(z) &= \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{z-x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy\Big)\, dx \\
&= \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(v-x)^2}{2}}\, dv\Big)\, dx \\
&= \int_{-\infty}^{z}\Big(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(v-x)^2}{2}}\, dx\Big)\, dv,
\end{aligned}$$

após a substituição $y = v - x$ e mudança na ordem de integração. A densidade será dada pelo integrando da expressão acima aplicado em $v = z$. Assim,

$$f_{X_1+X_2}(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}}\, dx = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-x^2+zx-\frac{z^2}{2}}\, dx.$$

Note que z está fixo dentro da integral e podemos reescreve-la como:

$$f_{X_1+X_2}(z) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} e^{-\frac{z^2}{4}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\left(\frac{1}{\sqrt{2}}\right)} e^{-\frac{1}{2}\left(\frac{x-z/2}{\sqrt{1/2}}\right)^2} dx.$$

O integrando corresponde à densidade de uma Normal com $\mu = z/2$ e $\sigma^2 = 1/2$. Logo a integral vale 1 e o termo restante é a densidade $N(0, 2)$. $\square$

Retomamos a apresentação das proposições.

Proposição 3.7: Distribuição do Mínimo e do Máximo

Considere o conjunto $X_1, X_2, \ldots, X_m$ de variáveis aleatórias independentes, cujas funções de distribuição são $F_1, F_2, \ldots, F_m$, respectivamente. As expressões da função de distribuição de $Y_1 = \min(X_1, X_2, \ldots, X_m)$ e $Y_m = \max(X_1, X_2, \ldots, X_m)$ são dadas, respectivamente, por:

$$F_{Y_1}(z) = 1 - \prod_{i=1}^{m}[1 - F_{X_i}(z)]\,;$$

$$F_{Y_m}(w) = \prod_{i=1}^{m} F_{X_i}(w)\,.$$

Demonstração:

Para obter a distribuição do mínimo, utilizamos um argumento lógico freqüentemente útil nesses casos: se o mínimo é maior que z, todas as variáveis também precisam ser maiores que z. Assim,

$$P(Y_1 > z) = P(\min(X_1, \ldots, X_m) > z) = P(X_1 > z, \ldots, X_m > z).$$

Pela independência entre as variáveis essa última probabilidade pode ser escrita como

$$P(Y_1 > z) = P(X_1 > z) \cdots P(X_m > z) = \prod_{i=1}^{m}[1 - F_{X_i}(z)].$$

Logo, tomando-se o complementar, a expressão de $F_{Y_1}(z)$ é obtida.

Para o máximo, observe que se o máximo é menor que z então todas as variáveis também são. Assim

$$P(Y_m \leq w) = P(\max(X_1, \ldots, X_m) \leq w) = P(X_1 \leq w, \ldots, X_m \leq w).$$

Pela independência entre as variáveis, temos

$$P(Y_m \leq w) = P(X_1 \leq w) \cdots P(X_m \leq w) = \prod_{i=1}^{m} F_{X_i}(w).$$

Se as variáveis forem identicamente distribuídas, as expressões podem ser apresentadas em forma de potências da função de distribuição comum. □

Proposição 3.8: Densidade do produto e do quociente

Sejam X e Y duas variáveis aleatórias contínuas com densidade conjunta $f_{X,Y}$. As funções densidade do produto e do quociente são, respectivamente, dadas por:

$$f_{XY}(u) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_{X,Y}(x, \frac{u}{x}) \, dx;$$
$$f_{X/Y}(v) = \int_{-\infty}^{\infty} |y| \, f_{X,Y}(vy, y) \, dy.$$

Demonstração:

Aplicamos o método direto, mas o resultado também poderia ser obtido via o método do Jacobiano a ser apresentado a seguir nesta seção.

Vamos iniciar obtendo a função de distribuição do produto XY. Com $B_u = \{(x, y) : xy \leq u\}$, vem $F_{XY}(u) = P(XY \leq u) = \iint\limits_{B_u} f_{X,Y}(x, y) dx\, dy$.

Assumimos $x \neq 0$ e precisamos separar os casos $x > 0$ e $x < 0$. Então,

$$F_{XY}(u) = \int_{-\infty}^{0} \Big[\int_{u/x}^{\infty} f_{X,Y}(x, y) dy\Big] dx + \int_{0}^{\infty} \Big[\int_{-\infty}^{u/x} f_{X,Y}(x, y) dy\Big] dx,$$

que, após a substituição $s = xy$, vem:

$$\begin{aligned} F_{XY}(u) &= \int_{-\infty}^{0} \Big[\int_{u}^{-\infty} f_{X,Y}(x, \frac{s}{x}) \frac{ds}{x}\Big] dx + \int_{0}^{\infty} \Big[\int_{-\infty}^{u} f_{X,Y}(x, \frac{s}{x}) \frac{ds}{x}\Big] dx \\ &= \int_{-\infty}^{u} \Big[\int_{-\infty}^{0} \frac{1}{-x} f_{X,Y}(x, \frac{s}{x}) \, dx\Big] ds + \int_{-\infty}^{u} \Big[\int_{0}^{\infty} \frac{1}{x} f_{X,Y}(x, \frac{s}{x}) \, dx\Big] ds \\ &= \int_{-\infty}^{u} \Big[\int_{-\infty}^{\infty} \frac{1}{|x|} f_{X,Y}(x, \frac{s}{x}) \, dx\Big] ds. \end{aligned}$$

Aplicando a derivada, obtemos a densidade

$$\frac{\partial F_{XY}(u)}{\partial u} = f_{XY}(u) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_{X,Y}(x, \frac{u}{x}) \, dx.$$

Para o quociente, o cálculo é similar. Com $B_v = \{(x, y) : x/y \leq v\}$, vem $F_{X/Y}(v) = P(X/Y \leq v) = \iint\limits_{B_v} f_{X,Y}(x, y) dx\, dy$. Assumimos $y \neq 0$ e separamos os casos $y > 0$ e $y < 0$. Assim,

$$F_{X/Y}(v) = \int_{-\infty}^{0}\Big[\int_{vy}^{\infty} f_{X,Y}(x,y)dx\Big]dy + \int_{0}^{\infty}\Big[\int_{-\infty}^{vy} f_{X,Y}(x,y)dx\Big]dy,$$

que, após a substituição $s = x/y$, resulta em:

$$\begin{aligned} F_{X/Y}(v) &= \int_{-\infty}^{0}\Big[\int_{v}^{-\infty} f_{X,Y}(sy,y)\,y\,ds\Big]dy + \int_{0}^{\infty}\Big[\int_{-\infty}^{v} f_{X,Y}(sy,y)\,y\,ds\Big]dy \\ &= \int_{-\infty}^{v}\Big[\int_{-\infty}^{0} (-y) f_{X,Y}(sy,y)\,dy\Big]ds + \int_{-\infty}^{v}\Big[\int_{0}^{\infty} y\, f_{X,Y}(sy,y)\,dy\Big]ds \\ &= \int_{-\infty}^{v}\Big[\int_{-\infty}^{\infty} |y| f_{X,Y}(sy,y)\,dy\Big]ds. \end{aligned}$$

Derivando, obtemos: $\dfrac{\partial F_{X/Y}(v)}{\partial v} = f_{X/Y}(v) = \displaystyle\int_{-\infty}^{\infty} |y|\, f_{X,Y}(vy, y)\,dy.$ □

Para as variáveis contínuas, pode-se usar o *Método do Jacobiano*, utilizado em Cálculo Diferencial, para obter a distribuição de transformações de variáveis aleatórias. O método é apresentado a seguir sem demonstração formal.

Seja $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ um vetor aleatório contínuo com densidade $f_{\mathbf{X}}(\mathbf{x})$ conhecida. O vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)$ é uma transformação g (multidimensional) do vetor $\mathbf{X}$. Em notação simplificada, podemos escrever $\mathbf{Y} = g(\mathbf{X}) = (g_1(\mathbf{X}), g_2(\mathbf{X}), \cdots, g_m(\mathbf{X}))$, de modo que $Y_i = g_i(X_1, X_2, \ldots, X_m)$ para $i = 1, 2, \ldots, m$.

Supomos que todas as funções g_i, $i = 1, \cdots, m$, têm derivadas parciais contínuas e que o Jacobiano da transformação seja diferente de zero para todos os pontos $(x_1, x_2, \ldots, x_m)$; isto é

$$J_g(\mathbf{x}) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m} \end{vmatrix} \neq 0.$$

Vamos também supor que o conjunto de equações:

$$\begin{aligned} y_1 &= g_1(x_1, x_2, \ldots, x_m); \\ y_2 &= g_2(x_1, x_2, \ldots, x_m); \\ \vdots \quad & \quad \vdots \\ y_m &= g_m(x_1, x_2, \ldots, x_m), \end{aligned}$$

determina, de modo único, os x_i's em função de $(y_1, y_2, \ldots, y_m)$. Em outras palavras, isto significa dizer que g é bijetora. Representamos a solução do sistema por $\mathbf{x} = h(\mathbf{y})$, ou seja,

$$\begin{aligned} x_1 &= h_1(y_1, y_2, \ldots, y_m); \\ x_2 &= h_2(y_1, y_2, \ldots, y_m); \\ \vdots \quad & \quad \vdots \\ x_m &= h_m(y_1, y_2, \ldots, y_m). \end{aligned}$$

Nessas condições, as variáveis $Y_1, Y_2, \ldots, Y_m$ têm densidade conjunta dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(h(\mathbf{y}))\, |J_g(h(\mathbf{y}))|^{-1},$$

em que, a densidade $f_{\mathbf{X}}(.)$ e o Jacobiano $J_g(.)$ são aplicados nos pontos $x_i = h_i(\mathbf{y}), i = 1, \cdots, m$.

Uma das condições para a aplicação do método do Jacobiano foi que a função $g : \mathbf{X} \rightarrow \mathbf{Y}$ fosse bijetora (um a um). Se esse não for o caso, podemos particionar o domínio em subconjuntos disjuntos de modo que, restrita a cada subconjunto, ela seja biunívoca. Nesse caso, aplicamos o método Jacobiano a cada uma das partições e somamos o resultado para obter a densidade desejada.

Assim, denotando por $\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_k$ a partição do domínio e sendo $g^{(i)}$ a restrição de g ao subconjunto $\mathcal{D}_i$, temos

$$f_{\mathbf{Y}}(\mathbf{y}) = \sum_{i=1}^{k} f_{\mathbf{X}}(h^{(i)}(\mathbf{y}))\, |J_{g^{(i)}}(h^{(i)}(\mathbf{y}))|^{-1},$$

em que $h^{(i)}$ é a inversa de $g^{(i)}$ (referente ao subconjunto $\mathcal{D}_i$).

A figura a seguir ilustra a repartição do domínio em 4 subconjuntos.

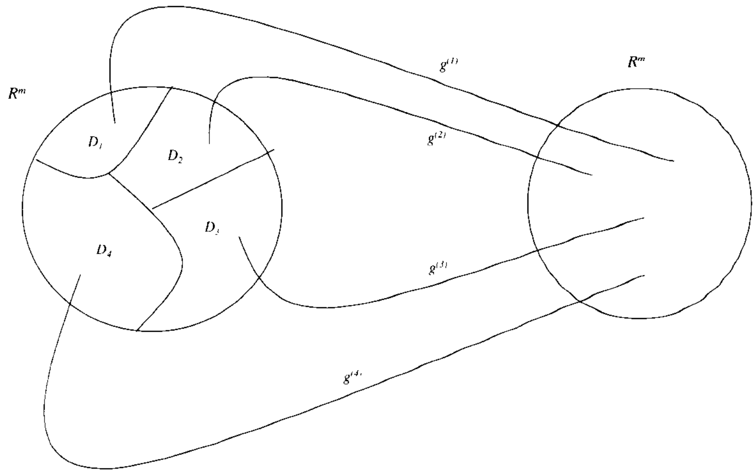

Figura 3.8: Repartição do Domínio- 4 subconjuntos.

Observação 1: Alguns autores, preferem calcular o Jacobiano referente à função inversa de g, denotada aqui por h. O determinante a ser calculado seria

$$J_h(\mathbf{y}) = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \cdots & \frac{\partial h_1}{\partial y_m} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} & \cdots & \frac{\partial h_2}{\partial y_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_m}{\partial y_1} & \frac{\partial h_m}{\partial y_2} & \cdots & \frac{\partial h_m}{\partial y_m} \end{vmatrix}.$$

Se usarmos $J_h(\mathbf{y})$ ao invés de $J_g(h(\mathbf{y}))$, a fórmula para o cálculo da densidade é modificada coerentemente tendo em vista que

$$J_h(\mathbf{y}) = \Big(J_g(h(\mathbf{y}))\Big)^{-1}.$$

Para evitar confusão, é conveniente sempre usar o mesmo Jacobiano.

Observação 2: Se o vetor $\mathbf{Y}$ tiver dimensão menor que a de $\mathbf{X}$, acrescentamos variáveis artifíciais convenientes até a igualdade nas dimensões. Após aplicar o método do Jacobiano, calculamos a marginal adequada da densidade conjunta obtida e determinamos a desejada densidade de $\mathbf{Y}$.

Exemplo 3.27: Considere o vetor aleatório $\mathbf{X}$ com densidade

$$f_{\mathbf{X}}(\mathbf{x}) = \lambda^3\, e^{-\lambda(x_1+x_2+x_3)} I_{(0,\infty)}(x_1) I_{(0,\infty)}(x_2) I_{(0,\infty)}(x_3),\ \lambda > 0.$$

Estamos interessados em $\mathbf{Y} = (Y_1, Y_2, Y_3)$ tal que

$$\begin{aligned} Y_1 &= X_1; \\ Y_2 &= X_1 + X_2; \\ Y_3 &= X_1 + X_2 + X_3. \end{aligned}$$

Seguindo nossa notação temos

$$\begin{aligned} g_1(X_1, X_2, \dots, X_m) &= X_1; \\ g_2(X_1, X_2, \dots, X_m) &= X_1 + X_2; \\ g_3(X_1, X_2, \dots, X_m) &= X_1 + X_2 + X_3; \end{aligned}$$

e, portanto, o Jacobiano da transformação é dado por

$$J_g(\mathbf{x}) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_3} \\ \frac{\partial g_3}{\partial x_1} & \frac{\partial g_3}{\partial x_2} & \frac{\partial g_3}{\partial x_3} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 1.$$

Observe que, como o Jacobiano não depende de $\mathbf{x}$, podemos já estabelecer $|J_g(h(\mathbf{y}))|^{-1} = 1$.

Resolvendo para $\mathbf{x}$ o sistema abaixo

$$\begin{aligned} y_1 &= x_1; \\ y_2 &= x_1 + x_2; \\ y_3 &= x_1 + x_2 + x_3, \end{aligned}$$

obtemos uma única solução dada por $\mathbf{x} = (y_1, y_2 - y_1, y_3 - y_2)$, correspondendo à função h mencionada anteriormente. Ou seja, $h_1(\mathbf{y}) = y_1, h_2(\mathbf{y}) = y_2 - y_1$ e $h_3(\mathbf{y}) = y_3 - y_2$. Então,

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \lambda^3\, e^{-\lambda(y_1+(y_2-y_1)+(y_3-y_2))} I_{(0,\infty)}(y_1) I_{(0,\infty)}(y_2 - y_1) I_{(0,\infty)}(y_3 - y_2) \\ &= \lambda^3\, e^{-\lambda y_3}\, I_A(y_1, y_2, y_3), \end{aligned}$$

com $A = \{(y_1,\ y_2, y_3) : 0 < y_1 < \ y_2 < y_3\}$. □

os casos $y > 0$ e $y < 0$. Assim,

$$F_{X/Y}(v) = \int_{-\infty}^{0}\Big[\int_{vy}^{\infty} f_{X,Y}(x,y)dx\Big]dy + \int_{0}^{\infty}\Big[\int_{-\infty}^{vy} f_{X,Y}(x,y)dx\Big]dy,$$

que, após a substituição $s = x/y$, resulta em:

$$\begin{aligned}F_{X/Y}(v) &= \int_{-\infty}^{0}\Big[\int_{v}^{-\infty} f_{X,Y}(sy,y)\,y\,ds\Big]dy + \int_{0}^{\infty}\Big[\int_{-\infty}^{v} f_{X,Y}(sy,y)\,y\,ds\Big]dy\\ &= \int_{-\infty}^{v}\Big[\int_{-\infty}^{0}(-y)f_{X,Y}(sy,y)\,dy\Big]ds + \int_{-\infty}^{v}\Big[\int_{0}^{\infty} y\,f_{X,Y}(sy,y)\,dy\Big]ds\\ &= \int_{-\infty}^{v}\Big[\int_{-\infty}^{\infty}|y|f_{X,Y}(sy,y)\,dy\Big]ds.\end{aligned}$$

Derivando, obtemos: $$\frac{\partial F_{X/Y}(v)}{\partial v} = f_{X/Y}(v) = \int_{-\infty}^{\infty}|y|\,f_{X,Y}(vy,y)\,dy. \qquad \square$$

Para as variáveis contínuas, pode-se usar o *Método do Jacobiano*, utilizado em Cálculo Diferencial, para obter a distribuição de transformações de variáveis aleatórias. O método é apresentado a seguir sem demonstração formal.

Seja $\mathbf{X} = (X_1, X_2, \dots, X_m)$ um vetor aleatório contínuo com densidade $f_{\mathbf{X}}(\mathbf{x})$ conhecida. O vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_m)$ é uma transformação g (multidimensional) do vetor $\mathbf{X}$. Em notação simplificada, podemos escrever $\mathbf{Y} = g(\mathbf{X}) = (g_1(\mathbf{X}), g_2(\mathbf{X}), \cdots, g_m(\mathbf{X}))$, de modo que $Y_i = g_i(X_1, X_2, \dots, X_m)$ para $i = 1, 2, \dots, m$.

Supomos que todas as funções g_i, $i = 1, \cdots, m$, têm derivadas parciais contínuas e que o Jacobiano da transformação seja diferente de zero para todos os pontos $(x_1, x_2, \dots, x_m)$; isto é

$$J_g(\mathbf{x}) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m}\\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_m}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m}\end{vmatrix} \neq 0.$$

Vamos também supor que o conjunto de equações:

Exemplo 3.28: (DeGroot e Schervich [02])

Considere o vetor aleatório $\mathbf{X}$ com densidade

$$f_{\mathbf{X}}(\mathbf{x}) = 4x_1x_2\, I_{(0,1)}(x_1)I_{(0,1)}(x_2).$$

Estamos interessados em $\mathbf{Y} = (Y_1, Y_2)$ tal que

$$Y_1 = X_1/X_2;$$
$$Y_2 = X_1X_2.$$

Seguindo nossa notação temos

$$g_1(X_1, X_2) = X_1/X_2;$$
$$g_2(X_1, X_2) = X_1X_2.$$

e, portanto, o Jacobiano da transformação g é dado por

$$J_g(\mathbf{x}) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} 1/x_2 & -x_1/x_2^2 \\ x_2 & x_1 \end{vmatrix} = 2x_1/x_2.$$

Resolvendo para $\mathbf{x}$ o sistema abaixo

$$y_1 = x_1/x_2;$$
$$y_2 = x_1x_2,$$

obtemos uma única solução dada por

$$x_1 = h_1(y_1, y_2) = (y_1y_2)^{1/2} \text{ e } x_2 = h_2(y_1, y_2) = (y_2/y_1)^{1/2}.$$

Observe que h é a função inversa de g. Os valores possíveis de y_1 e y_2 estão no conjunto A:

$$A = \left\{(y_1, y_2) \in \mathbb{R}^2 : y_1 > 0, y_2 > 0, (y_1y_2)^{1/2} < 1, (y_2/y_1)^{1/2} < 1\right\}.$$

Assim,

$$f_{\mathbf{X}}(h(\mathbf{y})) = 4(y_1y_2)^{1/2}(y_2/y_1)^{1/2} = 4y_2 \text{ e } J_g(h(\mathbf{y})) = \frac{2(y_1y_2)^{1/2}}{(y_2/y_1)^{1/2}} = 2y_1;$$

então a densidade de $\mathbf{Y}$ é dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(h(\mathbf{y}))|J_g(h(\mathbf{y}))|^{-1} = 4y_2\,|2y_1|^{-1}I_A(y_1,y_2) = \frac{2y_2}{y_1}\,I_A(y_1,y_2).$$

O leitor pode verificar que a função obtida satisfaz as propriedades de função densidade conjunta. □

Exemplo 3.29: (Ross [98])

Sejam X_1, X_2 e X_3 variáveis aleatórias independentes com densidade $N(0,1)$. Vamos obter a densidade conjunta de Y_1, Y_2 e Y_3 sendo $Y_1 = X_1 + X_2 + X_3$, $Y_2 = X_1 - X_2$ e $Y_3 = X_1 - X_3$.

A densidade conjunta de $\mathbf{X} = (X_1, X_2, X_3)$ é dada por

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{3}\frac{1}{\sqrt{2\pi}}e^{-\frac{x_i^2}{2}}, \; -\infty < x_i < \infty, i = 1,2,3.$$

Na notação que estamos utilizando, temos $\mathbf{Y} = g(\mathbf{X})$, ou

$$\begin{aligned} Y_1 &= g_1(X_1,X_2,X_3) = X_1 + X_2 + X_3;\\ Y_2 &= g_2(X_1,X_2,X_3) = X_1 - X_2;\\ Y_3 &= g_3(X_1,X_2,X_3) = X_1 - X_3; \end{aligned}$$

e, portanto, o Jacobiano da transformação é dado por

$$J_g(\mathbf{x}) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_3}\\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_3}\\ \frac{\partial g_3}{\partial x_1} & \frac{\partial g_3}{\partial x_2} & \frac{\partial g_3}{\partial x_3} \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1\\ 1 & -1 & 0\\ 1 & 0 & -1 \end{vmatrix} = 3;$$

como não depende de $\mathbf{x}$, podemos já estabelecer $|J_g(h(\mathbf{y}))|^{-1} = 1/3$.

O sistema

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3;\\ y_2 &= x_1 - x_2;\\ y_3 &= x_1 - x_3, \end{aligned}$$

tem solução única dada por

$$x_1 = h_1(y_1,y_2,y_3) = \frac{y_1+y_2+y_3}{3};$$

$$x_2 = h_2(y_1, y_2, y_3) = \frac{y_1 - 2y_2 + y_3}{3};$$
$$x_3 = h_3(y_1, y_2, y_3) = \frac{y_1 + y_2 - 2y_3}{3}.$$

Então, sendo $\mathbf{Y} = (Y_1, Y_2, Y_3)$ e $h(\mathbf{y}) = (h_1(\mathbf{y}), h_2(\mathbf{y}), h_3(\mathbf{y}))$, temos

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{X}}(h(\mathbf{y}))\,|J_g(h(\mathbf{y}))|^{-1} \\ &= \frac{1}{3(2\pi)^{3/2}}\,e^{-\frac{1}{2}\left[(\frac{y_1+y_2+y_3}{3})^2+(\frac{y_1-2y_2+y_3}{3})^2+(\frac{y_1+y_2-2y_3}{3})^2\right]} \\ &= \frac{1}{3(2\pi)^{3/2}}\,e^{-\frac{1}{2}\left[\frac{y_1^2}{3}+\frac{2y_2^2}{3}+\frac{2y_3^2}{3}-\frac{2y_2y_3}{3}\right]}, -\infty < y_i < \infty, i = 1, 2, 3. \end{aligned}$$

A partir dessa conjunta, as densidade marginais das variáveis de interesse podem ser obtidas de modo usual. □

Exemplo 3.30: Jacobiano sem bijeção (Mood, Graybill e Boes [74])

Vamos obter a densidade de $Y = g(X) = X^2$, sendo X uma variável contínua com densidade $f(x) = \frac{1}{2}\,e^{-|x|}$, $-\infty < x < \infty$.

A função g não é bijetora no domínio dos reais e, assim, para a aplicação do método do Jacobiano será necessário fazer uma separação no domínio. Considere que os conjuntos $\mathcal{D}_1 = \{x \in \mathbb{R} : x < 0\}$ e $\mathcal{D}_2 = \{x \in \mathbb{R} : x \geq 0\}$ representam a repartição mencionada. Dessa forma, $g^{(1)} : \mathcal{D}_1 \to \mathbb{R}^+$ e $g^{(2)} : \mathcal{D}_2 \to \mathbb{R}^+$ são as respectivas restrições de g aos subconjuntos $\mathcal{D}_1$ e $\mathcal{D}_2$. Temos, portanto,

$$h^{(1)}(y) = (g^{(1)})^{-1}(y) = -\sqrt{y} \text{ e } h^{(2)}(y) = (g^{(2)})^{-1}(y) = \sqrt{y}\,.$$

Estamos no caso univariado e, portanto, não precisamos calcular o determinante. Basta derivar a função g para obter $J_g(x) = 2x$. Logo,

$$J_g(h^{(1)}(y)) = -2\sqrt{y} \text{ em } \mathcal{D}_1 \text{ e } J_g(h^{(2)}(y)) = 2\sqrt{y} \text{ em } \mathcal{D}_2.$$

Segue, então

$$f_Y(y) = \sum_{i=1}^{2} f_X(h^{(i)}(y))\,|J_g(h^{(i)}(y))|^{-1} = \frac{1}{2\sqrt{y}}\,e^{-\sqrt{y}}\,I_{(0,\infty)}(y).$$

□

Exercícios da Seção 3.3:

1. Seja X uma variável aleatória com função de distribuição F e sejam a e b constantes reais. Determine a função de distribuição das seguintes variáveis aleatórias:
a. $-X$.
b. $aX + b$, com $a \neq 0$.
c. $XI_{(a,b)}(X)$, supondo $1 < a < b$.
d. Comente as diferenças nas soluções se X é discreta ou contínua.

2. Seja X uma variável aleatória com função de distribuição F. Considere uma função $h : X \to \mathbb{R}$, bijetora. Expresse, em função de F, a função de distribuição de $h(X)$, nos seguintes casos:
a. A função h é monótona não decrescente.
b. A função h é monótona não crescente.

3. A variável X tem função de distribuição F. Para $Y = X^2$, obtenha $F_Y(y)$.

4. Sejam X e $Y \sim N(0, 1)$, independentes. Obtenha a densidade de $2X + Y$.

5. Suponha que X e Y são variáveis independentes e com densidade Exponencial de parâmetros λ_1 e λ_2, respectivamente. Mostre que, para $k > 0$, $P(X - kY > 0) = \lambda_2/(k\lambda_1 + \lambda_2)$.

6. Sejam X e $Y \sim Exp(1)$, independentes. Obtenha a densidade de $X - Y$.

7. A densidade conjunta de (X, Y) é $f(x, y) = 2(x + y)\, I_{\{0<x<y<1\}}(x, y)$.
a. Determine a densidade de $X + Y$.
b. Calcule $P(X + Y > 1)$.
c. Obtenha $P(X + Y > 1 | \frac{1}{3} \leq X \leq \frac{2}{3})$.

8. A densidade conjunta de (X, Y) é dada por:

$$f(x,y) = \begin{cases} \frac{1}{2\pi} e^{-\frac{x^2}{4\pi} - y}, & -\infty < x < \infty, y \geq 0; \\ 0, & \text{caso contrário.} \end{cases}$$

a. Verifique se X e Y são independentes.
b. Determine a densidade conjunta de Y e $X + Y$.

9. Sendo X e Y variáveis independentes com distribuição $Exp(\lambda)$, obtenha a conjunta entre $X + Y$ e X/Y e demonstre que são independentes.

10. Sendo $X \sim N(0,1)$, verifique que X^2 tem distribuição Qui-quadrado com 1 grau de liberdade usando:
a. O método direto.
b. O método do Jacobiano.

3. 4 Exercícios

1. Sendo X e Y variáveis aleatórias em $(\Omega, \mathcal{F}, \mathcal{P})$, mostre que $\min(X, Y)$ e $\max(X, Y)$ também são variáveis aleatórias.

2. Para X e Y variáveis aleatórias em $(\Omega, \mathcal{F}, \mathcal{P})$, verifique que os conjuntos $[X < Y], [X = Y]$ e $[X > Y]$ pertencem a σ-álgebra $\mathcal{F}$.

3. Mostre que $A_1, A_2, \cdots, A_n$ são eventos independentes se e só se $I_{A_1}, I_{A_2}, \cdots, I_{A_n}$ forem variáveis aleatórias independentes.

4. Para as variáveis discretas, as expressões de soma e diferença de duas variáveis são similares ao caso contínuo. Verifique as expressões abaixo:

$$p_{X+Y}(z) = \sum_x P(X = x, Y = z - x) = \sum_x p_{X,Y}(x, z - x);$$

$$p_{X-Y}(z) = \sum_x P(X = x, Y = x - z) = \sum_x p_{X,Y}(x, x - z).$$

Se X e Y forem independentes, essas expressões se transformam em

$$p_{X+Y}(z) = \sum_x p_X(x)\, p_Y(z - x);$$

$$p_{X-Y}(z) = \sum_x p_X(x)\, p_Y(x - z).$$

5. Seja X uma variável aleatória com função de distribuição F_X. Seja $Y = h(X)$ com função de distribuição F_Y. Mostre que:
a. $F_{X,Y}(x, y) = \min\{F_X(x), F_Y(y)\}$, se h é monótona crescente.
b. $F_{X,Y}(x, y) = \max\{F_X(x) + F_Y(y) - 1, 0\}$, se h é monótona decrescente.

6. Mostre ou dê um contra-exemplo:
a. Se $F_X(z) = F_Y(z), \forall\, z \in \mathbb{R}$ então $P(X = Y) = 1$.
b. Se $Y = X + 1$ então $F_X(z) = F_Y(z + 1), \forall\, z \in \mathbb{R}$.

7. Mostre que, para variáveis contínuas X e Y temos

$$F_X(x) + F_Y(y) - 1 \leq F_{X,Y}(x, y) \leq \sqrt{F(x)F(y)}\,; \forall\, x, y \in \mathbb{R}.$$

8. Seja $F(.)$ a função de distribuição de alguma variável aleatória. Para as funções G, definidas abaixo, avalie quais podem ser função de distribuição conjunta:

a. $G(x, y) = F(x) + F(y)$; $x, y \in \mathbb{R}$.

b. $G(x, y) = F(x)F(y)$; $x, y \in \mathbb{R}$.

c. $G(x, y) = F(x)/F(y)$; $x, y \in \mathbb{R}$.

d. $G(x, y) = \max(F(x), F(y))$; $x, y \in \mathbb{R}$.

e. $G(x, y) = \min(F(x), F(y))$; $x, y \in \mathbb{R}$.

9. Considere $X \sim Poisson(2)$. Defina Y pelo truncamento de X que impede valores superiores a 2. Assim, Y tem o valor 2 sempre que $X \geq 2$. Obtenha a função de distribuição de Y.

10. Sendo $F(x, y)$ a distribuição conjunta de X e Y, defina $Z = \max(X, Y)$ e $W = \min(X, Y)$. Mostre que:
a. $F_Z(z) = F(z, z)$, $\forall\, z \in \mathbb{R}$.
b. $F_W(w) = F_X(w) + F_Y(w) - F(w, w)$, $\forall\, w \in \mathbb{R}$.

11. Seja $X \sim B(n = 5, p = 1/2)$ e defina $Y = \frac{1}{2}(\max(X, 4) + \min(X, 2))$.
a. Obtenha $p(y)$ e F_Y.
b. Calcule $P(Y \leq 2)$ e $P(Y > 2 | Y \leq 4)$.

12. Para $X \sim Poisson(2)$, defina $Y = \max(\min(X, 4), \max(X, 2))$.
a. Obtenha $p(y)$ e F_Y.
b. Calcule $P(Y > 1)$ e $P(Y > 1 | Y \leq 3)$.

13. Suponha que X e Y são variáveis aleatórias, independentes e identicamente distribuídas, com valores -1 e 1 com mesma probabilidade. Verifique se X e $Z = XY$ são independentes.

14. Uma urna contém bolas numeradas de 1 a 3. Duas retiradas são feitas sem reposição. Defina as variáveis X e Y como, respectivamente, o menor e o maior valor observado. Obtenha a conjunta dessas variáveis e verifique se são independentes.

15. Sejam X e Y duas variáveis independentes com distribuição Geométrica de parâmetro p. Determine $P(X = Y)$.

16. Para X e Y independentes com distribuição Binomial de parâmetros (n_1, p) e (n_2, p), respectivamente, verifique que:
a. A soma $X + Y$ também segue o modelo Binomial.

b. A distribuição condicional de X, dada a soma $X + Y$, é Hipergeométrica com parâmetros que não dependem de p.

17. Obtenha a conjunta de X e Y e a marginal de Y, sendo dados:

$$p_X(x) = x/3\,, x = 1, 2;\; p_{Y|X}(y|x) = \binom{x}{y}\left(\frac{1}{2}\right)^x, y \in \mathbb{N}, y \leq x \text{ e } x = 1, 2.$$

18. Mostre que, sendo X e Y variáveis independentes com distribuição Poisson de parâmetros λ_1 e λ_2, respectivamente, temos que:
a. A soma $X + Y$ é Poisson.
b. A distribuição condicional de X, dada a soma $X + Y$, é Binomial.

19. Defina Y como a restrição de uma variável Exponencial (2), em valores superiores a uma dada constante a ($a > 0$).
a. Determine a densidade de Y.
b. Calcule $P(Y > a + 1 | Y < a + 2)$.

20. Considere $X \sim U_c(0, 1)$. Para que função g, estritamente decrescente, temos $g(X)$ com distribuição Exponencial de parâmetro 1?

21. Seja $X \sim N(0, 1)$ e defina $Y = e^X$. A variável Y segue o modelo *Lognormal* de parâmetros 0 e 1. Obtenha a densidade de Y.

22. Indique para quais valores de α, a função abaixo é densidade.

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

23. A densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x, y) = \frac{1}{4}[\,1 + xy(x^2 - y^2)]\, I_{(-1,1)}(x) I_{(-1,1)}(y).$$

a. Verifique que f satisfaz as propriedades de função densidade conjunta.
b. Calcule $P(X > 0 | Y > 0)$.

24. As variáveis X e Y têm a seguinte densidade conjunta:

$$f_{X,Y}(x, y) = e^{-y}(1 - e^{-x}) I_{(0,y]}(x) I_{[0,\infty)}(y) + e^{-x}(1 - e^{-y}) I_{(0,x]}(y) I_{[0,\infty)}(x).$$

a. Mostre que f satisfaz as propriedades de função densidade conjunta.
b. Determine $P(X > 1, Y < 2)$.
c. Obtenha as marginais.
d. As variáveis são independentes? Justifique.

25. A função densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x,y) = 2\, I_{(0,y)}(x) I_{(0,1)}(y).$$

a. Verifique que f satisfaz as propriedades de função densidade conjunta.
b. Obtenha as marginais.

26. A função densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x,y) = e^{-(x+y)} I_{(0,\infty)}(x) I_{(0,\infty)}(y).$$

Determinar a função de distribuição de $X - Y$.

27. A densidade conjunta de X e Y é dada por:

$$f_{X,Y}(x,y) = \frac{1}{2}\, xy\, I_{(0,2)}(x) I_{(0,x)}(y).$$

a. Verifique que f satisfaz as propriedades de função densidade conjunta.
b. Obtenha as marginais. As variáveis são independentes?

28. Sejam $Y \sim Exp(1)$ e $X | (Y = y) \sim Poisson(y)$. Mostre que

$$P(X = n) = \left(\frac{1}{2}\right)^{n+1}, \; n = 0, 1, 2, \cdots.$$

29. A variável X tem densidade Uniforme Contínua em $[0,1]$. A variável Y, condicionada à X é Binomial, isto é,

$$p_{Y|X}(y|x) = \binom{n}{y} x^y (1-x)^{n-y}, y = 0, 1, \cdots, n.$$

Obtenha a marginal de Y.
Sugestão: Use a função Beta,

$$\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\, dx;\; a, b > 0.$$

Temos ainda a relação $\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

30. As variáveis X e Y são independentes e têm distribuição Uniforme Contínua em $[0, 1]$. Obtenha as densidades de:
a. $X + Y$.
b. $X - Y$.
c. $|X - Y|$.
d. X/Y.
e. $X/(X + Y)$.

31. A função densidade conjunta de X e Y é

$$f_{X,Y}(x, y) = 8xy\, I_{\{0<x<y<1\}}(x, y).$$

a. Obtenha as marginais.
b. Determine $P(X > 1/2 | Y < 3/4)$.

32. A densidade conjunta das variáveis X e Y é dada por

$$f_{X,Y}(x, y) = \frac{1}{8}(6 - x - y)\ I_{(0,2)}(x)\, I_{(2,4)}(y).$$

a. Obtenha as marginais.
b. Determine as densidades condicionais de X dado Y e de Y dado X.

33. A densidade condicional de Y dado X é dada por

$$f_{Y|X}(y|x) = I_{(x,x+1)}(y).$$

Calcule a marginal de Y e a densidade condicional de X dado Y, se temos $f_X(x) = I_{(0,1)}(x)$.

34. Seja $X \sim U_c[0, 2]$ e defina $Y = \min(X, 1)$.
a. Classifique Y e obtenha sua função de distribuição.
b. Represente F_Y através de funções de distribuição discreta e absolutamente contínua.

35. Para $X \sim Exp(2)$, defina $Y = \max(X, 2)$.
a. Obtenha F_Y.
b. Apresente F_Y como combinação linear de funções de distribuição discreta, singular e absolutamente contínua.

36. Seja $X \sim Exp(\lambda)$ e defina Y da seguinte forma:

$$Y = \begin{cases} 1, & \text{se } X < 1; \\ X, & \text{se } 1 \leq X < 2; \\ 2, & \text{se } X \geq 2. \end{cases}$$

a. Classifique Y e obtenha sua função de distribuição.
b. Decomponha a função de distribuição nas suas partes discreta e contínua.

37. Sendo X uma variável aleatória contínua com densidade:

$$f(x) = \begin{cases} \frac{1}{4}x, & 0 \leq x < 2; \\ \frac{1}{8}, & 2 \leq x < 6; \\ 0, & \text{caso contrário.} \end{cases}$$

a. Determine a função de distribuição de $Y = \min(3, X)$.
b. Calcule $P(Y < 3 | Y \geq 2)$ e $P(Y \leq 3 | Y > 2)$.
c. Faça a decomposição de F_Y nas suas partes discreta, contínua e singular.

38. Seja X uma variável aleatória com densidade:

$$f(x) = \begin{cases} \frac{1}{4}x \text{ se } & 0 \leq x \leq 2; \\ \frac{1}{8} \text{ se } & 2 < x \leq 6; \\ 0 \text{ se } & \text{caso contrário.} \end{cases}$$

Defina uma nova variável $Y = \max\{\min(3, X), 2\}$. Obtenha $F^{ac}(y)$ e $F^{d}(y)$.

39. Sejam X, Y e Z variáveis aleatórias independentes com X e Y, tendo distribuição Uniforme em $[0, 2]$ e Z assumindo os valores $\frac{1}{2}$ e $\frac{3}{2}$ com probabilidade $\frac{1}{3}$ e $\frac{2}{3}$, respectivamente. Seja $W = \max(X, Y, Z, 1)$.
a. Obtenha a função de distribuição de W.
b. Calcule $P(1 < W \leq \frac{7}{4} | W \geq \frac{3}{2})$.
c. Determine as partes discreta, absolutamente contínua e singular da função de distribuição de W.

40. A densidade conjunta de X, Y e Z é dada por

$$f(x, y, z) = \frac{1}{16}[4(xy + xz + yz) + 4(x + y + z) + 7], \text{ para } x, y, z \in [0, 1].$$

a. O que pode ser dito da independência de X, Y e Z?
b. Calcule $P(X \leq \frac{1}{2}, Y > \frac{1}{3})$.

41. As variáveis aleatórias X_1 e X_2 são independentes e têm densidade comum $Exp(\lambda)$.

a. Obtenha a função densidade conjunta das variáveis $Y_1 = \min(X_1, X_2)$ e $Y_2 = \max(X_1, X_2)$.
b. Determine a densidade de $Y_2 - Y_1 | Y_1$.

42. Sejam X e Y independentes com distribuição $U_c(0, 1)$. Defina $W = X - Y$ e $Z = X + Y$.

a. Obtenha a densidade conjunta de W e Z.
b. Determine $P(Z > 3W)$.

43. As variáveis aleatórias contínuas X e Y são independentes, identicamente distribuídas e têm densidade $Exp(1)$. Através do método do Jacobiano, obtenha a conjunta e as marginais de $U = 2X + Y$ e $V = 2X - Y$ e verifique se são independentes.

44. Considere as variáveis X e Y independentes e identicamente distribuídas com densidade Normal Padrão. Defina $U = X + Y$ e $V = X/Y$.

a. Obtenha a densidade conjunta de U e V.
b. Mostre que V tem densidade Cauchy Padrão.

45. Se X_1 e X_2 tem função densidade conjunta:

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2}(x_1^2 - x_2^2)\, I_A(x_1, x_2)$$

com $A = \{(x_1, x_2) \in \mathcal{R}^2; x_1 \in (0, 2), |x_2| < x_1, |x_2| < 2 - x_1\}$, obtenha a função densidade conjunta de $W_1 = \frac{1}{2}(X_1 + X_2)$ e $W_2 = \frac{1}{2}(X_1 - X_2)$.

46. As variaveis aleatórias X_1, X_2 e X_3 são independentes e identicamente distribuídas com densidade Normal Padrão. Considere as seguintes novas variáveis: $W_1 = X_1$, $W_2 = (X_1 + X_2)/2$ e $W_3 = (X_1 + X_2 + X_3)/3$.

a. Determine a densidade de W_3, pelo método direto.
b. Obtenha a densidade conjunta de W_1, W_2 e W_3.
c. Obtenha a densidade marginal conjunta de W_1 e W_2.
d. Confirme a resposta de (a), calculando a densidade de W_3 a partir da conjunta obtida em (b).

47. Para algum $\alpha > 0$, as variáveis aleatórias contínuas X e Y têm a seguinte densidade conjunta:

$$f_{X,Y}(x,y) = 2\alpha^{-2} e^{-\frac{(x+y)}{\alpha}} I_{\{0<x<y<\infty\}}(x,y).$$

Usando o método do Jacobiano, obtenha a conjunta e as marginais de $U = X/Y$ e Y. Elas são independentes?

48. Sejam X_1, X_2 e X_3, variáveis independentes e identicamente distribuídas com densidade Exponencial de parâmetro 1. Defina $W_1 = X_1/(X_1 + X_2)$, $W_2 = (X_1 + X_2)/(X_1 + X_2 + X_3)$ e $W_3 = X_1 + X_2 + X_3$.

a. Obtenha a densidade conjunta de W_1, W_2 e W_3 e verifique se são independentes.

b. Mostre que W_1 é $U_c(0,1)$.

c. Verifique que W_3 segue o modelo Gama com parâmetros $(\alpha = 3, \beta = 1)$.

49. Considere as variáves aleatórias X e $Y \sim U_c(-1,1)$, independentes. Defina as variáveis $W = X - Y$ e $Z = X + Y$.

a. Obtenha a densidade conjunta de W e Z.

b. Determine a densidade de $Z|W$.

50. Sendo X e Y variáveis independentes com distribuição Uniforme contínua em $(1, \beta)$, obtenha a densidade de XY via método do Jacobiano.

51. Sejam X e Y variáveis aleatórias contínuas com função densidade conjunta

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

Obtenha a densidade de $X + Y$ via o método do Jacobiano.

52. Sendo X e Y variáveis independentes com distribuição Uniforme Contínua em (α, β), obtenha a densidade de $X + Y$, usando os procedimentos abaixo:

a. Através do método direto, obtendo primeiro a função de distribuição.

b. Via a expressão da convolução de densidades.

c. Via método do Jacobiano.

53. Sejam $X_1 \sim Gama(\alpha, \lambda)$ e $X_2 \sim Gama(\beta, \lambda)$ duas variáveis independentes. Defina novas variáveis $W = X_1/(X_1 + X_2)$ e $T = X_1 + X_2$. Mostre que:

a. W tem densidade Beta de parâmetros (α, β).

b. T tem densidade Gama de parâmetros $(\alpha + \beta, \lambda)$.
c. W e T são independentes.

54. Sendo $X \sim N(0, 1)$ e $Y \sim \chi^2(r)$ (modelo Qui-quadrado com r graus de liberdade), defina a variável $T = X/\sqrt{Y/r}$. Verifique que T segue o modelo *t-Student* com r graus de liberdade, cuja densidade é dada por:

$$f_T(t) = \frac{\Gamma[\frac{1}{2}(r+1)]}{\sqrt{\pi r}\ \Gamma(\frac{r}{2})}\left(1+\frac{t^2}{r}\right)^{-\frac{1}{2}(r+1)} I_{(-\infty,\infty)}(t).$$

55. Considere X_1, X_2 e X_3 variáveis contínuas, independentes e com mesma densidade f que tem valores positivos apenas em (a, b), $a, b \in \mathbb{R}$. Sejam Y_1, Y_2 e Y_3 as correspondentes *Estatísticas de Ordem*, isto é, temos Y_1 como $\min(X_1, X_2, X_3)$, Y_2 como a segunda menor após Y_1, e assim por diante.

a. Via método do Jacobiano, mostre que a conjunta de Y_1, Y_2 e Y_3 é dada por:

$$f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3) = 6\, f(y_1) f(y_2) f(y_3)\, I_{\{a<y_1<y_2<y_3<b\}}(y_1, y_2, y_3).$$

b. Qual seria a expressão do item (a) para n variáveis?

56. Sendo $X \sim \chi^2(m)$ e $Y \sim \chi^2(n)$ independentes, seja $W = \frac{(X/m)}{(Y/n)}$.

a. Verifique que W segue o modelo *F-Snedecor* com (m, n) graus de liberdade, cuja densidade é dada por:

$$f_W(w) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\,\Gamma(\frac{n}{2})}\left(\frac{m}{n}\right)^{m/2} \frac{w^{(m-2)/2}}{(1+\frac{m}{n}x)^{(m+n)/2}}\, I_{(0,\infty)}(w).$$

b. Qual seria a distribuição de $1/W$?

c. Mostre que $V = \frac{\frac{m}{n}W}{1+\frac{m}{n}W}$ tem distribuição Beta.

57. Considere $X_1,\ X_2,\ \cdots,\ X_n$ variáveis independentes com densidade *Exp*(λ_i), $i = 1, 2, \cdots, n$.

a. Verifique que $Y_1 = \min(X_1,\ X_2,\ \cdots,\ X_n)$ é *Exp*$\left(\sum_{i=1}^{n}\lambda_i\right)$.

b. Mostre que $P[X_k = \min(X_1,\ X_2,\ \cdots,\ X_n)] = \lambda_k/\left(\sum_{i=1}^{n}\lambda_i\right)$.

c. Calcule a função de distribuição de $Y_n = \max(X_1,\ X_2,\ \cdots,\ X_n)$.

58. Sejam $Y_1, Y_2, \cdots, Y_n$ as estatísticas de ordem de uma sequência de n variáveis independentes da $U_c(0,1)$. Obtenha a conjunta de $V_n = Y_n$ e $V_i = Y_i/Y_{i+1}$ para $i = 1, 2, \cdots, n-1$. Elas são independentes?

59. Seja $X \sim Exp(1)$, obtenha a densidade de $Y = \cos X$.
Sugestão: Particione o domínio em subconjuntos $D_i = ((i-1)\pi, i\pi]$ *para* $i \geq 1$. *Ao obter as inversas, separe os casos de par e ímpar para* i.

60. Sejam X_1 e X_2 independentes com distribuição $N(0, \sigma^2)$. Defina as seguintes variáveis aleatórias $U = X_1^2 + X_2^2$ e $V = X_1/X_2$.

a. Sem fazer "muita" conta indique os modelos de U e V.
b. Via método do Jacobiano, obtenha a densidade conjunta de U e V.
c. As variáveis U e V são independentes?

61. (Problema da agulha de Buffon) Considere uma mesa plana com linhas horizontais separadas pela distância $2a$. Uma agulha de tamanho $2c$ $(c < a)$ é aleatoriamente jogada nessa mesa. Qual é a probabilidade da agulha tocar uma das linhas?
Sugestão: Defina X *como a distância do centro da agulha até a linha mais próxima e* θ *como o ângulo entre a agulha e a perpendicular do centro da agulha até a linha mais próxima. Faça uma interpretação da aleatoriedade do lançamento da agulha em termos de* X *e* θ.

b.

$x \in \mathbb{R}$	$x \leq 0$ ou $x > 1$	$0 < x \leq \frac{1}{2}$	$\frac{1}{2} \leq x < 1$
$f(x)$	0	x	$3x^2$

29. a. Note que $-1 \leq \frac{(x-\alpha)}{\beta} \leq 1$ e $\beta > 0$ e, portanto, $f(x) \geq 0$, $\forall x \in \mathbb{R}$. Também com um pouco de trabalho algébrico verificamos que $\int_{-\infty}^{\infty} f(x)dx = 1$.
b. Sendo F a função de distribuição de X, temos:

$x \in \mathbb{R}$	$x < \alpha - \beta$	$\alpha - \beta \leq x < \alpha + \beta$	$x \geq \alpha + \beta$
$F(x)$	0	$\frac{1}{2} + \frac{3}{4}\left(\frac{x-\alpha}{\beta}\right) - \frac{1}{4}\left(\frac{x-\alpha}{\beta}\right)^3$	1

c. $s = \alpha$.

31. a.

$x \in \mathbb{R}$	$x < 0$	$0 \leq x < 1$	$1 \leq x < 2$	$x \geq 2$
$F(x)$	0	$x - \frac{1}{2}x^2$	$1 - x + \frac{1}{2}x^2$	1

b. $5/8$. **c.** $1/9$.

33. a. $f(x) = 2\alpha(\alpha^2 - x^2)/(\alpha^2 + x^2)^2$. **b.** 0.

35. a. X é do tipo mista. O gráfico é omitido.
b. $P(X > -1) = 5/8$ e $P(X \leq 4|X > 0) = 1 - e^{-4} \simeq 0,9817$.
c. $F(x) = \frac{1}{2}F^d(x) + \frac{1}{2}F^{ac}(x)$ com

$x \in \mathbb{R}$	$x < -2$	$-2 \leq x < 0$	$x \geq 0$
$F^d(x)$	0	$\frac{1}{2}$	1

e

$x \in \mathbb{R}$	$x < -2$	$-2 \leq x < 0$	$x \geq 0$
$F^{ac}(x)$	0	$\frac{x+2}{4}$	$1 - \frac{1}{2}e^{-x}$

37. a. No. Acertos é $B(n = 3,\ p = \alpha\beta + (1 - \alpha)(1 - \beta))$. **b.** $\beta = 1$. **c.** Ele tem razão.

39. a. $P(X < 3|X \geq 0) = 4/5$ e $P(X \leq 3|X > 1) = 1$.
b. $F(x) = \frac{1}{3}F^d(x) + \frac{2}{3}F^{ac}(x)$ com

$x \in \mathbb{R}$	$x < 0$	$0 \leq x < 1$	$x \geq 1$
$F^d(x)$	0	$\frac{1}{2}$	1

e

$x \in \mathbb{R}$	$x < -2$	$-2 \leq x < 0$	$0 \leq x < 1$	$1 \leq x < 3$	$x \geq 3$
$F^{ac}(x)$	0	$\frac{x+2}{8}$	$\frac{1}{8}(x^2 + x + 2)$	$\frac{1}{16}(-x^2 + 8x + 1)$	1

41. Estude o quociente da aplicação da função de probabilidade nos pontos x e $x - 1$. Se $(n + 1)p$ não for inteiro, o máximo ocorre em $[n + 1)p]$ (maior inteiro contido). Se $(n + 1)p$ for inteiro temos duas respostas: $(n + 1)p$ e $(n + 1)p - 1$.

43. Do gráfico, $x = \mu$ é o ponto de máximo. De modo analítico, resolva $f' = 0$ e estude o sinal de f''.

Capítulo 3

Seção 3.2

1. Aplique a propriedade (FC4).
2. Calcule $P(-1 < X \leq \frac{1}{2}, -1 < Y \leq \frac{1}{2})$ e verifique que ocorre um absurdo.
3. a. $f_X(x) = \frac{1}{2}I_{[-1,1]}(x)$ e $f_Y(y) = I_{[0,1]}(y)$. São independentes pois $f_X f_Y = f_{(X,Y)}$.
b. Pela independência é a própria f_Y. **c.** Pela independência, $F_{(X,Y)} = F_X F_Y$. Logo

$x, y \in \mathbb{R}$	$x < -1$ ou $y < 0$	$-1 \leq x < 1, 0 \leq y < 1$	$x \geq 1, y \geq 1$
$F(x, y)$	0	$\frac{(x+1)y}{2}$	1

4. a. $f_X(x) = \frac{2\sqrt{1-x^2}}{\pi}I_{[-1,1]}(x)$ e $f_Y(y) = \frac{2\sqrt{1-y^2}}{\pi}I_{[-1,1]}(y)$. Não independentes: $f_X f_Y \neq f_{(X,Y)}$.

b. $f_{X|(Y=\frac{1}{2})}(x) = \frac{\sqrt{3}}{3} I_{[-\frac{\sqrt{3}}{2},\frac{\sqrt{3}}{2}]}(x)$.

5. a. $f_X(x) = \frac{1}{4}I_{[-2,0)}(x) + \frac{1}{2}I_{[0,1]}(x)$ e $f_Y(y) = \frac{1}{4}I_{[-2,0)}(y) + \frac{1}{2}I_{[0,1]}(x)$. Note que X e Y são identicamente distribuídos, mas não são independentes pois $f_X f_Y \neq f_{(X,Y)}$.

b. Para calcular a função de distribuição é preciso considerar vários intervalos de x e y.

$x, y \in \mathbb{R}$	$x < -2$ ou $y < -2$ ou $(-2 \leq x < 0, y < 0)$	$-2 \leq x < 0, 0 \leq y < 1$
$F(x,y)$	0	$\frac{(x+2)y}{4}$

$x, y \in \mathbb{R}$	$-2 \leq x < 0, y \geq 1$	$0 \leq x < 1, -2 \leq y < 0$	$0 \leq x < 1, 0 \leq y < 1$	$0 \leq x < 1, y \geq 1$
$F(x,y)$	$\frac{(x+2)}{4}$	$\frac{x(y+2)}{4}$	$\frac{x}{2} + \frac{y}{2}$	$\frac{(x+1)}{2}$

$x, y \in \mathbb{R}$	$x \geq 1, -2 \leq y < 0$	$x \geq 1, 0 \leq y < 1$	$x \geq 1, y \geq 1$
$F(x,y)$	$\frac{(y+2)}{4}$	$\frac{(y+1)}{2}$	1

c. Cuidado com o valor de x. $F(y|x)$ é zero para $x < -2$ ou $x > 1$. Para outros valores de x:

$$\underset{x\in[-2,0]}{F(y|x)} = \begin{cases} 0, & \text{se } y < 0; \\ y, & \text{se } 0 \leq y < 1; \\ 1, & \text{se } y \geq 1. \end{cases} \qquad \underset{x\in[0,1]}{F(y|x)} = \begin{cases} 0, & \text{se } y < -2; \\ \frac{(y+2)}{2}, & \text{se } -2 \leq y < 0; \\ 1, & \text{se } y \geq 0. \end{cases}$$

6. a. $X \sim \textit{Poisson}(\lambda p)$. **b.** Para $y \geq x$, inteiro, $P(Y = y|X = x) = P(W = y - x)$, sendo $W \sim \textit{Poisson}(\lambda(1-p))$.

7. a. $f(x,y) = \frac{1}{8}I_{(0,4)}(x)\, I_{(0,x)}(y)$. **b.** $f_Y(y) = \frac{4-y}{8}I_{(0,4)}(y)$.

8. a.

$x \in \mathbb{R}$	$x < 0$	$0 \leq x < 2$	$x \geq 2$
$F_X(x)$	0	$\frac{8x^2-x^4}{16}$	1

$y \in \mathbb{R}$	$y < 0$	$0 \leq y < 2$	$y \geq 2$
$F_Y(y)$	0	$\frac{y^4}{16}$	1

b. $f(x,y) = \frac{1}{2}xy\, I_{[0,y)}(x)\, I_{[0,2)}(y)$. **c.** $f_X(x) = \frac{4x-x^3}{4}I_{[0,2)}(x)$ e $f_Y(y) = \frac{y^3}{4}I_{[0,2)}(y)$.

9. a. Com y fixado, $f(.|y)$ é a expressão de uma *Poisson*(y). **b.** $f(x,y) = \frac{y^x e^{-2y}}{x!}, x \geq 0$ (inteiro) e $y > 0$. **c.** $f_X(x) = \frac{1}{2^{x+1}}$, $x \geq 0$, inteiro.

10. a. Temos $F(x,y) = \begin{cases} 0, & \text{se } x < 0 \text{ ou } y < 0\,; \\ y, & \text{se } x \geq 0, 0 \leq y < 1; \\ 1, & \text{se } x \geq 0, y \geq 1. \end{cases}$

b. X é a constante 0, logo sua função de distribuição tem um salto no zero. **c.** $Y \sim U_c(0,1)$ logo é contínua. **d.** Não, $F(x,y)$ induz probabilidade zero em todos os pontos de $\mathbb{R}^2$ mas tem discontinuidades.

Seção 3.3

1. a. $F_{-X}(x) = 1 - F(-x^-)$. **b.** $F_{aX+b}(x) = F\big(\frac{x-b}{a}\big)$. **c.** Seja $Y = XI_{(a,b)}(X)$ temos:

$y \in \mathbb{R}$	$y < 0$	$0 \leq y < a$	$a \leq y < b$	$y \geq b$
$F_Y(y)$	0	$1 - F(b^-) - F(a)$	$1 + F(y) - F(b^-) - 2F(a)$	1

d. A questão relevante é avaliar $F(x^-)$ e $F(x)$. Se X é contínua então elas são iguais para todo $x \in \mathbb{R}$. No caso discreto, poderia haver diferença se x fosse um valor assumido pela variável.

2. a. $P(h(X) \leq y) = F(h^{-1}(y))$. **b.** $P(h(X) \leq y) = 1 - F([h^{-1}(y)]^-)$.

3. $F_Y(y) = P(X^2 \leq y) = P\Big(-\sqrt{y} \leq X \leq \sqrt{y}\Big) = F\Big(\sqrt{y}\Big) - F\Big((-\sqrt{y})^-\Big)$.

4. Pela independência: $P(2X+Y\le z)=\int_{-\infty}^{\infty}\int_{-\infty}^{\frac{z-y}{2}} f_X(x)f_Y(y)\,dx\,dy$. Um caminho é resolver a integral e depois derivar. Outro, mais simples é reescrever a integral na forma: $\int_{-\infty}^{z} h(u)\,du$. A densidade será $\frac{1}{\sqrt{2\pi}}\frac{e^{-\frac{z^2}{10}}}{\sqrt{5}}$.

5. Note que X e kY também são independentes e a densidade de kY é $\frac{\lambda_2}{k}e^{-\frac{\lambda_2}{k}y}, y>0$. Use a expressão da densidade da diferença de variáveis para obter $\frac{\lambda_1\lambda_2}{k\lambda_1+\lambda_2}e^{-\lambda_1 w}, w>0$. Para completar a demonstração, calcule $P(X-kY>0)=\int_0^{\infty} f_{X-kY}(w)\,dw$.

6. $f_{X-Y}(w)=\frac{1}{2}e^{-w}I_{[0,\infty)}(w)+\frac{1}{2}e^{w}I_{(-\infty,0)}(w)$.

7. **a.** $f_{X+Y}(z)=z^2I_{(0,1)}(z)+z(2-z)I_{[1,2)}(z)$. **b.** $2/3$. **c.** $1/8$.

8. **a.** São independentes. $X\sim N(0,2\pi)$ e $Y\sim Exp(1)$. **b.** Calcule primeiro $F_{X+Y|Y=y}(w|y)$ e, então obtenha $f_{X+Y|Y=y}(w|y)$. Logo,

$$f_{Y,X+Y}(y,w)=f_{X+Y|Y=y}(w|y)f_Y(y)=\frac{1}{2\pi}e^{-y}e^{-\frac{(w-y)^2}{4\pi}}I_{(-\infty,\infty)}(w)I_{(0,\infty)}(y).$$

9. Sejam $Z=X+Y$ e $W=X/Y$. Via Jacobiano vem $f_{Z,W}(z,w)=\frac{\lambda^2 z\,e^{-\lambda z}}{(w+1)^2}I_{(0,\infty)}(z)I_{(0,\infty)}(w)$. Obtenha as marginais: $f_Z(z)=\lambda^2 z\,e^{-\lambda z}I_{(0,\infty)}(z)$ e $f_W(w)=\frac{1}{(w+1)^2}I_{(0,\infty)}(w)$.

10. **a.** Lembrar que X é contínua. $F_Y(y)=F_X\big(\sqrt{y}\big)-F_X\big(-\sqrt{y}\big)$. Derive aplicando a regra da cadeia e obtenha a densidade da Qui-quadrado com 1 gl: $f_Y(y)=\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{y}}e^{-\frac{1}{2}y}$. **b.** Faça a partição de domínio em $x<0$ e $x\ge 0$. Proceda como no Exemplo 3.30.

Seção 3.4

1. Precisamos mostrar que $[\min(X,Y)\le t]$ e $[\max(X,Y)\le t]$ pertencem a $\mathcal{F}$, $\forall t\in\mathbb{R}$. Para o mínimo trabalhe com $[\min(X,Y)>t]=[X>t]\cap[Y>t]$. Para máximo use que $[\max(X,Y)\le t]=[X\le t]\cap[Y\le t]$.

3. Note que $A_i=\{w; I_{A_i}(w)=1\}$. Verifique o resultado para a intersecção dos A_i's e correspondente intersecção dos $[I_{A_i}\text{'s}=1]$. Os outros casos seguem similarmente.

5. **a.** $F_{X,Y}(x,y)=P[X\le x,h(X)\le y)]=P(X\le x,X\le h^{-1}(y))=\min\{F_X(x),F_X(h^{-1}(y)\}$ $=\min\{F_X(x),F_{h(X)}(y)\}$. **b.** $F_{X,Y}(x,y)=P[X\le x,h(X)\le y)]=P(X\le x,X\ge h^{-1}(y))$. Temos dois casos a considerar. Se $x\le h^{-1}(y)$ então a probabilidade é zero. Se $x>h^{-1}(y)$ então a probabilidade é igual a $F_X(x)-F_X\{[(h^{-1}(y)]^-\}=F_X(x)-\big(1-F_{h(X)}(y)\big)$. Daí segue o resultado.

7. Note que $P(X\le x)=P(X\le x,Y\le y)+P(X\le x,Y>y)$. Desenvolvendo obtenha que $F_{X,Y}(x,y)\ge F_X(x)+F_Y(y)-1$. Por outro lado, $F_{X,Y}(x,y)\le F_X(x)$ e também $F_{X,Y}(x,y)\le F_Y(y)$. Daí $F_{X,Y}(x,y)\le\sqrt{F_X(x)F_Y(y)}$.

9.

$y\in\mathbb{R}$	$y<0$	$0\le y<1$	$1\le y<2$	$y\ge 2$
$F_Y(y)$	0	e^{-2}	$3e^{-2}$	1

11. **a.**

y	2	5/2	3	7/2
$p(y)$	1/32	5/32	25/32	1/32

b.

$y\in\mathbb{R}$	$y<2$	$2\le y<5/2$	$5/2\le y<3$	$3\le y<7/2$	$y\ge 7/2$
$F_Y(y)$	0	1/32	3/16	31/32	1

13. X e XY têm a mesma distribuição: -1 e 1 com probabilidade $1/2$. A conjunta (X,XY) tem probabilidade $1/4$ em cada um dos pares possíveis. Então X e XY são independentes.

15. $p/2-p$.

17. $p_{X,Y}(x,y) = \binom{x}{y}\left(\frac{1}{2}\right)^x\frac{x}{3}$, $x = 1, 2$ e $y \leq x, y \in \mathbb{N}$. $p_Y(y) = \frac{1}{6}\left[\left[\binom{1}{y} + \binom{2}{y}\right]\right], y = 0, 1, 2.$

19. a. $f(y) = 2e^{-2(y-a)}I_{(a,\infty)}(y)$. **b.** $0,1192$.

21. $f(y) = \frac{1}{y\sqrt{2\pi}}e^{-\frac{1}{2}(\log y)^2}I_{(0,\infty)}(y)$.

23. a. Omitida. **b.** $1/2$.

25. a. Omitida. **b.** $f_X(x) = 2(1-x)\,I_{(0,1)}(x)$, $f_Y(y) = 2y\,I_{(0,1)}(y)$.

27. a. Omitida. **b.** Não são independentes: $f_X(x) = \frac{x^3}{4}\,I_{(0,2)}(x)$, $f_Y(y) = (y - \frac{y^3}{4})I_{(0,2)}(y)$, .

29. $p_Y(y) = 1/(n+1)$, $n = 0, 1, \cdots$.

31. a. $f_X(x) = 4x(1-x^2)I_{(0,1)}(x)$, $f_Y(y) = 4y^3\,I_{(0,1)}(y)$. **b.** $25/81$.

33. Marginal de Y: $f_Y(y) = y\,I_{(0,1]}(y) + (2-y)I_{(1,2]}(y)$. Condicional de X dado Y: $f_{X|Y}(x|y) = \frac{1}{y}\,I_{(0,y)}(x)\,I_{(0,1]}(y) + \frac{1}{(2-y)}I_{(y-1,1)}(x)I_{(1,2]}(y)$.

35. a. $F_Y(y) = (1 - e^{-2y})\,I_{[2,\infty]}$.

b. $\alpha_d = 1 - e^{-4}$, $\alpha_{ac} = e^{-4}$ e $\alpha_s = 0$. Temos $F_Y(y) = \alpha_d F^d(y) + \alpha_{ac}F^{ac}(y)$ com

$y \in \mathbb{R}$	$y < 2$	$y \geq 2$
$F^d(y)$	0	1

e

$y \in \mathbb{R}$	$y < 2$	$y \geq 2$
$F^{ac}(y)$	0	$1 - e^{4-2y}$

.

37. a.

$y \in \mathbb{R}$	$y < 0$	$0 \leq y < 2$	$2 \leq y < 3$	$y \geq 3$
$F_Y(y)$	0	$\frac{y^2}{8}$	$\frac{1}{2} + \frac{1}{8}(y-2)$	1

.

b. $P(Y < 3|Y \geq 2) = 1/4$ e $P(Y \leq 3|Y > 2) = 1$.

c. $\alpha_d = 3/8$, $\alpha_{ac} = 5/8$ e $\alpha_s = 0$. Temos $F_Y(y) = \alpha_d F^d(y) + \alpha_{ac}F^{ac}(y)$ com

$y \in \mathbb{R}$	$y < 3$	$y \geq 3$
$F^d(y)$	0	1

e

$y \in \mathbb{R}$	$y < 0$	$0 \leq y < 2$	$2 \leq y < 3$	$y \geq 3$
$F^{ac}(y)$	0	$\frac{y^2}{5}$	$\frac{4}{5} + \frac{1}{5}(y-2)$	1

39. a.

$w \in \mathbb{R}$	$w < 1$	$1 \leq w < 3/2$	$3/2 \leq w < 2$	$w \geq 2$
$F_W(y)$	0	$\frac{w^2}{12}$	$\frac{w^2}{4}$	1

.

b. $P(1 < W \leq \frac{7}{4}|W \geq \frac{3}{2}) = 37/52$.

c. $\alpha_d = 11/24$, $\alpha_{ac} = 13/24$ e $\alpha_s = 0$. Temos $F_Y(y) = \alpha_d F^d(y) + \alpha_{ac}F^{ac}(y)$ com

$w \in \mathbb{R}$	$w < 1$	$1 \leq w < \frac{3}{2}$	$w \geq \frac{3}{2}$
$F^d(w)$	0	$\frac{2}{11}$	1

e

$w \in \mathbb{R}$	$w < 1$	$1 \leq w < \frac{3}{2}$	$\frac{3}{2} \leq w < 2$	$w \geq 2$
$F^{ac}(w)$	0	$\frac{2(w^2-1)}{13}$	$\frac{5}{26} + \frac{6}{13}(w^2 - \frac{9}{4})$	1

.

41. a. $f_{Y_1,Y_2}(y_1,y_2) = 2\lambda^2 e^{-\lambda(y_1+y_2)}I_{(0,y_2)}(y_1)I_{(0,\infty)}(y_2)$. **b.** $f_{Y_2-Y_1|Y_1}(w) = \lambda e^{-\lambda w}I_{(0,\infty)}(w)$.

43. Conjunta: $f_{U,V}(u,v) = \frac{1}{4}e^{-\frac{(3u-v)}{4}}I_{(-u,u)}(v)I_{(0,\infty)}(u)$. As marginais são as seguintes: $f_U(u) = e^{-\frac{u}{2}}(1 - e^{-\frac{u}{2}})I_{(0,\infty)}(u)$ e $f_V(v) = \frac{1}{3}e^{v}I_{(-\infty,0)}(v) + \frac{1}{3}e^{-\frac{1}{2}v}I_{[0,\infty)}(v)$. Não são independentes.

45. $f_{W_1,W_2}(w_1,w_2) = 4w_1w_2I_{[0,1]}(w_1)I_{[0,1]}(w_2)$.

47. Conjunta: $f_{U,Y}(u,y) = \frac{2y}{\alpha^2}e^{-\frac{y(u+1)}{\alpha}}I_{(0,1)}(u)I_{(0,\infty)}(y)$. Marginais: $f_U(u) = \frac{2}{(u+1)^2}I_{(0,1)}(u)$ e $f_Y(y) = \frac{2}{\alpha}e^{-\frac{y}{\alpha}}(1 - e^{-\frac{y}{\alpha}})I_{(0,\infty)}(y)$. Não são independentes.

49. a. $f_{W,Z}(w,z) = \frac{1}{8}I_A(w,z)$, $A = \{(w,z);\ -1 < \frac{w+z}{2} < 1, -1 < \frac{z-w}{2} < 1\}$.

b. $f_W(w) = \frac{2+w}{4}I_{(-2,0)}(w) + \frac{2-w}{4}I_{(0,2)}(w)$;
$f_{Z|W}(z|w) = \frac{1}{2(2+w)}I_{(-2-w,2+w)}(z) + \frac{1}{2(2-w)}I_{(-2+w,2-w)}(z)$.

51. Sejam $U = X + Y$ e $V = X$. Da conjunta $f_{U,V}(u,v) = \frac{1}{2|x|} I_{[-1,1]}(v) I_{[v,v+|v|)}(u)$, vem a marginal

$u \in \mathbb{R}$	$u < -1$ ou $u > 2$	$-1 \leq u < 0$	$0 \leq u < 1$	$1 \leq u \leq 2$
$f_U(u)$	0	$-\frac{1}{2}\ln(-u)$	$\frac{1}{2}\ln 2$	$-\frac{1}{2}[\ln(u) - \ln(2)]$

53. Via Jacobiano, obtemos a conjunta abaixo e daí segue a verificação dos itens (a), (b) e (c).
$f_{W,T}(w,t) = \left[\frac{1}{B(\alpha,\beta)} w^{\alpha-1}(1-w)^{\beta-1} I_{(0,1)}(w)\right]\left[\frac{1}{\Gamma(\alpha+\beta)} t^{\alpha+\beta-1} e^{-\lambda t} I_{(0,\infty)}(t)\right]$

55. a. Use o método do Jacobiano com separação de domínio. Como as variáveis são contínuas, independentes e identicamente distribuídas, vem $f_{X_1,X_2,X_3}(x_1,x_2,x_3) = f(x_1)f(x_2)f(x_3)$, com $a < x_i < b, i = 1,2,3$ e todos os x_i's diferentes. Entretanto, a transformação com as estatísticas de ordem não é bijetora. Por exemplo, com valores hipotéticos, se $(Y_1, Y_2, Y_3) = (0,1,2)$, a amostra poderia ter sido $(X_1, X_2, X_3) = (0,2,1)$ ou, ainda, $(X_1, X_2, X_3) = (2,0,1)$ ou $\cdots$. Portanto, para obter bijeção, precisamos particionar o domínio em conjuntos $S_{i,j,k}$ definidos por:

$$S_{i,j,k} = \{(x_1,x_2,x_3); a < x_i < x_j < x_k < b, \text{ todos diferentes}\}.$$

Por exemplo, para $S_{2,3,1}$ temos a transformação: $y_1 = x_2, y_2 = x_3, y_3 = x_1$. Existem $3! = 6$ desses conjuntos, cujos correspondentes Jacobianos são iguais a 1 ou -1. Daí o resultado segue, fazendo a soma das aplicações da transformação em cada subconjunto da partição.

b. $n! \prod_{i=1}^{n} f(y_i)\, I_{(a<y_1<\cdots<y_n<b)}(y_1,\ldots,y_n)$.

57. a. $P(Y_1 > t) = P(X_1 > t, X_2 > t, \cdots, X_n > t) = \prod_{i=1}^{n} P(X_i > t) = e^{-\sum_{i=1}^{n}\lambda_i} \Rightarrow Y_1 \sim Exp(\sum_{i=1}^{n}\lambda_i)$.

b. $[X_k = \min(X_1,\cdots,X_n)] = [X_k \leq \min_{i \neq k}(X_1,\cdots,X_n)]$. Note que essas duas variáveis são independentes. Use a expressão da diferença para achar a densidade de $X_k - \min_{i \neq k}(X_1,\cdots,X_n)$. Integre na região $(-\infty, 0)$ para obter a probabilidade de interesse. **c.** $P(Y_n \leq t) = \prod_{i=1}^{n}(1\text{-}e^{-\lambda_i t})$.

59. $f_Y(y) = \frac{e^{-\arccos y} + e^{-(2\pi - \arccos y)}}{(1-e^{-2\pi})(\sqrt{1-y^2})}, -1 < y < 1$.

61. $2c/\pi a$.

Capítulo 4

Seção 4.2

1. $X \sim Hgeo(m,n,r)$. Vamos usar as seguintes igualdades combinatórias:

$$k\binom{m}{k} = m\binom{m-1}{k-1} \quad \text{e} \quad r\binom{n}{r} = n\binom{n-1}{r-1}.$$

Note que $\binom{x}{k} = 0$ se $k < 0$ ou $k > x$. Para facilitar a demonstração, consideramos o valor de k entre 0 e r, desprezando os termos que serão zerados. Enfatizamos que os valores de k estão efetivamente entre $\max(0, r-(n-m))$ e $\min(r,m)$.