Variáveis Aleatórias Discretas

Prof. José Roberto Silva dos Santos

Depto. de Estatística e Matemática Aplicada - UFC

Fortaleza, 11 de maio de 2022

Sumário

Sumário

Motivação

- Para descrever um experimento aleatório é conveniente associar valores numéricos aos seus resultados. Isso permite um melhor tratamento matemático e o desenvolvimento de modelos gerais.
- Alguns experimentos possuem resultados que já são descritos numericamente; por exemplo, se observamos o número de chamadas telefônicas que chegam a uma central, o tempo de duração de uma lâmpada, o número de pontos que aparecem na face superior de um dado, etc.
- No entanto, em um grande número de situações, os pontos do espaço amostral são níveis de uma variável qualitativa e não valores numéricos. Por exemplo, quando retiramos peças em uma linha de produção os possíveis resultados são: *peça boa* ou *peça defeituosa*.

- Notação $C = cara$, $K = coroa$.
- $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$
- $\omega_1 = (C, C)$; $\omega_2 = (C, K)$; $\omega_3 = (K, C)$; $\omega_4 = (K, K)$
- Defina X: O número de coroas observadas.

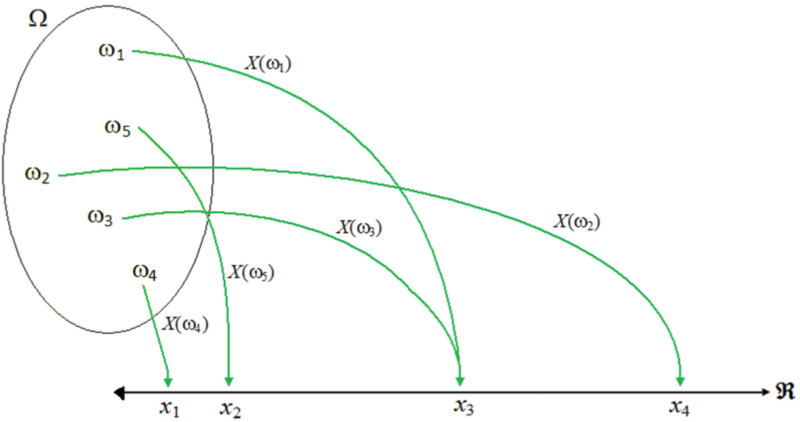
$$\begin{aligned} \omega_1 \to X &= 0 \\ \omega_2 \,\text{ou}\, \omega_3 \to X &= 1 \\ \omega_4 \to X &= 2 \end{aligned}$$

Definição

Sejam ε um experimento aleatório e Ω o espaço amostral associado a este. Uma *função* X, que associe a cada elemento $\omega \in \Omega$ um número real, $X(\omega)$, é denominada variável aleatória (v.a).

Notação:

- Letras maiúsculas (variável aleatória), X, Y, Z, etc.
- Letras minúsculas, (valor assumido pela variável aleatória), x, y, z etc.


Ω
ω1
ω5
ω2
ω3
ω4
X(ω1)
X(ω2)
X(ω3)
X(ω5)
X(ω4)
x1
x2
x3
x4
ℜ

Definição

As variáveis aleatórias que assumem valores em um conjunto enumerável serão denominadas *discretas* e aquelas que assumem valores em um intervalo da reta real serão denominadas *contínuas*.

(a) Número de peças defeituosas entre n peças retiradas de uma linha de produção.
$X = \{x|x \in \{0, 1, 2, \ldots, D\}\}$ onde D é o total de peças defeituosas

(b) Número de veículos que passam por um posto de pedágio durante uma hora.
$X = \{x|x \in \{0, 1, 2, \ldots\}\}$

(c) Sexo de um indivíduo selecionado ao acaso de uma população
$X = \{x|x = 0 \text{ ou } x = 1\}$

(d) Tempo de duração de um componente de um circuito.
$X = \{x|x \in [0, \infty)\}$

(e) Peso de animais sujeitos a uma dieta de engorda.
$X = \{x|x \in [0, \infty)\}$

Sumário

1. Introdução
2. Distribuição de Probabilidade de Variáveis Aleatórias
   - Distribuição de probabilidade
3. Função de Distribuição Acumulada

Definição

Seja X uma variável aleatória discreta. Portanto, X assume no máximo um número infinito enumerável de valores $x_1, x_2, \ldots$. A cada possível resultado x_i associaremos um número $p(x_i) = \mathbb{P}(X = x_i)$, denominado probabilidade de x_i satisfazendo as seguintes condições:

(a) $p(x_i) \geq 0$ para todo i,

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

A função $p(x_i)$ é chamada *função de probabilidade* e a coleção de pares $[x_i, p(x_i)], \quad i = 1, 2, \ldots,$ é denominada *distribuição de probabilidade* de X.

- Considere uma urna com duas bolas brancas e três bolas vermelhas. Dessa urna, são extraídas, sem reposição, duas bolas. Seja X: "número de bolas vermelhas obtidas nas duas extrações". Determine a distribuição de probabilidade de X.

- Exemplo: Número de filhos - Com dados do último censo, a assistente social de um centro de saúde constatou que para as famílias de uma determinada região:
  - 20% não têm filhos
  - 30% têm 1 filho
  - 35% têm 2 filhos
  - 15% têm igualmente 3, 4 ou 5 filhos

- Vamos definir a variável aleatória N = número de filhos por família.
- Suponha que uma família será selecionada aleatoriamente nessa região e o número de filhos averiguado, queremos estudar a v.a N.
- N pode assumir valores em $\{0, 1, 2, 3, 4, 5\}$
  - $\mathbb{P}(N = 0) = 0,2$
  - $\mathbb{P}(N = 1) = 0,3$
  - $\mathbb{P}(N = 2) = 0,35$
  - $\mathbb{P}(N = 3) = \mathbb{P}(N = 4) = \mathbb{P}(N = 5) = \frac{0,15}{3} = 0,05$

- Finalmente, determinamos a distribuição de probabilidade de N:

N	0	1	2	3	4	5
$p(n) = \mathbb{P}(N = n)$	0,20	0,30	0,35	0,05	0,05	0,05

- Tal que $\sum_{n=1}^{6} p_n = 1$

- Considere o experimento que consiste em lançar um dado sucessivamente e observar a face de cima até que face 6 apareça pela primeira vez. Seja X a variável aleatória número de lançamentos até observar face 6.

(a) Determine a distribuição de probabilidade de X.

(b) Calcule a probabilidade do experimento terminar após um número par de repetições.

Sumário

Função de distribuição acumulada

- O conceito de função de distribuição acumulada (ou distribuição cumulativa de probabilidade) que introduziremos aplica-se tanto a variáveis aleatórias discretas quanto a variáveis aleatórias contínuas.
- A função de distribuição acumulada nos dá uma outra maneira de descrever como as probabilidades são associadas aos valores de uma variável aleatória.

Continuidade da Probabilidade

- Sejam $A_1, A_2, \ldots$ uma sequência de eventos. Por $A_n \uparrow A$, denotamos que

$$A_1 \subset A_2 \subset A_3 \subset \cdots \quad \text{e} \quad A = \bigcup_{n=1}^{\infty} A_n$$

Assim, $A_n \uparrow A$ significa que a ocorrência de A_n implica a ocorrência de A_{n+1} para todo n e A é o evento de que pelo menos um dos A_n's ocorre.

- Por $A_n \downarrow A$, denotamos que

$$A_1 \supset A_2 \supset A_3 \supset \cdots \quad \text{e} \quad A = \bigcap_{n=1}^{\infty} A_n$$

Dessa forma, $A_n \downarrow A$ significa que a ocorrência de A_{n+1} implica a ocorrência de A_n para todo n e A é o evento de que todos os A_n's ocorrem.

Continuidade da Probabilidade

Teorema:

- **Continuidade à esquerda:** Se $A_n \uparrow A$, então $\mathbb{P}(A_n) \uparrow \mathbb{P}(A)$ quando $n \to \infty$
- **Continuidade à direita:** Se $A_n \downarrow A$, então $\mathbb{P}(A_n) \downarrow \mathbb{P}(A)$ quando $n \to \infty$

Função de distribuição acumulada

Definição

A função de distribuição acumulada (f.d.a) de uma variável aleatória (discreta ou contínua) X é definida como

$$F(x) = \mathbb{P}(X \leq x) \text{ para todo } x$$

Propriedades

A função de distribuição acumulada satisfaz as seguintes condições:

(F1) *Monotonicidade.* Se $x \leq y$, então $F(x) \leq F(y)$.

(F2) *Continuidade à direita.* Se $x_n \downarrow x$, então $F(x_n) \downarrow F(x)$.

(F3) $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$ e $F(\infty) = \lim_{x\to\infty} F(x) = 1$.

Função de distribuição acumulada

Teorema

- Se X for uma variável aleatória discreta

$$F(x) = \sum_{\{j : x_j \leq x\}} p(x_j).$$

- Por outro lado,

$$p(x_i) = F(x_i) - F(x_i^-)$$

em que $F(x_i^-)$ é o limite de $F(x)$ com x tendendo a x_i pela esquerda (isto é, por valores inferiores à x_i).

- Um grupo de 1000 crianças foi analisado para determinar a efetividade de uma vacina contra um tipo de alergia. As crianças recebiam uma dose de vacina e após um mês passavam por um novo teste. Caso ainda tivessem tido alguma reação alérgica, recebiam outra dose. Ao fim de 5 doses, foram consideradas imunizadas.
- Variável de interesse: X = número de doses

Doses (X)	1	2	3	4	5
Freq.	245	288	256	145	66

- Uma criança é sorteada ao acaso, qual será a probabilidade dela ter recebido 2 doses?
  $\mathbb{P}\left(X = 2\right) = \frac{288}{1000} = 0,288$

Doses (X)	1	2	3	4	5
$p(x)$	0,245	0,288	0,256	0,145	0,066

- Qual a probabilidade da criança ter recebido até duas doses?
  $F(2) = \mathbb{P}\left(X \leq 2\right) = \mathbb{P}\left(X = 1\right) + \mathbb{P}\left(X = 2\right) = 0,245 + 0,288 = 0,533$

- Notemos que a f.d.a. de X = número de doses é definido para qualquer valor real, logo:

$$F(x) = \begin{cases} 0; & x < 1 \\ 0,245; & 1 \leq x < 2 \\ 0,533; & 2 \leq x < 3 \\ 0,789; & 3 \leq x < 4 \\ 0,934; & 4 \leq x < 5 \\ 1; & x \geq 5 \end{cases}$$

- f.d.a. de X = número de doses: