Avaliação de desempenho de sistemas

Distribuições de probabilidade discretas

Distribuição de Bernoulli

X indica a ocorrência ou não de um evento de interesse A

$P[A] = p$ $\qquad$ $P[A^c] = 1 - p$

$\Omega_X = \{0, 1\}$

x	$p_X(x)$
0	$1 - p$
1	p

$p_X(x) = p^x (1-p)^{1-x}$

pmf

Distribuição de Bernoulli

- Valor esperado

$$E[X] = 0 \cdot P[X=0] + 1 \cdot P[X=1] =$$
$$= (0 \cdot (1 - p)) + (1 \cdot p) = p$$

- Variância

$$E[X^2] = 0^2 \cdot P[X=0] + 1 \cdot P[X=1]$$
$$= (0 \cdot (1 - p)) + (1 \cdot p) = p$$
$$V[X] = E[X^2] - (\mathrm{E}[X])^2 = p - p^2 = p \cdot (1 - p)$$

Distribuição Geométrica

X conta a quantidade de vezes até sair cara

$A = \{Ca\} \quad P[A] = 1/2$

$A^c = \{Co\} \quad P[A^c] = 1/2$

x	$p_X(x)$		
1	$(1/2)^1$	Ca	A
2	$(1/2)^2$	Co Ca	$A^c\,A$
3	$(1/2)^3$	Co Co Ca	$A^c\,A^c\,A$
x	$(1/2)^x$		

$p_X(x) = (1/2)^x \qquad$ *pmf*

Distribuição Geométrica

X conta quantas vezes sorteia A^c até sortear A

$A = \{...\}$ $P[A] = p$

$A^c = \{...\}^c$ $P[A^c] = (1 - p)$

x	$p_X(x)$		
1	$(1-p)^0\, p$	$\{...\}$	A
2	$(1-p)^1\, p$	$\{...\}^c\, \{...\}$	$A^c\, A$
3	$(1-p)^2\, p$	$\{...\}^c\, \{...\}^c\, \{...\}$	$A^c\, A^c\, A$
x	$(1-p)^{x-1}\, p$		

$p_X(x) = (1-p)^{x-1}\, p$ *pmf*

Distribuição Geométrica

- Função de probabilidade

$$p_X(x) = (1 - p)^{x\text{-}1}$$

sum[p*(1-p)^(x-1)] {x from 1 to infinity}

- Valor esperado

$$E[X] = \frac{1}{p}$$

sum[x*p*(1-p)^(x-1)] {x from 1 to infinity}

- Variância

$$V[X] = \frac{(1-p)}{p^2}$$

sum[(x-1/p)^2*p*(1-p)^(x-1)] {x from 1 to infinity}

Distribuição Geométrica

scipy.stats.geom

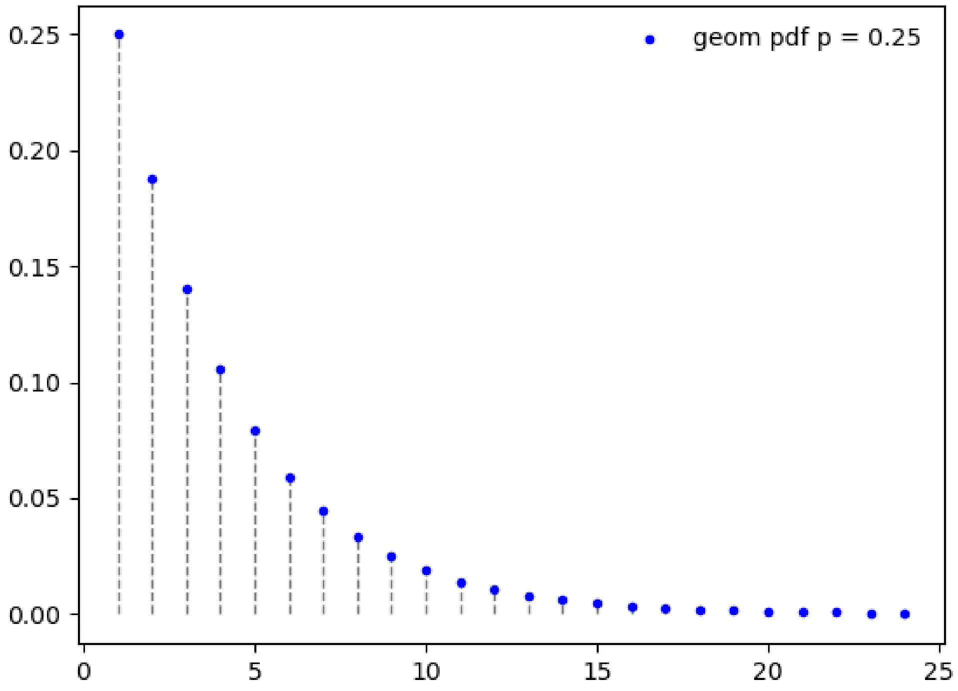
pmf	função de distribuição de probabilidade
cdf	função de distribuição acumulada
ppf	Função inversa da cdf
rvs	gerador de valores aleatórios

Distribuição geométrica

- Plotar a gráfico da pdf

  plotGeoPMF(p)

  Função em plot_PMF_.py

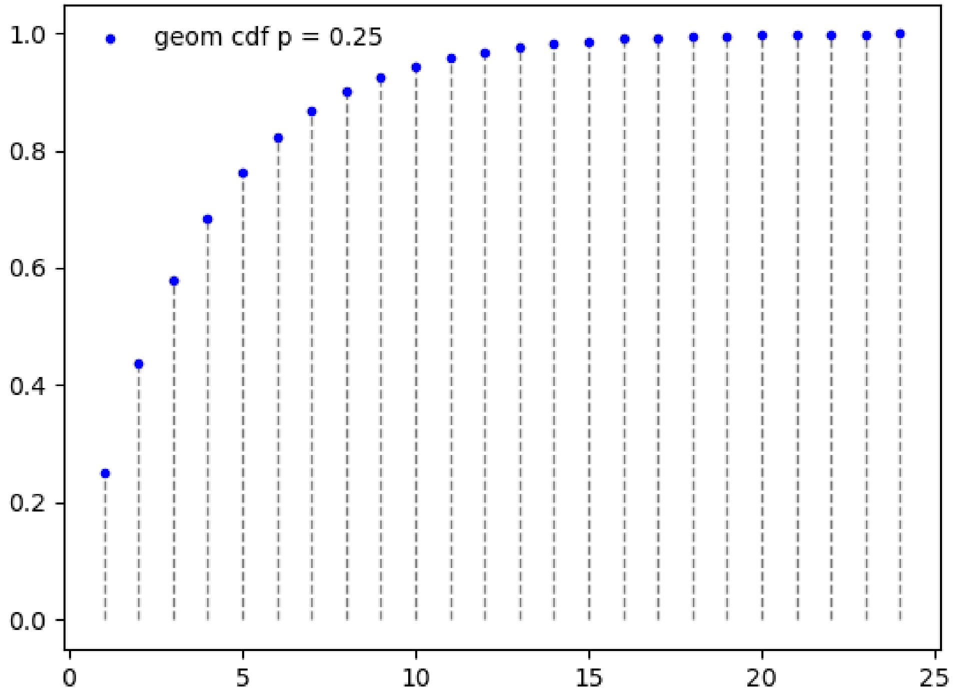


Distribuição geométrica

- Plotar a gráfico da cdf

  plotGeoCDF(p)

  Função em plot_PMF_CDF.py



Exemplo 1

A probabilidade de uma bateria acabar em um dia de uso é 0,03. Qual a probabilidade de ela acabar no terceiro dia?

- $P[X=x] = p_X(x) = (1 - p)^{x-1}\, p$

- Probabilidade de acabar $p = 0.03$
- Probabilidade de não acabar $(1 - p) = 0.97$

- Acabar no terceiro dia corresponde a $x = 3$
- $P[X=3] = (0.97^2 \cdot 0.03) = 0.028227$

- from scipy.stats import geom
- geom.pmf(3, 0.03)

Exemplo 2

- Supondo que a probabilidade de uma bateria acabar em um dia de uso é 0,03.
- Qual a probabilidade de ela acabar até o terceiro dia?
  - Acabar até o terceiro dia: acabar ou no primeiro, ou no segundo, ou no terceiro.

  - $P[X \leq 3] = P[X = 1] + P[X = 2] + P[X = 3]$
  - $P[X \leq 3] = 0{,}03 + 0{,}029099 + 0{,}028227 = 0{,}087327$

  - from scipy.stats import geom
  - geom.cdf(3, 0.03)

Distribuição de Poisson

- Função de probabilidade

$$p_X(x) = P[X = x] = e^{-L} \cdot \frac{L^x}{x!} \qquad x = 0,1,2,\ldots$$

sum[exp(-L)*L^x/factorial(x)] {x from 0 to infinity}

- Valor esperado

$$E[X] = L$$

sum[x*exp(-L)*L^x/factorial(x)] {x from 0 to infinity}

- Variância

$$V[X] = L$$

sum[(x-L)^2*exp(-L)*L^x/factorial(x)] {x from 0 to infinity}

Distribuição de Poisson

scipy.stats.poisson

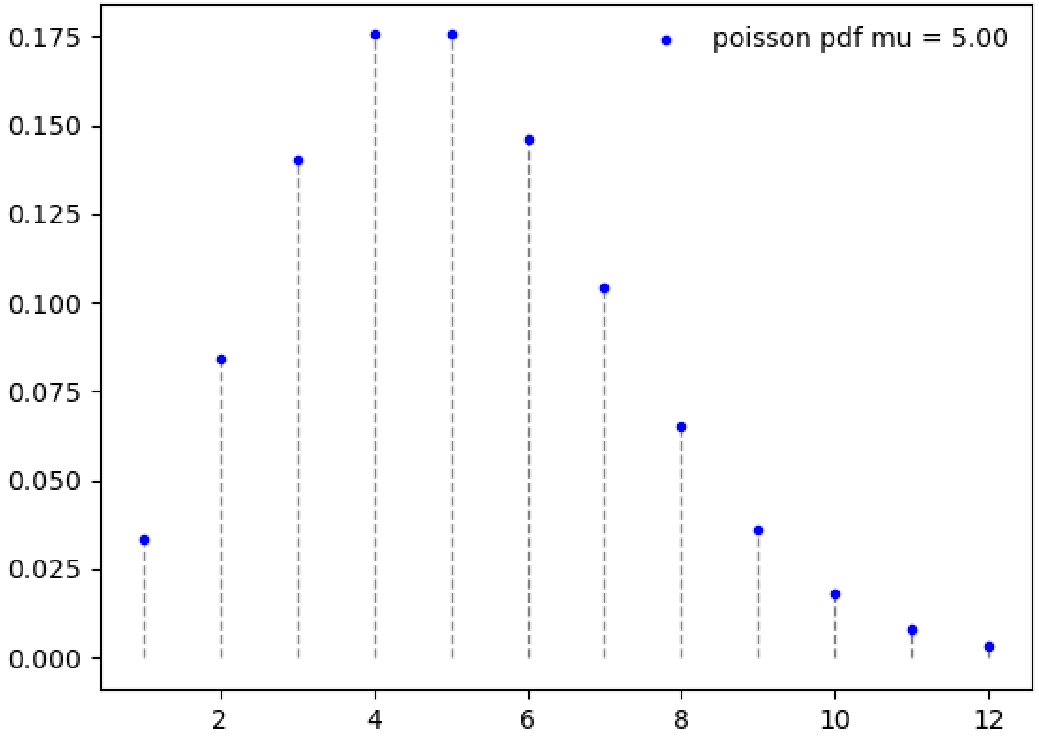
pmf	função de distribuição de probabilidade
cdf	função de distribuição acumulada
ppf	Função inversa da cdf
rvs	gerador de valores aleatórios

Distribuição de Poisson

- Plotar a gráfico da pdf

  plotPoissPMF(mu)

  Função em plot_PMF_CDF.py

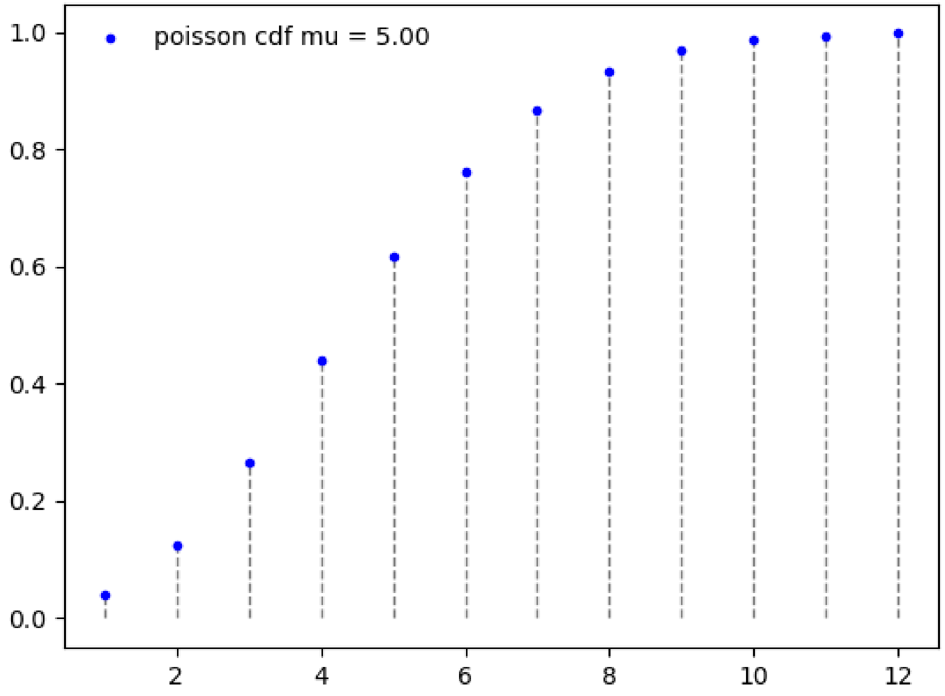


Distribuição de Poisson

- Plotar a gráfico da cdf

  plotPoissCDF(mu)

  Função em plotPMF.py



Exemplo 3

- Se o número médio de acidentes que ocorrem em uma estrada a cada dia é 3, e o número de acidentes por dia é uma variável aleatória com distribuição de Poisson:

  a) Qual a probabilidade de que não ocorram acidentes hoje?

  $E[X] = 3 = L$

  $$P[X = 0] = e^{-3}\frac{3^0}{0!} = e^{-3}$$

  b) Qual a probabilidade de ocorrer um ou dois acidentes em um dia?

  $P[X = 1] + P[X=2] = e^{-3} \cdot 3^1/1! + e^{-3} \cdot 3^2/2! = 15/2 \cdot e^{-3} = 0.3734$

Exemplo 3

```
>> import numpy as np
>> from scipy.stats import poisson
>> np.sum(poisson.pmf([1, 2], 3))
p =

    0.3734
```

Exemplo 4

- Um departamento de qualidade realiza testes em hard disk selecionados aleatoriamente. Os defeitos de fabricação ocorrem segundo uma distribuição de Poisson. A política definida é parar o processo de fabricação para calibração se um inspetor encontrar mais do que quatro defeitos em um disco. Qual é a probabilidade de o processo de fabricação ser interrompido se o número médio de defeitos é dois?
- O processo vai ser interrompido se o número de defeitos encontrado for maior ou igual a 5

pmf(5) + pmf(6) + ... = 1 – (pmf(0) + pmf(1) + ... + pmf(4))

```
import numpy as np
from scipy.stats import poisson
```

1 - np.sum(poisson.pmf([0, 1, 2, 3, 4], 2)) = 0,05265

1 - poisson.cdf(4, 2) = 0,05265

Avaliação de desempenho de sistemas

Distribuições de probabilidade discretas