Universidade Federal do Rio Grande do Norte
Instituto Metrópole Digital
IMD0601 - Bioestatística

Distribuição de probabilidade

Prof. Dr. Tetsu Sakamoto
Instituto Metrópole Digital - UFRN
Sala A224, ramal 182
Email: tetsu@imd.ufrn.br

Baixe a aula (e os arquivos)

- Para aqueles que não clonaram o repositório:

```
> git clone https://github.com/tetsufmbio/IMD0601.git
```

- Para aqueles que já tem o repositório local:

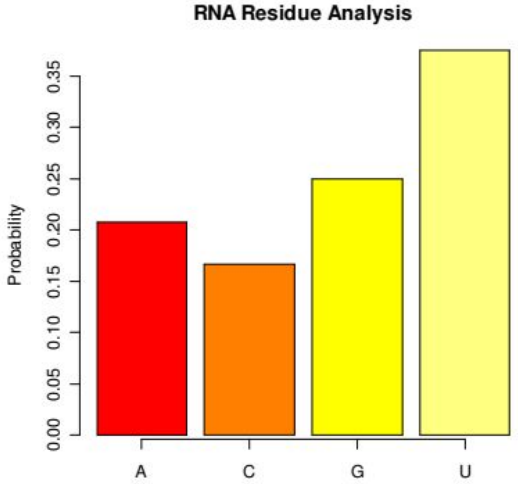
```
> cd /path/to/IMD0601

> git pull
```

Distribuição de probabilidades

Variáveis aleatórias ➔ Função que mapeia o espaço amostral em números reais (f: Ω ➔ R)

Função de distribuição de probabilidade ➔ Função que mapeia a variável aleatória ao valor de probabilidade.

Residue	Value of X (=x)	P (X=x)
A	0	5/24=0.208
C	1	4/24=0.167
G	2	6/24=0.25
U	3	9/24=0.375



Distribuição de probabilidades

Todos as variáveis aleatórias possuem uma função de distribuição de probabilidade associada;

Diferentes variáveis aleatórias possuem diferentes distribuições de probabilidade (infinitas);

Existem algumas formas comuns de distribuição dos dados;

Discreta	**Contínua**
Bernoulli	Normal
Binomial	Gama
Poisson	Qui-quadrado

Distribuição Bernoulli

Base para outras distribuições.

Experimentos com dois resultados possíveis (**Sucesso** e **Falha**);

nomes	falha	sucesso
valores	0	1
probabilidade	1-p	p

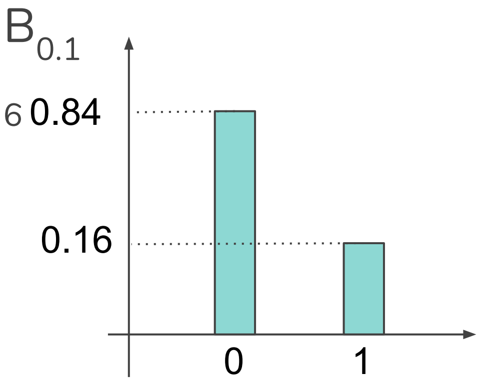
$$P(x) = p^x (1 - p)^{1-x}$$

Modelo probabilístico Bernoulli; x = {0, 1}

Distribuição Bernoulli

Considere o caso onde se quer ter um bebê de olhos azuis e o uso de Bernoulli para descrever as probabilidades. Digamos que a probabilidade de ter um bebê de olhos azuis (p) é 0.16. Para este experimento, a distribuição deste experimento será:

Outcome	Random Variable X =x	P(X=x)	Probability of outcome
Blue eyes	1	p	0.16
Not blue eyes	0	1-p	0.84



E se tivermos 10 bebês, qual a probabilidade de ter k bebês de olhos azuis?

Distribuição Binomial

E se tivermos 10 bebês, qual a probabilidade de ter k bebês de olhos azuis?

Experimentos independentes ➜ A probabilidade de um bebê ter olhos azuis não é influenciado pelo outro;

$p = 0.16$

$P(1110000000) = ?$ $\quad p^3(1-p)^7$

$P(0010010110) = ?$ $\quad p^4(1-p)^6$

$P(0000000111) = ?$ $\quad p^3(1-p)^7$

$$P(x) = p^k(1-p)^{n-k}$$

n = número de repetições do experimento;
k = número de sucesso.

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Distribuição Binomial

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

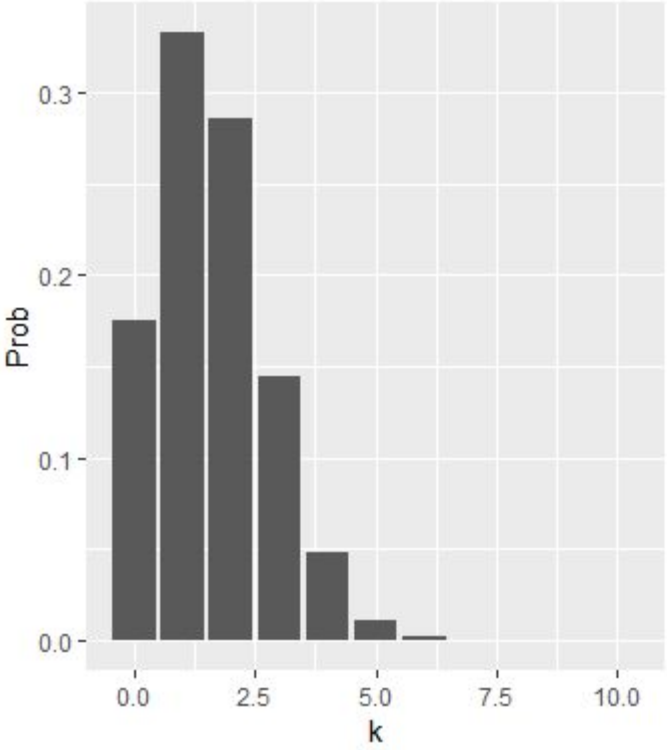
3 bebês de olhos azuis em 10 bebês (p=0.16):

$$P(3) = \binom{10}{3} 0.16^3 (1-0.16)^{10-3} = 0.1450$$

8 bebês de olhos azuis em 10 bebês (p=0.16):

$$P(8) = \binom{10}{8} 0.16^8 (1-0.16)^{10-8} = 1,3637.10^{-5}$$

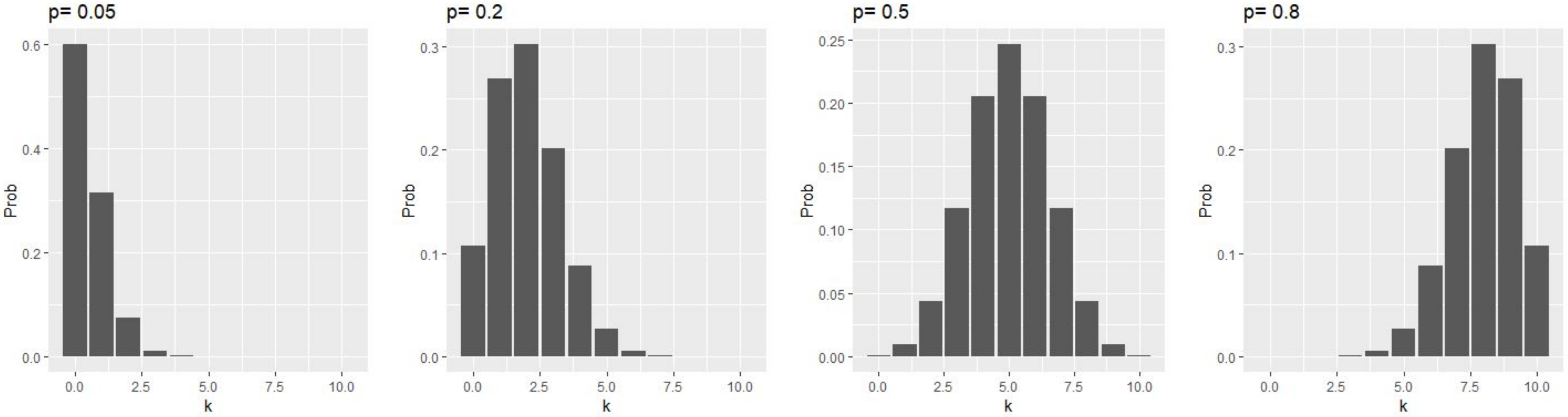
Distribuição binomial



k	P(k)
0	1.749012e-01
1	3.331452e-01
2	2.855530e-01
3	1.450428e-01
4	4.834760e-02
5	1.105088e-02
6	1.754108e-03
7	1.909233e-04
8	1.363738e-05
9	5.772436e-07
10	1.099512e-08

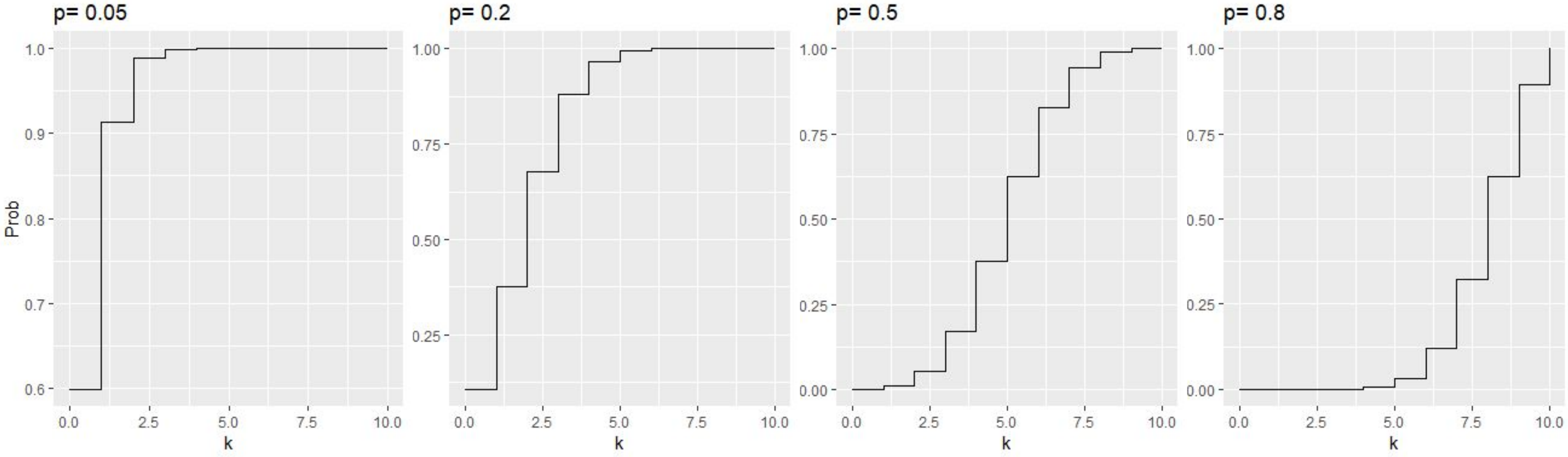
Distribuição binomial

Outros valores de p:



Distribuição binomial

Distribuição cumulativa



Distribuição de Poisson

Simeon D. Poisson (1781 - 1840) ➔ matemático físico francês;

Poisson é uma aproximação de $B_{p,n}$ para n muito grande e p muito pequeno.

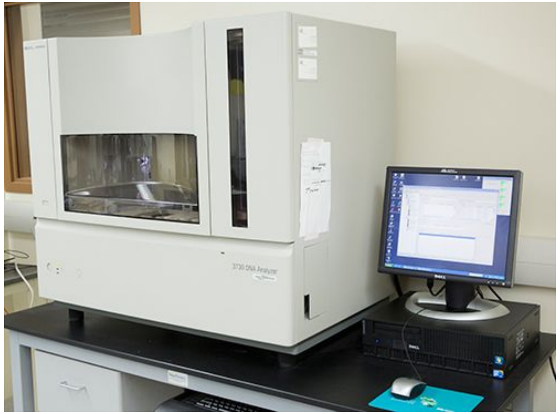
$$P_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad \lambda = n.p$$

Situações que seguem uma distribuição Poisson

P_λ aproxima de $B_{p,n}$ para p pequeno, n grande;

- Números de clicks em uma ad;
- Resposta a um spam;
- Número de erro de digitação;
- Número de acidentes ocorridos em um cruzamento em um período;
- Mutações em sequência de DNA;
- Erro produzido pela leitura de um sequenciador;
- Número de ocorrências de padrões de DNA.

Distribuição de Poisson



Sequenciador X:

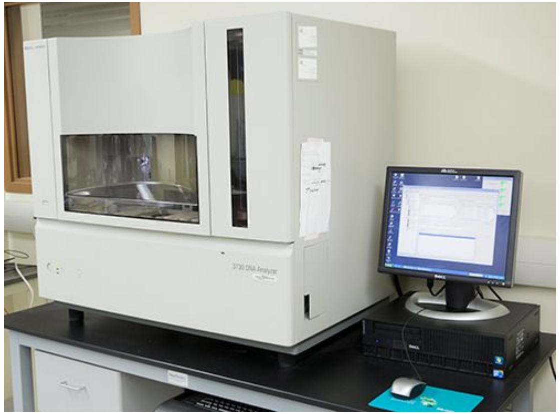
1 erro para cada 10.000 pares de base;

Sequenciar DNA de 2000 pares de base;

Qual a probabilidade de ter 0 erros usando este sequenciador?

Ou de 1 erro? Ou de 4 erros?

Distribuição de Poisson



Qual a probabilidade de ter 0 erros usando este sequenciador?

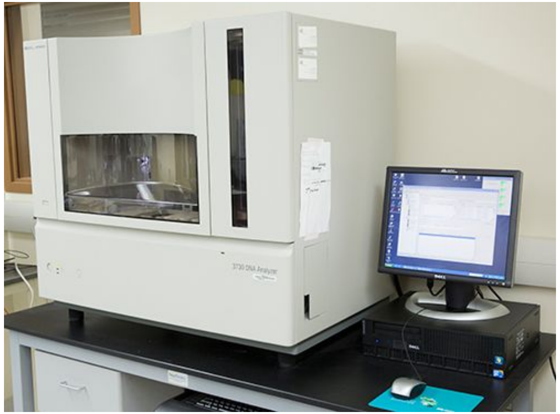
n = 2000; p = 1/10000;

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(0) = \binom{2000}{0} \frac{1}{10000}^{0} \left(\frac{9999}{10000}\right)^{2000}$$

$$P(0) = 0,8187$$

Distribuição de Poisson



Qual a probabilidade de ter 0 erros usando este sequenciador?

n = 2000; p = 1/10000;

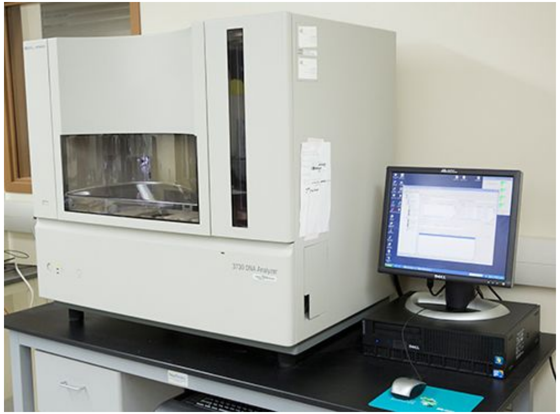
$$P_\lambda(k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

$$\lambda = n.p$$

$$\lambda = 2000 * \frac{1}{10000} = 0.2$$

$$P_{0.2}(0) = e^{-0.2}\frac{0.2^0}{0!} = 0.8187$$

Distribuição de Poisson



Qual a probabilidade de ter 1 erro usando este sequenciador?

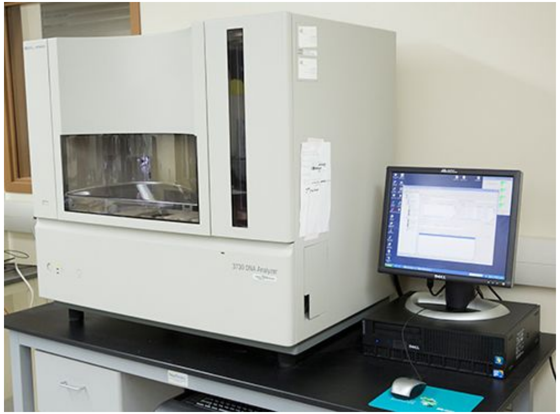
n = 2000; p = 1/10000;

$$P_\lambda(k) = e^{-\lambda}\frac{\lambda^k}{k!} \qquad \lambda = n.p$$

$$\lambda = 2000 * \frac{1}{10000} = 0.2$$

$$P_{0.2}(1) = e^{-0.2}\frac{0.2^1}{1!} = 0.1637$$

Distribuição de Poisson



Qual a probabilidade de ter 4 erros usando este sequenciador?

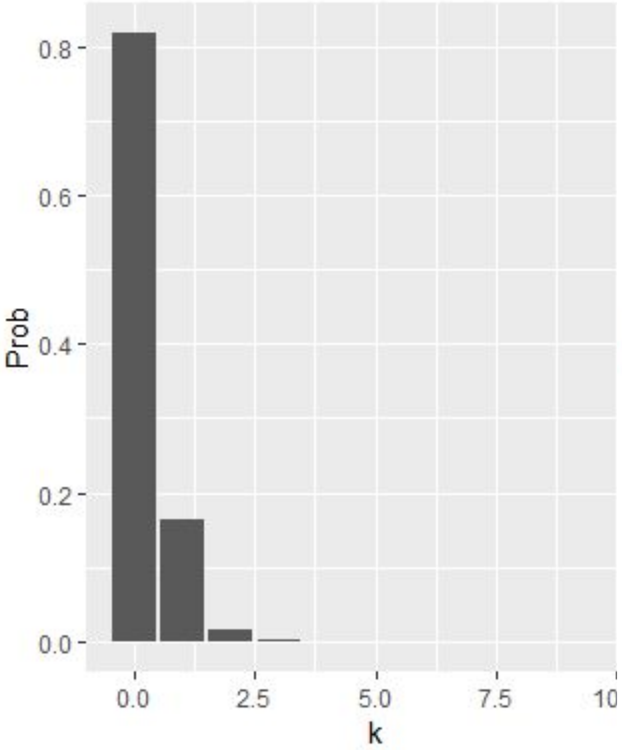
n = 2000; p = 1/10000;

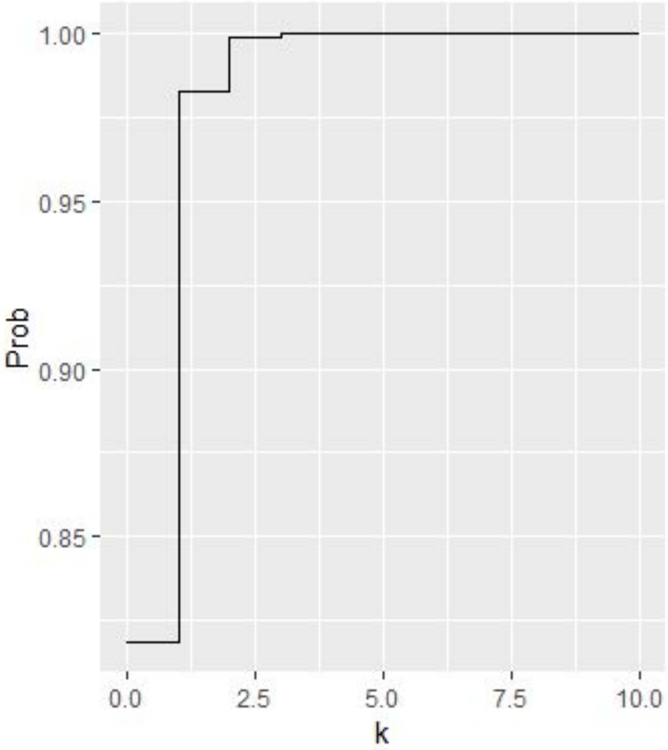
$$P_\lambda(k) = e^{-\lambda}\frac{\lambda^k}{k!} \qquad \lambda = n.p$$

$$\lambda = 2000 * \frac{1}{10000} = 0.2$$

$$P_{0.2}(4) = e^{-0.2}\frac{0.2^4}{4!} = 5.458 * 10^{-5}$$

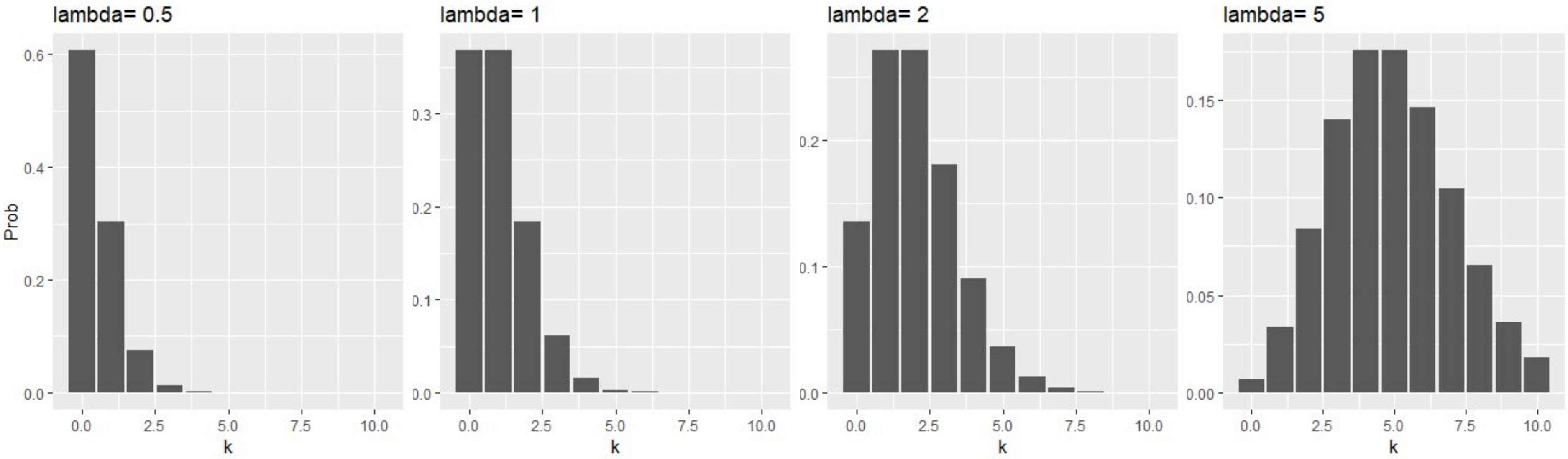
Distribuição de Poisson





k	P(k)
0	8.187308e-01
1	1.637462e-01
2	1.637462e-02
3	1.091641e-03
4	5.458205e-05
5	2.183282e-06
6	7.277607e-08
7	2.079316e-09
8	5.198290e-11
9	1.155176e-12
10	2.310351e-14

Distribuição de Poisson



Distribuição de Poisson