

Introdução à Probabilidade e à Estatística (BCN0406-15)

Prova 2 (A) – Gabarito

1. São dadas as seguintes probabilidades associadas à variável aleatória X :

x	-1	1	2
$p(x)$	1/2	1/3	1/6

a) Calcule a esperança e a variância de X .

b) Calcule a esperança e a variância de $\frac{1}{X}$.

Resposta a)

$$E[X] = \sum x_i p(x_i) = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} = \frac{1}{6}$$

$$Var(X) = E[X^2] - (E[X])^2 \quad E[X^2] = \sum x_i^2 p(x_i) = (-1)^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{1}{6} = \frac{3}{2}$$

$$Var(X) = \frac{3}{2} - \frac{1}{6^2} = \frac{53}{36} \approx 1,47$$

b)

$$E[1/X] = \sum \frac{1}{x_i} p(x_i) = \frac{1}{-1} \cdot \frac{1}{2} + \frac{1}{1} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{6} = -\frac{1}{12}$$

$$Var(1/X) = E[(1/X)^2] - (E[1/X])^2 \quad E[1/X^2] = \sum \frac{1}{x_i^2} p(x_i) = \frac{1}{(-1)^2} \cdot \frac{1}{2} + \frac{1}{1^2} \cdot \frac{1}{3} + \frac{1}{2^2} \cdot \frac{1}{6} = \frac{7}{8}$$

$$Var(1/X) = \frac{7}{8} - \left(-\frac{1}{12}\right)^2 = \frac{125}{144} \approx 0,87$$

2. Em 1986, o desastre da usina nuclear de Chernobil liberou uma grande quantidade de partículas radioativas de vários tipos, entre eles o Césio-137 ($^{137}_{55}\text{Cs}$), cuja taxa de decaimento é de $0,023 \text{ ano}^{-1}$.

a) Qual a probabilidade de uma partícula de Césio-137 não ter decaído nesses 33 anos desde o acidente?

b) Qual a probabilidade de uma partícula de Césio-137 ainda presente hoje decair até 2050 (ou seja, nos próximos 31 anos)?

Resposta

a) Estamos interessados no **tempo de decaimento** da partícula. Essa deve ser uma variável contínua, e estritamente positiva e, além disso, sem memória, por isso devemos usar uma variável aleatória **exponencial**. Pelo enunciado, o parâmetro (a taxa) dessa V.A. deve ser $\lambda = 0,023$, e conhecemos a função densidade de probabilidade $f(x)$ e também a função de distribuição acumulada $F(x)$ (ver tabela). A pergunta pede qual é a prob. de ainda não ter decaído, ou seja, o tempo de decaimento ser maior que 33 anos: $P(X > 33)$. Você pode fazer usando $P(X > 33) = \int_{33}^{\infty} f(x) dx$ ou notando que

$$P(X > 33) = 1 - P(X \leq 33) = 1 - F(33) = 1 - (1 - e^{-0,023 \cdot 33}) \approx 0,468$$

b) A pergunta é sobre a probabilidade de o átomo decair nos próximos 31 anos (ou seja, $P(X < 64)$), onde X é o tempo desde 1986) **dado que** ela sobreviveu até hoje (ou seja, $X > 33$). Isso é escrito como $P(X < 64 | X > 33)$. Você pode calcular dessa forma, mas o problema fica mais simples se lembrarmos que essa V.A. não tem memória, logo a distribuição de probabilidades de decaimento agora é exatamente a mesma de 33 anos atrás, pra partícula que sobreviveu, logo podemos dizer que o tempo de decaimento a partir de hoje também é uma V.A. exponencial de taxa 0,023. Portanto, queremos encontrar $P(X < 31) = F(31) = 1 - e^{-0,023 \cdot 31} \approx 0,51$.

3. No ano de 2010, a produção de lixo plástico na América teve a distribuição da tabela 1 ao lado, que mostra a quantidade de lixo plástico diário *per capita* produzido em cada país (em kg/pessoa/dia). *Fonte: ourworldindata.org.*

Tabela 1

País	Qtd.
Argentina	0.183
Bahamas	0.390
Brazil	0.165
Canada	0.093
Chile	0.119
Colombia	0.144
Costa Rica	0.258
Cuba	0.089
Ecuador	0.147
El Salvador	0.147
Guatemala	0.280
Haiti	0.090
Honduras	0.189
Jamaica	0.034
Mexico	0.087
Nicaragua	0.143
Panama	0.145
Peru	0.144
Suriname	0.163
United States	0.335
Uruguay	0.252
Venezuela	0.252

a) Calcule a média e a mediana desses dados.

b) Construa um histograma dos dados, usando 4 classes (*bins*). A área desse gráfico deve ser 1.

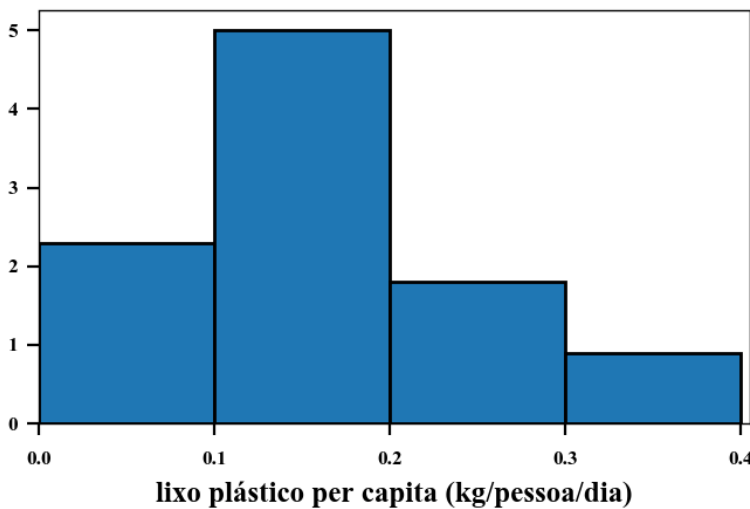
Resposta

a) A média é $\bar{x} = \frac{1}{n} \sum_i x_i = 0,175$. Temos 22 dados, logo a mediana é calculada tomando a média entre o 11º e o 12º dados, que valem ambos 0,147, logo a mediana é 0,147.

b) Os valores se encontram entre 0 e 0,4, logo dividimos esse intervalo em 4 sub-intervalos iguais (cada um de largura 0,1) e contamos quantos dados existem em cada um:

classe	# dados	y
[0; 0,1)	5	2,3
[0,1; 0,2)	11	5
[0,2; 0,3)	4	1,8
[0,3; 0,4)	2	0,9

Agora plotamos esses valores num gráfico de barras. A área embaixo desses gráfico será dado pela soma das alturas (22, dos 22 dados) pela largura da base, 0,1, logo a área total é de 2,2. Então precisamos reescalonar o eixo das ordenadas dividindo os valores por 2,2. O resultado aparece na terceira coluna acima. O histograma resultante aparece abaixo.



4. A função densidade de probabilidade de uma certa variável aleatória X é dada por

$$f(x) = \begin{cases} 0, & x \leq -1 \\ \frac{1+x}{2}, & -1 < x < 1 \\ 0, & x \geq 1 \end{cases}$$

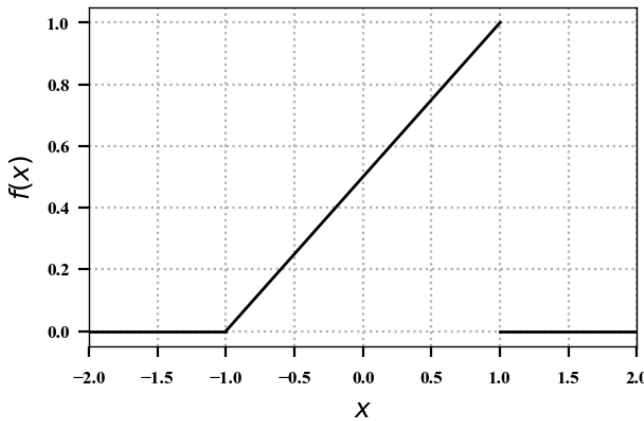
a) Esboce o gráfico da função densidade de probabilidade $f(x)$ da variável aleatória X .

b) Encontre a função de distribuição acumulada de probabilidade $F(x)$ da variável aleatória X e esboce o gráfico dessa função.

c) Calcule $P(X > 0.8)$. Para tanto, represente essa probabilidade nos gráficos acima.

Resposta

a) A função é constante igual a zero pra x menor que -1 ou maior que 1, e é linear—logo uma reta—entre -1 e 1. Como $f(-1) = 0$ e $f(1) = 1$, obtemos o gráfico abaixo.



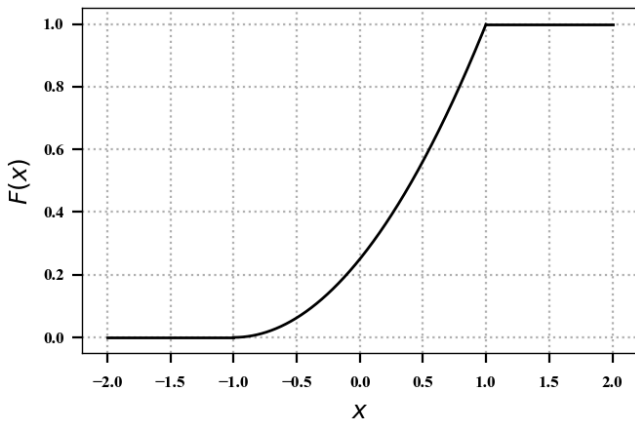
b) Por definição, $F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$. Temos três casos:

$$x < -1 \quad \int_{-\infty}^x f(t)dt = 0$$

$$-1 < x < 1 \quad \int_{-\infty}^x f(t)dt = \int_{-\infty}^{-1} 0 dt + \int_{-1}^x \frac{1+t}{2} dt = \frac{t+t^2/2}{2} \Big|_{-1}^x = \frac{x}{2} + \frac{x^2}{4} + \frac{1}{4}$$

$$x > 1 \quad \int_{-\infty}^x f(t)dt = \int_{-\infty}^{-1} 0 dt + \int_{-1}^1 \frac{1+t}{2} dt + \int_1^x 0 dt = 1$$

O resultado é constante igual a zero pra x menor que -1, constante igual a 1 pra $x > 1$, e uma função quadrática—uma parábola—entre -1 e 1, como na figura abaixo.



c) $P(X > 0,8) = 1 - P(X < 0,8) = 1 - F(0,8) = 1 - \frac{0,8}{2} + \frac{0,8^2}{4} + \frac{1}{4} = 0,19$. Outra forma de fazer isso é integrando a densidade de probabilidade de 0,8 a infinito: $P(X > 0,8) = \int_{0,8}^{\infty} f(t)dt = \int_{0,8}^1 f(t)dt = \frac{t+t^2/2}{2} \Big|_{0,8}^1 = 0,19$. No gráfico do item (a), isso corresponde à área embaixo do gráfico na região com x entre 0,8 e 1. No gráfico do item (b), isso corresponde à diferença entre o valor de $F(x)$ no ponto $x = 0,8$ e o máximo de 1.

5. O número de acidentes de trânsito graves por ano em uma cidade pode ser modelado como uma variável aleatória com distribuição de Poisson de esperança 4. Qual é a probabilidade de que em um determinado ano ocorra no máximo 1 acidente?

Resposta

A variável aleatória X que representa o número de acidentes é uma Poisson de esperança 4, do que sabemos que seu parâmetro $\lambda = 4$. Queremos calcular

$$P(X \leq 1) = P(X = 0) + P(X = 1) = e^{-4} \frac{4^0}{0!} + e^{-4} \frac{4^1}{1!} = e^{-4} + 4 \cdot e^{-4} \approx 0,09$$

6. Uma sorveteria atende 100 pessoas por dia. O consumo médio de sorvete de cada pessoa é de 80g, com desvio-padrão de 30g.

a) Assumindo que o consumo de cada pessoa seja uma variável aleatória normal, qual é a probabilidade de uma pessoa consumir mais que 100g?

b) Usando o Teorema do Limite Central, calcule a probabilidade de que a média do consumo dessas 100 pessoas seja maior que 100g.

c) Baseando-se nas respostas acima, é necessário que a sorveteria mantenha o estoque necessário pra servir 100g de sorvete para 100 pessoas, de $100 \times 100g = 10kg$ de sorvete?

Resposta

a) Pelo enunciado, o consumo de arroz de uma pessoa é uma V.A. normal com $\mu = 80$ e $\sigma = 30$. Daí temos que $P(X > 100) = 1 - P(X < 100) = 1 - F(100)$. Mudando de variáveis pra uma normal padrão com $Z = \frac{X-\mu}{\sigma}$, temos que $F(100) = \Phi\left(\frac{100-80}{30}\right) = \Phi(2/3)$. Usando a tabela, o valor mais próximo é de 0,726 (o valor exato seria próximo de 0,747), logo $P(X > 100) \approx 1 - 0,726 = 0,274$.

b) O Teorema do Limite Central diz que a variável aleatória \bar{X} da média de n variáveis aleatórias X independentes e identicamente distribuídas tende a uma V.A. normal com $\mu = \mu_X$ e $\sigma = \frac{\sigma_X}{\sqrt{n}}$, em que μ_X e σ_X são a esperança e desvio-padrão de cada experimento. Neste caso, temos $\mu = \mu_X = 80$ e $\sigma = \frac{30}{\sqrt{100}} = 3$. Como no item anterior, teremos

$$P(X > 100) = 1 - P(X < 100) = 1 - F(100) = 1 - \Phi\left(\frac{100 - 80}{3}\right) = 1 - \Phi(6,67) < 0,003,$$

onde o valor da tabela utilizado foi o maior de todos (o valor exato seria bem menor, próximo de 10^{-11}).

c) É raro ou comum que precisemos manter em estoque mais do que 100g de sorvete por pessoa? A resposta do item (a) diz que é comum que **uma** pessoa consuma mais do que 100g, mas a do item (b) diz que é muito raro que **a média de 100** pessoas o faça. O que precisamos pro nosso cálculo do estoque é justamente o consumo médio, não o individual, logo não precisamos nos preocupar muito com uma possibilidade de menos de 0,3%, e não precisamos de um estoque tão grande.

Dados:

Tabela 2: Variáveis aleatórias discretas

Variável aleatória	função de probabilidade $P(X = k)$	esperança	variância
Binomial	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	$e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ
Geométrica	$(1-p)^{k-1} p, k \geq 1$	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Tabela 3: Variáveis aleatórias contínuas

Variável aleatória	densidade de prob. $f(x)$	distribuição acumulada $F(x)$	esperança	variância
Uniforme	$\frac{1}{b-a}, x \in (a; b)$	$\frac{x-a}{b-a}, x \in (a; b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$\Phi\left(\frac{x-\mu}{\sigma}\right)$	μ	σ^2
Exponencial	$\lambda e^{-\lambda x}, x > 0$	$1 - e^{-\lambda x}, x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Tabela 4: Tabela de valores da função de distribuição acumulada $\Phi(z)$ da distribuição normal padrão, $Z = \frac{X-\mu}{\sigma}$, de esperança 0 e variância 1. Lembre-se que a normal é simétrica, logo $\Phi(-z) = 1 - \Phi(z)$.

z	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8
$\Phi(z)$	0.500	0.579	0.655	0.726	0.788	0.841	0.885	0.919	0.945	0.964	0.977	0.986	0.992	0.995	0.997