www.estudar.com.vc

Probabilidade

Resumo e Lista de Exercícios

Fuja do Nabo LIVE P3 2019.1

Resumo

1. Revisão: Variáveis Aleatórias Contínuas

Uma variável aleatória X é **contínua** se os valores que ela pode assumir pertencem a um **intervalo** (ou seja, existem infinitas possibilidades de valores x que ela pode assumir, desde que pertençam a esse intervalo).

a. Densidade de Probabilidade

Para as variáveis aleatórias contínuas, fala-se em **função densidade de probabilidade** (abreviada por f.d.p.), tal que:

I. A função $f(x)$ é sempre **positiva** (ou nula):

$$f(x) \geq 0 \text{ para todo } x;$$

II. A **integral** da f.d.p., de menos infinito até infinito, vale 1:

$$\int_{-\infty}^{\infty} f(x) = 1 \text{ (área sob o gráfico de } f(x));$$

III. E a probabilidade da variável aleatória X estar dentro de um intervalo $[a; b]$ é dada pela **integral da f.d.p., de a até b**:

$$P(a < X \leq b) = \int_a^b f(x)dx, \text{ com } a \leq b.$$

b. Distribuição Acumulada

Assim como no caso discreto, a **distribuição acumulada** representa a probabilidade de X assumir qualquer valor **menor ou igual** a x.

No entanto, no caso contínuo, a distribuição acumulada é dada por uma **função** $F(x)$, dada pela **integral** da f.d.p., de menos infinito até o valor x:

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

Assim, a **função distribuição acumulada** representa a **primitiva** da função densidade de probabilidade, tal que:

$$f(x) = \frac{dF(x)}{dx} = F'(x)$$

c. Propriedades

A probabilidade da variável aleatória X ser maior do que um valor a é a probabilidade **complementar** ao caso em que X é **menor ou igual** ao valor a (que é dado pela função distribuição acumulada em a):

$$P(X > a) = 1 - P(X \leq a) = 1 - F(a)$$

A probabilidade de X estar entre a e b é dada pela integral de $f(x)$, limitada pelos extremos a e b. Portanto, isso é igual à primitiva $F(x)$, calculada em b, menos a primitiva calculada em a:

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

Por fim, a probabilidade de X ser algum ponto c do intervalo é **nula**:

$$P(X = c) = \int_{c}^{c} f(x)dx = 0$$

2. Revisão: Medidas Descritivas

Em Probabilidade, dada uma distribuição contínua ou discreta, calculamos algumas **medidas descritivas** que identificam as principais métricas da população, em relação a sua **posição** ou sua **dispersão**.

a. Valor Médio ou Esperança

É o **valor esperado** (E) para a variável aleatória ou função.

No **caso contínuo**, o valor esperado da variável x é dado pela seguinte integral:

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

Enquanto o valor esperado de uma função $f(x)$ é:

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

O valor esperado pode ser interpretado como uma **medida da localização do centro** da variável aleatória.

A função de esperança é **linear**, e, por isso, ela apresenta as seguintes propriedades:

I. Dadas uma **constante** $\boldsymbol{a} \in \mathbb{R}$ e uma **constante** $\boldsymbol{b} \in \mathbb{R}$, a seguinte propriedade é válida:

$$E(aX + b) = a \cdot E(X) + b$$

II. O **valor esperado da soma** de duas variáveis aleatórias X_1 e X_2 é a **soma dos valores esperados** de X_1 e X_2:

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

III. Considere duas variáveis aleatórias **independentes**, X e Y. O **valor esperado do produto** das variáveis é o **produto dos valores esperados**:

$$E(XY) = E(X) \cdot E(Y)$$

b. Variância
A **variância** ($\sigma^2(X)$ ou $Var(X)$) é uma medida da **variabilidade** da distribuição de uma variável aleatória. O cálculo da variância utiliza o conceito de **valor esperado**:

$$\sigma^2(X) = E(X^2) - [E(X)]^2$$

Algumas propriedades da variância são:

I. Se a é uma **constante** real, então:

$$Var(X + a) = Var(X)$$

II. Se a é uma **constante** real, então:

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

III. Se X e Y são variáveis aleatórias **independentes**, a variância da soma das variáveis é a soma das variâncias:

$$Var(X+Y) = Var(X) + Var(Y)$$

c. Desvio Padrão

O **desvio padrão** (σ) mede a **dispersão** entre a variável aleatória e a média:

$$\sigma(X) = \sqrt{\sigma^2(X)}$$

d. Momento da Função

O **momento de ordem *k*** de uma função é definido pela seguinte equação, para o caso contínuo:

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

A **esperança** de uma função é o **momento de primeira ordem** dela, enquanto a **variância** é a **diferença** entre o **momento de segunda ordem** e o **quadrado do momento de primeira ordem**.

3. Distribuições Unidimensionais Contínuas

As variáveis aleatórias **contínuas** possuem diferentes tipos de distribuição; algumas são mais comuns e estudadas no curso de Probabilidade.

a. Uniforme

A **distribuição uniforme** tem como principal característica a **probabilidade igual** de ocorrer qualquer fenômeno com **mesmo comprimento**.

A **função de densidade de probabilidade** de uma variável aleatória que tem distribuição uniforme em um intervalo $[a; b]$ é dada por:

$$f(x) = \begin{cases} 0, x < a \\ \dfrac{1}{b-a}, a \leq x \leq b \\ 0, x > b \end{cases}$$

A notação utilizada para quando uma variável X tem distribuição uniforme é $X \sim U(a, b)$.

O **valor esperado** é:

$$E(X) = \frac{a+b}{2}$$

Por sua vez, a **variância** é:

$$\sigma^2(X) = \frac{(b-a)^2}{12}$$

b. Exponencial

A **distribuição exponencial** é caracterizada por ter uma taxa de falha constante. Seu parâmetro é um λ, e sua f.d.p. é dada por:

$$f(x) = \lambda e^{-\lambda x}$$

A notação para quando uma variável aleatória X tem distribuição exponencial é $X \sim Exp(\lambda)$.

O **valor esperado** é calculado da seguinte forma:

$$E(X) = \frac{1}{\lambda}$$

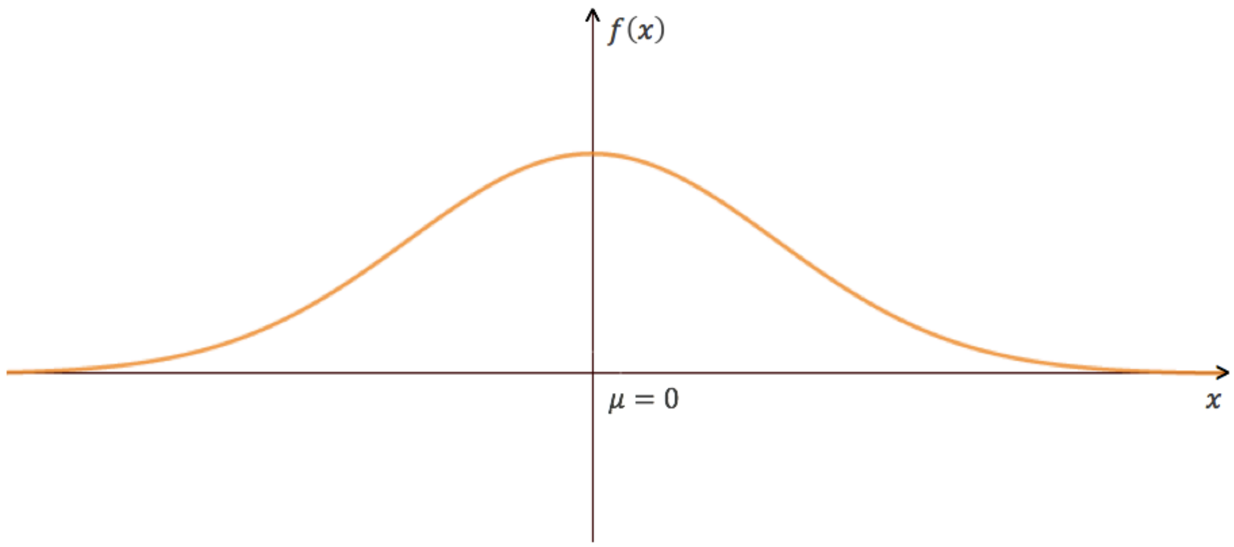
A **variância** é calculada da seguinte forma (note que, por isso, o **desvio padrão** é igual ao valor esperado):

$$\sigma^2(X) = \frac{1}{\lambda^2} \Rightarrow \sigma(X) = \frac{1}{\lambda}$$

c. Normal

A **distribuição Normal** é, talvez, a distribuição de probabilidades mais importante, visto que **qualquer distribuição se aproxima da Normal** quando há um **número grande de dados**.

A distribuição Normal **padrão** é **simétrica** e tem a seguinte forma:



A **média** e o **desvio padrão** da Normal padrão são:

$$\mu = 0 \text{ e } \sigma = 1$$

No entanto, uma variável X que apresenta distribuição Normal (ou seja, $X \sim N(\mu, \sigma)$), possui diferentes valor esperado e variância.

Nesse caso, o **valor esperado** é:

$$E(X) = \mu$$

E a **variância** é:

$$Var(X) = \sigma^2$$

A **função de distribuição de probabilidades** é dada como:

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

No entanto, a integral dessa função não é analiticamente resolvida. Para calcular a probabilidade de X estar em um determinado intervalo, cria-se uma variável z, tal que:

$$z = \frac{x-\mu}{\sigma}$$

E, com essa variável, a probabilidade é obtida pela **tabela Normal** (anexada no fim deste documento). A variável z tem distribuição Normal com média 0 e variância 1.

Para isso, considera-se que a probabilidade se divide **igualmente** entre os dois lados do gráfico da Normal; assim, a probabilidade de z ser negativo ou positivo é igual e vale 0,5:

$$P(z < 0) = P(z > 0) = 0{,}5$$

Assim, a probabilidade de $P(0 < z < z_0)$ é fornecida pela **tabela**, dado um valor de z_0. Utiliza-se a simetria da distribuição para calcular probabilidades de intervalos diferentes. Por exemplo:

$$P(z < 2{,}3) = P(z < 0) + P(0 < z < 2{,}3) = 0{,}5 + P(0 < z < 2{,}3)$$

d. Teorema do Limite Central

Este teorema diz que, dado um número de dados muito grande (ou seja, n grande), a distribuição da média de uma distribuição qualquer **tende a se aproximar de uma distribuição Normal**.

Neste caso, o **valor esperado** dessa Normal aproximada vale:

$$E(\bar{X}) = \mu$$

E a **variância** é calculada como:

$$Var(\bar{X}) = \frac{\sigma^2}{n}$$

Assim, com n grande, a notação para a média de uma variável aleatória X que passa a ter distribuição Normal é: pelo Teorema do Limite Central, $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

Outra consequência desse teorema é que a **soma de variáveis aleatórias** independentes e identicamente distribuídas se aproxima de uma distribuição normal.

Então, seja S_n o somatório de uma sequência de variáveis aleatórias independentes e identicamente distribuídas X_i (cuja média é μ e cujo desvio padrão é σ). Vale que:

$$z = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

4. Combinação Linear de Distribuições Normais

Suponha que existam duas variáveis X e Y com **distribuição Normal**, tal que:

$$X \sim N(\mu_x, \sigma_x)$$

$$Y \sim N(\mu_y, \sigma_y)$$

Seja também uma variável W definida por uma **combinação linear** das variáveis X e Y:

$$W = aX + bY, \;\; a, b \in \mathbb{R}$$

Consequentemente, a variável W também tem **distribuição Normal**, com média μ_W e desvio padrão σ_W:

$$W \sim N(\mu_W, \sigma_W)$$

A **média** é calculada pela combinação linear das médias μ_X e μ_Y:

$$\mu_W = \; a\mu_X + b\mu_Y$$

E a **variância** é calculada da seguinte forma:

$$\sigma_W^2 = a^2\sigma_X^2 + b^2\sigma_Y^2$$

5. Distribuições Multidimensionais

Distribuições multidimensionais ou conjuntas são utilizadas para situações em que **mais de um resultado** é observado em um experimento.

A probabilidade de um determinado evento envolve todas as variáveis, ou seja:

$$P(X_1 = x_1, X_2 = x_2, \dots)$$

No curso de Probabilidade, trabalha-se com apenas **duas variáveis**.

No **caso contínuo**, em que as variáveis X e Y são contínuas, a probabilidade é tal que:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x,y)\, dy\, dx$$

E são as seguintes propriedades que valem:

$$0 \leq f(x,y) < \infty$$

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

a. Probabilidades Marginais

Dada uma distribuição multidimensional, é possível retirar a distribuição unidimensional para uma variável.

No **caso contínuo**, isso ocorre pela seguinte integral:

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) \cdot dy$$

b. Independência

Podemos dizer que duas variáveis são independentes, no **caso contínuo**:

$$f(x, y) = f_x(x) \cdot f_y(y)$$

c. Distribuições Condicionais

A probabilidade de $X = x$, dado que $Y = y$, no **caso contínuo**, a função de distribuição de probabilidade condicional é:

$$f_{y|x}(y|x) = \frac{f(x, y)}{f_x(x)}$$

Se as variáveis são independentes, então: $P(X = x|Y = y) = P(X = x)$.

d. Esperança Condicional

O **valor esperado** de X, dado que $Y = y$, é calculado, no caso **contínuo**, por:

$$E(X) = E[X|Y = y] = \int_{-\infty}^{\infty} x \cdot f_x(x)\, dx$$

e. Média de uma Função sobre X e Y

O **valor esperado** de uma função $h(x,y)$, dada uma distribuição conjunta, é, no caso **contínuo**:

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

f. Covariância e Correlação de X e Y

A **covariância** é uma medida que estima variabilidade conjunta de duas variáveis X e Y. A covariância ($Cov[X,Y]$) é calculada por:

$$Cov[X,Y] = E[XY] - E[X]E[Y]$$

Uma importante propriedade da covariância é que, para $a > 0$ e $b > 0$, segue que:

$$Cov[aX, bY] = abCov[X,Y]$$

O **coeficiente de correlação** $\boldsymbol{\rho}$ estabelece uma razão entre a variação conjunta de X e Y (covariância) e o produto dos desvios padrões de cada variável:

$$\rho[X,Y] = \frac{Cov[X,Y]}{\sigma_x\sigma_y}$$

O coeficiente de correlação está entre -1 e 1:

$$-1 \leq \rho \leq 1$$

As seguintes propriedades são válidas, para $a > 0$ e $b > 0$:

$$\rho[aX, bY] = \rho[X, Y]$$

$$\rho[-X, Y] = \rho[X, -Y] = -\rho[X, Y]$$

$$\rho[X, X] = 1$$

Lista de Exercícios

1. Distribuição Uniforme

P3 2018.1 Probabilidade Poli-USP, Exercício 1

Experimentos com animais indicam que o tempo necessário para que um determinado remédio contra febre faça efeito segue uma distribuição uniforme no intervalo de 20 a 50 (em minutos). Qual a probabilidade da febre de uma pessoa que tomou esse remédio durar mais de 40 minutos, sabendo-se que sua febre não cedeu nos 30 primeiros após ter tomado o remédio?

A. $1/3$
B. $1/4$
C. 1
D. $1/2$
E. $2/3$

2. Distribuição Uniforme

P2 2016.1 Probabilidade Poli-USP, Exercício 10

A variável aleatória X tem distribuição uniforme no intervalo $[-1, 1]$. A função densidade de probabilidade da variável $Y = X^2$ no intervalo $[0,1]$ é:

A. $f(y) = \frac{1}{2\sqrt{y}}$
B. $f(y) = 2\sqrt{y}$
C. $f(y) = \sqrt{y}$
D. $f(y) = \frac{1}{4}$

E. $f(y) = 1$

3. Distribuição Exponencial

P2 2018.1 Probabilidade Poli-USP, Exercício 7

Um sistema é constituído de 3 componentes que funcionam de forma independente. Sabe-se que se um dos componentes falhar o sistema falha. O tempo de vida de cada um dos componentes segue uma distribuição exponencial cujos valores esperados são: $10t$, $10t/3$ e $10t/6$. Qual a probabilidade de o sistema não falhar antes de t?

A. e^{-1}
B. $e^{-0,9} + e^{-0,4} + e^{-0,7} - 3e^{-1}$
C. $(1 - e^{-0,1}) \cdot (1 - e^{-0,3}) \cdot (1 - e^{-0,6})$
D. $1 - e^{-1}$
E. $e^{-0,1} + e^{-0,3} + e^{-0,6}$

4. Distribuição Normal

P3 2018.1 Probabilidade Poli USP, Exercício 15

Um comprimido é produzido em série numa empresa farmacêutica com peso médio de $1{,}400\ mg$ e desvio padrão de $0{,}100\ mg$ com distribuição normal. Os comprimidos passam por um controle de qualidade e os comprimidos com peso menor que $1{,}350\ mg$ e peso maior que $1{,}519$ são eliminados do lote. Qual a probabilidade de nesse lote restante retirarmos ao acaso um comprimido com peso entre os valores $1{,}350$ e $1{,}400\ mg$?

A. $0{,}19$
B. $0{,}33$

C. 0,38
D. 0,50
E. 0,98

5. Distribuição Normal

P2 2016.1 Probabilidade Poli-USP, Exercício 1

A variável aleatória R tem função densidade de probabilidade $f(r) = re^{-\frac{r^2}{2}}$, caso $r \geq 0$, e $f(r) = 0$ para $r < 0$. A esperança $E(G)$, sendo $G = \frac{1}{R}$, é:

A. $\frac{1}{\sqrt{8\pi}}$
B. $\sqrt{2\pi}$
C. $\frac{\sqrt{\pi}}{2}$
D. $+\infty$
E. $\frac{1}{2}$

6. Teorema do Limite Central

P3 2018.1 Probabilidade Poli-USP, Exercício 3

Seja S a soma de 16 variáveis aleatórias independentes, todas com distribuição uniforme no intervalo $[6-\sqrt{3}, 6+\sqrt{3}]$. Quanto vale (aproximadamente) a probabilidade de S estar entre 92 e 100?

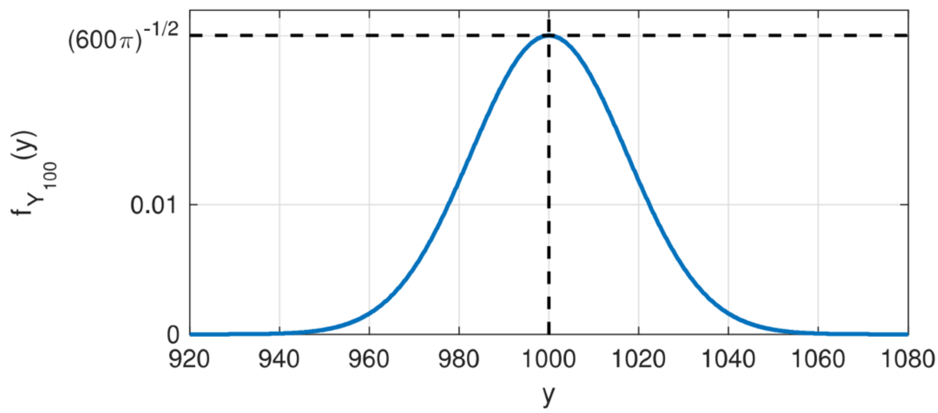
A. 0,92
B. 0,53

C. 0,68
D. 0,75
E. 0,10

7. Teorema do Limite Central

P3 2018.1 Probabilidade Poli-USP, Exercício 14

A variável aleatória Y_{100} é obtida a partir da soma de 100 variáveis aleatórias independentes, todas com distribuição uniforme em $[a, b]$ (ou seja, $Y_{100} = X_1 + X_2 + \cdots + X_{100}$). Uma aproximação para a função densidade de probabilidade de Y_{100} é mostrada na figura abaixo.



Os valores de a e b são respectivamente iguais a:

A. 2 e 18
B. 7 e 13
C. $1/\sqrt{6\pi}$ e $10 + 1/\sqrt{6\pi}$
D. 9,4 e 10,4
E. 2π e $2\pi + 10$

8. Distribuição Conjunta

P3 2018.1 Probabilidade Poli-USP, Exercício 2

Obtenha $P(X > Y)$, se a função densidade de probabilidade conjunta de variáveis aleatórias X e Y é:

$$f(x,y) = \begin{cases} C(x^2 + xy), & 0 < x < 1, 0 < y < 2; \\ 0, & caso\ contrário. \end{cases}$$

A. $1/2$
B. $5/8$
C. $31/40$
D. $9/40$
E. $1/3$

9. Distribuição Exponencial Conjunta

P3 2018.1 Probabilidade Poli-USP, Exercício 4

Considere duas variáveis aleatórias X e Y, contínuas, não negativas e independentes. Ambas têm distribuição exponencial com esperanças: $E[x] = \beta$ e $E[Y] = 2\beta$. Qual a probabilidade $P(0 < X < \beta, 0 < Y < 2\beta)$?

A. $1 - e^{-2\beta}$
B. $e^{-2\beta}$
C. $\left(1 - e^{-\beta}\right)^2$
D. e^{-2}
E. $(1 - e^{-1})^2$

10. Distribuição Marginal

P3 2018.1 Probabilidade Poli-USP, Exercício 6

A função densidade de probabilidade conjunta de duas variáveis aleatórias X e Y, para uma constante c, é dada por:

$$f(x,y) = \begin{cases} cy^2, & 0 < x < 2, 0 < y < 1; \\ 0, & caso\ contrário. \end{cases}$$

A densidade marginal de X, para $0 < x < 2$, é:

A. $1/3$
B. 1
C. $3x^2$
D. $1/2$
E. $3x$

11. Distribuição Condicional

P3 2018.1 Probabilidade Poli-USP, Exercício 8

A função densidade de probabilidade conjunta de duas variáveis aleatórias X e Y é dada por:

$$f(x,y) = \begin{cases} 2(x+y), & 0 < x < y < 1; \\ 0, & caso\ contrário. \end{cases}$$

A densidade condicional de X dado Y, para $0 < x < 1$, é:

A. $3y^2$

B. $2(x+y)/3y^2$
C. $2(x+y)$
D. $2(x+y)/(1+2y-3y^2)$
E. $1+2y-3y^2$

Gabarito

1. Alternativa D

2. Alternativa A

3. Alternativa A

4. Alternativa B

5. Alternativa C

6. Alternativa C

7. Alternativa B

8. Alternativa D

9. Alternativa E

10. Alternativa D

11. Alternativa B