Probabilidade e inferência estatística com aplicações em R

Aula 2 - Distribuições discretas

Felipe Barletta

Departamento de Estatística

15 setembro, 2020

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

Sumário

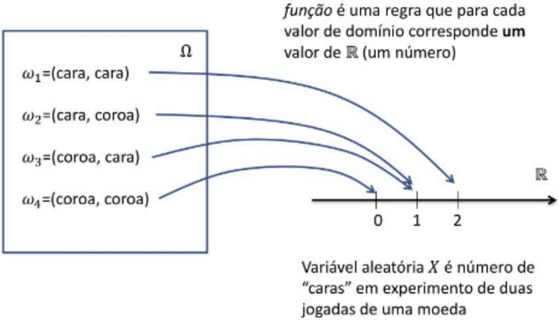
Variável aleatória

- **Variável aleatória** - Descrição numérica do resultado de um fenômeno aleatório.
- Em probabilidade, uma função X que associa a cada evento do espaço amostral um número real $X(\omega) \in \mathbb{R}$, é denominada uma variável aleatória (V.A.).
- Dessa forma, podemos atribuir probabilidades aos eventos desse espaço amostral, dando origem ao conceito de variável aleatória

Notação

- X denota a variável aleatória.
- x denota os valores realizados da V.A.
- Probabilidade de X assumir o valor x é denotada $P[X = x]$.

Variável aleatória



Distribuições de probabilidade

- Há diversos modelos probabilísticos
- Distribuições de probabilidade de V.A. (discretas ou contínuas)
- Descrição das probabilidades associadas com os possíveis valores de X
  - **Variáveis discretas** $\Rightarrow$ suporte em um conjunto de valores enumeráveis (finitos ou infinitos)
  - **Variáveis contínuas** $\Rightarrow$ suporte em um conjunto não enumerável de valores

Distribuições de probabilidade

Definição: distribuição discreta

O conjunto de todos os valores possíveis de uma variável discreta, com suas probabilidades associadas, é chamada de **distribuição de probabilidade discreta**, ou seja,

$$P[X = x_i] = p(x_i) = p_i, \quad i = 1, 2, \ldots.$$

i. A probabilidade de cada valor deve estar entre 0 e 1

$$0 \leq p(x_i) \leq 1, \quad i = 1, 2, \ldots.$$

ii. A soma de todas as probabilidades é igual a 1

$$\sum_i p(x_i) = 1.$$

Exemplo

Para obter um modelo para falhas de hardware em um sistema de computador, o número de acidentes ocorridos a cada semana foram observados durante um período de um ano.

Falhas(X)	Frequência (f_i)	Frequência relativa (fr_i)
0	9	9/52
1	14	14/52
2	13	13/52
3	9	9/52
4	4	4/52
5	2	2/52
6	1	1/52

UTFPR UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

Função de distribuição (ou Função de distribuição acumulada)

- Em muitas situações, é útil calcularmos a probabilidade **acumulada** até um certo valor.
- Definimos a **função de distribuição** ou **função acumulada de probablidade** de uma V.A. X pela expressão:

$$F(x) = P[X \leq x]$$

para qualquer número real x.

Modelo Uniforme discreto

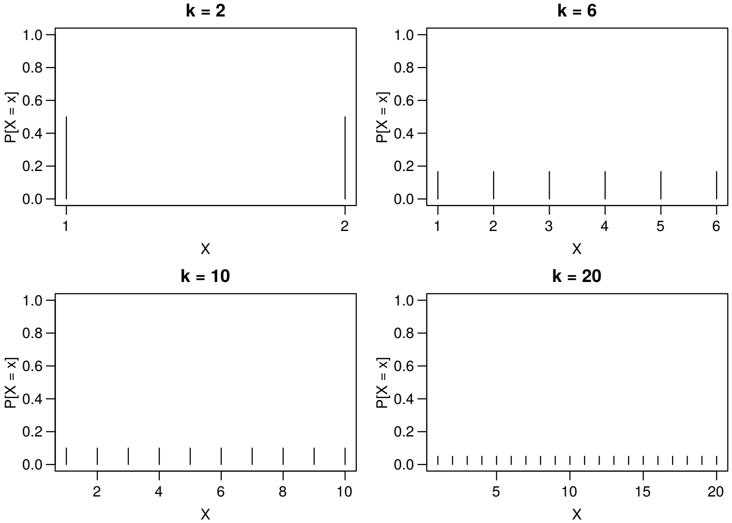
Definição:

Seja X uma v.a assumindo valores $1, 2, \ldots, k$. Dizemos que X segue o modelo **Uniforme Discreto** se atribui a mesma probabilidade $1/k$ a cada um desses k valores.
Então, sua função de probabilidade é dada por

$$P[X = j] = \frac{1}{k}, \quad j = 1, 2, \ldots, k.$$

Notação: $X \sim \mathsf{U}_D[1, k]$

Modelo Uniforme Discreto



Distribuições de probabilidade discreta

- Para resolução dos exercícios vamos explorar as seguintes funcionalidades do R para operações com distribuição de probabilidade:
  - *d:* calcula a densidade de probabilidade $p(x)$ no ponto
  - *p:* calcula a função de probabilidade acumulada $F(x)$ no ponto
  - *q:* calcula o quantil correspondente a uma dada probabilidade
  - *r:* gera uma amostra aleatória da distribuição

Modelo Bernoulli

Definição:

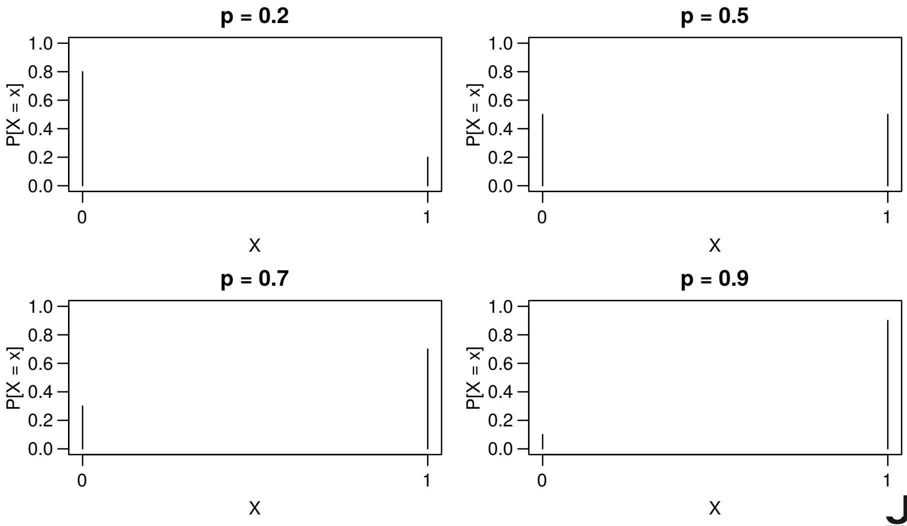
Uma variável aleatória X segue o modelo Bernoulli se assume apenas os valores 0 ("fracasso") ou 1 ("sucesso"). Sua função de probabilidade é dada por

$$P[X = x] = p^x (1 - p)^{1-x}, \qquad x = 0, 1$$

onde o parâmetro $0 \leq p \leq 1$ é a probabilidade de sucesso.

Notação: $X \sim \text{Ber}(p)$

Modelo Bernoulli



Modelo Binomial

Definição:

Seja um experimento realizado dentro das seguintes condições:

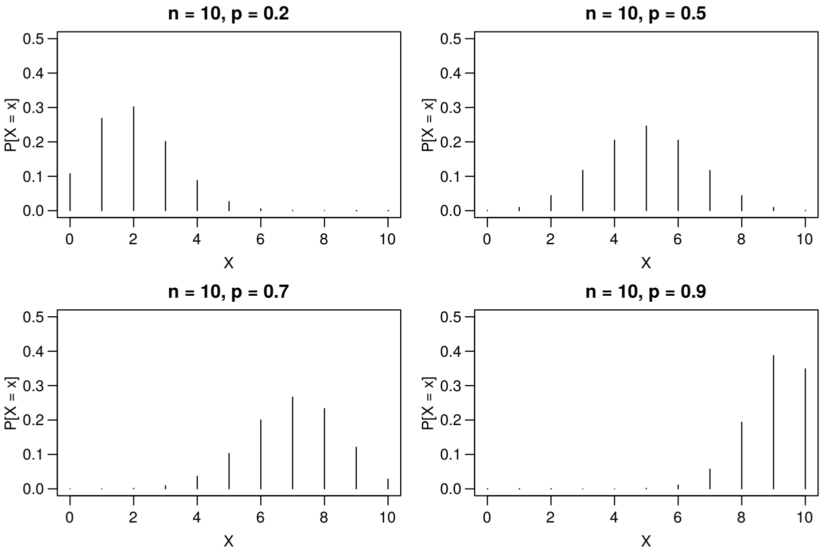
i. São realizados n "ensaios" de Bernoulli independentes.
ii. Cada ensaio só pode ter dois resultados possíveis: "sucesso" ou "fracasso".
iii. A probabilidade p de sucesso em cada ensaio é constante.

Vamos associar a v.a X o número de sucessos em n ensaios de Bernoulli. Portanto X poderá assumir os valores $0, 1, \ldots, n$.

$$P[X = x] = \binom{n}{x} p^x (1-p)^{n-x}$$

Notação: $X \sim B(n, p)$

Modelo Binomial



Modelo Geométrico

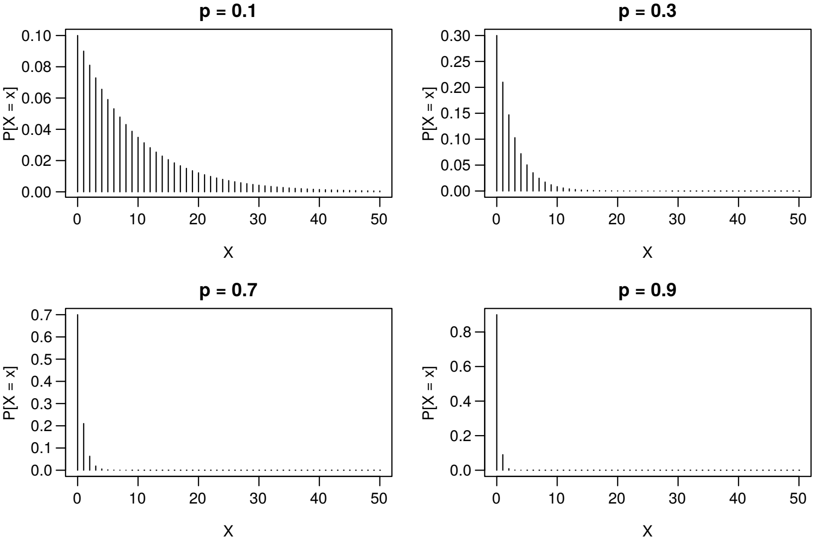
Definição:

Considere o número (k) de ensaios Bernoulli **que precedem o primeiro sucesso**. Nesse caso, dizemos que a v.a X tem distribuição Geométrica de parâmetro p, e sua função de probabilidade tem a forma

$$P[X = k] = p(1-p)^k, \quad k = 0, 1, 2, \ldots.$$

onde $0 \leq p \leq 1$ é a probabilidade de sucesso.
Notação: $X \sim \text{G}(p)$.

Modelo Geométrico



Distribuição Poisson

Definição:

Seja um experimento realizado nas seguintes condições:

i. As ocorrências são independentes.
ii. As ocorrências são aleatórias.
iii. A variável aleatória X é o número de ocorrências de um evento **ao longo de algum intervalo** (de tempo ou espaço).

Denominamos esse experimento de **processo de Poisson**.

$$P[X = k] = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

onde o parâmetro $\lambda > 0$ é a taxa média de ocorrências em um intervalo de tempo ou espaço.
Notação: $X \sim P(\lambda)$.

Modelo Poisson