# Métodos Estatísticos - LEI/LEICE/LEIRP - 2022-2023

Deolinda M. L. Dias Rasteiro

Capítulo II, cont.

# Conteúdos

# Distribuição de Bernoulli

Seja $A$ um acontecimento associado a uma determinada experiência aleatória, tal que $P(A) = p$, com $p \in ]0, 1[$.
Considere-se a variável aleatória $X$ definida do modo seguinte

$$X = \begin{cases} 1 & \text{se} \quad \text{A ocorre} \\ 0 & \text{se} \quad \text{A não ocorre} \end{cases}$$

A variável assim definida é uma v.a. discreta de **suporte** $S = \{0, 1\}$ e **função de probabilidade**

$$p(x) = \begin{cases} 1-p & \text{se} \quad x = 0 \\ p & \text{se} \quad x = 1 \\ 0 & \text{se} \quad x \in \overline{S} \end{cases} = \begin{cases} p^x \, (1-p)^{1-x} & \text{se} \quad x \in \{0, 1\} \\ 0 & \text{se} \quad x \in \overline{S} \end{cases}$$

- A função (lei) de probabilidade assim definida é denominada por **lei de Bernoulli de parâmetro p**, denotada por $\mathcal{B}(\mathbf{p})$ (ou **Bernoulli**(**p**)).
- Diz-se que a v.a. **X segue uma lei de Bernoulli** (ou **X tem distribuição de Bernoulli**) **de parâmetro p**, e escreve-se simbolicamente $\mathbf{X} \sim \mathcal{B}(\mathbf{p})$.

## Nota

A distribuição de Bernoulli está associada a experiências aleatórias com apenas dois resultados possíveis, usualmente denominados por **sucesso** e **insucesso**, com probabilidade de ocorrer sucesso $p$ (o seu parâmetro!) e insucesso $q = 1 - p$. Uma experiência aleatória assim caracterizada chama-se **prova (ou tentativa) de Bernoulli**.

## Propriedades

Se $X \sim \mathcal{B}(p)$ então

$$E(X) = p$$
$$V(X) = p(1-p)$$

# Distribuição Binomial

Seja $A$ um acontecimento associado a uma determinada experiência aleatória, tal que $P(A) = p$, com $p \in ]0, 1[$.
Suponha agora que esta mesma experiência é realizada $n$ vezes, com $n > 1$, sempre nas mesmas condições (note que os resultados das sucessivas experiências são independentes e logo $P(A) = p$ em qualquer realização).
Considere-se a variável aleatória
**X = "número de vezes que ocorre o acontecimento A, nas n realizações da experiência aleatória"**.

# Observações

- O suporte de $X$ é $S = \{0, 1, ..., n\}$
- Um acontecimento favorável a que ocorram exatamente $x$ vezes o acontecimento $A$, **e consequentemente** $n - x$ vezes $\overline{A}$, é $B = \underbrace{A \cap ... \cap A}_{x \ vezes} \cap \underbrace{\overline{A} \cap ... \cap \overline{A}}_{n-x \ vezes}$ e a sua probabilidade é

$$P(B) = \underbrace{P(A)...P(A)}_{x \ vezes} \underbrace{P(\overline{A})...P(\overline{A})}_{n-x \ vezes}$$

$$= p^x (1-p)^{n-x}$$

# Observações

- O número de maneiras distintas de ocorrer $x$ vezes o acontecimento $A$, nas $n$ realizações, é
  $C_x^n = \begin{pmatrix} n \\ x \end{pmatrix} = \dfrac{n!}{(n-x)!x!}$
- A probabilidade de ocorrer exatamente $x$ vezes o acontecimento $A$, nas $n$ realizações, é dada por

$$P(X = x) = \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x}, \; x \in \{0, 1, ..., n\}.$$

$X$ é uma v.a. discreta de **suporte** $S = \{0, 1, ..., n\}$ e **função de probabilidade** dada por

$$p(x) = P(X = x) = \begin{cases} \left(\begin{array}{c} n \\ x \end{array}\right) p^x (1-p)^{n-x} & \text{se} \quad x \in \{0, 1, ..., n\} \\ 0 & \text{se} \quad x \in \overline{S} \end{cases}$$

- $P(X = x) > 0, \forall x \in \{0, 1, ..., n\}$
- $\displaystyle\sum_{x=0}^{n} P(X = x) = 1$

Diz-se que a v.a. **X segue uma lei Binomial** (ou $X$ **tem distribuição Binomial**) **de parâmetros** $n$ e $p$, e escreve-se simbolicamente $X \sim \mathcal{B}(n, p)$.

Diz-se que a v.a. $X$ **segue uma lei Binomial** (ou $X$ **tem distribuição Binomial**) **de parâmetros** $n$ e $p$, e escreve-se simbolicamente $X \sim \mathcal{B}(n, p)$.

## Observação

A lei Binomial descreve a contagem do número de *sucessos* em $n$ repetições, independentes, de uma prova de Bernoulli com probabilidade de sucesso $p$. Note que se $X_i$ representar o número de sucessos (0 ou 1) obtidos na prova $i$, então $X_i \sim \mathcal{B}(p)$, $i = 1, 2, ..., n$, e o número de sucessos obtidos nas $n$ provas é a v.a.

$$X = \sum_{i=1}^{n} X_i \sim \mathcal{B}(n, p).$$

Uma das aplicações mais interessantes da lei Binomial é a contagem do número de elementos de determinado tipo numa amostra de dimensão $n$, quando selecionados com reposição.

# Nota

A família de distribuições Binomial depende apenas dos parâmetros $n$ e $p$. Esta está implementada na maioria das máquinas de calcular, assim como tabelada nos livros da área. Em particular, nas aulas de ME podem (ou devem, em falta de alternativa) ser consultadas as tabelas da unidade curricular onde constam as probabilidades

$$F(x) = P(X \leq x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}, \, x \in \{0, 1, ..., n\}$$

# Distribuição Binomial

## Exemplos:

- Considere um lote com 5 bolas, das quais 2 são brancas e 3 são vermelhas. Extração sucessiva, com reposição, de 3 bolas do lote.
  $X =$ "número de **bolas brancas** nas 3 extraídas"
  1)Qual é a probabilidade de saírem 2 bolas brancas?
  2)Qual é a probabilidade de saírem, no máximo, 2 bolas brancas?

# Distribuição Binomial

**Exemplos:**

- Um avião comercial tem 4 motores independentes e num voo a probabilidade de cada motor funcionar, sem avarias, é de 0.9. Qual a probabilidade do avião fazer uma viagem segura se para isso precisar de pelo menos 2 motores a funcionar corretamente?
  $X =$ "número de motores a funcionar sem avarias, em 4"
  $Y = 4 - X =$ "número de motores com avarias, em 4"

# Distribuição Hipergeométrica

Suponha que se tem um número finito de $N$ objetos, dos quais $M$ são de um tipo (*sucesso*) e os restantes $N - M$ de outro tipo (*insucesso*). Considere a experiência aleatória: extração, ao acaso, de um objeto. Suponha agora que esta mesma experiência é realizada $n$ vezes, isto é, extração sucessiva, e sem reposição, de $n$ dos $N$ objetos (note que os resultados das sucessivas experiências não são independentes).
Seja

**X = "número de objetos do tipo sucesso obtidos nas n extrações"**

# Distribuição Hipergeométrica

## Observações

- $x \leq n$ e $x \leq M \iff x \leq min(n, M)$

  $n - x \leq n$ e $n - x \leq N - M \iff x \geq 0$ e $x \geq n - (N - M)$

  $\iff x \geq max(0, n - (N - M))$

  Então
  $$S = \{\max(0, n - (N - M)), 1, \ldots, \min(n, M)\}$$

- O número de casos favoráveis à saída de $x$ sucessos, nas $n$ extrações, é $\left(\begin{array}{c} M \\ x \end{array}\right)\left(\begin{array}{c} N-M \\ n-x \end{array}\right)$

- O número de casos possíveis nas $n$ extrações, é $\left(\begin{array}{c} N \\ n \end{array}\right)$

- A probabilidade de obter exatamente $x$ sucessos, nas $n$ extrações, é dada por

$$P(X = x) = \frac{\left(\begin{array}{c} M \\ x \end{array}\right)\left(\begin{array}{c} N-M \\ n-x \end{array}\right)}{\left(\begin{array}{c} N \\ n \end{array}\right)}, \quad \text{para } x \in S$$

# Distribuição Hipergeométrica

$X$ é então uma v.a. discreta de Suporte $S = \{\max(0, n-(N-M)), 1, \ldots, \min(n, M)\}$ e função de probabilidade dada por

$$p(x) = P(X = x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & \text{se} \quad x \in S \\ 0 & \text{se} \quad x \in \overline{S} \end{cases} .$$

Diz-se que a v.a. **X segue a lei Hipergeométrica** (ou **X tem distribuição Hipergeométrica**) **de parâmetros n**, **N** e **M**, e escreve-se simbolicamente $\mathbf{X} \sim \mathcal{H}(\mathbf{n}, \mathbf{N}, \mathbf{M})$.

# Distribuição Hipergeométrica

## Propriedades

- Se $X \sim \mathcal{H}(n, N, M)$ então

$$E(X) = n\frac{M}{N}$$

$$V(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

# Distribuição Hipergeométrica

## Notas

A distribuição Hipergeométrica é usada na contagem do número de sucessos que ocorrem em $n$ realizações de uma experiência aleatória, cada uma com apenas dois resultados possíveis (sucesso ou insucesso)onde a probabilidade do sucesso é alterada em cada realização.

Uma das aplicações mais interessantes destas distribuições é na área da **Amostragem**, onde uma amostra de dimensão $n$ é selecionada numa população de dimensão $N$. Uma nova leitura dos parâmetros da Hipergeométrica, $(n, N, M)$, é

$n$: dimensão da amostra; $N$: dimensão da população (finita); $M$: número de sucessos na população

Sob certas condições, a distribuição $\mathcal{H}(n, N, M)$ pode ser aproximada pela distribuição $\mathcal{B}(n, p)$, com $p = \frac{M}{N}$.

# Distribuição Hipergeométrica

Na prática,

$$\text{se} \quad \frac{n}{N} \leq 0.1 \quad \text{então} \quad \mathcal{H}(n, N, M) \,\dot{\sim}\, \mathcal{B}\left(n, \tfrac{M}{N}\right)$$

$$\downarrow$$

"*aproximadamente (no limite)*"

Intuitivamente, se a dimensão da amostra, $n$, é muito pequena relativamente à dimensão da população, $N$, a seleção, embora seja feita sem reposição, não vai alterar "significativamente" a probabilidade de ocorrer um sucesso.
(Mais à frente voltaremos a esta aproximação !!)

# Distribuição de Poisson

Seja $\lambda$ um número real positivo. Diz-se que uma v.a. $X$ segue a lei de Poisson de parâmetro $\lambda$, e escreve-se simbolicamente $X \sim \mathcal{P}(\lambda)$, se $X$ for discreta de suporte $S = \mathbb{N}_0$, com função de probabilidade

$$p(x) = P(X = x) = \begin{cases} e^{-\lambda}\dfrac{\lambda^x}{x!} & \text{se} \quad x \in \mathbb{N}_0 \\ 0 & \text{se} \quad x \in \overline{S} \end{cases}.$$

- $P(X = x) > 0, \forall x \in \mathbb{N}_0$
- $\displaystyle\sum_{x=0}^{+\infty} P(X = x) = 1$

# Distribuição de Poisson

## Propriedades

- $E(X) = V(X) = \lambda$
- Sejam $X_1, X_2, ...X_k$ v.a.'s independentes tais que $X_i \sim \mathcal{P}(\lambda_i)$. Então
$$Y = \sum_{i=1}^{k} X_i \sim \mathcal{P}(\lambda_1 + \lambda_2 + ... + \lambda_k)$$

# Distribuição de Poisson

## Notas

Concluir que uma certa v.a. segue a lei de Poisson (ao contrário das leis Binomial e Hipergeométrica) não é imediato, e sai fora do âmbito do nosso curso. No entanto, esta distribuição é aplicada na contagem de eventos independentes que ocorrem durante um dado intervalo de tempo ou numa dada região espacial. Esta distribuição também é conhecida pela distribuição dos acontecimentos raros, pois a probabilidade de ocorrer mais do que um evento num intervalo muito pequeno é nula.

Prova-se que a distribuição $\mathcal{B}(n,p)$, com $p$ muito pequeno e $n$ suficientemente grande, pode ser aproximada pela distribuição $\mathcal{P}(\lambda)$, com $\lambda = np$. Na prática,

$$\text{se} \quad p \leq 0.1 \quad \text{então} \quad \mathcal{B}(n,p) \dot{\sim} \mathcal{P}(np)$$

# Distribuição de Poisson

**Exemplo:**

- Um material radioativo emite um certo tipo de partículas a uma taxa média de duas por milissegundo, e segue uma distribuição de Poisson.
  Qual é a probabilidade de:
  serem emitidas 3 partículas num milissegundo?
  serem emitidas 5 partículas em dois milissegundos?
  serem emitidas pelo menos 3 partículas em dois milissegundos?