

Distribuições Discretas: Poisson

Prof. Alejandra, Cristian, Paula e Rafael

Introdução à Probabilidade e Estatística



Estes slides tem por objetivo apenas complementar e esclarecer os textos principais indicados para a disciplina, e não devem ser usados como único texto de referência.

Eles tratam apenas dos principais aspectos da teoria abordada, deixando para os textos e vídeos os detalhes teóricos e exemplos mais elaborados.

Junto à essas notas o aluno deve ler os textos e ver os vídeos indicados no site do curso no Moodle, além da referência principal do curso.

- Uma v. a. X é chamada *Bernoulli* se sua função de probabilidade for

$$\begin{aligned}p(1) &= P(X = 1) = p \\ p(0) &= P(X = 0) = 1 - p,\end{aligned}$$

para algum $p \in (0,1)$.

- Uma v.a. Y cuja função de probabilidade é dada por

$$p(k) = P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n,$$

é chamada de *Binomial* de parâmetros n e p . Neste caso, $Y \sim \text{Bin}(n, p)$.

O valor esperado e a variância de Y são

$$E(Y) = np \quad \text{e} \quad \text{Var}(Y) = np(1-p),$$

respectivamente.

Suponha que estejamos interessados num evento de um experimento aleatório cuja probabilidade é pequena. Os eventos deste tipo são chamados **eventos raros**. Logo, ele provavelmente não será observado em uma realização só do experimento. Será necessário repeti-lo (de forma independente) um bom número de vezes para termos uma boa chance de observá-lo.

Seja A um evento tal que $p = P(A)$, onde p é próximo de 0. Se X é a v.a. que conta o número de ocorrências de A em n repetições independentes do experimento, então $X \sim \text{Bin}(n, p)$. Como $E(X) = np$, para termos uma chance razoável de observar A deveríamos ter um número de repetições do experimento próximo de $1/p$. Desta forma $E(X) = np$ ficaria próximo de 1.

Vamos inverter o raciocínio e imaginar que $p = \lambda/n$, onde λ é um parâmetro positivo, e n é um número natural bastante grande (de forma que p seja bem pequeno). Então $E(X) = \lambda$. Temos ainda que, para $i \geq 0$

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

que podemos reescrever como:

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

Como estamos pensando em n bastante grande vamos calcular

$$\lim_{n \rightarrow \infty} P(X=i).$$

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

Para tal vamos estudar o comportamento de cada um dos fatores da (1). Começamos com a quantidade entre chaves, que reescrevemos de forma conveniente como:

$$\frac{n(n-1)(n-2)\cdots(n-\{i-1\})}{n^i} = \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{i-1}{n}\right). \quad (2)$$

Como i está fixo, o limite de (2) quando $n \rightarrow \infty$ é o produto do limite de cada fator. Desde que

$$\lim_{n \rightarrow \infty} \left(1 - \frac{a}{n}\right) = 1, \quad (3)$$

onde a é uma constante real qualquer. Temos que

$$\lim_{n \rightarrow \infty} \frac{n(n-1)(n-2)\cdots(n-\{i-1\})}{n^i} = 1.$$

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

o limite em (3) e o fato de que i é um valor fixo temos que

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^i = 1.$$

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

tro lado, sabemos que

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}.$$

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

como o fator $\frac{\lambda^i}{i!}$ não depende de n temos que

$$\lim_{n \rightarrow \infty} P(X=i) = \frac{e^{-\lambda} \lambda^i}{i!}.$$

Neste ponto, precisaremos de uma nova série (soma infinita) para conseguir terminar nossos cálculos. Mais exatamente precisaremos saber que

$$\sum_{i=0}^{\infty} \frac{x^i}{i!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = e^x,$$

para $x \in \mathbb{R}$.

(Essa identidade, conhecida como série de Taylor da função e^x , está brevemente motivada no final destes slides, e os interessados estão convidados dar uma olhada!)

Dando sequência, concluímos do resultado acima que

$$\sum_{i=0}^{\infty} \frac{e^{-\lambda} \lambda^i}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1.$$

Por toda a argumentação feita anteriormente, a função

$$p(i) = \frac{\lambda^i}{i!} e^{-\lambda} \quad \text{para } i = 0, 1, 2, 3, \dots,$$

pode representar a função de probabilidade de uma certa variável aleatória.

Diremos que uma v.a. X tem *distribuição de Poisson* com parâmetro $\lambda \in (0, \infty)$ se

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

Alternativamente, podemos dizer que X é uma v.a. de Poisson.

Notação: neste caso escrevemos $X \sim \text{Poisson}(\lambda)$.

Observação!

A v.a. 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 .

Ilustração de $Y \sim \text{Poisson}(10)$

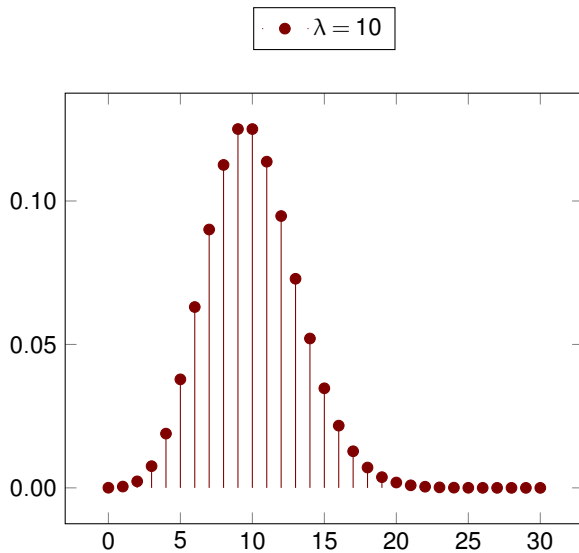
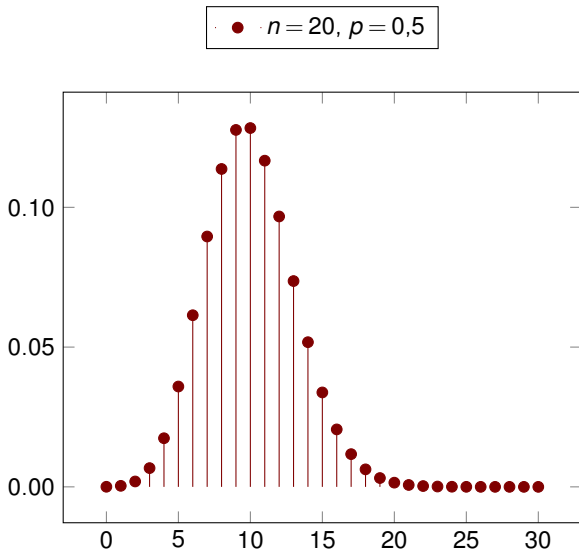


Ilustração de $X \sim \text{Bin}(200, 0,05)$



$$Y \sim \text{Poisson}(\lambda) \iff p(Y=i) = \frac{\lambda^i e^{-\lambda}}{i!}, \text{ para } i = 0, 1, 2, \dots$$

Suponha que os erros tipográficos, na edição de um livro de 300 páginas, estejam distribuídos aleatoriamente com uma média de 1 erro a cada 100 páginas. Qual a probabilidade de acharmos no livro todo

- 1 nenhum erro tipográfico?
- 2 pelo menos 2 erros tipográficos?
- 3 entre 1 e 3 erros tipográficos?

Solução: Vamos considerar a v. a. $Y :=$ número erros tipográficos no livro.

A probabilidade de ter um erro numa página é de $1/100 = 3/300$. Neste caso, podemos pelo menos de forma aproximada, dizer que $Y \sim \text{Poisson}(3)$. Assim

- 1 $p(0) = e^{-3} = 0,05$.
- 2 $P(Y \geq 2) = 1 - p(0) - p(1) = 1 - e^{-3} - 3e^{-3} = 0,80$.
- 3 $P(1 \leq Y \leq 3) = p(1) + p(2) + p(3) = 3e^{-3} + \frac{3^2}{2}e^{-3} + \frac{3^3}{6}e^{-3} = 0,60$.

Variáveis Aleatórias que obedecem à lei de Poisson

- 1 O número de pessoas, em uma certa população, que vivem mais de 100 anos.
- 2 Quantidade de números de telefone discados incorretamente em um dia.
- 3 O número de pacotes de biscoitos caninos vendidos em uma determinada loja em um dia.
- 4 Quantidade de clientes que entram em uma agência bancaria num dia.
- 5 O número de partículas α descarregadas por um material radioativo em um período de tempo fixo.

Observação!

Cada uma das variáveis aleatórias acima, e inúmeras outras, são aproximadas pela distribuição de Poisson por causa da aproximação de Poisson para a distribuição binomial.

Exercício

Uma máquina muito bem regulada tem probabilidade 0,01 de que um item produzido por ela apresente defeito. Determine a probabilidade de que uma amostra de 1000 itens contenha no máximo 1 item defeituoso. Justifique sua resposta.

$$Y \sim \text{Poisson}(\lambda) \iff P(Y = i) = \frac{\lambda^i e^{-\lambda}}{i!}, \text{ para } i = 0, 1, 2, \dots$$

Vimos que uma v.a. $Y \sim \text{Poisson}(\lambda)$ é uma aproximação de uma v.a. $X \sim \text{Bin}(n, p)$, quando n é grande, p é pequeno e $\lambda = np$. Além disso

$$E(Y) = np = \lambda$$

e

$$\text{Var}(Y) = np(1 - p) = \lambda(1 - p) \approx \lambda,$$

onde a aproximação vale pois p é pequeno.

Parece então razoável suspeitar que tanto o valor esperado quanto a variância de uma v.a. de Poisson são também iguais ao parâmetro λ .

Vamos verificar que de fato é isto o que acontece.

$$Y \sim \text{Poisson}(\lambda) \iff P(Y = i) = \frac{\lambda^i e^{-\lambda}}{i!}, \text{ para } i = 0, 1, 2, \dots$$

Para a esperança note que

$$\begin{aligned} E[Y] &= 0 \cdot e^{-\lambda} + 1 \cdot e^{-\lambda} \lambda + 2 \cdot e^{-\lambda} \frac{\lambda^2}{2!} + 3 \cdot e^{-\lambda} \frac{\lambda^3}{3!} + 4 \cdot e^{-\lambda} \frac{\lambda^4}{4!} + \dots \\ &= e^{-\lambda} \lambda + e^{-\lambda} \frac{\lambda^2}{1!} + e^{-\lambda} \frac{\lambda^3}{2!} + e^{-\lambda} \frac{\lambda^4}{3!} + \dots \\ &= \lambda \left(e^{-\lambda} + e^{-\lambda} \lambda + e^{-\lambda} \frac{\lambda^2}{2!} + e^{-\lambda} \frac{\lambda^3}{3!} + \dots \right) \\ &= \lambda [P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) + \dots] \\ &= \lambda. \end{aligned}$$

E, como suspeitamos, $E[Y] = \lambda$.

Variância de uma v.a. de Poisson

Para calcular a variância precisamos saber $E(Y^2)$. Assim

$$\begin{aligned}E[Y^2] &= 0^2 \cdot e^{-\lambda} + 1^2 \cdot e^{-\lambda} \lambda + 2^2 \cdot e^{-\lambda} \frac{\lambda^2}{2!} + 3^2 \cdot e^{-\lambda} \frac{\lambda^3}{3!} + \dots \\&= e^{-\lambda} \lambda + 2 \cdot e^{-\lambda} \frac{\lambda^2}{1!} + 3 \cdot e^{-\lambda} \frac{\lambda^3}{2!} + \dots \\&= \lambda \left(e^{-\lambda} + 2 \cdot e^{-\lambda} \frac{\lambda}{1!} + 3 \cdot e^{-\lambda} \frac{\lambda^2}{2!} + \dots \right)\end{aligned}$$

Ou seja,

$$E(Y^2) = \lambda \cdot \left[\sum_{j=0}^{\infty} (j+1) \frac{\lambda^j}{j!} e^{-\lambda} \right] = \lambda \left[\overbrace{\sum_{j=0}^{\infty} j \cdot \frac{\lambda^j}{j!} e^{-\lambda}}^{E(Y)} + \overbrace{\sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!}}^1 \right]$$

Encontramos assim que $E[Y^2] = \lambda^2 + \lambda$.

Desta forma, como $E[Y] = \lambda$, temos que

$$\text{Var}[Y] = E[Y^2] - E[Y]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda,$$

confirmando nossas suspeitas de que:

Teorema

Se Y é uma v.a. com $Y \sim \text{Poisson}(\lambda)$, então

$$E[Y] = \text{Var}[Y] = \lambda.$$

Exemplo

Suponha que o número de acidentes que ocorrem em uma rodovia em cada dia seja uma variável aleatória de Poisson com parâmetro 3. Determine:

- 1 a probabilidade de que 3 ou mais acidentes ocorram hoje;
- 2 a probabilidade de que 3 ou mais acidentes ocorram hoje, dado que pelo menos um acidente ocorra hoje;
- 3 o número esperado de acidentes que ocorrem nos 7 dias desta semana.

Solução: Faça N = total de acidentes em um dia. Assim $N \sim \text{Poisson}(3)$, e

1 $P(N \geq 3) = 1 - P(N < 3) = 1 - (P(N=0) + P(N=1) + P(N=2))$

$$\text{Assim } P(N \geq 3) = 1 - \left(e^{-3} + 3e^{-3} + \frac{9}{2}e^{-3} \right) = 0,57681.$$

- 2 Note que

$$P(N \geq 3 | N \geq 1) = \frac{P(\{N \geq 3\} \cap \{N \geq 1\})}{P(N \geq 1)} = \frac{P(N \geq 3)}{P(N \geq 1)} = \frac{0,57681}{1 - e^{-3}} = 0,60703.$$

- 3 Em uma dia temos $E[N] = 3$, e portanto em 7 dias a média é de $7 \cdot 3 = 21$ acidentes.

Outro uso da distribuição de probabilidades de Poisson é na contagem do total de ocorrências em um dado período de eventos que ocorrem em intervalos de tempo independentes. Por exemplo:

- Terremotos.
- Erupções de um vulcão.
- Entrada de pessoas em determinado estabelecimento (bancos, agências de correio, casas lotéricas, postos de gasolina, etc).
- Conflitos bélicos em uma dada região.
- Ligações telefônicas recebidas por uma central.

Informalmente falando, estes tipos de fenômeno tem as seguintes características:

- A probabilidade de que exatamente 1 evento ocorra em um intervalo de tempo pequeno, de tamanho h , é aproximadamente λh (diminui com h).
- A probabilidade de que 2 ou mais eventos ocorram no intervalo de tempo h é pequena em comparação com h .
- O total observado em certo intervalo é independente do total observado nos demais intervalos não superpostos a este.

Pode-se verificar que, nestas condições, o de eventos que ocorre em um intervalo de tempo t é uma v.a. de Poisson com parâmetro λt , para algum $\lambda > 0$.

Mais precisamente, se considerarmos o intervalo de tempo $[0, t]$ e $N(t)$ é a v.a. que representa o número de eventos que ocorrem neste intervalo então

$$N(t) \sim \text{Poisson}(\lambda t).$$

Suponha que terremotos ocorram numa região A com uma média de 2 por semana. Determine:

- 1 a probabilidade de que pelo menos 3 terremotos ocorram durante as próximas 2 semanas.
- 2 a distribuição de probabilidade da v.a. que conta o tempo, começando de agora, até a ocorrência do próximo terremoto.

Se $N(t)$ é a v.a. que conta a quantidade de terremotos na região A em t semanas então $N(t) \sim \text{Poisson}(2t)$. Desta forma, para $t = 2$ temos que

$$P(N(2) \geq 3) = 1 - p(0) - p(1) - p(2) = 1 - 13e^{-4}.$$

Seja X a v.a. que conta a quantidade de tempo até o próximo terremoto então

$$X > t \iff \{ \text{nenhum evento ocorrer nas próximas } t \text{ semanas} \}.$$

Desta forma

$$P(X > t) = P(N(t) = 0) = e^{-\lambda t}.$$

A Série de Taylor da Função Exponencial

Agora é hora de olharmos com calma para a segunda série (soma infinita) que usaremos neste curso.

Mas desta vez faremos um pouco diferente, e eu não direi com que série trabalharemos ainda... Vamos começar com um problema diferente, e bastante comum em cálculo: aproximação de funções.

Considere a função $f(x) = e^x$, e pergunte-se que polinômio de grau n melhor aproxima $f(x)$, ao menos para valores de x perto de 0?

Ou seja, como escolhemos $a_0, a_1, \dots, a_n \in \mathbb{R}$ tais que

$$e^x \approx a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots + a_n x^n,$$

para x próximo a 0?

Para conseguir responder esta pergunta, precisaríamos definir precisamente o que queremos dizer por "aproximadamente" ou "próximo", mas deixaremos estes detalhes para outros cursos (mais uma vez, façam Sequência e Séries! É muito bom!).

Deixando esta ideia de lado, vamos simplesmente pensar em como fazer as funções $f(x) = e^x$ e $p(x) = a_0 + a_1x + \dots + a_nx^n$ coincidirem da melhor maneira possível

Existem várias formas de se fazer isso, mas aqui queremos explorar a ideia de focar no $x = 0$ e fazer coincidir não apenas $f(0) = p(0)$, mas também todas as derivadas $f^{(k)}(0) = p^{(k)}(0)$, com $k = 1, \dots, n$.

Vamos seguir essa ideia e ver onde ela nos leva...

Começemos calculando as derivadas e ver o que encontramos...

Como $f(x) = e^x$, temos trivialmente que $f^{(k)}(x) = f(x) = e^x$ para qualquer $k \geq 1$! Assim

$$f^{(k)}(0) = f(0) = e^0 = 1, \quad k = 1, 2, 3, 4, \dots$$

Para $p(x)$ precisaremos fazer passo a passo, mas vamos ver o que encontramos...

Como $p(x) = a_0 + a_1 x + \dots + a_n x^n$, logo de cara encontramos que $p(0) = a_0$, o que nos leva a

$$a_0 = p(0) = f(0) = 1,$$

e podemos escrever

$$p(x) = 1 + a_1 x + \dots + a_n x^n.$$

Antes de calcular as derivadas, vamos lembrar que se $g(x) = x^m$, então $g'(x) = mx^{m-1}$, $g''(x) = m \cdot (m-1)x^{m-2}$, e em geral, se $k \leq m$

$$g^{(k)}(x) = m \cdot (m-1) \cdots (m-k+1) \cdot x^{m-k} = \frac{m!}{(m-k)!} x^{m-k},$$

enquanto para $k > m$, temos $g^{(k)} = 0$.

Deste modo, se $p(x) = 1 + a_1 x + \cdots + a_n x^n$, então

$$\begin{aligned} p^{(k)}(x) &= a_k \frac{k!}{(k-k)!} x^{k-k} + a_{k+1} \frac{(k+1)!}{(k+1-k)!} x^{k+1-k} + \cdots + a_n \frac{n!}{(n-k)!} x^{n-k} \\ &= a_k \cdot k! + a_{k+1} (k+1)! \cdot x + \cdots + a_n \frac{n!}{(n-k)!} x^{n-k} \\ &= a_k \cdot k! + x \cdot \left(a_{k+1} (k+1)! + \cdots + a_n \frac{n!}{(n-k)!} x^{n-k-1} \right), \end{aligned}$$

nos dando que $p^{(k)}(0) = a_k \cdot k!$.

Encontramos então que

- $f^{(k)}(0) = f(0) = 1$, para todo $k = 1, 2, \dots$,
- $p(0) = 1$ e $p^{(k)}(0) = a_k \cdot k!$, para todo $k = 1, \dots, n$.

Assim temos para $1 \leq k \leq n$

$$a_k \cdot k! = p^{(k)}(0) = f^{(k)}(0) = 1,$$

e portanto $a_k = \frac{1}{k!}$.

Finalmente, encontramos então que

$$e^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!},$$

e essa aproximação fica melhor sempre que n cresce. De fato, podemos mostrar, fazendo $n \rightarrow \infty$, que

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$