Métodos Estatísticos Básicos

Aula 8 - Variáveis aleatórias

Regis A. Ely

Departamento de Economia
Universidade Federal de Pelotas

25 de agosto de 2020

Conteúdo

Definições

- **Variável aleatória**: dado um experimento e um espaço amostral Ω, uma variável aleatória é uma função X, que associa um número real $X(\omega)$ a cada elemento $\omega \in \Omega$
- **Há duas interpretações de variável aleatória**:
    1. Realizamos um experimento, que resulta em $\omega \in \Omega$, e a seguir calculamos o número $X(\omega)$
    2. O número $X(\omega)$ é pensado como o próprio resultado do experimento, e a imagem de $X(\omega)$, denotada R_X, torna-se o espaço amostral

Lembre da definição de uma função:

- $\forall\ \omega \in \Omega, \exists\ y \in \mathbb{R}$ tal que $X(\omega) = y$
- $\forall\ y, z \in \mathbb{R}$ com $X(\omega) = y$ e $X(\omega) = z$, temos $y = z$

Exemplos de variáveis aleatórias

- *Ex 1*: No experimento de lançar duas moedas e observar os resultados, temos $\Omega = (H, H), (H, T), (T, H), (T, T)$. Podemos definir a variável aleatória X como sendo o número de caras obtidas, de modo que $X(H, H) = 2$, $X(H, T) = X(T, H) = 1$ e $X(T, T) = 0$. Note que ao aplicar a função X alteramos a observação do experimento
- *Ex 2*: Considere o experimento de lançar 3 moedas e observar a descrição detalhada de como e onde as moedas pousaram. Poderíamos avaliar:
  - $X(\omega) =$ nº de caras que aparecem
  - $Y(\omega) =$ distância máxima entre 2 moedas quaisquer
  - $Z(\omega) =$ distância mínima entre as moedas e a borda da mesa

Exemplos de variáveis aleatórias

- Podemos incluir a avaliação de $X(\omega)$ na descrição do nosso experimento, de modo que $R_X = \{0, 1, 2\}$ (ex. 1) é o nosso novo espaço amostral
- Podemos também relacionar certos eventos $A \subseteq \Omega$ a eventos de R_X. Seja $B \subseteq R_X$, podemos definir A como $A = \{\omega \in \Omega | X(\omega) \in B\}$. Dizemos então que A e B são equivalentes

Variáveis aleatórias discretas

- Uma *variável aleatória discreta* possui o conjunto imagem $X(\Omega)$ finito ou infinito enumerável
- A função de probabilidade de uma variável aleatória discreta X é uma função que associa para cada resultado $x_1, x_2, ... \in X$, um número $p(x_i) = P(X = x_i)$, tal que:
  1. $p(x_i) \geq 0$ para todo i
  2. $\sum_{i=1}^{\infty} p(x_i) = 1$
- Chamamos p de probabilidade, e a coleção de pares $[x_i, p(x_i)]$ para $i = 1, 2, ...$ de distribuição de probabilidade de X

Variáveis aleatórias discretas

- Seja $B \subseteq X(\Omega)$ tal que $B = \{x_{i1}, x_{i2}, ...\}$, então $P(B) = P[\omega | X(\omega) \in B] = P[\omega | X(\omega) = x_{ij}, j = 1, 2, ...] = \sum_{j=1}^{\infty} p(x_{ij})$
- Ou seja, a probabilidade de um evento B é igual a soma das probabilidades dos resultados individuais associados a B

Distribuição de Bernoulli

- Considere um evento A associado a um experimento e defina $P(A) = p$ e $P(\bar{A}) = 1 - p$. Agora considere a seguinte variável aleatória: $X = 0, \text{ se } \omega \notin A$ (fracasso), ou $X = 1, \text{ se } \omega \in A$ (sucesso)
- Qual a função de probabilidade desta variável aleatória?
- **Distribuição de Bernoulli**: $P(X = k) = p^k(1 - p)^{1-k}$ para $k \in 0, 1$

Distribuição Binomial

- Considere n repetições independentes do experimento de Bernoulli, sendo que $P(A)$ permanece a mesma para todas as repetições
- O espaço amostral deste novo experimento será formado por todas as sequências possíveis $\{a_1, a_2, ..., a_n\}$, onde cada a_i pertence a A ou $\bar{A}$
- A variável aleatória $X =$ nº de elementos favoráveis à A (número de sucessos), terá valores possíveis que vão de 0 até n. Já o número total de formas de se obter k sucessos em n repetições do experimento é $\binom{n}{k}$. Assim, a distribuição de X será:
- **Distribuição binomial**: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ para $k = 0, 1, ..., n$

Exemplo de distribuição binomial

Qual a probabilidade de obtermos menos de 3 caras em 5 lançamentos de uma moeda justa?

$$P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$P(X < 3) = \binom{5}{0}(1/2)^0(1/2)^5 + \binom{5}{1}(1/2)^1(1/2)^4 + \binom{5}{2}(1/2)^2(1/2)^3$$

$$P(X < 3) = 1/32 + 5(1/32) + 10(1/32) = 1/2$$

Exemplo no R

No `R` podemos calcular a probabilidade acima com a função `dbinom`:

```
dbinom(2, size=5, prob=0.5) +
  dbinom(1, size=5, prob=0.5) +
  dbinom(0, size=5, prob=0.5)
```

```
[1] 0.5
```

Variáveis aleatórias contínuas

- **Variável aleatória contínua**: quando a imagem da variável aleatória X gera um conjunto infinito não-enumerável de valores
  - Neste caso substituímos a probabilidade p, definida somente para $x_1, x_2, ...,$ por uma função f, definida para todos os valores de x
- **Função densidade de probabilidade**: X é uma variável aleatória contínua se existir uma função f, denominada função densidade de probabilidade (fdp) de X, que satisfaça:
  1. $f(x) \geq 0$ para todo x
  2. $\int_{-\infty}^{+\infty} f(x)dx = 1$
  3. Para quaisquer a, b com $-\infty < a < b < \infty$, teremos $P(a \leq x \leq b) = \int_a^b f(x)dx$

Distribuição uniforme

- **Distribuição uniforme**: variável aleatória contínua X com valores no intervalo $[a, b]$, sendo a e b finitos
- Se X for uniformemente distribuída, então terá fdp dada por:

$$f(x) = \begin{cases} \frac{1}{b-a} \text{ se } a \leq x \leq b \\ 0 \text{ caso contrário} \end{cases}$$

Exemplo de distribuição uniforme

- *Ex*: Um ponto é escolhido ao acaso no segmento de reta $[0, 2]$. Qual a probabilidade de que o ponto esteja entre 1 e $3/2$?

$f(x) = \frac{1}{2}$ para $0 < x < 2$. Logo,
$P(1 \leq x \leq 3/2) = \int_1^{3/2} \frac{1}{2} = \frac{1}{2}.\frac{3}{2} - \frac{1}{2}.1$ Assim, $P(1 \leq x \leq 3/2) = \frac{1}{4}$

Exemplo no R

Podemos calcular o exemplo anterior no `R` utilizando a função `punif`:

```
punif(3/2, min=0, max=2)-punif(1, min=0, max=2)
```

```
[1] 0.25
```

Este código calcula $P(X \leq 3/2) - P(X \leq 1)$ utilizando a chamada função de distribuição acumulada da distribuição uniforme

Função de distribuição acumulada

- A **Função de Distribuição Acumulada (fd)** de uma variável aleatória discreta ou contínua X é definida como $F(x) = P(X \leq x)$
  1. Se X for uma variável aleatória discreta, $F(x) = \sum_j p(x_j)$ para todo j tal que $x_j \leq x$
  2. Se X for uma variável aleatória contínua com fdp f, $F(x) = \int_{-\infty}^{x} f(s)ds$

Exemplo de distribuição acumulada

Ex: Seja X uma variável aleatória contínua com fdp $f(x) = 2x$ para $0 < x < 1$ e igual a zero para quaisquer outros valores valores de x. Nesse caso, a função de distribuição acumulada será dada por:

$$F(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ \int_0^x 2sds = x^2 & \text{se } 0 < x \leq 1 \\ 1 & \text{se } x > 1 \end{cases}$$

Distribuição uniforme acumulada

A função de distribuição acumulada de uma variável aleatória contínua uniformemente distribuída será:

$$F(x) = \begin{cases} 0 \text{ se } x < a \\ \frac{x-a}{b-a} \text{ se } a \leq x < b \\ 1 \text{ se } x > b \end{cases}$$

Exemplo no R

Podemos calcular a probabilidade de obtermos menos de 3 caras em 5 lançamentos de uma moeda justa através da função de distribuição binomial acumulada utilizando a função `pbinom` no `R`:

```
pbinom(2, size=5, prob=0.5)
```

```
[1] 0.5
```

Propriedades da função de distribuição acumulada

- A função F é não-decrescente, ou seja, se $x_1 \leq x_2$, teremos $F(x_1) \leq F(x_2)$
- $\underset{x \to -\infty}{lim} F(x) = 0$ e $\underset{x \to \infty}{lim} F(x) = 1$
- $f(x) = \frac{dF(x)}{dx}$ para todo X no qual F é derivável
- Se X é variável aleatória discreta com valores $x_1, x_2, \ldots$ tais que $x_1 < x_2 < \ldots$; então $p(x_i) = p(X = x_i) = F(x_i) - F(x_{i-1})$

Exemplo de propriedades da distribuição acumulada

Suponha que uma variável aleatória contínua tenha *fd* dada por:

$$F(x) = \begin{cases} 0 \text{ se } x \leq 0 \\ 1 - e^{-x} \text{ se } x > 0 \end{cases}$$

Nesse caso, $F'(x) = e^{-x}$ para $x > 0$, e a fdp será $f(x) = e^{-x}$ para $x > 0$, e zero para quaisquer outros valores