www.estudar.com.vc

Resumo e Lista de Exercícios

Probabilidade

Fuja do Nabo P2 2019.1

Resumo

1. Revisão: Variáveis Aleatórias Discretas

Uma Variável aleatória é **discreta** se ela é **enumerável** (ou seja, podemos listar cada valor que a variável pode assumir, por exemplo, x_1, x_2, até um x_n, que é o valor final assumido).

a. Distribuição de Probabilidade

A probabilidade de a variável aleatória X assumir um valor x, dentro de um espaço amostral S (cujos elementos são s) é:

$$P[X = x] = P\ (para\ todos\ s \in S : X(s) = x)$$

b. Distribuição Acumulada

A distribuição acumulada considera a possibilidade da variável aleatória X assumir qualquer valor **menor ou igual** a um x.

Isso é dado pela **somatória** de todas as probabilidades em que X assume valores y, para todo $y \leq x$:

$$F(x) = P(X \leq x) = \sum_{y:\ y \leq x} P[X = y]$$

c. Propriedades

A probabilidade da variável aleatória X assumir um valor x está entre 0 (no mínimo) e 1 (no máximo):

$$0 \leq P[X = x] \leq 1$$

Além disso, a soma de todas as probabilidades em que X assume valores x, para todos os x dentro do espaço amostral S, é 1:

$$\sum_{todos\ os\ possíveis\ x} P[X = x] = 1$$

2. 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 de a variável aleatória X estar dentro de um intervalo $[a; b]$ é dada pela **integral da f.d.p., de $\boldsymbol{a}$ até $\boldsymbol{b}$**:

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

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**:

$$P(X = c) = \int_c^c f(x)dx = 0$$

3. 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 de **distribuições discretas**, o valor esperado da variável é:

$$E[X] = \sum_{i=1}^{N} x_i \cdot P[X = x_i]$$

De forma similar, considerando a mesma distribuição, o valor esperado de uma função $\boldsymbol{g(x)}$ qualquer é:

$$E[g(X)] = \sum_{i=1}^{N} g(x_i) \cdot P[X = x_i]$$

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:

$$E(X_1 \pm X_2) = E(X_1) \pm E(X_2)$$

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**:

$$Var(X) = E(X^2) - [E(X)]^2$$

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:

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

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

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

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:

I. Para o caso discreto:

$$E[X^k] = \sum_{i=1}^{N} {x_i}^k \cdot P[X = x_i]$$

II. 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**.

4. Distribuições Unidimensionais Discretas

Existem algumas **distribuições** de variáveis aleatórias **discretas** que, por serem muito comuns, devem ser estudadas.

a. Bernoulli

Uma distribuição de Bernoulli ocorre quando um experimento pode resultar apenas em **falhas** ou **sucessos**.

Ou seja, o espaço amostral do experimento é $\{0, 1\}$, onde 1 ocorre com a probabilidade de sucesso p, e 0 ocorre com a probabilidade de fracasso $q = 1 - p$.

O valor esperado e a variância são:

$$E(X) = p$$

$$Var(X) = p(1-p)$$

b. Binomial
A **distribuição binomial** ocorre quando temos n experimentos que resultam somente em **sucessos** (S) ou **falhas** (F) (ou seja, n ensaios de *Bernoulli*).

Sendo a **probabilidade de sucesso** $P(S) = p$ (e, portanto, $P(F) = 1 - p$) e x o **número de sucessos**, a probabilidade de ocorrerem x sucessos é:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Usamos a notação $X \sim b(n, p)$ quando a variável X assume uma distribuição binomial.

O **valor esperado** de X é dado por:

$$E(X) = np$$

E a **variância** de X é:

$$\sigma^2(X) = np(1-p)$$

c. Geométrica

A **distribuição geométrica** ocorre quando existe uma sequência de ensaios de *Bernoulli* até que ocorra o **primeiro sucesso**.

Portanto, x ensaios são realizados até que ocorra o primeiro sucesso. A probabilidade de ocorrer sucesso no x-ésimo ensaio é:

$$P(X = x) = (1 - p)^{x-1}p$$

A notação utilizada para quando uma variável X tem distribuição geométrica é $X \sim Geo(p)$.

O **valor esperado**, nesta distribuição, é:

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

E a **variância** é:

$$\sigma^2(X) = \frac{1 - p}{p^2}$$

Observação: Às vezes, em vez de usar o número de ensaios até o primeiro sucesso x (incluindo o sucesso), usa-se o número de fracassos j. Assim, a probabilidade é dada por:

$$P(X = x) = (1 - p)^{j}p$$

d. Poisson

A **distribuição de Poisson** possui um parâmetro positivo λ que representa a **taxa de ocorrência** por unidade de tempo (enquanto Δt é um intervalo de tempo).

A probabilidade de a variável aleatória assumir um valor x é:

$$P(X = x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

Onde:

$$\mu = \lambda \Delta t$$

O **valor esperado** e a **variância** assumem o mesmo valor:

$$E(X) = \sigma^2(X) = \mu$$

A notação utilizada para quando uma variável X tem distribuição de Poisson é $X \sim Poisson(\mu)$.

Quando ocorre uma distribuição binomial, em que n é um número **grande** e p é um número **pequeno** ($n \to \infty$ e $p \to 0$), podemos **aproximar** essa distribuição por uma distribuição de Poisson em que $\mu = np$.

Observação: Se λ for dado como não como uma taxa, mas apenas como um parâmetro de Poisson que não depende do tempo, ou caso o intervalo de tempo não seja considerado na análise, então $\mu = \lambda$.

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:

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

Em Probabilidade, costuma-se trabalhar com apenas **duas variáveis**.

Dado um **caso discreto**, em que as variáveis X e Y são discretas, as seguintes propriedades se verificam:

$$0 \leq P(X = x, Y = y) \leq 1$$

$$\sum_x \sum_y P(X = x, Y = y) = 1$$

a. Probabilidades Marginais

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

Para isso, no **caso discreto**, para achar a probabilidade de $X = x$, devemos somar todas as probabilidades em que $Y = y$, para todos os y possíveis:

$$P(X=x)=\sum_{y} P(X=x, Y=y)$$

b. Independência

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

$$P(X=x, Y=y)=P(X=x)\cdot P(Y=y)$$

c. Distribuições Condicionais

A probabilidade de $X=x$, dado que $Y=y$, no **caso discreto**, é dada por:

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

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 **discreto**, por:

$$E(X)=E[X|Y=y]=\sum_{x} x\cdot P(X=x|Y=y)$$

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 **discreto**:

$$E[h(X,Y)] = \sum_x \sum_y h(x,y) \cdot P(X = x, Y = y)$$

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:

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

Se X e Y são **independentes**, então:

$$Cov[X,Y] = 0$$

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. Valor Esperado e Variância

P2 2017.1 Probabilidade Poli USP, Exercício 4 Adaptado

Seja uma variável aleatória X. Considere:

I. Se $X = 1$, então $E(X) = 1$
II. $E(X + 2X) = 3E(X)$
III. $V(X + 2X) = 5V(X)$
IV. $E(X^4) = V(X^2) + (V(X) + E(X)^2)^2$

Assinale a alternativa correta:

A. Apenas as afirmações II, III e IV são verdadeiras.
B. Apenas as afirmações I, III e IV são verdadeiras.
C. Apenas as afirmações I e II são verdadeiras.
D. Todas as afirmações são verdadeiras.
E. Apenas as afirmações I, II e IV são verdadeiras.

2. Valor Esperado e Variância

P2 2018.1 Probabilidade Poli USP, Exercício 2 Adaptado

A função distribuição cumulativa $F_X(x)$ de uma variável aleatória X é definida da seguinte forma:

$$F_X(x) = \begin{cases} 0, & para\ x < -2 \\ 2/9, & para - 2 \leq x < 0 \\ 7/9, & para\ 0 \leq x < 2 \\ 1, & para\ x \geq 2 \end{cases}$$

Dentre as afirmações a seguir, qual a verdadeira?

I. $E[X] + \sigma_X = 4/3$ e $E[X^2] = V[X]$.
II. $E[aX^2 - bX + c] = (20a - 10b + 9c)/9$, sendo a, b, c constantes.
III. $E[X^2] = 20/9$ e $E[X] = 10/9$.

a. Só **III**.
b. Só **II**.
c. Só **II** e **III**.
d. Só **I** e **III**.
e. Só **I**.

3. Valor Esperado e Variância

P2 2018.1 Probabilidade Poli USP, Exercício 10 Adaptado

Seja X_i, $i = 1,2,3$ uma sequência de variáveis aleatórias independentes (resultados de 3 experimentos idênticos, mas independentes). Sabe-se que $E\{X_i\} = \mu$ e $Var(X_i) = \sigma^2$, $i = 1,2,3$. Deseja-se estimar μ com o seguinte estimador:

$$\hat{\mu} = (X_1 + X_2 + X_3)/3$$

A variância de $\hat{\mu}$ é igual a:

a. $\sigma^2/9$
b. $\sigma^2/3$
c. $(\sigma^2 - \mu^2)/3$
d. $\sigma^2/3 + \mu^2$
e. $3\sigma^2$

4. Transformação de Variáveis Aleatórias

P3 2016.1 Probabilidade Poli USP, Exercício 1 Adaptado

Sejam duas variáveis aleatórias contínuas independentes X e Y sobre as quais sabe-se que $E(X) = 10$, $\sigma(X) = 10$, $\sigma\,(Y) = 10$ e $E(XY) = 60$. Define-se uma nova variável W dada por $W = X^2 + Y^2$. Calcule o valor de $E(W)$.

A. 96
B. 136
C. 256
D. 200
E. 336

5. Distribuição Binomial

P2 2016.1 Probabilidade Poli USP, Exercício 6 Adaptado

Uma máquina de fabricar sorvete é bastante defeituosa. A probabilidade de ela apresentar defeito em uma hora qualquer é q. Em função disto, deve-se programar o momento de parada do equipamento para a manutenção preventiva do equipamento. Qual a probabilidade de não acontecer defeito nas primeiras k horas?

A. $1 - \sum_{j=1}^{k}(1-q)^j q$
B. $\sum_{j=1}^{k}(1-q)^{j-1} 1$
C. qk
D. $(1-q)^k$
E. $(1-q)^k q$

6. Distribuição Binomial

P2 2016.1 Probabilidade Poli-USP, Exercício 2

Um escritório de uma empresa tem 8 vendedores que trabalham o mesmo número de horas no escritório e em serviço externo, com agenda aleatória. Qual o menor número de mesas de trabalho que deve existir no escritório de modo que cada um tenha uma mesa pelo menos 90% do tempo?

A. 8
B. 4
C. 6
D. 7
E. 5

7. Distribuição Geométrica

P2 2018.1 Probabilidade Poli-USP, Exercício 3

Considere duas variáveis aleatórias X e Y, independentes, e ambas com distribuição geométrica com parâmetro $p = 1/4$. Considere também duas variáveis $W = X + Y$ e $Z = X - Y + 3$. O valor de $P(W \geq Z)$ é:

a. $1/4$
b. $7/16$
c. $9/16$
d. $1/2$
e. $3/4$

8. Distribuição Geométrica

P2 2017.1 Probabilidade Poli-USP, Exercício 5

Considere duas variáveis aleatórias X e Y com distribuições geométricas, com mesmo parâmetro p, e tais que:

$$P(\{X = x\} \cap \{Y = y\}) = P(X = x)P(Y = y)$$

Para quaisquer x e y. Qual é o valor de $P(X + Y = n)$?

A. $np^2(1 - p)^n$
B. $p^n(1 - p)^n$
C. $(n + 1)p^2(1 - p)^n$
D. $\frac{1}{n}$
E. $\frac{(1-p)^n}{(np^2)}$

9. Distribuição de *Poisson*

P2 2016.1 Probabilidade Poli USP, Exercício 4 Adaptado

Considere uma variável aleatória distribuição de Poisson com parâmetro $4t$, em que t é um período de tempo fornecido, dado em horas. Esse tipo de variável é usado para representar solicitações de assistência em uma empresa de seguros: X é o número de pedidos de assistência em um intervalo de tempo t. Se os operadores da empresa tirarem meia hora de folga para almoço, qual a probabilidade de não perderem nenhum chamado de assistência?

A. e^2
B. $\frac{1}{8}$
C. $\frac{1}{e}$
D. $\frac{1}{e^2}$
E. $\frac{1}{\sqrt{e}}$

10. Distribuição Multidimensional

P2 2018.1 Probabilidade Poli USP, Exercício 15 Adaptado

Para X e Y, duas variáveis aleatórias discretas, que possuem distribuição conjunta de acordo com a tabela abaixo, quais são os valores de $P(X = 1), P(X = 2), P(X = 3)$?

$p(x,y)$	$y = 1$	$y = 2$	$y = 3$
$x = 1$	$8k$	$6k$	$7k$
$x = 2$	$5k$	$3k$	$6k$
$x = 3$	0	k	$4k$

A. $17/40, 1/4, 13/40$
B. $21/40, 7/20, 1/8$
C. $13/40, 1/4, 17/40$
D. $1/5, 1/8, 0$
E. $1/5, 3/20, 7/40$

11. Distribuição Multidimensional

P3 2016.1 Probabilidade Poli USP, Exercício 3 Adaptado

O eixo de um motor precisa ter um determinado comprimento e diâmetro. Uma vez que o motor esteja montado, é possível medir o diâmetro do eixo, mas não o seu comprimento. Uma empresa produz eixos, cujos diâmetros e comprimentos estão distribuídos como descrito na tabela abaixo.

Diâmetro (⇩) Comprimento (⇨)	$10\ cm$	$11\ cm$	$12\ cm$	$13\ cm$
$5\ mm$	0,01	0,02	0,05	0,01
$6\ mm$	0	0,05	0,4	0,05
$7\ mm$	0,02	0,03	0,02	0,04
$8\ mm$	0,05	0,1	0,1	0,05

Dado que, para um certo motor montado, notou-se que o eixo tem o diâmetro $D = 6\ mm$, qual é a probabilidade de o comprimento ser $C = 12\ cm$?

A. 0,7
B. 0,4
C. 0,5
D. 0,57
E. 0,8

12. Covariância

P3 2016.1 Probabilidade Poli USP, Exercício 4 Adaptado

Uma urna contém 3 bolas vermelhas e 2 bolas pretas. Duas bolas são retiradas sem reposição. Seja X o número de bolas vermelhas e Y o número de bolas pretas. Calcule a covariância $Cov(X,Y)$.

A. $-\frac{4}{5}$
B. $\frac{1}{2}$
C. $-\frac{9}{25}$
D. $-\frac{10}{25}$
E. $\frac{2}{25}$

13. Covariância

P2 2018.1 Probabilidade Poli USP, Exercício 13 Adaptado

Sejam X e Y duas variáveis aleatórias. Dentre as afirmações a seguir, quais são verdadeiras?

I. Se X e Y são independentes, então $E[XZ] = Cov[X, Y + Z] + E[X]E[Z]$;
II. $Cov[X+Y, X-Y] = V[X] - V[Y] + 2Cov[X,Y]$
III. $Cov[X+Y, X+Y] = V[X] + V[Y] + 2Cov[X,Y]$

a. Só **I**.
b. Só **II** e **III**.
c. Só **I** e **II**.
d. Só **I** e **III**.
e. **I**, **II** e **III**.

14. Covariância

P2 2018.1 Probabilidade Poli USP, Exercício 8 Adaptado

A tabela abaixo mostra a distribuição de probabilidade conjunta das variáveis aleatórias X e Y.

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	a	0,2	b
$X = 1$	0,1	0,2	0,1

Sabe-se que $Cov[X, Y] = 0$. Dentre as afirmações a seguir, quais são verdadeiras?

I. X e Y são variáveis independentes e $a = b$.
II. $\rho(X, Y) = 0$ e $a/b = 2$.
III. $E[Y] = 1$ e $E[X] = E[XY]$.

a. Só **I**.
b. Só **III**.
c. Só **I** e **II**.
d. Só **II**.
e. Só **II** e **III**.

Gabarito

1. Alternativa E.
2. Alternativa E.
3. Alternativa B.
4. Alternativa E.
5. Alternativa D.
6. Alternativa C.
7. Alternativa C.
8. Alternativa C.
9. Alternativa D.
10. Alternativa B.
11. Alternativa E.
12. Alternativa C.
13. Alternativa D.
14. Alternativa B.