Aula 6: Variáveis aleatórias

Disciplina: PGE950 - Probabilidade

Programa de Pós-graduação em Estatística da UFPE

Prof. Pablo M. Rodriguez

http://www.pablo-rodriguez.org

Universidade Federal de Pernambuco
CCEN
Centro de Ciências Exatas e da Natureza

Conteúdo desta aula

- Distribuição de uma variável aleatória
- Distribuições discretas

Distribuição de uma variável aleatória

Proposição 6.1

Seja X variável aleatória em $(\Omega, \mathcal{F}, P)$ e $\mathcal{B}$ σ-álgebra de Borel. Então,

$$\{X \in B\} \in \mathcal{F} \text{ para todo } B \in \mathcal{B}.$$

Ideia da prova da Prop. 6.1:

Verificamos para B um intervalo. Lembre que $\{X \leq x\} \in \mathcal{F}$, $x \in \mathbb{R}$.

- $B = (-\infty, b]$:
$$\{X \in B\} = \{X \geq b\} \in \mathcal{F}.$$
- $B = (a, \infty)$:
$$\{X \in B\} = \{X \leq a\}^c \in \mathcal{F}.$$
- $B = (a, b]$:

$$\{X \in B\} = \{a < X \leq b\} = \{X \leq b\} - \{X \leq a\} \in \mathcal{F}.$$

Observação!

Usamos a seguinte notação: $A - B := A \cap B^c$.

... continuação da ideia da prova da Prop. 6.1:

- $B = (a, b)$: Como $B = \bigcup_{n=1}^{\infty} \left(a, b - \frac{1}{n}\right]$, então

$$\{X \in B\} = \bigcup_{n=1}^{\infty} \left\{a < X \leq b - \frac{1}{n}\right\} \in \mathcal{F}.$$

- Da mesma forma, $\{X \in B\} \in \mathcal{F}$ para todo intervalo e para todo

$$B = \bigcup_{i=1}^{n} B_i,$$

em que $B_1, \ldots, B_n$ são intervalos disjuntos.

Observação!

$P(X \in B)$ é determinada pela função de distribuição de X.

Distribuição de uma variável aleatória

A probabilidade P_X definida em $\mathcal{B}$, a σ-álgebra de Borel, por

$$P_X(B) := P(X \in B)$$

é chamada distribuição de X.

Proposição 6.2

Seja X uma variável aleatória.

1. *Se X é discreta com valores $\{x_1, x_2, \dots\}$ então*

$$P_X(B) = \sum_{i:x_i \in B} P(X = x_i) = \sum_{i:x_i \in B} p(x_i), \text{ para todo } B \in \mathcal{B}.$$

2. *Se X é absolutamente contínua então*

$$P_X(B) = \int_B f(x)dx, \text{ para todo } B \in \mathcal{B}.$$

... Ideia da prova da Prop. 6.2: Vamos verificar o caso contínuo.

Se $B = (a, b)$, por exemplo, então

$$P_X(B) = P(a < X < b) = P(a < X \leq b) = F(b) - F(a)$$

mas

$$F(b) - F(a) = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = \int_{a}^{b} f(x)dx = \int_{B} f(x)dx.$$

Da forma análoga, verifica-se para qualquer intervalo. Agora, se

$$B = \bigcup_{i=1}^{n} B_i \text{ com } B_1, \ldots, B_n \text{ intervalos disjuntos, então}$$

$$P_X(B) = \sum_{i=1}^{n} P_X(B_i) = \sum_{i=1}^{n} \int_{B_i} f(x)dx = \int_{\cup_i B_i} f(x)dx.$$

Lembrete

Uma variável aleatória X é discreta se $X(\omega) \in \mathcal{S} \subset \mathbb{R}$ para todo $\omega \in \Omega$, para algum conjunto $\mathcal{S}$ finito ou enumerável. Neste caso

$$p(i) := P(X = i), i \in \mathcal{S},$$

é a função de probabilidade de X.

Distribuição Bernoulli

X tem *distribuição de Bernoulli* com parâmetro $p \in (0,1)$ se:

$$p(1) = p = 1 - p(0).$$

Notação: $X \sim Bernoulli(p)$.

Observação!

Um experimento com apenas dois resultados possíveis, sucesso ou fracasso, chama-se ensaio de Bernoulli.

Exemplo 6.1

Seja X a variável aleatória indicadora do evento A com $P(A) = p$:

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

Então $X \sim Bernoulli(p)$.

Distribuição Binomial

X tem *distribuição Binomial* com parâmetros $n \in \mathbb{N}$ e $p \in (0,1)$ se:

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}, \text{ para } i \in \{0, 1, \ldots, n\}.$$

Notação: $X \sim B(n,p)$.

Observação!

X pode ser interpretada como o número de sucessos obtidos quando n ensaios independentes de Bernoulli, com probabilidade de sucesso p, são realizados.

Note que, para $i \in \{0, 1, \ldots, n\}$, por exemplo a configuração

$$\sigma : \underset{1}{\underline{\checkmark}} \; \underset{2}{\underline{\times}} \; \underset{3}{\underline{\checkmark}} \; \underset{4}{\underline{\times}} \; \underset{5}{\underline{\checkmark}} \; \cdots\cdots \; \underset{n-3}{\underline{\times}} \; \underset{n-2}{\underline{\checkmark}} \; \underset{n-1}{\underline{\times}} \; \underset{n}{\underline{\checkmark}}$$

i sucessos (✔)

$n - i$ fracassos (✖)

é favorável para a ocorrência de $\{X = i\}$. Como temos $\binom{n}{i}$ formas diferentes de obter uma configuração destas e como

$$P(\sigma) = p^i (1 - p)^{n-i},$$

por independência!

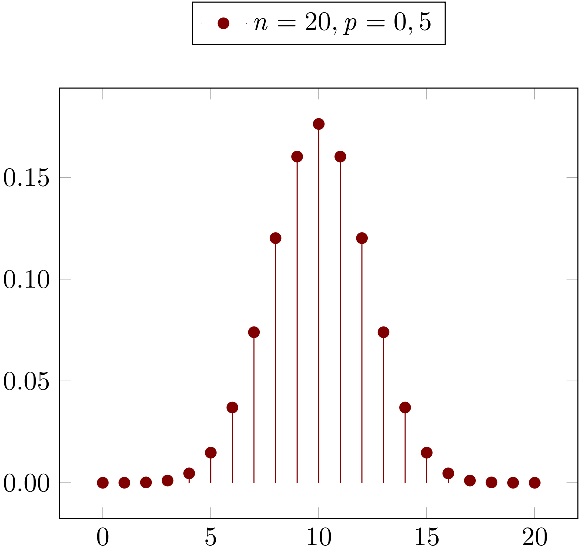
concluímos que

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}.$$

Universidade Federal de Pernambuco

CCEN Centro de Ciências Exatas e da Natureza

$p(i)$



Universidade Federal de Pernambuco
CCEN
Centro de Ciências Exatas e da Natureza

Se $X \sim B(n,p)$ podemos escrever

$$X = \sum_{i=1}^{n} X_i,$$

em que $X_1, \ldots, X_n$ são i.i.d. com $X_1 \sim Bernoulli(p)$.

Grafo aleatório de Erdős-Rényi

Considere n vértices e assuma que cada par de vértices é conectado, independentemente dos outros pares, com probabilidade p. O modelo resultante chama-se grafo aleatório de Erdős-Rényi e denota-se por $G(n, p)$. Se $i \in \{1, \ldots, n\}$ é um vértice do grafo e D_i denota o número de vértices conectados com i em $G(n, p)$ (i.e., D_i é o grau de i) então $D_i \sim B(n-1, p)$.

De fato, note que, fixando i, se

$$X_{ij} = \begin{cases} 1, & \text{se } j \text{ está conectado com } i, \\ 0, & \text{se } j \text{ não está conectado com } i, \end{cases}$$

então as variáveis aleatórias X_{ij}, $j \in \{1, \ldots, n\} \setminus \{i\}$, são i.i.d. com $X_{ij} \sim Bernoulli(p)$. Além disto:

$$D_i = \sum_{j \neq i} X_{ij}.$$

Distribuição de Poisson

X tem *distribuição de Poisson* com parâmetro $\lambda \in (0, \infty)$ se:

$$p(i) = \frac{e^{-\lambda}\lambda^i}{i!}, \text{ para } i \in \{0, 1, \ldots\}.$$

Notação: $X \sim Poisson(\lambda)$.

Observação!

X pode ser interpretada como uma aproximação para o número de sucessos obtidos quando n ensaios independentes de Bernoulli, com probabilidade de sucesso p, são realizados, assumindo valores grandes de n e pequenos de p.

Aproximação da Binomial para a Poisson

Seja $X \sim B(n,p)$, n grande, e seja $Y \sim Poisson(\lambda)$ com $\lambda = np$.

$$P(X=i) = \binom{n}{i} p^i (1-p)^{n-i} = \frac{n!}{(n-i)!i!}\left(\frac{\lambda}{n}\right)^i \left(1-\frac{\lambda}{n}\right)^{n-i}$$

que podemos reescrever:

$$P(X=i) = \left\{\frac{n(n-1)(n-2)\cdots(n-\{i-1\})}{n^i}\right\}\left(\frac{\lambda^i}{i!}\right)\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^i}.$$

Aproximação da Binomial para a Poisson

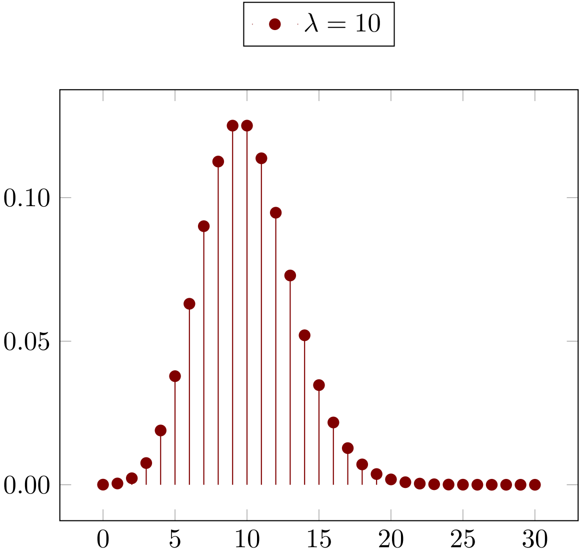
Como

$$1 - \frac{j}{n} \approx 1, \ \left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}, \ \left(1 - \frac{\lambda}{n}\right)^i \approx 1,$$

concluimos que

$$P(X = i) \approx \frac{e^{-\lambda}\lambda^i}{i!} = P(Y = i).$$

$p(i)$

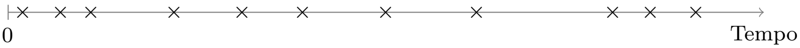


Universidade Federal de Pernambuco
CCEN Centro de Ciências Exatas e da Natureza

Processo de Poisson de parâmetro λ

Usado para representar, principalmente, os instantes de ocorrência de eventos de interesse (terremotos em uma região, ligações telefônicas recebidas por uma central, chegadas de clientes em um sistema etc).

Um Processo de Poisson de parâmetro λ unidimensional pode ser representado como uma sequência de pontos ou marcas em $\mathbb{R}^+$:



tais que, se $N(B)$ = número de pontos contidos no intervalo $B \subset \mathbb{R}^+$, então:

- $N(B) \sim Poisson(\lambda|B|)$;
- $B_1, B_2 \subset \mathbb{R}^+$ intervalos, $B_1 \cap B_2 = \emptyset$, então $N(B_1)$ e $N(B_2)$ são independentes.

Observação!

Podemos estender o anterior para qualquer $B \in \mathcal{B}$, com $|B|$ representando a medida de Lebesgue de B.

Universidade Federal de Pernambuco

CCEN
Centro de Ciências Exatas e da Natureza

Distribuição Geométrica

X tem *distribuição geométrica* com parâmetro $p \in (0,1)$ se:

$$p(i) = p(1-p)^{i-1}, \text{ para } i \in \mathbb{N}.$$

Notação: $X \sim Geom(p)$.

Observação!

X pode ser interpretada como o número de ensaios independentes de Bernoulli, com probabilidade de sucesso p, que devem ser realizados até obter o primeiro sucesso.

Pensando na interpretação de X, note que se

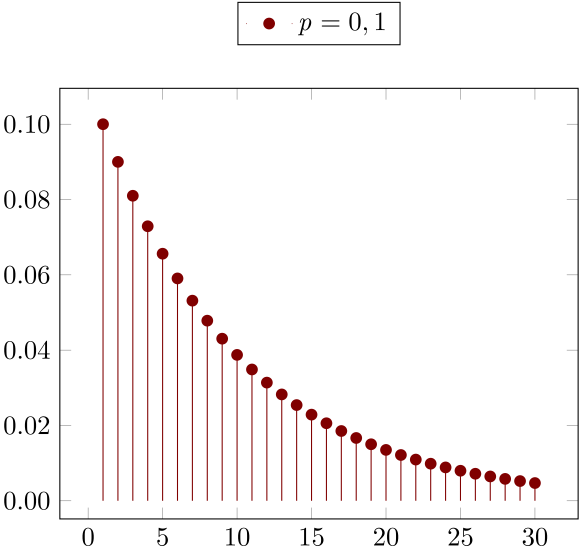
$$X_i = \begin{cases} 1, & \text{se o } i\text{-ésimo ensaio resulta em sucesso,} \\ 0, & \text{caso contrário,} \end{cases}$$

para todo $i \in \mathbb{N}$, então:

$$P(X = i) = P\left(\left\{\bigcap_{k=1}^{i-1} \{X_k = 0\}\right\} \cap \{X_i = 1\}\right) = (1 - p)^{i-1} p.$$

por independência!

$p(i)$



Problema da coleção de cupons

Existem n tipos de cupons e cada vez que um colecionador pega um cupom, este tem, independentemente das seleções anteriores, a mesma probabilidade de ser de qualquer um dos n tipos. Uma variável aleatória de interesse é o número de cupons que precisam ser recolhidos até obter uma coleção completa de pelo menos um cupom de cada tipo. Se X é esta variável aleatória então:

$$X = \sum_{i=1}^{n} T_i,$$

em que $T_1, T_2, \ldots, T_n$ são variáveis aleatórias independentes com

$$T_i \sim Geom\left(\frac{n-i+1}{n}\right),$$

para $i \in \{1, \ldots, n\}$.

Bom estudo!