MAT02025 - Amostragem 1

Conceitos básicos de probabilidade e inferência estatística: uma revisão

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Departamento de Estatística

Porto Alegre, 2023

Probabilidade

Introdução

- **Fenômeno aleatório:** situação ou acontecimento cujos resultados não podem ser previstos com certeza.
- **Fenômeno aleatório** *versus* **Fenômeno determinístico**.
- Exemplos de fenômenos aleatórios:
  - Uma partida de futebol.
  - Eleições para presidente.
  - O preço do combustível no próximo mês.
  - Vida (duração) da bateria de um dispositivo móvel.

Espaço amostral e eventos

- **Espaço amostral:** é o conjunto de **todos os resultados possíveis** de um certo fenômeno aleatório.
  - Ele será representado pela letra grega Ω **(ômega)**.

Espaço amostral e eventos

- **Eventos:** Os subconjuntos de Ω são denominados **eventos** e representamos pelas letras latinas maiúsculas $A, B, \ldots$.
  - O subconjunto vazio será denotado por $\emptyset$.
  - Dizemos que um evento **ocorre** quando um dos resultados que o compõem ocorre.

Probabilidade

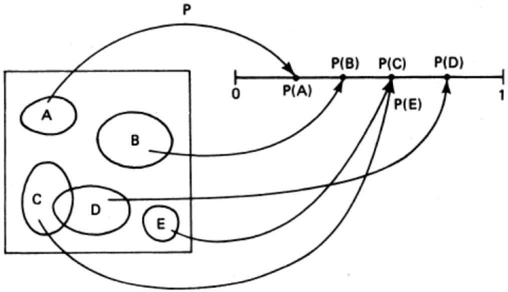
- Vamos considerar **probabilidade** como sendo uma função $\Pr(\cdot)$ que atribui valores numéricos aos eventos do espaço amostral, conforme a definição a seguir.

Probabilidade

Uma função $\Pr(\cdot)$ se satisfaz as condições:

- **(A1)** $0 \leq \Pr(A) \leq 1$, para qualquer evento $A \subset \Omega$;
- **(A2)** $\Pr(\Omega) = 1$;
- **(A3)** $\Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \Pr(A_i)$, com $A_i's$ disjuntos.
  - $\Pr(A \cup B) = \Pr(A) + \Pr(B)$, com A e B disjuntos.

Probabilidade



Variáveis aleatórias

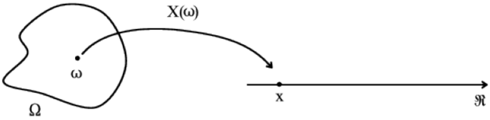
Variável aleatória

Definição

Sejam $\mathcal{E}$ um experimento e Ω um espaço amostral associado ao experimento. Uma função X, que associe a cada elemento $\omega \in \Omega$ um número real $X(\omega)$, é denominada variável aleatória.

- **Exemplo:** Suponha-se que uma lâmpada tenha sido posta em um soquete. O experimento termina quando a lâmpada se queima.
    - Qual será um possível resultado, ω?
    - Qual será espaço amostral consequente?
    - Qual será a variável aleatória X de interesse?
    - Quais serão os possíveis valores de X?
    - "Onde está a aleatoriedade de X"?

Variável aleatória



Variável aleatória discreta

Definição

Uma quantidade X, associada a cada possível resultado do espaço amostral, é denominada de **variável aleatória discreta**, se assume valores num conjunto enumerável, com certa probabilidade.

- **Exemplo:** Observa-se o sexo das crianças em famílias com três filhos. Denotamos m para o sexo masculino e f para o sexo feminino. Existem oito possibilidades para uma família de três filhos. Estas possibilidades são listadas no espaço amostral:

$$\Omega = \{(mmm), (mmf), (mfm), (fmm), (mff), (fmf), (ffm), (fff)\}.$$

Variável aleatória discreta

Definimos

- X : número de crianças do sexo masculino (m).
- A cada possível resultado do espaço amostral, X associa um valor numérico

Ω	mmm	mmf	mfm	fmm	mff	fmf	ffm	fff
X	3	2	2	2	1	1	1	0

- Note que X assume valores no conjunto $\{0, 1, 2, 3\}$ que é enumerável, portanto X é variável aleatória discreta.

Variável aleatória discreta

- **Pergunta:** com qual probabilidade X assume os valores $\{0, 1, 2, 3\}$?
- Cada resultado possível do espaço amostral tem probabilidade $\frac{1}{8}$ de acontecer, então:

$$\Pr(X = 0) = \Pr(fff) = \frac{1}{8}$$

- A probabilidade da variável aleatória X assumir o valor zero é a mesma probabilidade do evento (fff) ocorrer. Da mesma forma:

$$\begin{aligned}
\Pr(X = 1) &= \\
&= \\
&= \\
\Pr(X = 2) &= \qquad\qquad\qquad = \\
\Pr(X = 3) &= \qquad\qquad =
\end{aligned}$$

Variável aleatória discreta

Função discreta de probabilidade

- **Resumindo:**

X	0	1	2	3
$\Pr(X = x_i)$	$1/8$	$3/8$	$3/8$	$1/8$

Distribuição de probabilidade

Variável aleatória discreta

Função discreta de probabilidade

Definição

Seja X uma variável aleatória discreta e $x_1, x_2, x_3, \ldots,$ seus diferentes valores. A função que atribui a cada valor da variável aleatória sua probabilidade é denominada de **função discreta de probabilidade** ou, simplesmente **função de probabilidade**.

Variável aleatória discreta

Função discreta de probabilidade

- A notação a ser utilizada é:

$$\Pr(X = x_i) = p(x_i) = p_i, i = 1, 2, \ldots$$

ou ainda,

X	x_1	x_2	x_3	$\ldots$
p_i	p_1	p_2	p_3	$\ldots$

- Uma função de probabilidade satisfaz:
  1. $0 \leq p_i \leq 1$
  2. $\sum_i p_i = p_1 + p_2 + \ldots = 1$
- **No exemplo:**

$$\begin{aligned} p_1 + p_2 + p_3 + p_4 &= \Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2) + \Pr(X = 3) \\ &= 1/8 + 3/8 + 3/8 + 1/8 = 1. \end{aligned}$$

Variável aleatória discreta

Exemplo: lançamento de dois dados

- Dois dados são lançados, de forma independente.

Variável aleatória discreta

Exemplo: lançamento de dois dados

- O espaço amostral deste experimento é:

$$\Omega = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\}$$

- Seja X a variável *soma dos dois lançamentos*, ou seja,
  $X =$ "face do primeiro lançamento" $+$ "face do segundo lançamento" .

Variável aleatória discreta

Exemplo: lançamento de dois dados

- Quando o evento $(1,1)$ ocorre, X associa a este resultado o valor 2. Da mesma forma temos:

$$\Omega = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\} \overset{X}{\Rightarrow} \left\{ \begin{array}{cccccc} 2, & 3, & 4, & 5, & 6, & 7, \\ 3, & 4, & 5, & 6, & 7, & 8, \\ 4, & 5, & 6, & 7, & 8, & 9, \\ 5, & 6, & 7, & 8, & 9, & 10, \\ 6, & 7, & 8, & 9, & 10, & 11, \\ 7, & 8, & 9, & 10, & 11, & 12 \end{array} \right\}$$

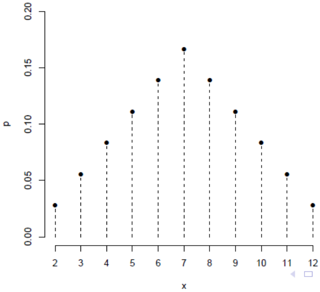
- X assume valores no conjunto $\{2,3,4,5,6,7,8,9,10,11,12\}$.
- A probabilidade dos possíveis resultados em Ω é
  $P[(1,1)] = P[(1,2)] = \ldots = P[(2,1)] = P[(2,2)] = \ldots = P[(6,6)] = 1/36$.
- $P[X=2] = P[(1,1)] = 1/36$,
  $P[X=3] = P[(1,2) \cup (2,1)] = 1/36 + 1/36 = 2/36$.

Variável aleatória discreta

Exemplo: lançamento de dois dados

- A função de probabilidade de X é dada por:

X	2	3	4	5	6	7	8	9	10	11	12
p_i	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36



Variável aleatória discreta

Exemplo: lançamento de dois dados

- **Observações:**
  - p_i pertence ao intervalo $(0, 1)$ para $i = 1, \ldots, 11$ (ex: $p_1 = 1/36 \in (0, 1)$)
  - $\sum_{i=1}^{11} p_i = p_1 + p_2 + \ldots + p_{11} = 1/36 + 2/36 + \ldots + 1/36 = 1$
  - A função de probabilidade de X satisfaz às condições 1 e 2, logo é de fato uma função de probabilidade.
- **Pergunta:** Qual a probabilidade da soma dos resultados dos dois lançamentos ser menor do que 6?

$$
\begin{aligned}
\Pr(X < 6) &= \Pr(X = 2) + \Pr(X = 3) + \Pr(X = 4) + \Pr(X = 5) \\
&= p_1 + p_2 + p_3 + p_4 \\
&= 1/36 + 2/36 + 3/36 + 4/36 = 10/36.
\end{aligned}
$$

Valor esperado, variância e desvio padrão

Média

- A **média**, **valor esperado** ou **esperança** de uma variável X é dada pela expressão:

$$\mathrm{E}\left(X\right) = \sum_{i=1}^{k} x_i p_i$$

- Uma notação alternativa é representar $\mathrm{E}\left(X\right)$ por μ_X ou simplismente μ.

Exemplo

- Considere a variável aleatória X com a seguinte função discreta de probabilidade:

X	-5	10	15	20
p_i	0,3	0,2	0,4	0,1

Temos,

$$\mu = \sum_i x_i p_i = (-5) \times 0,3 + 10 \times 0,2 + 15 \times 0,4 + 20 \times 0,1 = 8,5.$$

Variância

- Seja X uma variável aleatória com $\Pr(X = x_i) = p_i, i = 1, 2, \ldots, k$ e média μ. A **variância** de X é a ponderação pelas respectivas probabilidades, dos desvios relativos à média, elevados ao quadrado, isto é,

$$\mathrm{Var}\,(X) = \sum_{i=1}^{k} (x_i - \mu)^2 p_i.$$

- Muitas vezes, denotamos a variância por σ^2 e, se houver possibilidade de confusão, usamos σ_X^2.
  - Extraindo a raiz quadrada da variância obtemos o **desvio-padrão** que é representado por σ ou σ_X

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{k} (x_i - \mu)^2 p_i}.$$

Variância

- Note, que pela definição da variância, concluímos que a variância é o valor esperado de uma nova variável, o desvio quadrado. Isto é,

$$\mathrm{Var}\,(X) = \mathrm{E}\,[(X-\mu)^2],$$

a qual pode ser convenientemente reescrita na seguinte forma

$$\mathrm{Var}\,(X) = \mathrm{E}\,(X^2) - \mu^2 = \sum_{i=1}^{k} x_i^2 p_i - \mu^2.$$

- Esta última expressão é bastante útil e, para não criar confusão, explicitamos os seus termos.
  - O termo $\mathrm{E}\,(X^2)$ é o valor esperado da variável aleatória X^2;
  - $\mu^2 = (\mathrm{E}\,[X])^2$ indica o quadrado do valor esperado de X.

Exemplo

- Considere a variável aleatória X com a seguinte função discreta de probabilidade:

X	-5	10	15	20
p_i	0,3	0,2	0,4	0,1

Temos,

$$\mathrm{E}\,(X^2) = \sum_i x_i^2 p_i = (-5)^2 \times 0,3 + 10^2 \times 0,2 + 15^2 \times 0,4 + 20^2 \times 0,1 = 157,5.$$

$$Var(X) = 157,5 - 8,5^2 = 157,5 - 72,5 = 85,25.$$

Propriedades

- Seja X uma variável aleatória e

$$Y = aX + b,$$

então

$$\mathrm{E}\left(Y\right) = a\mathrm{E}\left(X\right) + b,$$

e

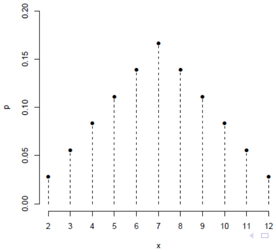
$$\mathrm{Var}\left(Y\right) = a^2\mathrm{Var}\left(X\right).$$

Exercício

Considere a seguinte função discreta de probabilidade:

- A função de probabilidade de X é dada por:

X	2	3	4	5	6	7	8	9	10	11	12
p_i	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$



- Calcule a média, a variância e o desvio padrão desta variável.

Para casa

- Revisar os tópicos discutidos nesta aula.
  - Ler os capítulos 5, 6, 10 e 11 do Livro "Estatística Básica"[1] (disponível no Sabi+).

[1] Morettin, P. A. e Bussab, W. O. **Estatística Básica**, Saraiva, 2010.

Próxima aula

- Revisão de conceitos básicos de **probabilidade e estatística** (continuação).

Por hoje é só!

Bons estudos!