A Distribuição Poisson Truncada

João Moraes

26 de Março de 2019

Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Departamento de Estatística
Programa de Educação Tutorial

Sumário

A Distribuição Poisson

História

Siméon Denis Poisson (1781 - 1840), matemático e físico francês, foi o responsável pela descoberta da distribuição que carrega seu nome. Seu principal trabalho em teoria da probabilidade foi publicado em 1837: *”Recherches sur la probabilité des jugements en matière criminelle et en matière civile”*.

Nele, Poisson introduziu o problema de se aplicar teoria da probabilidade nas decisões tomadas nos tribunais. Não se importando com a probabilidade de um julgamento em específico, mas considerando uma série de julgamentos. Ele descobriu que existia um certa relação e com isso esboçou uma fórmula para a distribuição de Poisson conhecida atualmente.

Além disso, sua investigação dos teoremas limites da probabilidade tornou-se o que hoje chamamos de **Lei dos Grandes Números**.

Teorema de Poisson

Teorema de Poisson

Se n ensaios independentes são realizados, resultando na ocorrência ou não-ocorrência de um evento A, e a probabilidade de ocorrência dos eventos não é a mesma em cada um dos ensaios, então, com probabilidade tão perto de 1 quanto desejado, pode-se afirmar que a frequência $\frac{m}{n}$ de ocorrência do evento A irá desviar arbitrariamente pouco da média aritmética $\hat{p}$ da probabilidade de ocorrência dos eventos nos ensaios individuais.

Em outras palavras:

$$\lim_{n\to\infty} P(|\frac{m}{n} - \hat{p}| < \epsilon) = 1.$$

Descobertas adicionais

Poisson também viu que, à medida que p, e consequentemente, $q = 1 - p$, desviava muito de $\frac{1}{2}$ e n aumentava significativamente, a aproximação da normal pela binomial:

$$(2npq\pi)^{-\frac{1}{2}} exp\{\frac{-(x-np)^2}{2npq}\}$$

se tornava menos precisa para a probabilidade de m ocorrências do evento A em n ensaios ($P_{m,n}$).

Descobertas Adicionais

Então Poisson demonstrou que, $n \to \infty$ e $P_n \to 0$, essa probabilidade era dada por:

$$P_{m,n} = \frac{\lambda^m e^{-\lambda}}{m!}$$

Com $\lambda = np_n$ sendo constante.

Trabalhos alheios

Trabalhos alheios

Embora Poisson tenha feito essa descoberta, seus outros trabalhos não estavam relacionados com essa formula. Foi **Laudislaus Von Bortkiewicz** (1868 - 1931), que encontrou uma correspondência entre a fórmula de Poisson e certos tipos de eventos com valores discretos, tais como:

- Suicídio de crianças prussianas
- Nascimentos de trigêmeos
- **Mortes por chutes de cavalos no exército prussiano**

Trabalhos alheios

Como as probabilidades desses eventos são muito pequenas num grande número de ocorrências, a Lei dos Grandes Números pode facilmente ser aplicada.

Na sua monografia de 1898 intitulada ***Das Gesetz der kleinen Zahlen*** (A Lei dos pequenos números), Bortkiewicz desenvolveu a atual distribuição de Poisson usando os estudos citados.

Variáveis de Poisson

Tendo isso em vista, definamos as seguintes variáveis aleatórias X e Y:

$$X \sim p(\lambda) = \frac{\lambda^x}{x!}e^{-\lambda} \quad e \quad Y \sim p(\mu) = \frac{\mu^y}{y!}e^{-\mu}$$

Assumindo que X e Y sejam independentes, fazemos a transformação:

$$R = \frac{X}{Y}$$

Poisson truncada

Poisson Truncada

Note que, para Y = 0, há uma indeterminação. Para evitar esse problema, devemos **truncar** o alcance de Y de uma forma que $P(Y = 0) = 0$, e consequentemente, $P(Y \geq 1) = 1$.

Sendo assim, consideremos uma variável de poisson Z:

$$p(z, \mu) = \frac{\mu^z}{z!} e^{-\mu}$$

Poisson Truncada

Note que:

$$
\begin{aligned}
1 &= P(Z = 0) + P(Z > 0) \\
&= P(Z = 0) + P(Z > 0) \\
&= e^{-\mu} + P(Z > 0)
\end{aligned}
$$

O que implica:

$$P(Z > 0) = 1 - e^{-\mu}$$

Poisson Truncada

Calculando então a seguinte probabilidade condicional, temos:

$$P(Z = k \mid Z > 0) = \frac{P(Z = k)}{1 - f(0; \mu)}$$

$$P(Z = k \mid Z > 0) = \frac{\mu^k e^{-\mu}}{k!(1 - e^{-\mu})}$$

$$P(Z = k \mid Z > 0) = \frac{\mu^k}{k!(e^{\mu} - 1)}$$

Prova que é uma f.p.

Para ser uma função de probabilidade, temos que provar que:

$$\sum_{z=1}^{\infty} \frac{\mu^z}{z!(e^\mu - 1)} = 1$$

$$\sum_{z=1}^{\infty} \frac{\mu^k}{z!(e^\mu - 1)} = \frac{1}{(e^\mu - 1)} \sum_{z=1}^{\infty} \frac{\mu^z}{z!}$$

Sabemos que:

$$e^\mu = \sum_{z=0}^{\infty} \frac{\mu^z}{z!} = \frac{\mu^0}{0!} + \sum_{z=1}^{\infty} \frac{\mu^z}{z!} = 1 + \sum_{z=1}^{\infty} \frac{\mu^z}{z!} \qquad (1)$$

Continuando

$$
\begin{aligned}
\sum_{z=1}^{\infty} \frac{\mu^k}{z!(e^{\mu}-1)} &= \frac{1}{(e^{\mu}-1)} \sum_{z=1}^{\infty} \frac{\mu^z}{z!} \\
&= \frac{1}{(e^{\mu}-1)} \frac{(e^{-\mu}-1)}{1} \\
&= 1
\end{aligned}
$$

A média de Z é dada por:

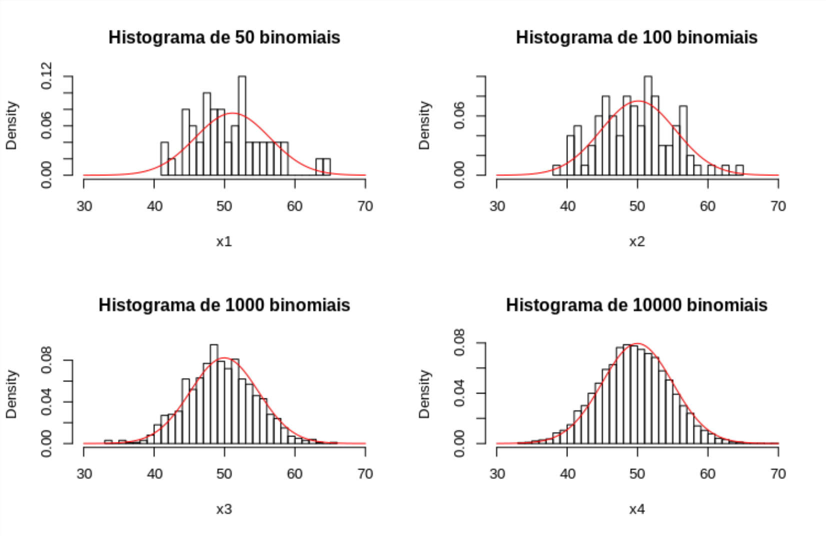
$$
\begin{aligned}
E(Z) &= \sum_{z=1}^{\infty} \frac{z\mu^z e^{-\mu}}{z!(1-e^{-\mu})} \\
&= \sum_{z=1}^{\infty} \frac{e^{-\mu}\mu\mu^{z-1}}{(x-1)!(1-e^{-\mu})} \\
&= \frac{e^{-\mu}\mu}{1-e^{-\mu}} \sum_{t=0}^{\infty} \frac{\mu^t}{t!} \\
&= \frac{e^{-\mu}\mu e^{\mu}}{(1-e^{-\mu})} \\
&= \frac{\mu}{(1-e^{-\mu})}
\end{aligned}
$$

Média e variância

A variância é dada por:

$$\begin{aligned} Var(x) &= E(X^2) - E(X)^2 \\ &= \frac{\mu}{(1 - e^{-\mu})^2} \end{aligned}$$

Distribuições


Histograma de 50 binomiais
Density
0.00 0.06 0.12
30 40 50 60 70
x1
Histograma de 100 binomiais
Density
0.00 0.06
30 40 50 60 70
x2
Histograma de 1000 binomiais
Density
0.00 0.04 0.08
30 40 50 60 70
x3
Histograma de 10000 binomiais
Density
0.00 0.04 0.08
30 40 50 60 70
x4

Distribuições