Estatítica para Ciência de Dados

Aula 3: Modelos Probabilísticos

Francisco A. Rodrigues
ICMC/USP
francisco@icmc.usp.br

Aula 3: Modelos Probabilísticos

- **Valor esperado e variância**
  Modelos discretos
  Modelos contínuos

Probabilidades

- **Definição**
- O valor esperado de uma variável aleatória é definido por:
- Caso discreto:
  - $E[X] = \sum_k kP(X = k)$
- Caso contínuo:
  - $E[X] = \int_{-\infty}^{\infty} xf(x)dx$
- A variância de X:

$$V(X) = \sigma^2 = E[(X - E[X])^2]$$

Probabilidades

- **Exemplo:**
- Seja a variável aleatória X com distribuição abaixo. Calcule E[X] e V(X).

$$P(X=0) = 0.2$$
$$P(X=1) = 0.2$$
$$P(X = 2) = 0.6$$

Probabilidades

- **Exemplo:**
- A variável aleatória X tem função densidade de probabilidade dada por:

$$\begin{cases} \dfrac{x^2}{3} \, se -1 < x < 2 \\ 0 \;\; caso\ contrário \end{cases}$$

Calcule a variância e o valor esperado de X.

Probabilidades

WolframAlpha computational intelligence.

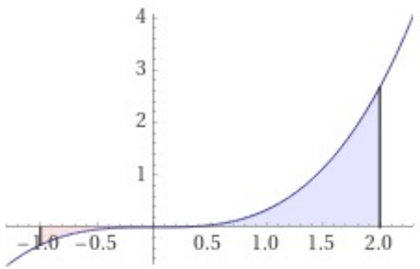
integrate x*x^2/3 from -1 to 2

Extended Keyboard   Upload

Definite integral:

$$\int_{-1}^{2} \frac{x\,x^2}{3}\,dx = \frac{5}{4} = 1.25$$

Visual representation of the integral:



integrate (x^2)*x^2/3 from -1 to 2 - (integrate x*x^2/3 from -1 to 2)^2

Extended Keyboard   Upload

Input:

$$\int_{-1}^{2} x^2 \times \frac{x^2}{3}\,dx - \left(\int_{-1}^{2} x \times \frac{x^2}{3}\,dx\right)^2$$

Result:

$$\frac{51}{80} = 0.6375$$

Computation result:

$$\int_{-1}^{2} \frac{x^2\,x^2}{3}\,dx - \left(\int_{-1}^{2} \frac{x\,x^2}{3}\,dx\right)^2 = \frac{51}{80}$$

Probabilidades

- **Distribuição Binomial**
- Seja uma v.a. baseada em n repetições de um processo de Bernoulli. Então:

$$P(X = k) = \frac{n!}{(n-k)!\,k!} p^k (1-p)^{\{n-k\}}, \mathrm{k} = 0,1,2, \ldots, n.$$

$$E[X] = np$$

$$V(X) = np(1-p)$$

Probabilidades

- **Exemplo:** Em uma urna há 8 bolas brancas e 4 pretas. Retira-se 5 bolas com reposição. Calcule a probabilidade de que:

  a) saiam duas bolas brancas.

  b) saiam ao menos 3 pretas.

Probabilidades

Exemplo: Em uma urna há 8 bolas brancas e 4 pretas. Retira-se 5 bolas com reposição. Calcule a probabilidade de que:
a) saiam duas bolas brancas.

Vamos construir uma função para calcular o valor exato.

In [6]:
```
import math
def binomial(n,p,k):
    C = (math.factorial(n)/(math.factorial(n-k)*math.factorial(k)))
    pk = C*(p**k)*(1-p)**(n-k)
    return pk
```

O valor teórico:

In [7]:
```
n = 5
p = 8/12
k = 2
print(binomial(n,p,k))
```

```
0.16460905349794244
```

Probabilidades

b) saiam ao menos 3 pretas.

Valor teórico.

In [17]:

```
n = 5
p = 4/12
k = 2
pk = 0
for k in range(3,6):
    pk = pk + binomial(n,p,k)
print('Valor teórico:', pk)
```

```
Valor teórico: 0.20987654320987656
```

Probabilidades

- **Distribuição de Poisson**
- Uma v.a. tem distribuição de Poisson com parâmetro $\lambda > 0$ (taxa) se sua função de probabilidade é dada por:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$
$$E[X] = \lambda$$
$$V(X) = \lambda$$

Probabilidades

- **Exemplo:**

Em uma central telefônica, chegam 300 mensagens por hora. Qual é a probabilidade de que:

a) Em um minuto não ocorra nenhuma chamada.

b) Em dois minutos ocorram duas chamadas.

Probabilidades

Exemplo: Em uma central telefônica, chegam 300 mensagens por hora. Qual é a probabilidade de que:

```
import numpy as np
import math
def Poisson(lbd, k):
    pk = np.exp(-lbd)*(lbd**k)/math.factorial(k)
    return pk
```

a) Em um minuto não ocorra nenhuma chamada.

```
lbd = 5 #numero de chamadas por minuto
k = 0
print("P(k = 0) = ",Poisson(lbd,k))
```

```
P(k = 0) =  0.006737946999085467
```

b) Em dois minutos ocorram duas chamadas.

```
lbd = 10 #numero de chamadas por 2 minutos
k = 2
print("P(k = 2) = ",Poisson(lbd,k))
```

```
P(k = 2) =  0.0022699964881242427
```

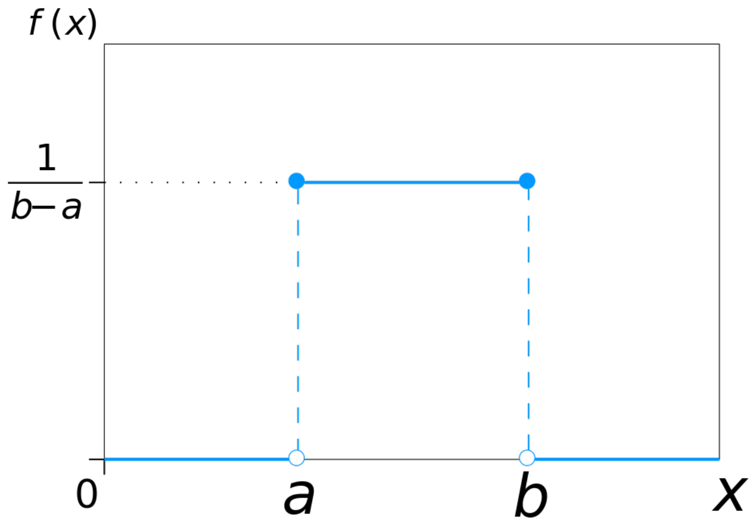
Probabilidades

- **Distribuição Uniforme**
- Uma v.a. contínua X segue o modelo uniforme no intervalo [a,b] se sua função densidade de probabilidade é dada por:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & x \notin [a,b] \end{cases}$$

$$E[X] = \frac{a+b}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$



Probabilidades

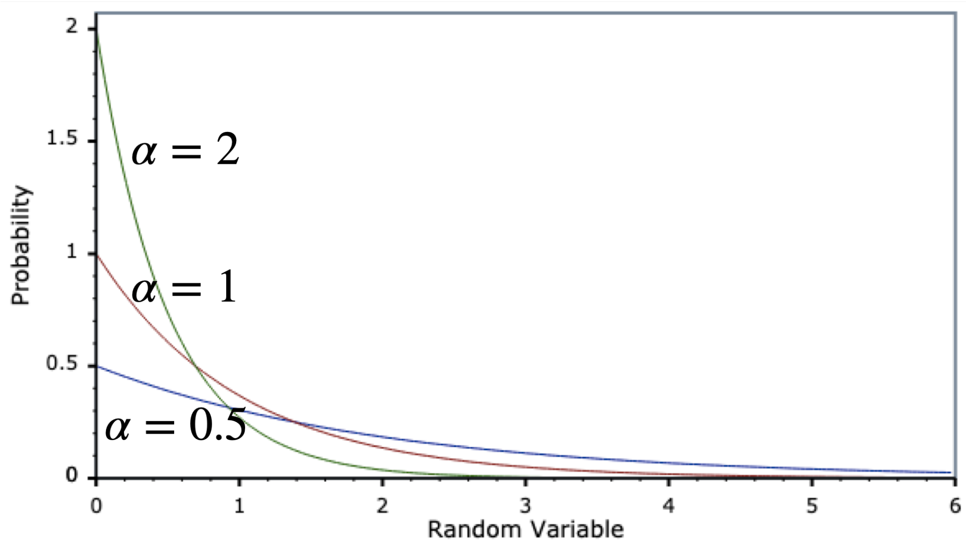
- **Distribuição exponencial**

Uma v.a. contínua X segue o modelo exponencial se sua função densidade de probabilidade é dada por:

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

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

$$V(X) = \frac{1}{\alpha^2}$$

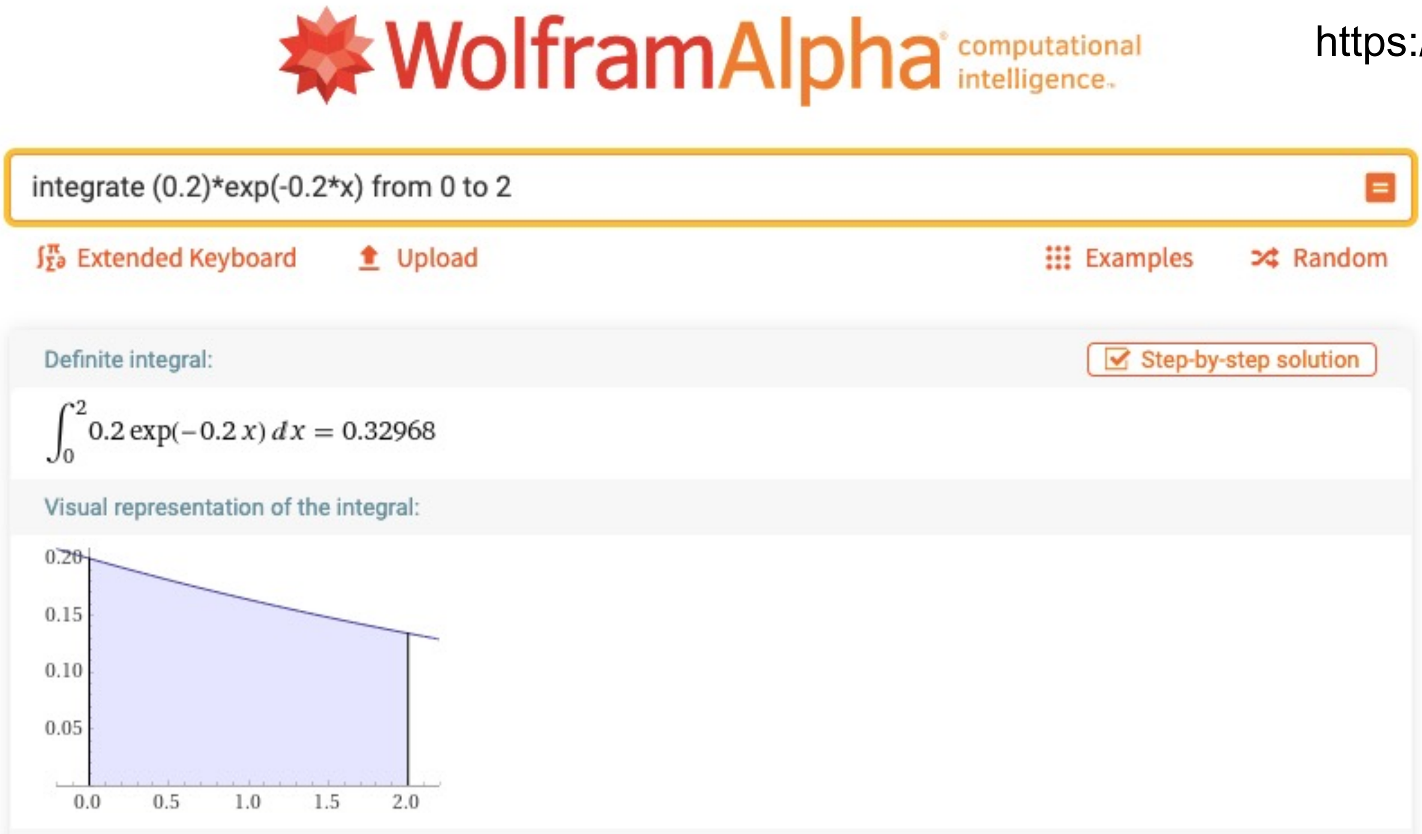


- **Exemplo:**

O intervalo de tempo, em minutos, entre emissões consecutivas de uma fonte radioativa é uma v.a. contínua que segue uma distribuição exponencial com parâmetro α. Qual é a probabilidade de que ocorra uma emissão em um intervalo inferior a 2 minutos?

Probabilidades

https://www.wolframalpha.com

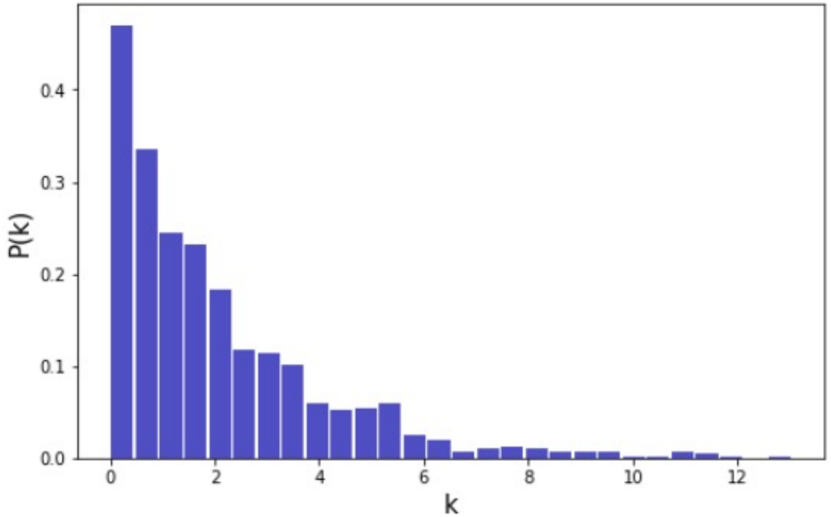


Probabilidades

Modelo exponencial

```python
from scipy.stats import expon

alpha = 2
X = expon.rvs(scale=alpha,size=1000)
plt.figure(figsize=(8,5))
P, bins, ignored = plt.hist(X, bins='auto', density=True, color='#0504aa',alpha=0.7,
                            rwidth=0.9)
plt.xlabel('k', fontsize = 15)
plt.ylabel('P(k)',fontsize = 15)
plt.show(True)
```



Sumário

- Valor esperado e variância
  Modelos discretos
  Modelos contínuos

Leitura Complementar

- Morettin e Bussab, **Estatística Básica**, Saraiva, 2017.