

A distribuição de Poisson

Prof. Dr. Sadraque E. F. Lucena

Este material foi desenvolvido como parte de um exercício da disciplina ESTAT0016 – Tópicos Especiais em Estatística, com o intuito de explorar a aplicação do Quarto, uma ferramenta de publicação técnica de código aberto projetada para auxiliar os cientistas de dados na compartilhamento de suas análises. Como exemplo, realizaremos uma breve exploração da distribuição de Poisson.

A distribuição de Poisson

A distribuição de Poisson é uma distribuição de probabilidade discreta que descreve o número de eventos que ocorrem em um intervalo fixo de tempo ou espaço. Ela é aplicada em situações em que os eventos são independentes uns dos outros, a taxa média de ocorrência é constante e a probabilidade de mais de um evento acontecer em um intervalo infinitesimal é praticamente nula.

A distribuição de Poisson é frequentemente utilizada em problemas do mundo real, como prever o número de chamadas recebidas em um call center em uma hora, o número de defeitos em um lote de produtos, ou a contagem de eventos raros em geral.

Vejamos a definição.

Seja X uma variável aleatória com distribuição de Poisson com parâmetro λ , denotada como $X \sim \text{Poisson}(\lambda)$. A função de probabilidade de X é então dada por

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!},$$

onde

- λ representa a média de ocorrências no intervalo especificado;
- x é o número de eventos, $x = 0, 1, 2 \dots$

A média e a variância da distribuição de Poisson são iguais a taxa média de ocorrência λ , isto é, $E(X) = V(X) = \lambda$.

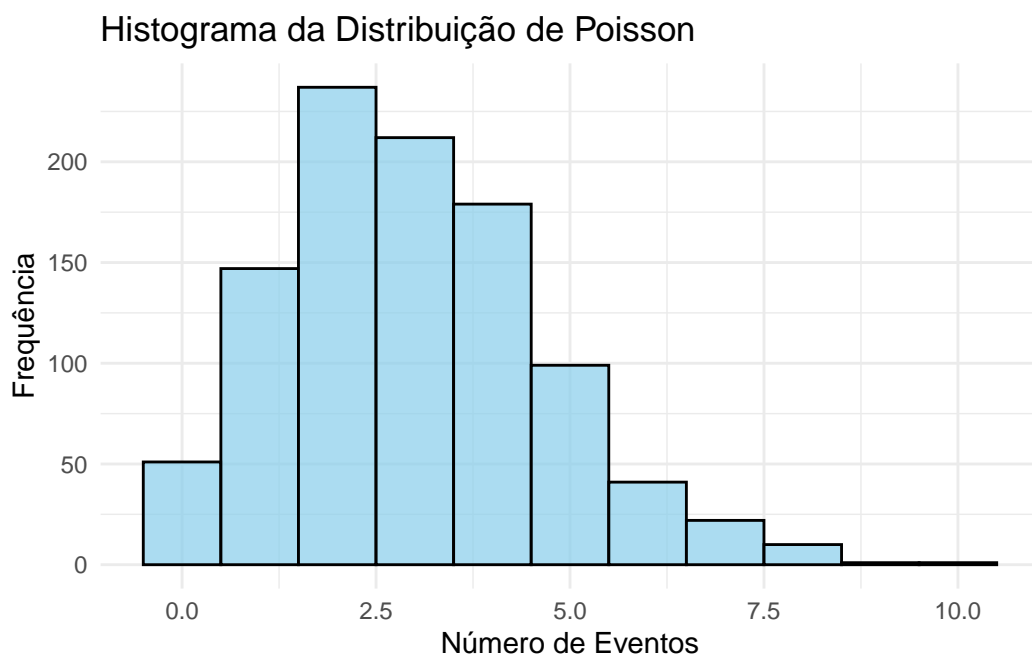
Uma característica fundamental de uma variável com distribuição de Poisson é que os valores que ela assume apenas valores não negativos. O código abaixo representa graficamente a distribuição dos valores de $X \sim \text{Poisson}(\lambda = 3)$.

```
# semente
set.seed(123)

# Parâmetros
lambda <- 3
n_sim <- 1000

# Simulação de dados
dados_poisson <- rpois(n_sim, lambda)

# Criando o histograma
library(ggplot2)
ggplot(data.frame(x = dados_poisson), aes(x)) +
  geom_histogram(binwidth = 1, fill = "skyblue", color = "black", alpha = 0.7) +
  labs(title = "Histograma da Distribuição de Poisson",
       x = "Número de Eventos",
       y = "Frequência") +
  theme_minimal()
```



Agora vamos apresentar novamente o histograma, mudando a cor para vermelho, mas sem mostrar o código.

