

Aula de hoje

1



Aula sobre

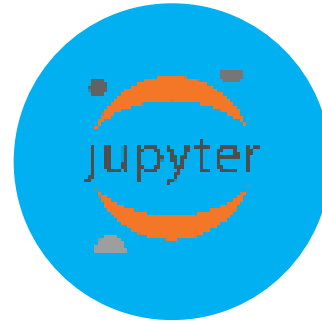
Modelo probabilístico discreto:

1) Poisson

1º: uso do notebook

**Como calcular probabilidades no
Jupyter utilizando a distribuição
Poisson.**

Aula12_Exemplo



2º: uso do notebook

Aula11_Atividade

3º: uso do notebook

Desenvolvimento do notebook

Aula12_Exercício

Há APS6 para Aulas 11 e 12



Insper

Ciência dos dados

Modelos Probabilísticos

Discretos

Poisson

Aula de hoje

Ao final desta aula, o aluno deve ser capaz de:

- Especificar as distribuições de probabilidades adequadas para variáveis aleatórias discretas considerando modelos probabilísticos discretos já bem definidos na literatura estatística.

Conversa entre usuário e chatbot

O número de conversas que se iniciam entre usuário e chatbot em um site e-commerce ocorre a uma taxa média de 4 a cada minuto.

- a. Defina a variável aleatória neste problema e os respectivos valores que essa pode assumir.
- b. Calcule a probabilidade de ocorrerem exatamente 3 início de conversa a cada 1 minuto qualquer do dia.
- c. Calcule a probabilidade de ocorrerem exatamente 3 início de conversa a cada dois minutos quaisquer do dia?

Distribuição de Poisson

Experimento de Poisson

Modela o número de ocorrências (sucessos) num espaço contínuo que pode ser de: tempo, área, distância, etc.

Exemplo:

Número de aviões que decolam por hora

Número de usuários conectados no Facebook por minuto

Número de acidentes a cada 100 km de uma rodovia

Número de clientes na fila por minuto (teoria das filas)

Consumo de um produto em um mês (gestão de estoques)

Propriedades (suposições):

- Probabilidade de uma ocorrência é a mesma para dois intervalos quaisquer de igual tamanho.
- A ocorrência ou não num dado intervalo é independente da ocorrência ou não em outro intervalo.



Siméon Denis Poisson

1781-1840

França

Distribuição de Poisson

Seja **X** o número de ocorrências de um evento num intervalo de tempo ou espaço, então

$$\mathbf{X \sim Poisson(\mu) \Rightarrow P(X = x) = \frac{e^{-\mu} \mu^x}{x!},}$$

em que

$x = 0, 1, 2, 3, \dots$ e

μ é o número médio de eventos ocorrendo no intervalo considerado.

Distribuição de Poisson

Quando **X** tem distribuição de Poisson com taxa μ , ou seja,

$$\mathbf{X \sim Poisson(\mu)}$$

com $x = 0, 1, 2, 3, \dots$

Então, prova-se que

$$\mathbf{E(X) = \mu} \quad \mathbf{e} \quad \mathbf{Var(X) = \mu}$$

Conversa entre usuário e chatbot

O número de conversas que se iniciam entre usuário e chatbot em um site e-commerce ocorre a uma taxa média de 4 a cada minuto.

- a. Defina a variável aleatória neste problema e os respectivos valores que essa pode assumir.
- b. Calcule a probabilidade de ocorrerem exatamente 3 inícios de conversa a cada 1 minuto qualquer do dia?
- c. Calcule a probabilidade de ocorrerem exatamente 3 inícios de conversa a cada dois minutos quaisquer do dia?

Validação das propriedades

O atendimento e-commerce recebe em média 4 conversas entre usuário e chatbot a cada minuto. Logo, $\mu = 4$.

- Probabilidade de ocorrer uma conversa é a mesma para dois intervalos de tempo de igual tamanho e a ocorrência ou não de uma conversa num dado instante de tempo é independente da ocorrência ou não de uma conversa em outro intervalo de tempo.
- Seja X a variável aleatória que denota o número de conversas entre usuário e chatbot por minuto, então

$$X \sim \text{Poisson}(4)$$

Vamos calcular no Jupyter

O número de conversas que se iniciam entre usuário e chatbot em um site e-commerce ocorre a uma taxa média de 4 a cada minuto. Calcule a probabilidade de:

- a) ocorrer 3 conversas entre usuário e chatbot em um minuto qualquer do dia?
- b) ocorrer menos de 3 conversas entre usuário e chatbot em um minuto qualquer do dia?
- c) ocorrer menos de 4 conversas entre usuário e chatbot em um minuto qualquer do dia, sabendo que houve pelo menos 1 conversa entre usuário e chatbot?

Distribuição de Poisson

Como calcular no Python: $X \sim \text{Poisson}(\mu)$

```
In [ ]: from scipy import stats
```

$P(X = k) \Rightarrow \text{stats.poisson.pmf}(k, \mu)$

$P(X \leq k) \Rightarrow \text{stats.poisson.cdf}(k, \mu)$

$E(X) \Rightarrow \text{stats.poisson.mean}(\mu)$

$\text{Var}(X) \Rightarrow \text{stats.poisson.var}(\mu)$

$\text{DP}(X) \Rightarrow \text{stats.poisson.std}(\mu)$

<https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.poisson.html>

Exemplo

- Download do notebook pelo Github
- Fazer individual e discutir em grupo
- Usar arquivo:

Aula12_Exemplo_...ipynb

Atividade

- Download do notebook pelo Github
- Fazer individual e discutir em grupo
- Usar arquivo:

Aula12_Atividade_....ipynb

Exercício

- Download do notebook pelo Github
- Fazer individual e discutir em grupo:
- Usar arquivo:

`Aula12_Exercicio_....ipynb`

Há APS6 nesse notebook.

Exercícios

Binomial e Poisson

Exercício 1

Um pendrive permite 60 minutos de gravação de um vídeo em qualidade full HD. Sabe-se que defeitos de fabricação surgem a uma taxa de 1,2 a cada 10 horas de gravação.

- a) Qual é a probabilidade de um pendrive apresentar defeito de fabricação?
- b) Sabendo que um pendrive apresentou defeito, qual a probabilidade de que ele tenha no máximo dois defeitos?
- c) Uma pessoa compra uma embalagem com 10 pendrives. Qual é a probabilidade de pelo menos 3 apresentarem defeito de fabricação?
- d) Quais suposições você fez para responder os itens (a) e (c)?

Exercício 2

Uma das responsabilidades da área de Tecnologia da Informação (TI) de uma empresa é controlar o fluxo de mensagens eletrônicas e detectar a presença de SPAMs (mensagens não desejadas).

Os responsáveis pela área de TI de uma empresa detectaram que um funcionário recebe no seu e-mail individual da empresa, em média, 0,75 SPAM por dia. Assuma que os SPAMs são enviados de forma independente para cada funcionário da empresa.

- a) Qual a probabilidade de um funcionário receber SPAM em um dia?
- b) Sabendo que um funcionário recebeu SPAM em uma semana, qual a probabilidade de receber entre 10 e 15 SPAMs nessa semana?
- c) Selecionando 40 funcionários aleatoriamente (e considerando que os funcionários recebem ou não SPAMs de forma independente um do outro), qual é a probabilidade de vinte deles receberem SPAM no seu e-mail individual em um dia qualquer?

Exercício 3 – Montgomery e Runger (adaptado)

Uma rede congestionada de computadores tem 1% de probabilidade de perder um bloco de dados, e perdas de blocos são eventos independentes. Uma mensagem de *e-mail* requer 100 blocos de dados.

- a) Qual a distribuição do número de blocos de dados que devem ser reenviados em uma mensagem de *e-mail* ? Inclua os valores dos parâmetros.
- b) Qual é a probabilidade de no mínimo um bloco ter de ser reenviado?
- c) Sabendo que no mínimo um bloco precisa ser reenviado, qual a probabilidade de dois ou mais serem reenviados?
- d) Quais são a média e o desvio padrão do número de blocos que devem ser reenviados?

Exercício 4

O número médio de clientes satisfeitos com o atendimento de uma loja, em amostras de 40 clientes escolhidos ao acaso, é de 5,5.

Qual é a probabilidade de, numa amostra de 40 clientes escolhidos ao acaso, encontrarmos pelo menos 2 satisfeitos?