

ME414 - Estatística para Experimentalistas

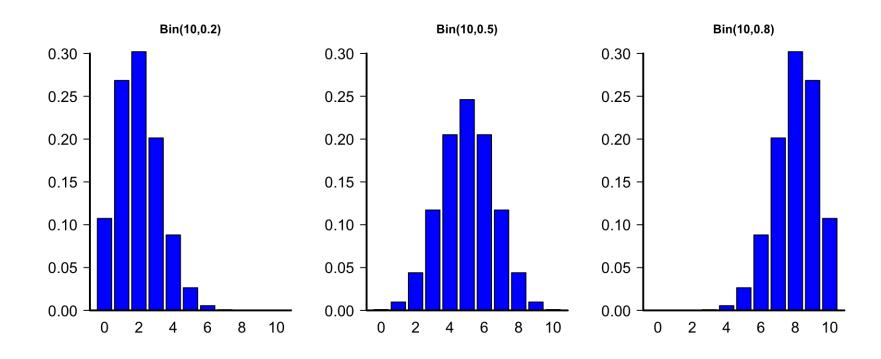
Parte 12

- Variáveis aleatórias discretas: v.a. com valores possíveis enumeráveis. Soma das probabilidades de todos os valores possíveis igual a 1.
- · Variáveis aleatórias contínuas: v.a. com valores possíveis em um intervalo no conjunto de números reais.
- · Exemplo: tempo para finalizar um experimento, peso de uma pessoa, duração de uma chamada telefônica, tempo de vida de uma lâmpada, etc...
- · Para esses tipos de quantidades, não é possível associar frequências pontuais tais que a soma de todas seja igual a 1.
- · Surge então o conceito de "função de densidade de probabilidade" (f.d.p.).
- Para cada v.a. contínua, associamos uma função de densidade de probabilidade.



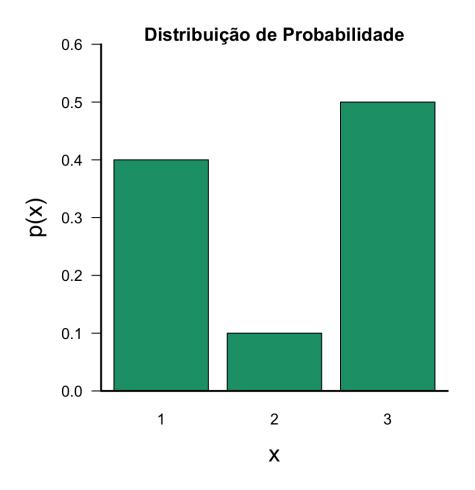
Exemplo: v.a. discreta

Distribuição de probabilidade de uma Bin(10, p), com p = 0.2, 0.5 e 0.8.



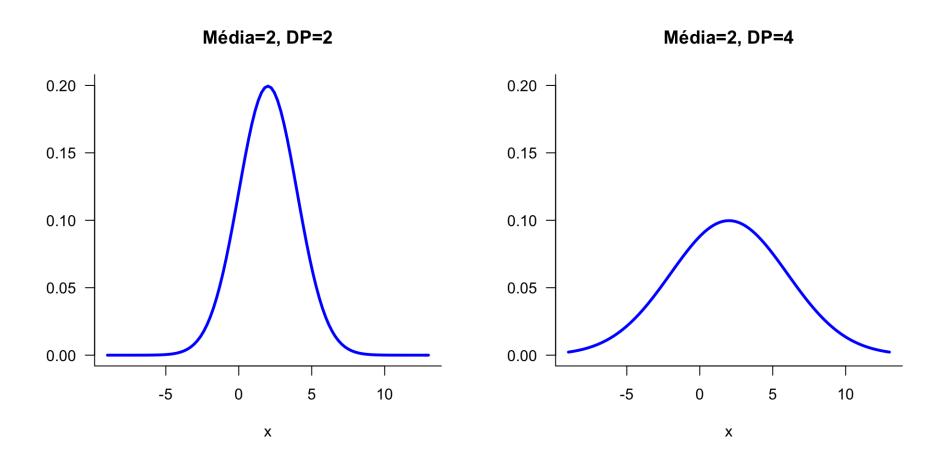


Exemplo: v.a. discreta





Exemplo: v.a. contínua





Definição: a função de densidade de probabilidade de uma v.a. X é uma função f que verifica:

$$f(x) \ge 0, \quad \forall x \in \mathbb{R}$$

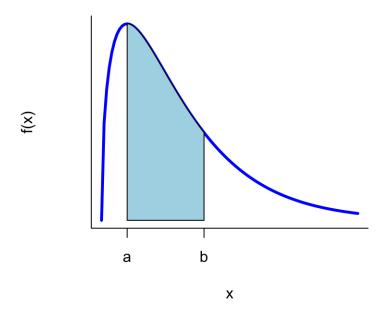
$$\int_{-\infty}^{\infty} f(x)dx = 1 \text{ (} f \text{ \'e integr\'avel)}$$

Toda v.a. X à qual seja possível associar uma função de densidade de probabilidade será chamada de v.a. contínua.



A probabilidade de que uma v.a. X contínua pertença a um intervalo da reta (a,b], a < b é dada por:

$$P(a < X \le b) = \int_{a}^{b} f(x)dx.$$



Obs: Quando X é v.a. contínua $P(a < X \le b) = P(a \le X \le b) = P(a < X < b) = P(a \le X < b)$



Notação: se X v.a. contínua com função de densidade de probabilidade f, no lugar de f denotaremos f_X .

A probabilidade de que uma v.a. X contínua pertença ao intervalo da reta $(-\infty,x]$ daremos o nome de **função de distribuição acumulada** (f.d.a.), e a denotaremos por F_X

$$F_X(x) = \int_{-\infty}^x f(u)du = P(X \le x).$$



Exemplo: X v.a. contínua com função de densidade:

$$f_X(x) = \begin{cases} x, & 0 \le x < 1 \\ 2 - x, & 1 \le x \le 2 \\ 0, & \text{caso contrário} \end{cases}$$

Pela definição de função de densidade:

$$f(x) \ge 0, \quad \forall x \in \mathbb{R}$$

$$\int_0^2 f(x)dx = \int_0^1 xdx + \int_1^2 (2-x)dx = 1$$

Podemos também calcular
$$P(0 < X \le 0.8) = \int_0^{0.8} x dx = 0.32$$



Propriedade: toda v.a. X contínua (ou seja, que possui f_X como função de densidade) tem probabilidade pontual nula: P(X=x)=0.

A função de distribuição acumulada $F(x) = P(X \le x)$:

caso discreto:

$$F(x) = P(X \le x) = \sum_{x_i \le x} P(X = x_i)$$

caso contínuo:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f_X(u) du$$



Dada a função

$$f(x) = \begin{cases} 0 & \text{se } x < 0 \\ 2e^{-2x} & \text{se } x \ge 0 \end{cases}$$

- 1. Mostre que esta é uma função de densidade de probabilidade.
- 2. Calcule a probabilidade de X > 10.

Fonte: Morettin & Bussab, Estatística Básica 5^a edição, pág 166.



Exemplo (solução - item 1)

Uma função de densidade de probabilidade deve satisfazer as seguintes propriedades:

 $1. f(x) \ge 0$ para todo $x \in \mathbb{R}$.

$$2. \int_{-\infty}^{\infty} f(x)dx = 1.$$

Note que e^{-x} é positiva para qualquer x e, consequentemente, $2e^{-2x}$ também.

Resta mostrar que sua integral é 1:

$$\int 2e^{-2x}dx = -e^{-2x}$$



Exemplo (solução)

1. Note que a função está definida nesta forma para $x \ge 0$; para x < 0, ela é 0. Então a integral é:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{\infty} 2e^{-2x}dx$$
$$= \left[-e^{-2x}\right]_{0}^{\infty}$$
$$= \lim_{x \to \infty} -e^{-2x} - \left(-e^{-0}\right) = 1$$

2. A probabilidade é dada por:

$$P(X > 10) = \int_{10}^{\infty} 2e^{-2x} dx = \lim_{x \to \infty} -e^{-2x} - \left(-e^{-2 \times 10}\right) = \frac{1}{e^{20}}$$



Esperança

Definição: seja X uma v.a. contínua com densidade f_X , a **esperança** de X é dada por:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Definição: seja X uma v.a. contínua com densidade f_X , o k-ésimo momento de X é dado por:

$$\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k f_X(x) dx.$$



Variância

Definição: seja X v.a. com valor esperado $\mathbb{E}(X)$, definimos por **variância**, a quantidade:

$$Var(X) = \mathbb{E}([X - \mathbb{E}(X)]^2) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2.$$

E definimos como desvio-padrão:

$$\sigma = \sqrt{Var(X)}$$

Note que assim como no caso discreto, ambas as quantidades oferecem medidas de dispersão da variável X em relação ao valor esperado $\mathbb{E}(X)$.



Para a função f_X seguinte, calcular $\mathbb{E}(X)$ e Var(X):

$$f_X(x) = \begin{cases} x, & 0 \le x < 1 \\ 2 - x, & 1 \le x \le 2 \\ 0, & \text{caso contrário} \end{cases}$$

Solução:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x^2 dx + \int_1^2 x (2 - x) dx = 1$$

$$Var(X) = \int_{-\infty}^{\infty} [x - \mathbb{E}(X)]^2 f_X(x) dx = \int_{-\infty}^{\infty} (x - 1)^2 f_X(x) dx = \frac{1}{6}$$



Uma variável aleatória X tem distribuição triangular no intervalo [0,1] se sua f.d.p. for dada por

$$f(x) = \begin{cases} 0 & \text{se } x < 0 \\ Cx & \text{se } 0 \le x \le 1/2 \\ C(1-x) & \text{se } 1/2 \le x \le 1 \\ 0 & \text{se } x > 1 \end{cases}$$

- 1. Qual valor deve ter a constante C?
- 2. Faça o gráfico de f(x).
- 3. Determine $P(X \le 1/2)$, P(X > 1/2) e $P(1/4 \le X \le 3/4)$.

Fonte: Morettin & Bussab, Estatística Básica 5^a edição, pág 166.



Item 1 - Devemos escolher C de modo que f(x) satisfaça:

 $f(x) \ge 0$ para todo $x \in \mathbb{R}$.

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

Pela primeira condição, temos que C>0. Agora, para que C satisfaça a segunda condição, devemos integrar f(x):

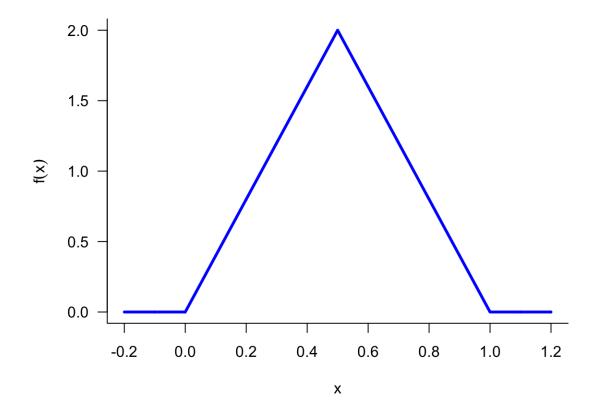
$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{1/2} Cxdx + \int_{1/2}^{1} C(1-x)dx + \int_{1}^{\infty} 0dx$$

$$= C \int_{0}^{1/2} xdx + C \int_{1/2}^{1} (1-x)dx = C \left(\left[\frac{x^{2}}{2} \right]_{0}^{1/2} + \left[x - \frac{x^{2}}{2} \right]_{1/2}^{1} \right)$$

$$= C \left(\frac{1}{8} + 1 - \frac{1}{2} - \frac{1}{2} + \frac{1}{8} \right) = \frac{C}{4} \Rightarrow C \text{ deve ser igual a 4.}$$



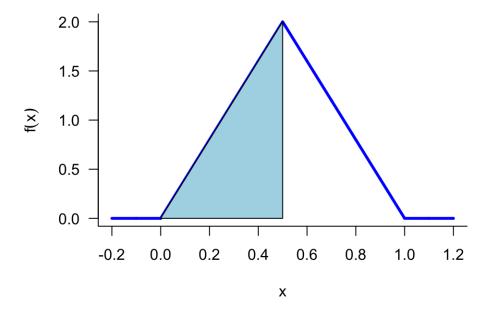
Item 2 - Função de densidade f(x)





Item 3 - Para encontrarmos as probabilidades dos eventos, basta integrar nas regiões correspondentes:

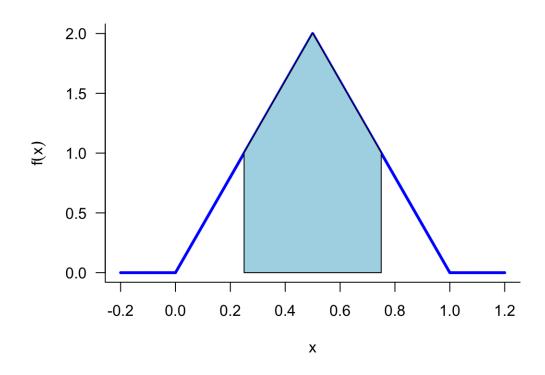
$$P(X \le 1/2) = \int_0^{1/2} f(x)dx = \int_0^{1/2} 4xdx = 1/2.$$



Note que $P(X > 1/2) = 1 - P(X \le 1/2) = 1 - 1/2 = 1/2$.



$$P(1/4 \le X \le 3/4) = \int_{1/4}^{3/4} f(x)dx = \int_{1/4}^{1/2} 4xdx + \int_{1/2}^{3/4} 4(1-x)dx = \frac{3}{4}.$$





Calcule a esperança, a variância e a f.d.a. da variável aleatória X com a densidade triangular em [0,1].

Fonte: Morettin & Bussab, Estatística Básica 5^a edição, pág 171.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{0}^{1/2} 4x^{2} dx + \int_{1/2}^{1} 4x(1-x) dx$$
$$= \left[\frac{4x^{3}}{3} \right]_{0}^{1/2} + \left[\frac{2}{3} x^{2} (3-2x) \right]_{1/2}^{1} = \frac{1}{2}$$



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

$$= \int_{0}^{1/2} 4\left(x - \frac{1}{2}\right)^2 x dx + \int_{1/2}^{1} 4\left(x - \frac{1}{2}\right)^2 (1 - x) dx$$

$$= \left[x^4 - \frac{4}{3}x^3 + \frac{1}{2}x^2\right]_{0}^{1/2} + \left[-x^4 + \frac{8}{3}^3 - \frac{5}{2}x^2 + x\right]_{1/2}^{1}$$

$$= \frac{1}{24}$$



A função de distribuição acumulada de uma variável aleatória contínua é dada por

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

Temos que para $x \in [0, 1/2)$, F(x) é dada por

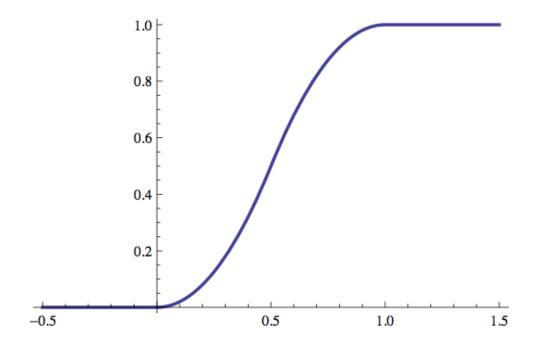
$$F(x) = \int_0^x 4tdt = 2x^2$$

Para $x \in [1/2, 1]$, a acumulada é dada por

$$F(x) = \int_0^{1/2} 4t dt + \int_{1/2}^x 4(1-t)dt = -2x^2 + 4x - 1$$



Para valores de $x \ge 1$, a acumulada assume valor 1. O gráfico de F(x) é dado por:





Distribuição Uniforme

Uniforme

· Dizemos que a v.a. X tem distribuição Uniforme no intervalo [a,b], a < b se a função de densidade de probabilidade f_X é dada por:

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{caso contrário} \end{cases}$$

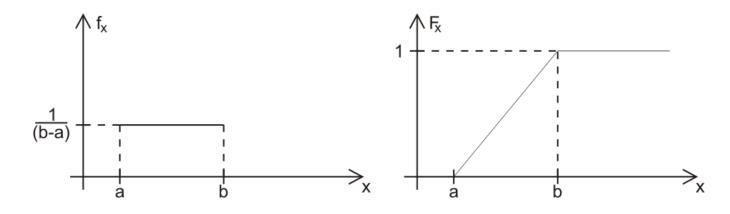
- · Notação: $X \sim U[a, b]$ ou $X \sim U(a, b)$.
- · Cálculo da função de distribuição acumulada:

$$F_X(x) = \begin{cases} 0, & x < a \\ \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$



Uniforme

Gráficos da função de densidade de probabilidade e da função de distribuição acumulada:





Esperança e Variância

· Cálculo da $\mathbb{E}(X)$:

$$\mathbb{E}(X) = \int_a^b x \frac{1}{(b-a)} dx = \frac{(b+a)}{2}$$

· Cálculo da Var(X):

$$\mathbb{E}(X^2) = \int_a^b x^2 \frac{1}{(b-a)} dx = \frac{(a^2 + ab + b^2)}{3}$$

$$Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

$$= \frac{(a^2 + ab + b^2)}{3} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12}$$



Exemplo: peça de aço

A dureza H de uma peça de aço pode ser pensada como uma v.a. com distribuição uniforme no intervalo [50,70] da escala de Rockwel. Calcule a probabilidade de que uma peça tenha dureza entre 55 e 60.

A densidade de uma U[50, 70] é dada por:

$$f_H(h) = \begin{cases} \frac{1}{70-50} = \frac{1}{20}, & 50 \le x \le 70\\ 0, & \text{caso contrário} \end{cases}$$

Portanto, a probabilidade de uma peça ter dureza entre 55 e 60 é:

$$P(55 < H < 60) = \int_{55}^{60} \frac{1}{20} dh = \frac{1}{20} (60 - 55) = \frac{1}{4}.$$



Seja X uma variável aleatória distribuída uniformemente, com média 15 e variância 25/3.

- Encontre a função de densidade de X.
- · Qual é a probabilidade que X > 14?



Lembre-se que a esperança de uma v.a. uniforme em [a,b] é dada por

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

e sua variância por

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

Temos, portanto, o seguinte sistema:

$$\begin{cases} \frac{a+b}{2} &= 15\\ \frac{(b-a)^2}{12} &= \frac{25}{3} \end{cases} \implies \begin{cases} a+b &= 30\\ (b-a)^2 &= 100 \end{cases}$$



Ou simplesmente (você é capaz de dizer por que tomamos a raiz positiva apenas, neste sistema não-linear?)

$$\begin{cases} a+b &= 30 \\ b-a &= 10 \end{cases}$$

O sistema tem solução a=10, b=20, o que nos mostra que $X\sim U[10,20]$ e

$$f_X(x) = \begin{cases} \frac{1}{10}, & 10 \le x \le 20\\ 0, & \text{caso contrário} \end{cases}$$

Então, a probabilidade de X>14 é dada por

$$P(X > 14) = \int_{14}^{20} \frac{1}{10} dx = \frac{(20 - 14)}{10} = 0.6$$



Distribuição Exponencial

Exponencial

· Dizemos que uma v.a. X possui distribuição exponencial com parâmetro λ ($\lambda>0$) se a função de densidade de probabilidade f_X é dada por:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0\\ 0, & \text{caso contrário} \end{cases}$$

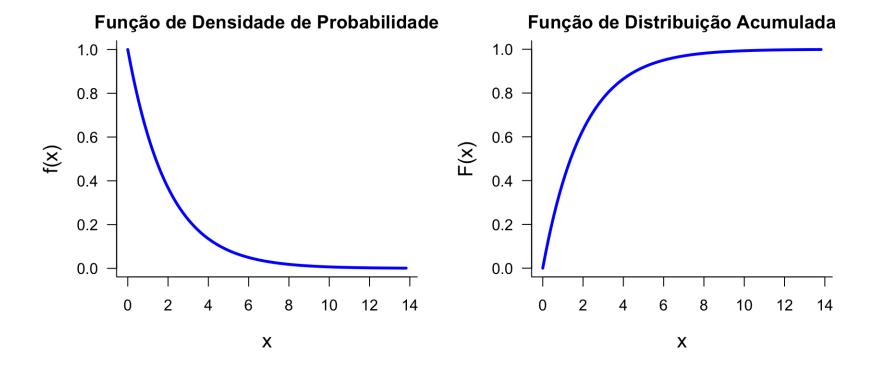
- · Notação: $X \sim exp(\lambda)$.
- · Cálculo da função de distribuição acumulada:

$$F_X(x) = \begin{cases} \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, & x \ge 0\\ 0, & \text{caso contrário} \end{cases}$$



Distribuição Exponencial

Gráficos da função de densidade de probabilidade (esquerda) e da função de distribuição acumulada (direita) de $X \sim Exp(0.5)$:





Esperança e Variância

· Cálculo da $\mathbb{E}(X)$:

$$\mathbb{E}(X) = \int_0^\infty x \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

· Cálculo da Var(X):

$$\mathbb{E}(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$



Exemplo: componente eletrônico

O tempo de vida, X, em horas, de um componente eletrônico segue uma distribuição exponencial de tal forma que $P(X \le 1000) = 0.75$.

Qual é o tempo médio de vida do componente?



Exemplo: componente eletrônico

Sabemos que se $X \sim exp(\lambda)$, então

$$F(x) = P(X \le x) = 1 - e^{-\lambda x}$$

 $e \mathbb{E}(X) = \lambda^{-1}$.

Basta então observarmos que

$$P(X \le 1000) = 1 - e^{-\lambda 1000} = 0.75 \quad \Leftrightarrow \quad \lambda = \frac{\ln(4)}{1000} = 0.0013863$$

Concluimos então que o tempo médio de vida, $\mathbb{E}(X)$, é igual a 1/0.0013863 = 721.35 horas, e que 75% dos componentes duram 1000 horas ou menos.



Exemplo: tubos de TV

Um antiga fábrica de tubos de TV determinou que a vida média dos tubos de sua fabricação é de 800 horas de uso contínuo e segue uma distribuição exponencial.

Qual a probabilidade de que a fábrica tenha que substituir um tubo gratuitamente, se oferece uma garantia de 300 horas de uso?



Exemplo: tubos de TV

X: vida útil do tubo de TV.

Como X tem distribuição exponencial com parâmetro λ e $\mathbb{E}[X] = 800$, temos:

$$\mathbb{E}[X] = \frac{1}{\lambda} = 800 \qquad \Longrightarrow \qquad \lambda = \frac{1}{800}.$$

Então, a f.d.p. é dada por
$$f_X(x) = \begin{cases} \frac{1}{800} e^{-\frac{x}{800}}, & x \ge 0 \\ 0, & \text{caso contrário} \end{cases}$$

Se X < 300, a fábrica tem que substituir gratuitamente.

$$P(X < 300) = \int_0^{300} \frac{1}{800} e^{-\frac{x}{800}} dx = \left[-e^{-\frac{x}{800}} \right]_0^{300} = 0.3127.$$



Exemplo: produto alimentício

A f.d.p.

$$f_X(x) = \begin{cases} 2e^{-2x}, & x \ge 0\\ 0, & \text{caso contrário} \end{cases}$$

representa a distribuição do índice de acidez (X) de um determinado produto alimentício.

O produto é consumível se este índice for menor do que 2.

O setor de fiscalização apreendeu 30 unidades do produto. Qual a probabilidade de que pelo menos 10% da amostra seja imprópria para consumo?



Exemplo: produto alimentício

Produto liberado para consumo se: $P(X < 2) = \int_0^2 2e^{-2x} dx = \left[-e^{-2x} \right]_0^2 = 0.98$.

Produto não consumível com probabilidade 1 - P(X < 2) = 0.02 = p.

Y: número de unidades impróprias para consumo na amostra de 30.

 $Y \sim \text{Bin}(30, 0.02).$

Probabilidade de que pelo menos 10% de uma amostra de 30 unidades seja imprópria:

$$P(Y \ge 3) = 1 - P(Y < 3)$$

$$= 1 - [P(Y = 0) + P(Y = 1) + P(Y = 2)]$$

$$= 1 - \left[\binom{30}{0} (0.02)^0 (0.98)^{30} + \binom{30}{1} (0.02)^1 (0.98)^{29} + \binom{30}{2} (0.02)^2 (0.98)^{28} \right]$$

$$= 0.022$$



Leituras

- Ross: seções 6.1 e 6.2.
- · Magalhães: capítulo 6.

Slides produzidos pelos professores:

- Samara Kiihl
- · Tatiana Benaglia
- Larissa Matos
- Benilton Carvalho



