Variáveis Aleatórias

Introdução às Probabilidades e Estatística

Sumário

1. Variáveis Aleatórias

Variáveis Aleatórias

- Muitas vezes o resultado de uma experiência aleatória são **numéricos**.

  Experiência aleatória: Lançamento de um dado
  $\Omega = \{1, 2, 3, 4, 5, 6\}$

- Noutros casos os resultados são **não numéricos**, o que dificulta a aplicação de procedimentos matemáticos.

  Experiência aleatória: Retirada ao acaso de 2 peças de uma caixa com 4 peças boas (B) e 5 peças defeituosas (D).
  $\Omega = \{BB, BD, DB, DD\}$

É frequente não estarmos interessados em todos os resultados detalhados de uma dada experiência aleatória mas apenas num valor numérico.

E o que podemos fazer para lidar com essa situação?

Informalmente, surge então a necessidade de definirmos uma **função real** X que transforme cada resultado ω da experiência aleatória num número real $X(\omega)$ que denominamos de **variável aleatória**.

Variáveis Aleatórias (cont.)

Definição (Variável aleatória)

Seja Ω e $\mathcal{A}$ uma $\sigma-$álgebra de Ω. Dizemos que uma função real X é uma **variável aleatória (v.a.)** se associar cada resultado $\omega \in \Omega$ a um número real $X(\omega) \in \mathbb{R}$ e garantir que para qualquer $x \in \mathbb{R}$, o conjunto

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \leq x\}$$

pertence à σ-álgebra $\mathcal{A}$.

Observações:

1. Geralmente são representadas por letras maiúsculas X, Y, Z, W;
2. Um valor particular representa-se por uma letra minúscula;
3. $P(X = x) = P(X(\omega) = x) = P(\{\omega \in \Omega : X(\omega) = x\})$.

Ciências ULisboa

Variáveis Aleatórias (cont.)

Exemplo 1: Considere a seguinte experiência aleatória. Retirada ao acaso de 2 peças de uma caixa com 4 peças boas (B) e 5 peças defeituosas (D).

- $\Omega = \{BB, BD, DB, DD\}$
- X = "número de peças defeituosas nas duas extrações"
- Quais são os possíveis valores de X?
- $X(\Omega) = \{0, 1, 2\}$

Ω	$X(\Omega)$
$\{BB\}$	0
$\{BD\}$	1
$\{DB\}$	1
$\{DD\}$	2

Ciências ULisboa

Variáveis Aleatórias (cont.)

Classificação de X

Seja $X(\Omega)$ o contradomínio de uma variável aleatória X.

- X é **discreta** se $X(\Omega)$ é finito ou infinito numerável.
- X é **contínua** se $X(\Omega)$ é infinito não numerável.

Ciências ULisboa

Variáveis Aleatórias (cont.)

Exemplos:

- X - “número de carros que passam por dia numa rua”
- $X(\Omega) = \{0, 1, 2, 3, \ldots\} = \mathbb{N}_0$

- X - “tempo de espera entre a passagem de dois carros”
- $X(\Omega) = [0, +\infty[$

Ciências ULisboa

Funções Caracterizadoras das Variáveis Aleatórias

- Variável aleatória **discreta**
  - Função massa de probabilidade (f.m.p): $f_X(x)$;
  - Função de distribuição (f.d.): $F_X(x)$.
- Variável aleatória **contínua**
  - Função densidade de probabilidade (f.d.p.): $f_X(x)$;
  - Função de distribuição (f.d.): $F_X(x)$.

Funções Caracterizadoras das Variáveis Aleatórias (cont.)

Exemplo 1 (cont.) Quais são as probabilidades associadas aos possíveis valores de X no exemplo 1?

- $P(X=0) = P(\{BB\}) = \frac{4}{9} \times \frac{3}{8} = \frac{12}{72}$
- $P(X=1) = P(\{BD, DB\}) = \frac{4}{9} \times \frac{5}{8} + \frac{5}{9} \times \frac{4}{8} = \frac{40}{72}$
- $P(X=2) = P(\{DD\}) = \frac{5}{9} \times \frac{4}{8} = \frac{20}{72}$

Note que $P(X=0) + P(X=1) + P(X=2) = P(\Omega) = 1$

Ciências ULisboa

2. Variáveis Aleatórias Discretas

Variável Aleatória Discreta

Definição (Variável aleatória discreta)

Seja Ω um espaço de resultados. Chama-se **variável aleatória discreta** a uma função

$$X : \Omega \to \mathbb{R}$$

com contradomínio $X(\Omega) = \{x_1, x_2, \ldots, x_n, \ldots\}$, finito ou infinito numerável, tal que

- $P(X = x) \neq 0$ se $x \in \{x_1, x_2, \ldots, x_n, \ldots\}$
- $P(X = x) = 0$ se $x \in \mathbb{R} \setminus \{x_1, x_2, \ldots, x_n, \ldots\}$

Ciências ULisboa

Variáveis Aleatórias Discretas (cont.)

Exemplos:

- Número de chamadas atendidas por um call center em uma hora;
- Número de alunos presentes em uma aula em um determinado dia;
- Número de falhas em um conjunto de componentes eletrônicos durante um teste de qualidade;
- Número de fósseis encontrados por metro cúbico de rocha;
- Número de amostras de solo contaminadas em um determinado local de mineração;
- Número de sismos registados numa região específica durante um ano;
- Número de clientes que visitam uma loja, por dia;
- Número de doentes atendidos na urgência de um hospital, entre as 8:00 e as 18:00.

Variáveis Aleatórias Discretas (cont.)

Definição (Função massa de probabilidade)

Seja X uma variável aleatória discreta tomando valores em $X(\Omega) = \{x_1, x_2, \ldots\}$. Designa-se por **função de probabilidade** ou **função massa de probabilidade (f.m.p.)** uma função f tal que

$$f_X(x) = P(X = x) = \begin{cases} P(X = x_i), & \text{se } x = x_i \in X(\Omega) \\ 0, & \text{se } x \neq x_i \end{cases}$$

e que satisfaz:

1. $f_X(x) \geq 0, \forall x \in \mathbb{R}$
2. $\sum_{x\in\mathbb{R}} f_X(x) = \sum_{x\in X(\Omega)} f_X(x) = 1$

Exemplo 1 (cont.):

Função massa de probabilidade

$$f_X(x) = P(X = x) = \begin{cases} \frac{3}{18}, & \text{se } x = 0 \\ \frac{10}{18}, & \text{se } x = 1 \\ \frac{5}{18}, & \text{se } x = 2 \\ 0, & \text{outros valores de x} \end{cases}$$

x	0	1	2
$P(X = x)$	$\frac{3}{18}$	$\frac{10}{18}$	$\frac{5}{18}$

Variáveis Aleatórias Discretas (cont.)

Definição (Função de Distribuição)

Seja X uma variável aleatória. Define-se **função de distribuição (f.d.)** de uma variável aleatória X à função de valor real $F_X : \mathbb{R} \to [0,1]$ tal que

$$F_X(x) = P(X \leq x), \forall x \in \mathbb{R}$$

Algumas propriedades da função de distribuição

1. $0 \leq F_X(x) \leq 1, \ \forall x \in \mathbb{R}$;
2. $\lim\limits_{x \to -\infty} F_X(x) = 0 \quad \text{e} \quad \lim\limits_{x \to +\infty} F_X(x) = 1$;
3. $F_X(x)$ é uma função contínua à direita:
   $\lim\limits_{x \to a^+} F_X(x) = F_X(a), \ \forall a \in \mathbb{R}$;
4. $F_X(x)$ é uma função monótona não decrescente:
   $\forall x_1, x_2 \in \mathbb{R}, \text{se } x_1 < x_2 \text{ então } F_X(x_1) \leq F_X(x_2)$;
5. $P[X = a] = F_X(a) - F_X(a^-) = F_X(a) - \lim\limits_{x \to a^-} F_X(x), \ \forall a \in \mathbb{R}$;
6. $P[x_1 < X \leq x_2] = F_X(x_2) - F_X(x_1), \ \forall x_1, x_2 \in \mathbb{R}, \text{com } x_1 < x_2$.

Essas propriedades são válidas quer para v.a discretas, quer para v.a. contínuas.

Ciências ULisboa

Variáveis Aleatórias Discretas (cont.)

Definição (Função de Distribuição de uma v.a. discreta)

Seja X uma variável aleatória discreta com função massa de probabilidade $f_X(x)$. A **função de distribuição** de X é dada por

$$F_X(x) = P(X \leq x) = \sum_{s \leq x} f_X(s)$$

Mais algumas propriedades da função de distribuição de uma v.a. discreta. Considere $a < b$.

- $P(X < x) = F_X(x^-)$;
- $P(X > x) = 1 - F_X(x)$;
- $P(X \geq x) = 1 - F_X(x^-)$;
- $P(a < X \leq b) = F_X(b) - F_X(a)$;
- $P(a < X < b) = F_X(b^-) - F_X(a)$;
- $P(a \leq X < b) = F_X(b^-) - F_X(a^-)$;
- $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$.

Ciências ULisboa

Variáveis Aleatórias Discretas (cont.)

Exemplo 1 (cont.) Calcule a função de distribuição da v.a. definida no exemplo 1.

Sabemos que

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i), \quad \forall x \in \mathbb{R}$$

- Se $x < 0$:

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) = 0$$

- Se $x = 0$:

$$F_X(0) = P(X \leq 0) = \sum_{x_i \leq 0} f(x_i) = f(0) = \frac{3}{18}$$

- Se $x = 1$:

$$F_X(1) = P(X \leq 1) = f(0) + f(1) = \frac{3}{18} + \frac{10}{18} = \frac{13}{18}$$

- Se $x = 2$:

$$F_X(2) = P(X \leq 2) = f(0) + f(1) + f(2) = \frac{3}{18} + \frac{10}{18} + \frac{5}{18} = 1$$

- Se $x > 2$:

$$F_X(x) = 1$$

Variáveis Aleatórias Discretas (cont.)

Exemplo 1 (cont.)

Tem-se então

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0, \\ \frac{3}{18}, & \text{se } 0 \leq x < 1, \\ \frac{13}{18}, & \text{se } 1 \leq x < 2, \\ 1, & \text{se } x \geq 2 \end{cases}$$

Variáveis Aleatórias Discretas (cont.)

Exercício 1: Um laboratório de geofísica está analisando uma remessa de 3 amostras de um mineral radioativo. A variável aleatória X representa o número de amostras que apresentam atividade radioativa acima de um certo limiar. A função massa de probabilidade de X é dada pela tabela abaixo:

x	0	1	2	3
$P(X = x)$	0.75	κ	0.08	0.15

(a) Determine o valor de κ.

(b) Determine a probabilidade de não mais de duas amostras apresentarem atividade radioativa acima do limiar.

(c) Determine $P(X = 3|X > 0)$ e $P(1 \leq X \leq 3|X \leq 2)$.

Respostas: a) $\kappa = 0.02$; b) 0.85; c) 0.6; 0.1176471

Variáveis Aleatórias Discretas (cont.)

Exercício 2: A procura diária de uma amostra de mineral é uma v.a. X com a seguinte função de probabilidade:

$$f(x) = \kappa \frac{2^x}{x!}, \quad x = 1, 2, 3, 4$$

(a) Determine o valor de κ.

(b) Determine a função de distribuição de X.

Respostas: $\kappa = \frac{1}{6}$; $F(x) = \begin{cases} 0, & \text{se} \quad x < 1 \\ \frac{1}{3}, & \text{se} \quad 1 \leq x < 2 \\ \frac{2}{3}, & \text{se} \quad 2 \leq x < 3 \\ \frac{8}{9}, & \text{se} \quad 3 \leq x < 4 \\ 1, & \text{se} \quad x \geq 4 \end{cases}$

3. Parâmetros das v.a. Discretas

Parâmetros das v.a. Discretas

- Para descrevermos completamente o comportamento probabilístico de uma v.a. discreta X, é necessário recorrer à função de distribuição ou à função massa de probabilidade.
- No entanto, podemos optar por uma caracterização parcial de X através de algumas medidas populacionais ou parâmetros populacionais (valores fixos), que resumem informações relevantes sobre o seu comportamento.

Veremos os parâmetros de:

- localização central $\rightarrow$ valor esperado, moda e mediana;
- ordem $\rightarrow$ quantil de probabilidade (ou quantil de ordem) p;
- dispersão $\rightarrow$ variância, desvio padrão e coeficiente de variação.

Parâmetros das v.a. Discretas (cont.)

Definição (Valor Esperado)

O **valor esperado** ou **valor médio** ou **média** de uma v.a. discreta X com função massa de probabilidade f_X, representa-se por $E(X)$, μ_X ou μ e é dado por

$$\mu_X = E(X) = \sum_x x\, f_X(x)$$

Observações:

- Quando X discreta toma um número infinito de valores, ou seja, quando o contradomínio de X é um conjunto infinito numerável, $E(X)$ vai existir se a série for absolutamente convergente;
- $E(X)$ é um valor numérico nas mesmas unidades que a v.a. X;
- $E(X)$ não pertence necessariamente ao conjunto de valores possíveis de X.

Parâmetros das v.a. Discretas (cont.)

Valor Esperado de uma função de X

Seja X uma v.a. discreta com função massa de probabilidade f_X e seja $Y = g(X)$ uma função da v.a. X. Então, caso $E(Y)$ exista,

$$E(Y) = E[g(X)] = \sum_x g(x) f_X(x)$$

Observações:

- De modo geral:

$$E[g(X)] \neq g[E(X)]$$

- Por exemplo:

$$E(|X|) \neq |E(X)| \text{ e } E\left(\frac{1}{X}\right) \neq \frac{1}{E(X)}$$

Ciências ULisboa

Parâmetros das v.a. Discretas (cont.)

Propriedades do Valor Esperado

Sendo X e Y duas variáveis aleatórias e κ uma constante real, tem-se

1. $E(\kappa) = \kappa$
2. $E(\kappa X) = \kappa E(X)$
3. $E(X \pm Y) = E(X) \pm E(Y)$
4. $E(XY) = E(X)E(Y)$ se X e Y são independentes

Parâmetros das v.a. Discretas (cont.)

Exemplo 1 (cont.): Calcule o valor esperado da v.a. X do exemplo 1.

- **F.m.p de X**

x	0	1	2
$P(X=x)$	$\frac{3}{18}$	$\frac{10}{18}$	$\frac{5}{18}$

- **Valor esperado de X**

$$E(X) = \sum_{i=1}^{3} x_i \, f_X(x_i) = 0 \cdot \frac{3}{18} + 1 \cdot \frac{10}{18} + 2 \cdot \frac{5}{18} = \frac{20}{18} \approx 1.11 \text{ peças defeituosas}$$

O número médio de peças defeituosas nas 2 extrações é de 1.11 peças.

Note que $E(X)$ não pertence ao conjunto de valores possíveis de X.

- **Como podemos calcular $E(X^2)$?**

$$E(X^2) = \sum_{i=1}^{3} x_i^2 \, f_X(x_i) = 0^2 \cdot \frac{3}{18} + 1^2 \cdot \frac{10}{18} + 2^2 \cdot \frac{5}{18} = \frac{30}{18} \approx 1.66$$

Ciências ULisboa

Parâmetros das v.a. Discretas (cont.)

Moda de uma v.a. discreta

A moda da v.a. discreta X será representada por $mo = mo(X)$ e corresponde ao ponto máximo da f.m.p de X, isto é,

$$mo \in X(\Omega) : P(X = mo) = \max_x P(X = x)$$

ou, equivalentemente,

$$mo = \arg \max_{x \in X(\Omega)} P(X = x)$$

Observação:

A moda existe sempre e nem sempre é única. (*unimodal, bimodal, multimodal*)

A identificação do ponto de máximo da f.m.p passa pela determinação de

$$mo = mo(X) \in X(\Omega) : \begin{cases} P(X = mo) & \geq P(X = mo - 1) \\ P(X = mo) & \geq P(X = mo + 1) \end{cases}$$

Parâmetros das v.a. Discretas (cont.)

Exemplo 1 (cont.) Obtenha a moda da v.a. X do exemplo 1.

- **V.a.**
  $X =$ "número de peças defeituosas nas duas extrações"
- **F.m.p de X**

x	0	1	2
$P(X = x)$	$\frac{3}{18}$	$\frac{10}{18}$	$\frac{5}{18}$

- **Moda de X**
  $mo = mo(X) = 1$ já que $P(X = 1) = \max_x P(X = x)$.
  Logo a v.a. X é unimodal.

Parâmetros das v.a. Discretas (cont.)

Exemplo 2: Admita que o número de partículas minerais detetadas em intervalos de 12 segundos durante a monitorização de uma erupção vulcânica é uma variável aleatória X com função massa de probabilidade (f.m.p.) dada por:

$$P(X = x) = \begin{cases} \frac{e^{-0.8} \cdot 0.8^x}{x!}, & x = 0, 1, 2, \ldots \\ 0, & \text{c.c.} \end{cases}$$

Obtenha o valor esperado e a moda de X.

- **V.a.**
  $X =$ "número de partículas minerais detetadas num intervalo de 12 segundos"
- **F.m.p. de** X

$$P(X = x) = \frac{e^{-0.8} \cdot 0.8^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Parâmetros das v.a. Discretas (cont.)

Exemplo 2 (cont.):

- **Valor esperado de** X

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{+\infty} x \times P(X = x) \\
&= \sum_{x=0}^{+\infty} x \times \frac{e^{-0.8} \times 0.8^{x}}{x!} \\
&= \sum_{x=1}^{+\infty} \frac{e^{-0.8} \times 0.8^{x}}{(x-1)!} \\
&= 0.8 \times e^{-0.8} \sum_{x=0}^{+\infty} \frac{0.8^{x}}{x!} \\
&= 0.8 \times e^{-0.8} e^{0.8} \\
&= 0.8
\end{aligned}
$$

Parâmetros das v.a. Discretas (cont.)

Exemplo 2 (cont.):

- **Moda de X**

$$mo = mo(X) \in X(\Omega) : \begin{cases} P(X = mo) \geq P(X = mo - 1) \\ P(X = mo) \geq P(X = mo + 1) \end{cases}$$

$$\begin{cases} \frac{P(X=mo)}{P(X=mo-1)} \geq 1 \\ \frac{P(X=mo+1)}{P(X=mo)} \leq 1 \end{cases}$$

$$\begin{cases} \frac{\frac{e^{-0.8} \cdot 0.8^{mo}}{mo!}}{\frac{e^{-0.8} \cdot 0.8^{mo-1}}{(mo-1)!}} \geq 1 \\ \frac{\frac{e^{-0.8} \cdot 0.8^{mo+1}}{(mo+1)!}}{\frac{e^{-0.8} \cdot 0.8^{mo}}{mo!}} \leq 1 \end{cases}$$

$$\begin{cases} \frac{0.8}{mo} \geq 1 \\ \frac{0.8}{mo+1} \leq 1 \end{cases}$$

$$\begin{cases} mo \leq 0.8 \\ mo + 1 \geq 0.8 \end{cases}$$

ou seja, $mo = mo(X) = 0$.

Definição (Mediana)

A **mediana** de uma v.a. X, com função de distribuição F_X, é todo e qualquer ponto

$$me = me(X) : P(X \leq me) \geq \frac{1}{2} \text{ e } P(X \geq me) \geq \frac{1}{2}$$

que é equivalente:

$$me : \frac{1}{2} \leq F_X(me) \leq \frac{1}{2} + P(X = me)$$

Parâmetros das v.a. Discretas (cont.)

Exemplo 1 (cont.) Obtenha a mediana da v.a. X do exemplo 1

- **V.a.**
  X = "número de peças defeituosas nas duas extrações"
- **F.d de X**

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0, \\ \frac{3}{18}, & \text{se } 0 \leq x < 1, \\ \frac{13}{18}, & \text{se } 1 \leq x < 2, \\ 1, & \text{se } x \geq 2 \end{cases}$$

- **Mediana de X**

Candidato a me	$\frac{1}{2} \leq F_X(me) \leq \frac{1}{2} + P(X = me)$	Obs.
0	$\frac{1}{2} \leq F_X(0) = \frac{3}{18} \leq \frac{1}{2} + P(X = 0) = \frac{1}{2} + \frac{3}{18} = \frac{13}{8}$	Prop. falsa
0.5	$\frac{1}{2} \leq F_X(0.5) = \frac{3}{18} \leq \frac{1}{2} + P(X = 0.5) = \frac{1}{2} + 0 = \frac{1}{2}$	Prop. falsa
1	$\frac{1}{2} \leq F_X(1) = \frac{13}{18} \leq \frac{1}{2} + P(X = 1) = \frac{1}{2} + \frac{10}{18} = \frac{19}{18}$	Prop. verdadeira
1.5	$\frac{1}{2} \leq F_X(1.5) = \frac{13}{18} \leq \frac{1}{2} + P(X = 1.5) = \frac{1}{2} + 0 = \frac{1}{2}$	Prop. falsa

Portanto, $me = 1$.

Ciências ULisboa

Parâmetros das v.a. Discretas (cont.)

Definição (Quantil de ordem p)

Seja X uma v.a. discreta, com função de distribuição F_X, e $p \in [0, 1]$. O **quantil de ordem p** (ou **quantil de probabilidade p**) da v.a. X é todo número real Q_p tal que

$$Q_p : \quad P(X \leq Q_p) \geq p \quad \text{e} \quad P(X \geq Q_p) \geq 1 - p$$

que é equivalente:

$$Q_p : \quad p \leq F_X(Q_p) \leq p + P(X = Q_p).$$

Q_p também é representado por $F_X^{-1}(p)$.

Quantis especiais (Quartis)

- $Q_{1/4} = F^{-1}(1/4) \rightarrow 1^\circ$ quartil
- $Q_{2/4} = F^{-1}(2/4) \rightarrow 2^\circ$ quartil ou mediana
- $Q_{3/4} = F^{-1}(3/4) \rightarrow 3^\circ$ quartil
- $Q_{1/100} = F^{-1}(1/100) \rightarrow 1^\circ$ percentil

Exemplo 1 (cont.) Obtenha o 3° quartil do número de peças defeituosas nas duas extrações.

- **V.a.**
  X = "número de peças defeituosas nas duas extrações"
- **F.d de X**

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0, \\ \frac{3}{18}, & \text{se } 0 \leq x < 1, \\ \frac{13}{18}, & \text{se } 1 \leq x < 2, \\ 1, & \text{se } x \geq 2 \end{cases}$$

- **Terceiro quartil de X**

Candidato a $Q_{3/4}$	$\frac{3}{4} \leq F_X(Q_{3/4}) \leq \frac{3}{4} + P(X = Q_{3/4})$	Obs.
1	$\frac{3}{4} \leq F_X(1) = \frac{13}{18} \leq \frac{3}{4} + P(X = 1) = \frac{3}{4} + \frac{10}{18} = \frac{47}{36}$	Prop. falsa
2	$\frac{3}{4} \leq F_X(2) = 1 \leq \frac{3}{4} + P(X = 2) = \frac{3}{4} + \frac{5}{18} = \frac{37}{36}$	Prop. verdadeira

Portanto, $Q_{3/4} = 2$.

Parâmetros das v.a. Discretas (cont.)

Definição (Variância de uma v.a. X)

A **variância** de uma v.a. discreta X, com função massa de probabilidade f_X é dada por

$$V(X) = E\left\{[X - E(X)]^2\right\} = \sum_x [x - E(X)]^2 f_X(x),$$

caso $E(X^2) < +\infty$.

Observações:

- $V(X)$ também pode ser representada por σ^2, σ_X^2, $Var(X)$.
- A variância é um valor numérico expresso no quadrado das unidades da v.a. X.
- É um parâmetro de dispersão que indica a variabilidade média da distribuição de probabilidade relativamente ao valor médio.

Parâmetros das v.a. Discretas (cont.)

Definição (Desvio padrão de uma v.a. X)

O **desvio padrão** de X é a raiz quadrada positiva da variância de X

$$\sigma_X = +\sqrt{V(X)}.$$

- O desvio padrão é utilizado para medir a dispersão dos valores de X em torno do seu valor médio;
- Ele é expresso nas unidades da variável.
- É um parâmetro de dispersão que indica a variabilidade média da distribuição de probabilidade relativamente ao valor médio.

Parâmetros das v.a. Discretas (cont.)

Propriedades da variância

Seja X uma variável aleatória e $a, b \in \mathbb{R}$.

1. $V(X) \geq 0$
2. $V(X) = E(X^2) - [E(X)]^2$ (fórmula de König)
3. $V(b) = 0$
4. $V(X + b) = V(X)$
5. $V(aX) = a^2 V(X)$
6. $V(aX + b) = a^2 V(X)$

Ciências ULisboa

Parâmetros das v.a. Discretas (cont.)

Exemplo 1 (cont.) Obtenha a variância e o desvio padrão do número de peças defeituosas nas duas extrações.

x_i	$f_X(x_i)$	$x_i f_X(x_i)$	$x_i^2 f_X(x_i)$
0	$\frac{3}{18}$	0	0
1	$\frac{10}{18}$	$\frac{10}{18}$	$\frac{10}{18}$
2	$\frac{5}{18}$	$\frac{10}{18}$	$\frac{20}{18}$
		$\mu_X = 0 + \frac{10}{18} + \frac{10}{18} = \frac{20}{18}$	$\sigma_X^2 = \frac{30}{18} - \left(\frac{20}{18}\right)^2 \approx 0.4309$

Estamos a usar $V(X) = E(X^2) - [E(X)]^2$ para o cálculo da variância.

O desvio padrão será $\sigma_X = \sqrt{0.4309} \approx 0.6564$.

Parâmetros das v.a. Discretas (cont.)

Exemplo 3: Considerem-se duas v.a. X e Y cujas funções de probabilidade são, respectivamente:

$$X : \begin{array}{c|ccc} x & -1 & 1 & \text{c.c.} \\ \hline f_X(x) & \frac{1}{2} & \frac{1}{2} & 0 \end{array} \qquad Y : \begin{array}{c|ccc} y & -1000 & 1000 & \text{c.c.} \\ \hline f_Y(y) & \frac{1}{2} & \frac{1}{2} & 0 \end{array}$$

Note que

- $E(X) = -1 \times \frac{1}{2} + 1 \times \frac{1}{2} = 0$
- $E(Y) = -1000 \times \frac{1}{2} + 1000 \times \frac{1}{2} = 0$
- $E(X^2) = (-1)^2 \times \frac{1}{2} + 1^2 \times \frac{1}{2} = 1$
- $E(Y^2) = (-1000)^2 \times \frac{1}{2} + 1000^2 \times \frac{1}{2} = 1000^2$

Logo

- $V(X) = 1 - 0^2 = 1$
- $V(Y) = 1000^2 - 0^2 = 1000^2$
- $\sigma_X = 1$
- $\sigma_Y = 1000$

4. Variáveis Aleatórias Contínuas

Variáveis Aleatórias Contínuas

Variável Aleatória Contínua

Seja Ω um espaço de resultados. Chama-se **variável aleatória contínua** a uma função

$$X : \Omega \to \mathbb{R}$$

com contradomínio $X(\Omega)$ infinito não numerável.

Variáveis Aleatórias Contínuas (cont.)

Exemplos:

- o tempo de reparação de uma peça mecânica (em segundos);
- a temperatura (em graus centígrados);
- a concentração de poluentes (em partes por milhão);
- a altura de um indivíduo (em metros);
- o tempo entre avarias de um certo tipo de equipamento (em meses);
- a profundidade em metros de um poço perfurado até atingir o aquífero;
- a concentração de um determinado mineral (como ouro ou ferro) em partes por milhão (ppm) numa amostra de solo ou rocha;
- o tempo necessário para que um mineral específico reaja completamente com um ácido durante um teste de laboratório, medido em minutos.

Variáveis Aleatórias Contínuas (cont.)

Definição (Variável Aleatória Contínua e função densidade de probabilidade)

Seja X uma v.a. e $F_X(x) = P(X \leq x)$, $x \in \mathbb{R}$, a sua função de distribuição. A v.a. X diz-se **contínua** se F_X for absolutamente contínua, isto é, se e somente se existe uma função **não negativa** $f_X : \mathbb{R} \to \mathbb{R}^+$, tal que, para todo $x \in \mathbb{R}$ se tem

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt.$$

A função f_X chama-se **função densidade de probabilidade (f.d.p.)** de X e satisfaz:

1. $f_X(x) \geq 0, \forall x \in \mathbb{R}$;
2. $\int_{-\infty}^{+\infty} f_X(x)\, dx = 1$.

Variáveis Aleatórias Contínuas (cont.)

Função de distribuição

Seja X uma v.a. contínua. A **função de distribuição** da v.a. X, é uma função $F_X : \mathbb{R} \to [0, 1]$ definida por

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt.$$

- Ela representa a probabilidade acumulada até ao ponto x.
- A função densidade de probabilidade $f_X(x) \approx F'_X(x) = \frac{dF_X(x)}{dx}$ mede, na vizinhaça do ponto $X = x$, a concentração de probabilidade por unidade da v.a. X.
  Note que, para $\varepsilon > 0$

$$P\left(x - \frac{\varepsilon}{2} < X < x + \frac{\varepsilon}{2}\right) = \int_{x-\frac{\varepsilon}{2}}^{x+\frac{\varepsilon}{2}} f_X(t)\, dt \approx \varepsilon \times f_X(x)$$

  ou

$$f_X(x) \approx \frac{P\left(x - \frac{\varepsilon}{2} < X < x + \frac{\varepsilon}{2}\right)}{\varepsilon}$$

  para cada intervalo de comprimento pequeno ε. Assim, $\varepsilon \times f_X(x)$ aproxima a probabilidade de a v.a. contínua X pertencer ao intervalo $[x - \frac{\varepsilon}{2}, x + \frac{\varepsilon}{2}]$.

Ciências ULisboa

Variáveis Aleatórias Contínuas (cont.)

Cálculo de probabilidades

- $P(X = x) = P(x \leq X \leq x) = \int_x^x f_X(t)\,dt = 0, \quad \forall x \in \mathbb{R};$
- $P(X > x) = \int_x^{+\infty} f_X(t)\,dt = 1 - P(X \leq x) = 1 - F_X(x) = 1 - \int_{-\infty}^x f_X(t)\,dt;$
- $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx, \quad \forall a, b \in \mathbb{R}.$

Ciências ULisboa

Variáveis Aleatórias Contínuas (cont.)

Exemplo 1: Seja X uma variável aleatória contínua que representa a densidade mineral, em gramas por centímetro cúbico (g/cm³), medida em uma amostra de rocha. A função densidade de probabilidade de X é dada por

$$f(x) = \begin{cases} k, & 1 \leq x \leq 1.8 \\ 0, & \text{caso contrário,} \end{cases}$$

onde k é uma constante desconhecida.

(a) Determine o valor de k de modo a que f seja a $f.d.p.$ da v.a. X.

(b) Qual a probabilidade da densidade mineral na amostra de rocha se situar entre 1.2 e 1.6 g/cm³?

(c) Determine a função de distribuição de X.

(d) Qual a probabilidade da densidade mineral na amostra de rocha ser no máximo 1.2 g/cm³?

(e) Determine $P(X > 1.6 | X > 1.2)$.

Exemplo 1 (cont.):

(a) Para que f seja uma f.d.p ela deve satisfazer:

1. $f(x) \geq 0, \quad \forall x \in \mathbb{R} \Rightarrow k > 0$
2. $\int_{-\infty}^{+\infty} f(x)\, dx = 1$

$$\int_{-\infty}^{+\infty} f(x)\, dx = 1 \Leftrightarrow \int_{1}^{1.8} k\, dx = 1 \Leftrightarrow 0.8k = 1 \Leftrightarrow k = 1.25$$

Portanto a f.d.p é dada por

$$f(x) = \begin{cases} 1.25, & 1 < x < 1.8 \\ 0, & \text{caso contrário.} \end{cases}$$

(b)

$$P(1.2 < X < 1.6) = \int_{1.2}^{1.6} f(x)\, dx = \int_{1.2}^{1.6} 1.25\, dx = [1.25x]_{1.2}^{1.6} = 1.25 \times 0.4 = 0.5$$

Ciências ULisboa

Exemplo 1 (cont.):

(c)

- Para $x < 1$:

$$F(x) = \int_{-\infty}^{x} f(s)\, ds = \int_{-\infty}^{x} 0\, ds = 0$$

- Para $1 \leq x < 1.8$:

$$F(x) = \int_{-\infty}^{x} f(s)\, ds = \int_{-\infty}^{1} f(s)\, ds + \int_{1}^{x} f(s)\, ds = \int_{-\infty}^{1} 0\, ds + \int_{1}^{x} 1.25\, ds$$

$$= 0 + (1.25)\; s\big|_{1}^{x} = (1.25)(x-1)$$

- Para $x \geq 1.8$:

$$F(x) = \int_{-\infty}^{x} f(s)\, ds = \int_{-\infty}^{1} 0\, ds + \int_{1}^{1.8} 1.25\, ds + \int_{1.8}^{x} 0\, ds$$

$$= (1.25)\; s\big|_{1}^{1.8} = (1.25)(1.8-1) = 1.25 \times 0.8 = 1$$

Exemplo 1 (cont.):

(c) Portanto,

$$F(x) = \begin{cases} 0, & \text{se } \ x < 1 \\ (1.25)(x-1), & \text{se } \ 1 \leq x < 1.8 \\ 1, & \text{se } \ x \geq 1.8. \end{cases}$$

(d)

$$P(X \leq 1.2) = F(1.2) = (1.25)(1.2-1) = (1.25)(0.2) = 0.25$$

Variáveis Aleatórias Contínuas (cont.)

Exemplo 1 (cont.):

(e)

$$P(X > 1.6|X > 1.2) = \frac{P(X > 1.6 \wedge X > 1.2)}{P(X > 1.2)} = \frac{P(X > 1.6)}{P(X > 1.2)} =$$
$$= \frac{\int_{1.6}^{+\infty} f(x)\,dx}{\int_{1.2}^{+\infty} f(x)\,dx} = \frac{\int_{1.6}^{1.8} 1.25\,dx}{\int_{1.2}^{1.8} 1.25\,dx} = \frac{0.2}{0.6} =$$
$$= 0.333$$

ou

$$P(X > 1.6|X > 1.2) = \frac{P(X > 1.6 \wedge X > 1.2)}{P(X > 1.2)} = \frac{P(X > 1.6)}{P(X > 1.2)} =$$
$$= \frac{1 - P(X \leq 1.6)}{1 - P(X \leq 1.2)} = \frac{1 - F_X(1.6)}{1 - F_X(1.2)} =$$
$$= \frac{1 - (1.25)(1.6 - 1)}{1 - (1.25)(1.2 - 1)} =$$
$$= 0.333$$

5. Parâmetros das v.a. Contínuas

Parâmetros das v.a. Contínuas

Definição (Valor Esperado)

O **valor esperado** de uma v.a. contínua X, com função densidade de probabilidade f_X é dada por

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

- $E(X)$ existe se o integral convergir absolutamente.

Parâmetros das v.a. Contínuas (cont.)

Valor Esperado de uma função de X

Seja X uma v.a. contínua com função densidade de probabilidade f_X e seja $Y = g(X)$ uma função da v.a. X. Então, caso $E(Y)$ exista,

$$E(Y) = E[g(X)] = \int_{-\infty}^{+\infty} g(x)\, f_X(x)\, dx$$

Parâmetros das v.a. Contínuas (cont.)

Definição (Moda de uma v.a. contínua)

A moda da v.a. contínua X corresponde ao ponto máximo da f.d.p de X, isto é,

$$mo : f_X(mo) = \max_x f_X(x),$$

ou, equivalentemente,

$$mo = \arg\max_x f_X(x).$$

Observação:

A **moda** de uma v.a. contínua obtém-se em geral, identificando

$$mo : \begin{cases} \dfrac{df_X(x)}{dx}\Big|_{x=mo} = 0 & \text{(ponto de estacionaridade)} \\ \dfrac{d^2 f_X(x)}{dx^2}\Big|_{x=mo} < 0 & \text{(ponto de máximo)} \end{cases}$$

Alternativamente, a sua identificação passa pela simples análise do comportamento monótono da f.d.p.

Ciências ULisboa

Parâmetros das v.a. Contínuas (cont.)

Definição (Mediana de uma v.a. contínua)

Representaremos a **mediana** da v.a. X por $me = me(X)$ onde

$$me : F_X(me) = \frac{1}{2}$$

Definição (Quantil de probabilidade p de uma v.a. contínua)

O **quantil de probabilidade** (ou **quantil de ordem**) **p** $(0 < p < 1)$ da v.a. contínua X define-se como

$$Q_p : F_X(Q_p) = p.$$

Parâmetros das v.a. Contínuas (cont.)

Definição (Variância de uma v.a. contínua)

A **variância** da v.a. contínua X é igual a

$$V(X) = E\{[X - E(x)]^2\} = \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(x)\, dx$$

caso exista $E(X)$ e o integral seja absolutamente convergente.

Observação:

$$\begin{aligned} V(X) &= E(X^2) - E^2(X) \\ &= \int_{-\infty}^{+\infty} x^2\, f_X(x)\, dx - \left(\int_{-\infty}^{+\infty} x\, f_X(x)\, dx\right)^2 \end{aligned}$$

Definição (Desvio padrão de uma v.a. contínua)

O **desvio padrão** da v.a. X é a raiz quadrada positiva da variância de X

$$\sigma_X = +\sqrt{V(X)}.$$

Parâmetros das v.a. Contínuas (cont.)

Exemplo 1: Em uma pesquisa geológica, deseja-se estudar a distribuição do teor de ouro (em gramas por tonelada) encontrado em amostras de solo de uma mina. O teor de ouro em uma amostra é uma variável aleatória contínua X com função de densidade de probabilidade (f.d.p.) dada por:

$$f(x) = \begin{cases} \frac{3}{8}x^2, & 0 \leq x \leq 2 \\ 0, & \text{caso contrário} \end{cases}$$

(a) Calcule o valor esperado do teor de ouro nas amostras de solo.

(b) Calcule a variância do teor de ouro nas amostras de solo.

(c) Calcule o desvio padrão do teor de ouro nas amostras de solo.

(d) Determine a moda do teor de ouro nas amostras de solo.

(e) Determine a mediana do teor de ouro nas amostras de solo.

(f) Determine o terceiro quartil do teor de ouro nas amostras de solo.

(a),(b),(c)

$$E(X) = \int_{-\infty}^{+\infty} x\, f(x)\, dx = \int_0^2 x \left(\frac{3}{8}x^2\right)\, dx = \frac{3}{8}\int_0^2 x^3\, dx = \frac{3}{8} \times \left[\frac{x^4}{4}\right]_0^2$$
$$= \frac{3}{8} \times \frac{16}{4} = \frac{3}{2}$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2\, f(x)\, dx = \int_0^2 x^2 \left(\frac{3}{8}x^2\right)\, dx = \frac{3}{8}\int_0^2 x^4\, dx = \frac{3}{8} \times \left[\frac{x^5}{5}\right]_0^2$$
$$= \frac{3}{8} \times \frac{32}{5} = \frac{12}{5}$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{12}{5} - \left(\frac{3}{2}\right)^2 = \frac{3}{28}$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{\frac{3}{28}} = 0.32732$$

Parâmetros das v.a. Contínuas (cont.)

(d) Como a f.d.p f é crescente dentro do intervalo $[0,2]$, a moda ocorre no valor $x = 2$.

(e) A função de distribuição de X é dada por

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{x^3}{8}, & 0 \leq x \leq 2 \\ 1, & x > 2 \end{cases}$$

Assim,

$$\begin{aligned} me = me(X) \in (0,2): \quad F_X(me) &= \frac{1}{2} \\ \frac{(me)^3}{8} &= \frac{1}{2} \\ (me)^3 &= 4 \\ me &\approx 1.5874 \end{aligned}$$

Calculamos que $\frac{1}{2}$ do teor de ouro encontrado em amostras de solo não excederão 1.5874 gramas por tonelada e a outra metade excederá 1.5874 gramas por tonelada .

(f) Queremos encontrar

$$
\begin{aligned}
Q_{3/4}: \quad F_X(Q_{3/4}) &= \frac{3}{4} \\
\frac{(Q_{3/4})^3}{8} &= \frac{3}{4} \\
(Q_{3/4})^3 &= 6 \\
Q_{3/4} &\approx 1.8171
\end{aligned}
$$

Note que $P(X \leq Q_{3/4}) = \frac{3}{4}$ e $P(X \geq Q_{3/4}) = \frac{1}{4}$. Ou seja, calculamos que $\frac{3}{4}$ (resp. $\frac{1}{4}$) do teor de ouro encontrado em amostras de solo não excederão (resp. excederá) 1.8171 gramas por tonelada.

Referências

- Manuel Cabral Morais (2020): Probabilidade e Estatística: Teoria, Exemplos & Exercícios. IST Press, 1a edição.