Variáveis aleatórias contínuas

Caracterização de uma v.a. contínua

Função densidade de probabilidade

A *função densidade de probabilidade* (f.d.p.) de uma variável aleatória X é uma função $f(x) \geq 0$ satisfazendo, para qualquer subconjunto de números reais B,

$$P(X \in B) = \int_B f(x)dx$$

Em particular, se $B = [a, b]$

$$P(X \in B) = P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Propriedades da f.d.p.

(i) $f(x) \geq 0$

(ii) Se $B = (-\infty, \infty) = \mathbb{R}$

$$P(X \in B) = P(-\infty \leq X \leq \infty) = 1$$

logo $f(x)$ deve satisfazer $\int_{-\infty}^{\infty} f(x)dx = 1$.

(iii) Se $B = [a, b]$ com $a = b$ então

$$P(X \in B) = P(X = a) = \int_a^a f(x)dx = 0.$$

Valor esperado

Se X é uma variável aleatória contínua com função densidade de probabilidade $f(x)$, então

$$EX = \int_{-\infty}^{\infty} xf(x)dx$$

Proposição

(i) Se X é uma v.a. não negativa com f.d.p $f(x)$, então, para qualquer função g tomando valores em $\mathbb{R}$, vale que

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

(ii) Se a e b são constantes reais, então

$$E[aX + b] = aEX + b.$$

Variância de uma v.a. contínua

Variância

Se X é uma variável aleatória com f.d.p $f(x)$ e $EX = \mu$, então a *variância* de X é definida como

$$Var(X) = E[(X-\mu)^2]x = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

e o *desvio padrão* de X é definido como $DP(X) = \sqrt{Var(X)}$

Alternativamente,

$$Var(X) = E[X^2] - (EX)^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2.$$

Proposição

Seja X é uma variável aleatória contínua. Para quaisquer a e b constantes reais, vale que

$$Var(aX+b) = a^2 Var(X)$$

Relação entre f.d.p e a média e a variância

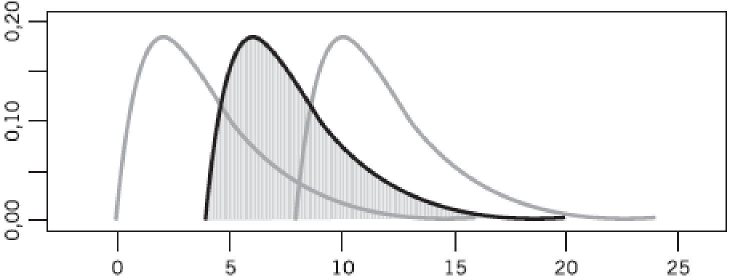


Figura 3.6 – Três funções densidade com médias diferentes e com a variância constante

Relação entre f.d.p e a média e a variância

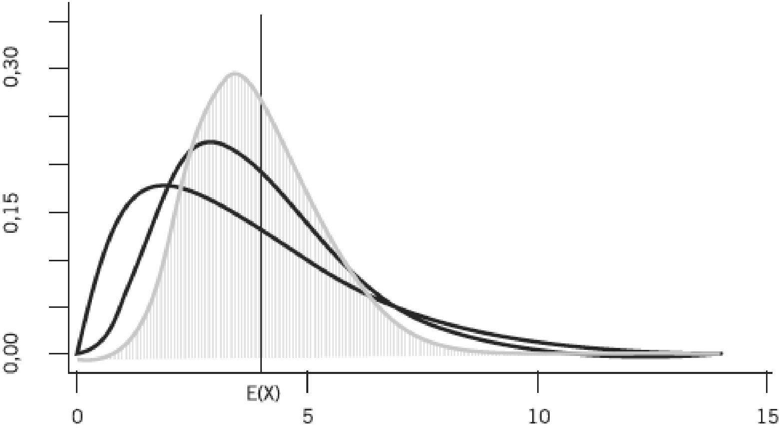


Figura 3.7 – Três funções densidade com variâncias diferentes e com a média constante E(X)

Função de Distribuição de uma v.a.

Dada uma v.a. X **contínua** com f.d.p $f(x)$, a *função de distribuição acumulada* (f.d.a.) de X é definida como

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy$$

Proposição

Dada uma v.a. X (contínua ou discreta) com f.d.a F temos que

(i) F é uma função não decrescente ($x < y$ então $F(x) \leq F(y)$);

(ii) $\lim_{x \to -\infty} F(x) = 0$;

(iii) $\lim_{x \to \infty} F(x) = 1$;

(iv) Se $a < b$ então $P(a < X < b) = F(b) - F(a)$.

Relação entre f.d.a. e f.d.p

Por definição, a relação entre a função distribuição acumulada F e a função densidade de probabilidade f é dada por

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

Derivando em ambos os lados da igualdade temos

$$\frac{d}{da}F(a) = f(a)$$

1. Considere $X \sim Unif(0,5)$ e defina

$$Y = \begin{cases} 3, & se\ X < 2 \\ 7, & se\ 2 \leq X < 4 \\ 8, & se\ X > 4. \end{cases}$$

Defina

(a) a função de probabilidade de Y;

(b) o valor esperado e o desvio padrão de Y.

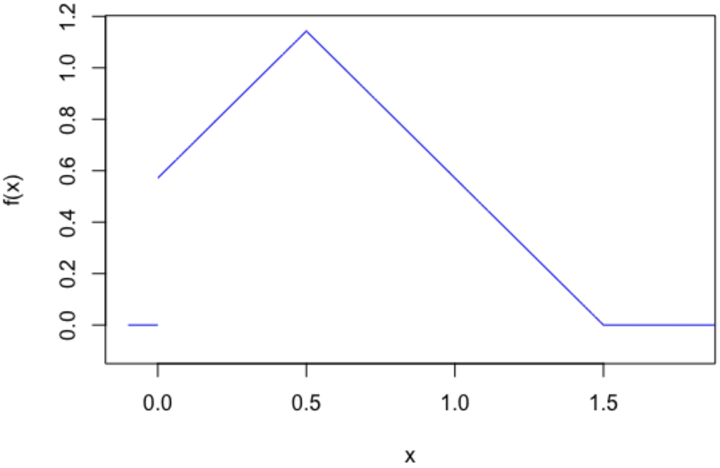
2. Seja X uma v.a. contínua com f.d.p. dada por

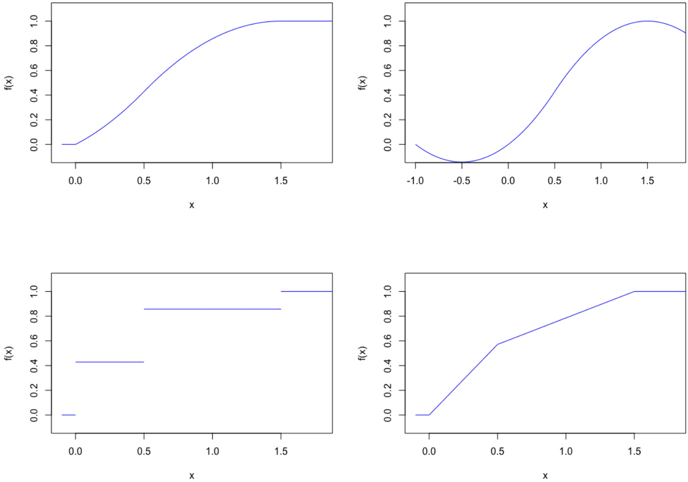
$$f(x) = \begin{cases} c, & se\ 0 < x < \ln(5)/6 \\ 6e^{-6x}, & se\ x \geq \ln(5)/6 \end{cases}$$

(a) Determine o valor de c.

(b) Determine a função de distribuição acumulada de X.

4 Determine a função de distribuição acumulada da variável aleatória X cuja a função densidade de probabilidade é dada abaixo




f(x)
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5
x
f(x)
0.0 0.2 0.4 0.6 0.8 1.0
-1.0 -0.5 0.0 0.5 1.0 1.5
x
f(x)
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5
x
f(x)
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5
x

Variável Aleatória Uniforme

Uma variável aleatória X é dita *uniformemente* distribuída ao longo do intervalo (a, b) se a sua função densidade de probabilidade é dada por

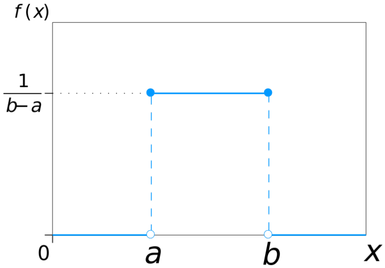
$$f(x) = \begin{cases} \dfrac{1}{b-a}, & se \ a < x < b \\ 0 & \text{caso contrário} \end{cases}.$$

Nessa caso, denotamos $X \sim Unif(a, b)$, e

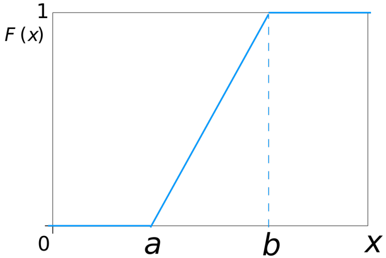
$$EX = \frac{b+a}{2} \qquad e \qquad Var(X) = \frac{(b-a)^2}{12}.$$

Além disso,

$$F(x) = \begin{cases} 0 & se \ x \leq a \\ \dfrac{x-a}{b-a} & se \ a < x < b \\ 1 & se \ x \geq b \end{cases}$$



f.d.p



f.d.a

Variável Aleatória Exponencial

Uma variável aleatória contínua X é dita *exponencial* se sua função densidade de probabilidade é dada, para algum $\lambda > 0$, por

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & se\ x \geq 0 \\ 0 & se\ x \leq 0 \end{cases}.$$

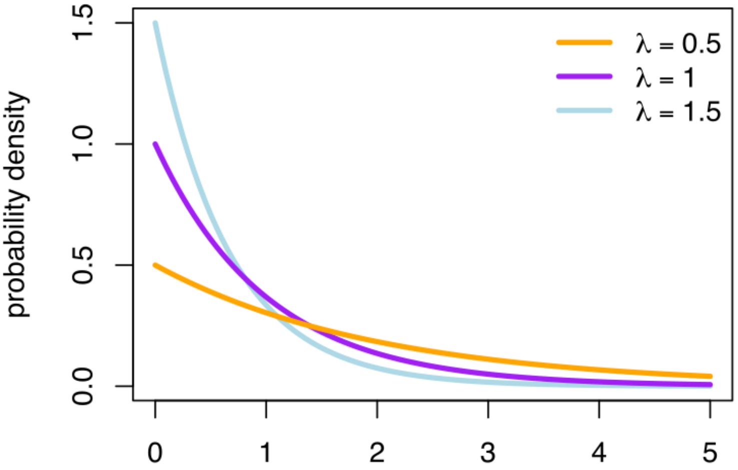
Nesse caso, denotamos $X \sim Exp(\lambda)$ e

$$EX = \frac{1}{\lambda} \qquad e \qquad Var(X) = 1/\lambda^2$$

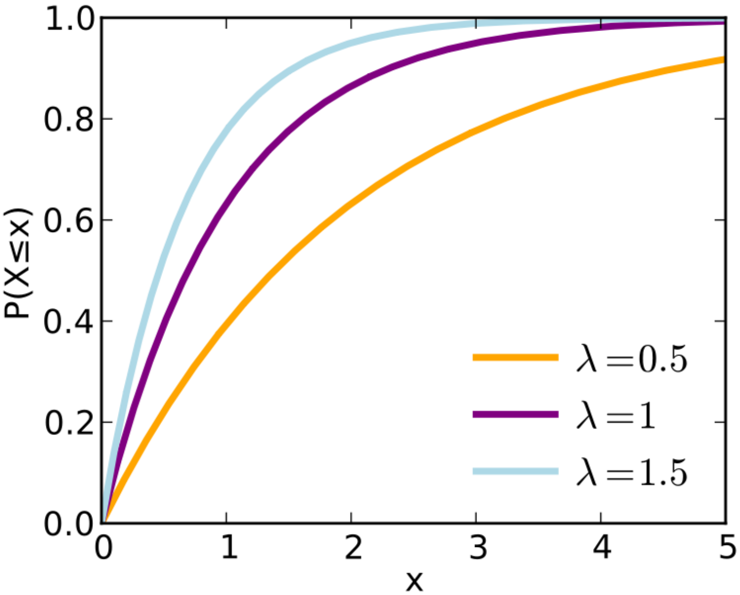
Além disso,

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & se\ x \geq 0 \\ 0 & se\ x \leq 0 \end{cases}.$$

f.d.p.



f.d.a.



Variável Aleatória Normal

Dizemos que X é uma variável aleatória *normal* (ou X é normalmente distribuída) com parâmetros μ e σ^2, se a função densidade de X é dada, para todo $x \in \mathbb{R}$ por

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

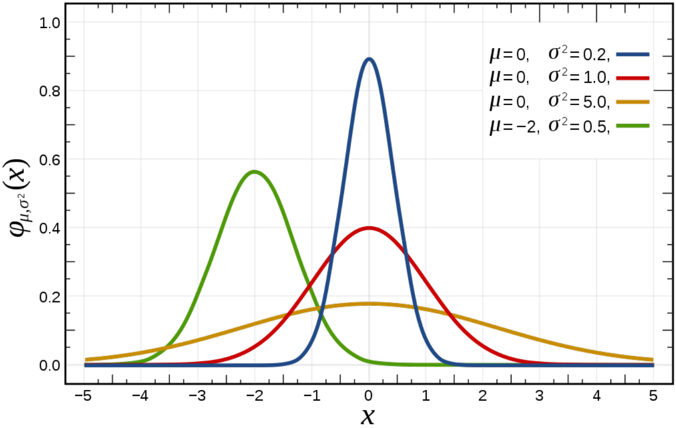
Nesse caso, denotamos $X \sim N(\mu, \sigma^2)$ e

$$EX = \mu \qquad e \qquad Var(X) = \sigma^2.$$

No caso especial em que $\mu = 0$ e $\sigma^2 = 1$ dizemos que X é uma v.a. *Normal padrão* e geralmente denotamos por Z e sua f.d.p por

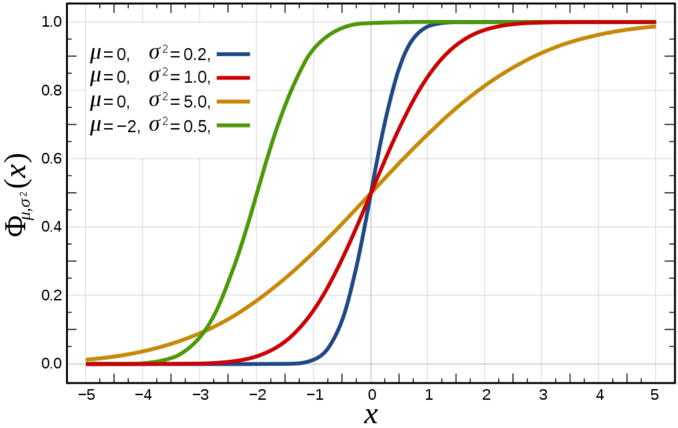
$$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

f.d.p



$\mu=0, \quad \sigma^2=0.2,$
$\mu=0, \quad \sigma^2=1.0,$
$\mu=0, \quad \sigma^2=5.0,$
$\mu=-2, \quad \sigma^2=0.5,$
1.0
0.8
0.6
0.4
0.2
0.0
$\varphi_{\mu,\sigma^2}(x)$
−5
−4
−3
−2
−1
0
1
2
3
4
5
x

f.d.a.



1.0
0.8
0.6
0.4
0.2
0.0
$\Phi_{\mu,\sigma^2}(x)$
$\mu=0, \quad \sigma^2=0.2,$
$\mu=0, \quad \sigma^2=1.0,$
$\mu=0, \quad \sigma^2=5.0,$
$\mu=-2, \quad \sigma^2=0.5,$
−5
−4
−3
−2
−1
0
1
2
3
4
5
x

Propriedades

Proposição

Se $X \sim N(\mu, \sigma^2)$ e $Y = aX + b$ então $Y \sim N(a\mu + b, a^2\sigma^2)$

Em particular se $X \sim N(\mu, \sigma^2)$ então $Z = \frac{X-\mu}{\sigma}$ é um Normal padrão

Proposição

Se Z é uma v.a. Normal padrão, então sua distribuição é simétrica em torno de zero. Isto é, para qualquer $x \in \mathbb{R}$,

$$P(Z \leq -x) = P(Z \geq x).$$

Em particular, se $\Phi(x)$ é a f.d.a de Z, dada por

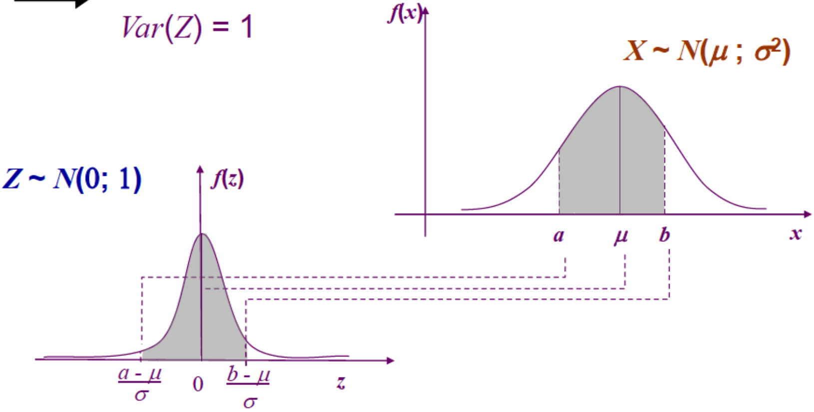
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2} dz,$$

então $\Phi(x)$ satisfaz, para qualquer valor $x \in \mathbb{R}$,

$$\Phi(-x) = 1 - \Phi(x)$$

Se $X \sim N(\mu\,;\,\sigma^2)$, definimos $\boxed{Z = \dfrac{X-\mu}{\sigma}}$

$$\Rightarrow \quad \begin{aligned} E(Z) &= 0 \\ Var(Z) &= 1 \end{aligned}$$



5 Uma barra de chocolates é vendida em embalagens de 90g. O fabricante exige que cada embalagem tenha pelo menos 85g, para evitar ser multado, e no máximo 100g, para evitar prejuízos. Se o fabricante compra uma máquina que produz chocolates com uma distribuição Normal de média 90g e desvio padrão 5g, a probabilidade de um chocolate estar dentro da exigência NÃO pode ser escrita como

(a) $P(-2 \leq Z \leq 1)$

(b) $\varphi(1) - \varphi(2)$

(c) $\varphi(2) - 1 + \varphi(1)$

(d) $P(-1 \leq Z \leq 2)$

Aproximação normal para a distribuição binomial

Teorema

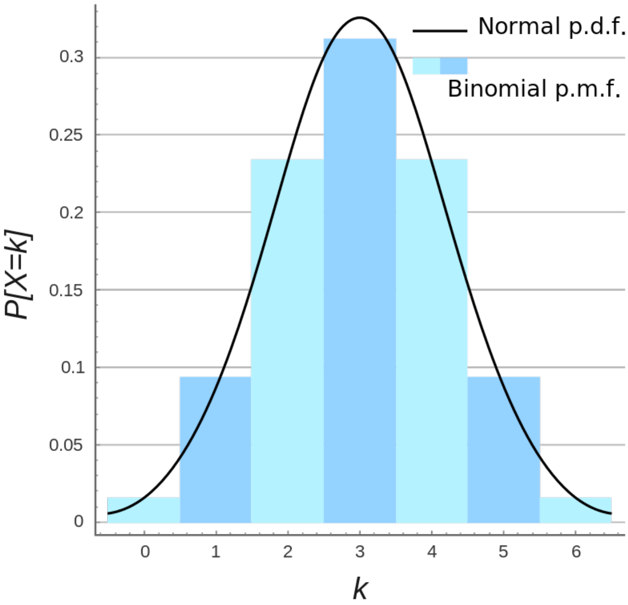
Se X, representa o número de sucessos que ocorrem quando n tentativas independentes, cada uma com probabilidade de sucesso p, são realizadas, então, para qualquer $a < b$, quando $n \to \infty$

$$P\left[a \leq \frac{X - np}{\sqrt{np(1-p)}} \leq b\right] \to P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$$

Note que
$X \sim Bin(n, p)$ implica que $EX = np$ e $Var(X) = np(1-p)$.

Ou seja, o teorema afirma é que quando $n \to \infty$, a distribuição de uma $Bin(n, p)$ se aproxima de uma $N(np, np(1-p))$.

Caso particular do Teorema do Limite Central.



Normal p.d.f.
Binomial p.m.f.
0.3
0.25
0.2
0.15
0.1
0.05
0
P[X=k]
0
1
2
3
4
5
6
k

6. Em um estudo sobre células cancerígenas observou-se que, sob determinadas condições, a proporção de células cancerígenas é de 1 a cada 100 células. Em uma lâmina com 7000 células, determine

(a) a probabilidade aproximada de pelo menos 2 células sejam cancerígenas.

(b) Que hipótese você está assumindo?

Teorema

Seja X uma variável aleatória contínua com função densidade de probabilidade f_X. Suponha que $g(x)$ seja uma função estritamente monotônica (crescente ou decrescente) e derivável (portanto contínua). Então a variável aleatória Y definida por $Y = g(X)$ tem uma função densidade de probabilidade dada por

$$f_Y(y) = \begin{cases} f_X[g^{-1}(y)] \left| \frac{d}{dy} g^{-1}(y) \right|, & se\ y = g(x)\ para\ algum\ x \\ 0 & se\ y \neq g(x)\ para\ todo\ x \end{cases}$$

Valor esperado e variância

Proposição

• Se X é uma v.a. discreta com função de distribuição $p(x_i)$ e $Y = g(X)$ então

$$EY = E(g(X)) = \sum_i g(x_i)p(x_i)$$

$$Var(Y) = Var(g(X)) = \sum_i \left(g(x_i) - E[g(X)]\right)^2 p(x_i)$$

• Se X é uma v.a. contínua com f.d.p $f(x)$ e $Y = g(X)$ então

$$EY = E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

$$Var(Y) = Var(g(X)) = \int_{-\infty}^{\infty} \left(g(x) - E[g(X)]\right)^2 f(x)dx$$

7. Seja X um v.a. com f.d.p

$$f(x) = \begin{cases} \frac{3x^2}{2} & -1 \leq x \leq 1 \\ 0 & c.c. \end{cases}$$

Determine

(a) a f.d.p de $Y = e^{-X}$;

(b) $E[3X^2]$;

(c) a f.d.p de $W = 3X^2$ e EW

Vetores aleatórios

X e Y v.a. discretas

As probabilidades associadas aos pares de valores nas variáveis X e Y,

$$p(x, y) = P(X = x, Y = y)$$

denotam a probabilidade do evento $\{X = x\} \cap \{Y = y\}$. O conjunto dos valores $p(x, y)$ é denominado a *distribuição de probabilidade conjunta de X e Y*. Sua representação pode ser feita através da tabela

Distribuição Marginal

Se X e Y são v.a. discretas com função de distribuição conjunta $p(x, y)$, a distribuição de X, (respec. de Y) pode ser obtida com

$$p_X(x) = \sum_y p(x, y) \quad \left(respec. \;\; p_Y(y) = \sum_x P(x, y) \right)$$

Se X e Y são v.a. contínuas com f.d.p conjunta $f(x, y)$, a f.d.p de X, (respec. de Y) pode ser obtida com

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \left(respec. \;\; f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx \right)$$

Essas distribuições recebem o nome de *marginais*.

FDA

Para quaisquer variáveis aleatórias X e Y, a *função distribuição (acumulada) conjunta* de X e Y é definida como

$$F(a,b) = P(X \leq a, Y \leq b) \quad -\infty \leq a, b \leq \infty$$

• Se X e Y são v.a. discretas com função de probabilidade conjunta $p(x_i, y_j)$ então

$$F(a,b) = \sum_{k:x_k \leq a} \sum_{\ell:y_\ell \leq b} p(x_k, y_\ell)$$

• Se X e Y são v.a. contínuas com função densidade de probabilidade conjunta $f(x,y)$ então

$$F(a,b) = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x,y)dxdy \quad \Longrightarrow f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

Independência entre duas v.a.

Duas v.a. X e Y são ditas independentes se para quaisquer dois conjuntos A e B vale que

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

Tomando $A = (-\infty, a]$ e $B = (-\infty, b]$ temos, para todo a e b,

$$F(a, b) = P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b) = F_X(a)F_Y(b)$$

De forma equivalente

- Se X e Y são v.a. discretas a condição de independência pode ser reescrita como

$$p(x, y) = p_X(x)p_Y(y) \;\; \textit{para todo } x \textit{ e } y$$

- Se X e Y são v.a. contínuas a condição de independência pode ser reescrita como

$$f(x, y) = f_X(x)f_Y(y) \;\; \textit{para todo } x \textit{ e } y$$

Independência

De forma geral, n variáveis aleatórias, $X_1, \ldots, X_n$ são ditas independentes se para quaisquer conjuntos $A_1, \ldots, A_n$ vale que

$$\begin{aligned} P(X_1 \in A_1, X_2 \in A_2, \ldots X_n \in A_n) &= P(X_1 \in A_1)P(X_2 \in A_2)\ldots P(X_n \in A_n) \\ &= \prod_{i=1}^{n} P(X_i \in A_i) \end{aligned}$$

Tomando $A_i = (-\infty, a_i], i = 1 \ldots, n$, temos

$$F(a_1, \ldots, a_n) = \prod_{i=1}^{n} P(X_i \leq a_i) = \prod_{i=1}^{n} F_{X_i}(a_i)$$

• Se $X_1, \ldots, X_n$ são v.a. discretas com função de probabilidade $p_{X_i}, i = 1, \ldots, n$, a condição de independência pode ser reescrita como

$$p(x_1, \ldots, x_n) = p_{X_1}(x_1) p_{X_2}(x_2) \ldots p_{X_n}(x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$$

• Se $X_1, \ldots, X_n$ são v.a. contínuas com função densidade de probabilidade $f_{X_i}, i = 1, \ldots, n$, a condição de independência pode ser reescrita como

$$f(x_1, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

Soma de v.a. independentes

Proposição

Dadas duas v.a. contínuas X e Y e independentes com f.d.p f_X e f_Y a função de distribuição acumulada de $Z = X + Y$, chamada de *convolução*, é dada por

$$F_Z(z) = \int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy = \int_{-\infty}^{\infty} F_Y(z-x) f_X(x) dx.$$

Além disso a f.p.d de Z é dada por

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy = \int_{-\infty}^{\infty} f_Y(z-x) f_X(x) dx.$$

8. Sejam X e Y variáveis aleatórias com distribuição conjunta dada por

$Y \setminus X$	2	3	4	$p_Y(y)$
-1	7/30		7/30	7/10
1	3/20	3/20		
$p_X(x)$				

(a) Complete a tabela de dupla entrada de maneira que as funções de probabilidade conjunta e marginais estejam bem definidas.

(b) Determine a f.d.a conjunta de X e Y.

(c) Verifique se X e Y são independentes

9. Sejam X e Y v.a's com f.d.p conjunta

$$f(x,y) = \begin{cases} \dfrac{4x}{125}, & se\ 0 \leq x < 5/2, 0 < y < 5, \\ \dfrac{4(5-x)}{125}, & se\ 5/2 \leq x < 5, 0 < y < 5, \end{cases}$$

(a) Determine a f.d.a conjunta de X e Y.

(b) Determine a f.d.p marginal de X.

(c) Determine a f.d.p marginal de Y.

(d) As variáveis X e Y são independentes?

Método Jacobiano

De forma geral, se jam X_1 e X_2 v.a.'s contínuas com f.d.p conjunta $f(x_1, x_2)$ e g_1 e g_2 duas funções em $\mathbb{R}$. Considere as variáveis $Y_1 = g_1(X_1, X_2)$ e $Y_2 = g_2(X_1, X_2)$.
Suponha que

(i) As equações $y_1 = g_1(x_1, y_1)$ e $y_2 = g_2(x_1, x_2)$ podem ser unicamente solucionadas para x_1 e x_2 em termos de y_1 e y_2, com soluções dadas por, digamos, $x_1 = h_1(y_1, y_2)$ e $x_2 = h_2(y_1, y_2)$.

(ii) As funções g, e g, têm derivadas parciais contínuas em todos os pontos (x_1, x_2), e são tais que, para todo (x_1, x_2), o determinante 2 X 2

$$J(x_1, x_2) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} = \frac{\partial g_1}{\partial x_1}\frac{\partial g_2}{\partial x_2} - \frac{\partial g_1}{\partial x_2}\frac{\partial g_2}{\partial x_1} \neq 0$$

Nessas condições

$$f_{(Y_1, Y_2)(y_1, y_2)} = f_{(X_1, X_2)}\left(h_1(y_1, y_2), h_2(y_1, y_2)\right)|J(h_1(y_1, y_2), h_2(y_1, y_2))|^{-1}$$

Distribuição do produto e quociente de variáveis aleatórias

Sejam X e Y v.a. contínuas com f.d.p conjunta f. A função densidade do produto e do quociente entre X e Y são, respectivamente, dadas por

$$f_{XY}(u) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{u}{x}\right) dx;$$

$$f_{X/Y}(v) = \int_{-\infty}^{+\infty} |y| f\left(vy, y\right) dy;$$

Valor esperado de um vetor aleatório

Seja $(X_1, \ldots, X_n)$ um vetor aleatório com função de probabilidade conjunta $p(x_1, \ldots, x_n)$ e $g : \mathbb{R}^n \to \mathbb{R}$. O *valor esperado* de $g(X_1, \ldots, X_n)$ é dado por

$$E[g(X_1, \ldots, X_n)] = \sum_{x_1} \cdots \sum_{x_n} g(x_1, \ldots, x_n) p(x_1, \ldots, x_n)$$

Analogamente,

Valor esperado de um vetor aleatório

Seja $(X_1, \ldots, X_n)$ um vetor aleatório com função densidade de probabilidade conjunta $f(x_1, \ldots, x_n)$ e $g : \mathbb{R}^n \to \mathbb{R}$. O *valor esperado* de $g(X_1, \ldots, X_n)$ é dado por

$$E[g(X_1, \ldots, X_n)] = \int \cdots \int g(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1, \ldots, dx_n$$

Proposição

Se X e Y são v.a. com EX e EY bem definidos, temos

$$E[aX + bY] = aEX + bEY\,, \quad a, b \in \mathbb{R}$$

De forma geral, se $X_1 \dots X_n$ são v.a. com esperança finita então

$$E\left[\sum_{k=1}^{n} a_k X_k\right] = \sum_{k=1}^{n} a_k EX_k,\; a_k \in \mathbb{R}, k = 1, \dots, n.$$

10. Considere um jogo de azar com uma moeda honesta (X_1) e outra com probabilidade $3/5$ de dar coroa (X_2). O apostador ganha $R\$3,00$ se sair duas caras e paga $R\$1,00$ caso contrário.

(a) Determine a função de distribuição conjunta das moedas.

(b) Use o item (a) para determinar o lucro médio por jogada.

11. Sejam X e Y v.a. independentes com distribuição normal padrão. Defina $T = X + Y$ e $U = X/Y$ e determine a f.d.p conjunta de T e U. As variáveis T e U são independentes?