Estatística aplicada

Lino Costa

Departamento de Produção e Sistemas
Escola de Engenharia
lac@dps.uminho.pt

Ano letivo 2015/2016

Sumário

Variáveis aleatórias e distribuições de probabilidade

1. Variável aleatória
2. Distribuições de probabilidade discretas
   - Função de probabilidade
   - Função de distribuição acumulada
3. Distribuições de probabilidade contínuas
   - Função densidade de probabilidade
   - Função de distribuição acumulada

Variável aleatória

Uma variável aleatória, representada por X (em maiúsculas), associa um número real aos resultados de uma experiência aleatória.
De notar que um valor medido da variável aleatória X é representado em minúsculas (por exemplo, $x = 10$).

Tipo de variáveis aleatórias

- **Variável aleatória discreta**: apresenta uma gama de valores finita (ou infinita contável). Por exemplo: o número de pontos resultante de um lançamento de um dado $X = 1, 2, 3, 4, 5, 6$.
- **Variável aleatória contínua**: define-se num intervalo de valores reais para uma gama de valores infinita. Por exemplo: duração de uma lâmpada $X \geq 0$.

Variável aleatória

Exemplo 1

Considere o lançamento de 2 dados. Liste os elementos do espaço amostral, os valores da variável aleatória X que representa a soma dos pontos e as probabilidades correspondentes.

- $S = \{(1,1),(1,2),(2,1),\ldots,(5,6),(6,5),(6,6)\}$
-

elementos	X	probabilidade
$(1,1)$	2	$1/36$
$(1,2),(2,1)$	3	$2/36$
$(1,3),(3,1),(2,2)$	4	$3/36$
$(1,4),(4,1),(2,3),(3,2)$	5	$4/36$
$(1,5),(5,1),(2,4),(4,2),(3,3)$	6	$5/36$
$(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)$	7	$6/36$
$(2,6),(6,2),(3,5),(5,3),(4,4)$	8	$5/36$
$(3,6),(6,3),(4,5),(5,4)$	9	$4/36$
$(4,6),(6,4),(5,5)$	10	$3/36$
$(5,6),(6,5)$	11	$2/36$
$(6,6)$	12	$1/36$

- $f(x) = P(X = x) = \frac{6-|x-7|}{36} \quad x = 2,3,4,5,6,7,8,9,10,11,12$

Variável aleatória

Exemplo 2

Duas meias são selecionadas aleatoriamente de uma gaveta contendo 5 meias castanhas e 3 verdes. Liste os elementos do espaço amostral, os valores da variável aleatória W que representa o número de meias castanhas selecionadas e as probabilidades correspondentes.

- $S = \{CC, CV, VC, VV\}$
-

elemento	W	probabilidade
CC	2	$5/8 \times 4/7 = 20/56$
CV	1	$5/8 \times 3/7 = 15/56$
VC	1	$3/8 \times 5/7 = 15/56$
VV	0	$3/8 \times 2/7 = 6/56$

- $f(w) = P(W = w) = \begin{cases} 6/56 & w = 0 \\ 30/56 & w = 1 \\ 20/56 & w = 2 \end{cases}$

Distribuição de probabilidade

A distribuição de probabilidade indica como é que as probabilidades se encontram distribuídas para os diversos valores da variável aleatória discreta X.

Tipos de funções para uma variável aleatória discreta X

- **função de probabilidade**: indica a probabilidade de uma valor de X, i.e., $P(X = x)$
- **função de distribuição acumulada**: indica a soma das probabilidades dos valores de X que são menores ou iguais a um valor especificado, i.e., $P(X \leq x)$.

Função de probabilidade

Se X é uma variável aleatória discreta, a função dada por $f(x) = P(X = x)$, para cada valor de x na gama de valores de X, é chamada **função de probabilidade** de X.

Propriedades

$f(x)$ é função de probabilidade de uma variável aleatória discreta X se e só se os seus valores satisfazem as seguintes condições:

1. $f(x) \geq 0$ para qualquer valor do seu domínio;
2. $\sum_x f(x) = 1$.

Função de probabilidade

Exemplo 3

Encontre a fórmula para a função de probabilidade do número total de caras obtidas no lançamento de 4 moedas equilibradas.

- $$S = \{HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT,\\ THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT\}$$

-

elementos	X	probabilidade
$HHHH$	4	$1/16$
$HHHT, HHTH, HTHH, THHH$	3	$4/16$
$HHTT, HTHT, HTTH, THHT, THTH, TTHH$	2	$6/16$
$HTTT, THTT, TTHT, TTTH$	1	$4/16$
$TTTT$	0	$1/16$

- $f(x) = P(X = x) = \frac{C_x^4}{16} = \frac{\binom{4}{x}}{16} = \frac{\frac{4!}{x!(4-x)!}}{16} \quad x = 0, 1, 2, 3, 4$

Função de probabilidade

Exemplo 4

Verifique se a função dada por

$$f(x) = \frac{x+2}{25} \quad x = 1, 2, 3, 4, 5$$

pode servir como função de probabilidade de uma variável aleatória.

- $f(x) \geq 0$ para $x = 1, 2, 3, 4, 5$
- $\displaystyle\sum_{x=1}^{5} f(x) = \frac{3}{25} + \frac{4}{25} + \frac{5}{25} + \frac{6}{25} + \frac{7}{25} = \frac{25}{25} = 1$
- $f(x)$ é uma função de probabilidade

Função de distribuição acumulada

A **função de distribuição acumulada** $F(x)$ de um variável aleatória discreta X é:

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \quad \text{onde} \quad -\infty < x < \infty$$

Propriedades

A função de distribuição acumulada $F(x)$ satisfaz as seguintes condições:

1. $0 \leq F(x) \leq 1$
2. $F(-\infty) = 0$
3. $F(\infty) = 1$
4. $F(a) \leq F(b)$ se $a \leq b$ ($F(x)$ é não decrescente)

Função de distribuição acumulada

Exemplo 5

Considere a experiência em que duas meias são selecionadas aleatoriamente de uma gaveta contendo 5 meias castanhas e 3 verdes. Encontre a função de probabilidade acumulada da variável W, número de meias castanhas extraídas da gaveta. Trace os gráficos de $f(w)$ e $F(w)$.

-

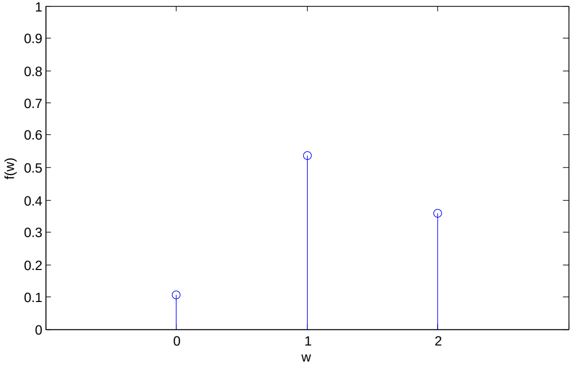
W	$f(w) = P(W = w)$	$F(w) = P(W \leq w)$	
0	$6/56$	$6/56$	$f(0)$
1	$30/56$	$36/56$	$f(0) + f(1)$
2	$20/56$	1	$f(0) + f(1) + f(2)$

- $F(w) = P(W \leq w) = \begin{cases} 0 & w < 0 \\ 6/56 & 0 \leq w < 1 \\ 36/56 & 1 \leq w < 2 \\ 1 & w \geq 2 \end{cases}$

Função de probabilidade

Exemplo 5

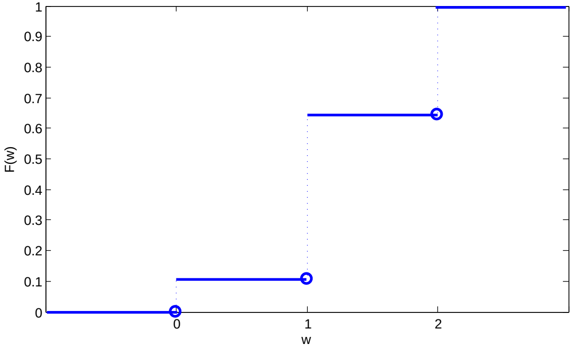
$$f(w) = P(W = w) = \begin{cases} 6/56 & w = 0 \\ 30/56 & w = 1 \\ 20/56 & w = 2 \end{cases}$$



Função de distribuição acumulada

Exemplo 5

$$F(w) = P(W \leq w) = \begin{cases} 0 & w < 0 \\ 6/56 & 0 \leq w < 1 \\ 36/56 & 1 \leq w < 2 \\ 1 & w \geq 2 \end{cases}$$



Distribuição de probabilidade

A distribuição de probabilidade indica como é que as probabilidades se encontram distribuídas na gama de valores da variável aleatória contínua X.

Tipos de funções para uma variável aleatória contínua X

- **função densidade de probabilidade**: descreve a densidade de probabilidade num intervalo de valores de X
- **função de distribuição acumulada**: representa a probabilidade de X ser menor ou igual a um valor especificado, i.e., $P(X \leq x)$.

Função densidade de probabilidade

Se X é uma variável aleatória contínua, a função dada por $f(x)$, para valores de x na gama de valores de X, é chamada **função densidade de probabilidade** de X.

Propriedades

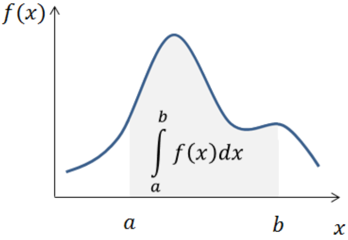
$f(x)$ é função densidade de probabilidade de uma variável aleatória contínua X se e só se os seus valores satisfazem as seguintes condições:

1. $f(x) \geq 0$ para qualquer valor do seu domínio;
2. $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$;
3. $P(a \leq X \leq b) = \displaystyle\int_{a}^{b} f(x)dx$;
4. $P(X = x) = 0$.

Função densidade de probabilidade

O integral de $f(x)$ entre a e b representa a probabilidade que X pertença ao intervalo $a \leq x \leq b$

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



Como $P(X = x) = 0$ então

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

Função densidade de probabilidade

Exemplo 6

A função densidade de probabilidade da variável aleatória X é dada por

$$f(x) = \begin{cases} ke^{-3x} & x > 0 \\ 0 & \text{outros valores} \end{cases}$$

Determine o valor de k. Calcule $P(X \geq 0.5)$ e $P(0.5 < X < 1)$.

- $\displaystyle\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} ke^{-3x}dx = \left[k\frac{e^{-3x}}{-3}\right]_0^{\infty} = \frac{k}{3} = 1 \Leftrightarrow k = 3$
- $\displaystyle P(X \geq 0.5) = \int_{0.5}^{\infty} 3e^{-3x}dx = \left[3\frac{e^{-3x}}{-3}\right]_{0.5}^{\infty} = 0 - (-0.223) = 0.223$
- $\displaystyle P(0.5 < X < 1) = \int_{0.5}^{1} 3e^{-3x}dx = \left[3\frac{e^{-3x}}{-3}\right]_{0.5}^{1} = -0.050 - (-0.223) = 0.173$

Função de distribuição acumulada

A **função de distribuição acumulada** $F(x)$ de um variável aleatória contínua X é:

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

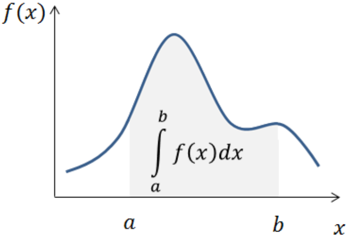
Propriedades

A função de distribuição acumulada $F(x)$ satisfaz as seguintes condições:

1. $0 \leq F(x) \leq 1$
2. $F(-\infty) = 0$
3. $F(\infty) = 1$
4. $F(a) \leq F(b)$ se $a \leq b$ ($F(x)$ é não decrescente)

Função de distribuição acumulada

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



Se a derivada existir então

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

Função de distribuição acumulada

Exemplo 7

Determine a função de distribuição acumulada correspondente à função densidade de probabilidade da variável aleatória X:

$$f(x) = \begin{cases} 3e^{-3x} & x > 0 \\ 0 & \text{outros valores} \end{cases}.$$

Utilizando a função de distribuição acumulada, calcule $P(X \geq 0.5)$ e $P(0.5 < X < 1)$.

- para $x > 0$, $F(x) = \int_0^x 3e^{-3x}dx = \left[3\frac{e^{-3x}}{-3}\right]_0^x = 1 - e^{-3x}$
- $F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-3x} & x \geq 0 \end{cases}$
- $P(X \geq 0.5) = 1 - P(X < 0.5) = 1 - F(0.5) = 1 - (1 - 0.223) = 0.223$
- $P(0.5 < X < 1) = F(1) - F(0.5) = 1 - 0.050 - (1 - 0.223) = 0.173$

Função de distribuição acumulada

Exemplo 8

Encontre a função densidade de probabilidade para a variável aleatória X cuja função de distribuição acumulada é dada por

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

- $f(x) = \frac{dF(x)}{dx}$
- para $x < 0$, $f(x) = 0$
- para $0 \leq x < 1$, $f(x) = \frac{dF(x)}{dx} = 1$
- para $x \geq 1$, $f(x) = 0$
- logo $f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{outros valores} \end{cases}$

Sumário

Esperança matemática

1. Valor esperado de uma distribuição de probabilidade discreta
   - média
   - variância
   - desvio padrão
2. Valor esperado de uma distribuição de probabilidade contínua
   - média
   - variância
   - desvio padrão

Valor esperado

Média de X (μ)

Se X é uma variável aleatória discreta e $f(x)$ a sua função de probabilidade, o valor esperado da variável aleatória X, representado por μ ou $E(X)$, é dado por

$$\mu = E(X) = \sum_x x f(x)$$

Propriedades

O valor esperado $E(X)$ tem as seguintes propriedades:

1. $E(kX) = kE(X)$ com k constante
2. $E(k) = k$ com k constante
3. $E(g(X)) = \sum_x g(x)f(x)$ sendo $g(X)$ uma qualquer função da variável aleatória X

Variância

Variância de X (σ^2)

Se X é uma variável aleatória discreta e $f(x)$ a sua função de probabilidade, a variância da variável aleatória X, representada por σ^2 ou $V(X)$, é dada por

$$\sigma^2 = V(X) = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2 = E(X^2) - (E(X))^2$$

Propriedades

A variância $V(X)$ tem as seguintes propriedades:

1. $V(kX) = k^2 V(X)$ com k constante
2. $V(k) = 0$ com k constante

Valor esperado e variância

Exemplo 1

Considere a experiência em que duas meias são selecionadas aleatoriamente de uma gaveta contendo 5 meias castanhas e 3 verdes. Encontre a média, a variância e o desvio padrão da variável W, número de meias castanhas extraídas da gaveta.

W	$f(w)$
0	$6/56$
1	$30/56$
2	$20/56$

- $\mu = E(W) = \sum_w w f(w) = 0 \times f(0) + 1 \times f(1) + 2 \times f(2) =$
  $= 0 + 30/56 + 40/56 = 70/56 = 1.25$
- $\sigma^2 = V(W) = E(W^2) - (E(W))^2$
- $E(W^2) = \sum_w w^2 f(w) = 0^2 \times f(0) + 1^2 \times f(1) + 2^2 \times f(2) =$
  $= 0 + 30/56 + 80/56 = 110/56 = 1.964$
- $\sigma^2 = V(W) = E(W^2) - (E(W))^2 = 110/56 - (70/56)^2 = 0.402$
- $\sigma = \sqrt{V(W)} = 0.634$

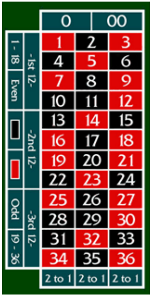
Valor esperado e variância

Exemplo 1

No jogo da roleta existem 36 números inscritos, alternadamente, em casas **vermelhas** e **negras**.
Além destes 36 números existem ainda duas casas **verdes**, com **0** e **00** inscritos.
Neste jogo existe a possibilidade de **apostar na cor negra ou na cor vermelha**:

- se **sair a cor em que apostou**, o jogador **recebe um prémio igual ao montante apostado**;
- se **sair uma das casas verdes**, o jogador **perde o dinheiro apostado** pois a banca recolhe tudo o que está em cima da mesa.

Suponha que um jogador **aposta 1000 euros na cor negra**. Calcule o valor esperado, a variância e o desvio padrão do ganho por parte do jogador.



Valor esperado e variância

Exemplo 2

Seja X a variável aleatória que representa o resultado desta experiência (o ganho do jogador quando aposta numa das cores), então as probabilidades são:

-

X	$f(X)$	
-1000	$20/38$	(o jogador perde e a banca ganha)
1000	$18/38$	(o jogador ganha e a banca perde)

- $\mu = E(X) = \sum_x x f(x) = -1000 \times 20/38 + 1000 \times 18/38 = -55$ euros, i.e., o ganho médio do jogador é -55 euros
- $\sigma^2 = V(X) = E(X^2) - (E(X))^2$
- $E(X^2) = \sum_x x^2 f(x) = (-1000)^2 \times 20/38 + 1000^2 \times 18/38 = 1000000$
- $\sigma^2 = V(X) = E(X^2) - (E(X))^2 = 1000000 - (-55)^2 = 996975$
- $\sigma = \sqrt{V(X)} = 998.49$ euros, i.e., o desvio padrão do ganho do jogador é 998.49 euros

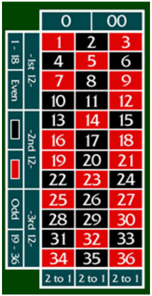
Valor esperado e variância

Exemplo 2

No jogo da roleta existe também a possibilidade de **apostar num de três grupos de 12 números dispostos em 3 colunas na mesa**:

- se **sair um dos números da coluna em que apostou**, o jogador **recebe um prémio igual ao dobro do montante apostado**;
- se **sair uma das casas verdes**, o jogador **perde o dinheiro apostado** pois a banca recolhe tudo o que está em cima da mesa.

Suponha que um jogador **aposta 1000 euros na primeira coluna**. Calcule para este tipo de aposta o valor esperado, a variância e o desvio padrão do ganho por parte do jogador.



Valor esperado e variância

Exemplo 2

Seja Y a variável aleatória que representa o resultado desta experiência (o ganho do jogador quando aposta numa das colunas), então as probabilidades são:

Y	$f(Y)$	
-1000	$26/38$	(o jogador perde e a banca ganha)
2000	$12/38$	(o jogador ganha e a banca perde)
- $\mu = E(Y) = \sum_y yf(y) = -1000 \times 26/38 + 2000 \times 12/38 = -55$ euros, i.e., o ganho médio do jogador é -55 euros
- $\sigma^2 = V(Y) = E(Y^2) - (E(Y))^2$
- $E(Y^2) = \sum_y y^2 f(y) = (-1000)^2 \times 26/38 + 2000^2 \times 12/38 = 1947368$
- $\sigma^2 = V(Y) = E(Y^2) - (E(Y))^2 = 1947368 - (-55)^2 = 1944343$
- $\sigma = \sqrt{V(Y)} = 1394.40$ euros, i.e., o desvio padrão do ganho do jogador é 1394.40 euros

O que será melhor apostar numa cor ou numa coluna?

- cor: $\mu = -55$ euros e $\sigma = 998.49$ euros
- coluna: $\mu = -55$ euros e $\sigma = 1394.40$ euros

Valor esperado

Média de X (μ)

Se X é uma variável aleatória contínua e $f(x)$ a sua função densidade de probabilidade, o valor esperado da variável aleatória X, representado por μ ou $E(X)$, é dado por

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

Propriedades

O valor esperado $E(X)$ tem as seguintes propriedades:

1. $E(kX) = kE(X)$ com k constante
2. $E(k) = k$ com k constante
3. $E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$ sendo $g(X)$ uma qualquer função da variável aleatória X

Variância

Variância de X (σ^2)

Se X é uma variável aleatória contínua e $f(x)$ a sua função densidade de probabilidade, a variância da variável aleatória X, representada por σ^2 ou $V(X)$, é dada por

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2 =$$
$$= E(X^2) - (E(X))^2$$

Propriedades

A variância $V(X)$ tem as seguintes propriedades:

1. $V(kX) = k^2 V(X)$ com k constante
2. $V(k) = 0$ com k constante

Valor esperado e variância

Exemplo 3

Considere a seguinte função densidade de probabilidade para a variável aleatória X

$$f(x) = \begin{cases} \frac{1}{8}(x+1) & 2 < x < 4 \\ 0 & \text{outros valores} \end{cases}$$

Calcule $E(X)$, $E(X^2)$, $V(X)$ e o desvio padrão.

- $\mu = E(X) = \int\limits_{-\infty}^{+\infty} x f(x) dx = \int\limits_{2}^{4} x \frac{1}{8}(x+1) dx = \frac{1}{8} \int\limits_{2}^{4} (x^2 + x) dx =$
  $= \frac{1}{8} \left[\frac{x^3}{3} + \frac{x^2}{2} \right]_2^4 = \frac{1}{8} \left(\frac{64}{3} + 8 - \frac{8}{3} - 2 \right) = \frac{37}{12} = 3.083$
- $E(X^2) = \int\limits_{-\infty}^{+\infty} x^2 f(x) dx = \int\limits_{2}^{4} x^2 \frac{1}{8}(x+1) dx = \frac{1}{8} \int\limits_{2}^{4} (x^3 + x^2) dx =$
  $= \frac{1}{8} \left[\frac{x^4}{4} + \frac{x^3}{3} \right]_2^4 = \frac{1}{8} \left(64 + \frac{64}{3} - 4 - \frac{8}{3} \right) = \frac{59}{6}$
- $\sigma^2 = V(X) = E(X^2) - (E(X))^2 = \frac{59}{6} - \left(\frac{37}{12} \right)^2 = \frac{47}{144} = 0.326$
- $\sigma = \sqrt{V(X)} = 0.571$