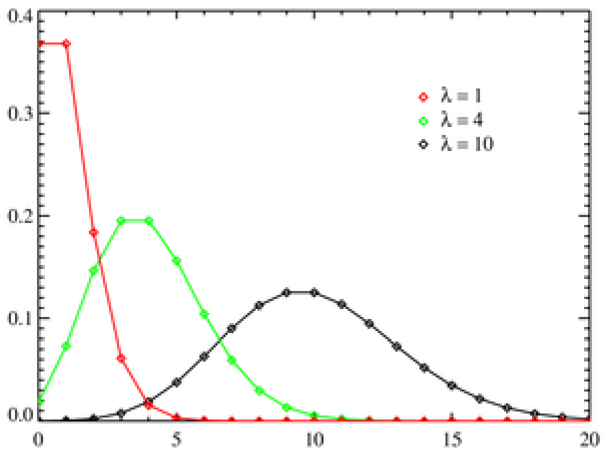
WIKIPÉDIA

Distribuição de Poisson

Origem: Wikipédia, a enciclopédia livre.

Na teoria da probabilidade e na estatística, a **distribuição de Poisson** é uma distribuição de probabilidade de variável aleatória discreta que expressa a probabilidade de uma série de eventos ocorrer num certo período de tempo se estes eventos ocorrem independentemente de quando ocorreu o último evento.



Função de probabilidade da distribuição de Poisson para vários valores de λ.

A distribuição foi descoberta por Siméon Denis Poisson (1781–1840) e publicada, conjuntamente com a sua teoria da probabilidade, em 1838 no seu trabalho *Recherches sur la probabilité des jugements en matières criminelles et matière civile* ("Inquérito sobre a probabilidade em julgamentos sobre matérias criminais e civis"). O trabalho focava-se em certas variáveis aleatórias *N* que contavam, entre outras coisas, o número de ocorrências discretas de um certo fenômeno durante um intervalo de tempo de determinada duração. A probabilidade de que existam exactamente *k* ocorrências (*k* sendo um inteiro não negativo, *k* = 0, 1, 2, ...) é

$$f(k;\lambda)=\frac{e^{-\lambda}\lambda^k}{k!},$$

onde

- *e* é base do logaritmo natural (*e* = 2.71828...),
- *k*! é o fatorial de *k*,
- λ é um número real, igual ao número esperado de ocorrências que ocorrem num dado intervalo de tempo. Por exemplo, se o evento ocorre a uma média de 4 minutos, e estamos interessados no número de eventos que ocorrem num intervalo de 10 minutos, usariámos como modelo a distribuição de Poisson com λ = 10/4 = 2.5.

Como função de *k*, esta é a função de probabilidade. A distribuição de Poisson pode ser derivada como um caso limite da distribuição binomial.

Índice

Processo de Poisson

A distribuição de Poisson aparece em vários problemas físicos, com a seguinte formulação: considerando uma data inicial (*t = 0*), seja *N(t)* o número de eventos que ocorrem até uma certa data *t*. Por exemplo, *N(t)* pode ser um modelo para o número de impactos de asteróides maiores que um certo tamanho desde uma certa data de referência.

Uma aproximação que pode ser considerada é que a probabilidade de acontecer um evento em qualquer intervalo não depende (no sentido de independência estatística) da probabilidade de acontecer em qualquer outro intervalo disjunto.

Neste caso, a solução para o problema é o processo estocástico chamado de **Processo de Poisson**, para o qual vale:

$$P[N(t) = K] = \frac{e^{-\lambda t}(\lambda t)^k}{k!},$$

em que λ é uma constante (de unidade inversa da unidade do tempo).

Ou seja, o número de eventos até uma época qualquer *t* é uma distribuição de Poisson com parâmetro *λ t*.

Propriedades

Média

O valor esperado de uma distribuição de Poisson é igual a *λ*. Esta propriedade pode ser derivada facilmente[1]:

	Em linguagem matemática	Em Português
	$E[X] = \sum_{k=0}^{\infty} k\mathbb{P}[X=k]$	Por definição, a esperança de uma variável aleatória X é igual à soma de cada uma das suas possíveis ocorrências ponderadas pela probabilidade de que estas ocorrências aconteçam.
	$E[X] = \sum_{k=0}^{\infty} k\left[\frac{e^{-\lambda}\lambda^k}{k!},\right]$	No caso de variáveis com distribuição, a probabilidade de que determinado evento ocorra é calculado por : $f(k;\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$. Portanto, este valor foi substituído na fórmula.
	$E[X] = \underbrace{0\left[\frac{e^{-\lambda}\lambda^0}{0!},\right]}_{k=0} + \underbrace{1\left[\frac{e^{-\lambda}\lambda^1}{1!},\right]}_{k=1} + \underbrace{2\left[\frac{e^{-\lambda}\lambda^2}{2!},\right]}_{k=2} + \ldots$	Esta expressão equivale à expressão da linha imediatamente superior; apenas se substituiu a expressão de somatório pela soma infinita para melhor compreensão. Note que como o primeiro termo é sempre igual a zero, podemos reescrever $E[X] = \sum_{k=0}^{\infty} k\left[\frac{e^{-\lambda}\lambda^k}{k!}\right] = \sum_{k=1}^{\infty} k\left[\frac{e^{-\lambda}\lambda^k}{k!}\right]$
	Como $\sum_{k=1}^{\infty} k\left[\frac{e^{-\lambda}\lambda^k}{k!}\right] = \sum_{k=1}^{\infty} \lambda\left[\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\right]$	Fazemos uma substituição para facilitar o cálculo.
	$E[X] = \lambda\sum_{k=1}^{\infty}\left[\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\right]$	Tomamos a substituição acima e tiramos a constante λ para fora do somatório (pois o primeiro termo da expressão imediatamente superior é igual à $\lambda * 1$.
	$E[X] = \lambda e^{-\lambda}\sum_{k=0}^{\infty}\left[\frac{\lambda^k}{(k)!}\right]$	Nova transformação para facilitar os cálculos...
	$E[X] = \lambda e^{-\lambda}\left[\frac{\lambda^0}{(0)!} + \frac{\lambda^1}{(1)!} + \frac{\lambda^2}{(2)!} + \frac{\lambda^3}{(3)!} + \ldots\right]$	Abrindo o somatório, verifica-se que a série converge para e^{λ}
	$E[X] = \lambda e^{-\lambda}e^{\lambda}$	Obtemos $e^{-\lambda}e^{\lambda} = e^0 = 1$
	$E[X] = \lambda$	Como queríamos demonstrar

Variância

A variância de uma distribuição de Poisson é igual a λ.

Soma de variáveis

A **soma** de duas variáveis de Poisson independentes é ainda uma variável de Poisson com parâmetro igual à soma dos respectivos parâmetros. Ou seja, se $X_i \sim \mathrm{Poisson}(\lambda_i)$ segue uma distribuição de Poisson com parâmetro λ_i e as variáveis aleatórias X_i são estatisticamente independentes, então

$Y = \sum_{i=1}^{N} X_i \sim \mathrm{Poisson}\left(\sum_{i=1}^{N} \lambda_i\right)$ também segue uma distribuição de Poisson cujo parâmetro é igual à soma dos λ_i .

Por exemplo, X_1 é uma variável aleatória que representa o número de óbitos por mil nascimentos na cidade "A" (distribuição de Poisson com média 1,2, digamos) e X_2 é uma variável aleatória que representa o número de óbitos por mil nascimentos na cidade "B" (variável de Poisson com média 3). Ao todo, o número de óbitos por mil nascimentos nas cidades "A" e "B" têm distribuição de Poisson com média $\sum_{i=1}^{2}\lambda_i = 1,2+3 = 4,2$.

Intervalo de confiança

Um método rápido e fácil para calcular um intervalo de confiança de aproximada de λ, é proposto na Guerriero (2012).[2] Dado um conjunto de eventos k (pelo menos 15 - 20) ao longo de um período de tempo T, os limites do intervalo confiança para a frequência são dadas por:

$$F_{low} = (1 - \frac{1.96}{\sqrt{k-1}})\frac{k}{T}$$
$$F_{upp} = (1 + \frac{1.96}{\sqrt{k-1}})\frac{k}{T}$$

em seguida, os limites do parâmetro λ são dadas por: $\lambda_{low} = F_{low}T; \lambda_{upp} = F_{upp}T$.

Exemplos

A distribuição de Poisson representa um modelo probabilístico adequado para o estudo de um grande número de fenômenos observáveis. Eis alguns exemplos:

- Chamadas telefônicas por unidade de tempo;
- Defeitos por unidade de área;
- Acidentes por unidade de tempo;
- Chegada de clientes a um supermercado por unidade de tempo;
- Número de glóbulos visíveis ao microscópio por unidade de área;
- Número de partículas emitidas por uma fonte de material radioativo por unidade de tempo.
- Batimentos cardíacos por unidade de tempo.

Ligações externas

- Calculadora - Distribuição de Poisson (http://www.elektro-energetika.cz/calculations/po.php?language=portugues)

Referências

1. Sayan Mukherjee. Lecture 6.5.- Poisson processes. In: PROBABILITY AND STATISTICS IN ENGINEERING. http://www.isds.duke.edu/courses/Fall06/sta113/poisson.pdf
2. V, Guerriero (2012). «Power Law Distribution: Method of Multi-scale Inferential Statistics» (http://www.sjmmf.org/paperInfo.aspx?ID=886). *J. Mod. Math. Fr*