Estatística Básica

Variáveis Aleatórias Discretas

Renato Dourado Maia

Instituto de Ciências Agrárias

Universidade Federal de Minas Gerais

Variável Aleatória

- Uma quantidade X, associada a cada possível resultado do espaço amostral, é denominada de variável aleatória (VA) discreta, se assume valores num conjunto enumerável, com certa probabilidade. Por outro lado, será denominada variável aleatória contínua, se seu conjunto de valores é qualquer intervalo dos números reais, o que seria um conjunto não enumerável.

Função Discreta de Probabilidade

- A função que atribui a cada valor da variável aleatória a sua probabilidade é denominada de função discreta de probabilidade ou, simplesmente, função de probabilidade. A notação a ser utilizada é:

$$P(X = x_i) = p(x_i) = p_i,\ i = 1, 2, \cdots$$

X	x_1	x_2	x_3	$\cdots$
p_i	p_1	p_2	p_3	$\cdots$

1) $0 \le p_i \le 1$

2) $\sum_{\mathrm{i}} p_i = 1$

Função de Probabilidade

- Variáveis aleatórias são completamente caracterizadas pela sua função de probabilidade e uma parte importante da Estatística é, justamente, obter, para uma dada variável de interesse, a função de probabilidade que melhor represente o seu comportamento na população.

Exemplo 1

- Com dados do último censo, a assistente social de um Centro de Saúde constatou que para as famílias da região, 20% não têm filhos, 30% têm um filho, 35% têm dois, e as restantes se dividem igualmente entre três, quatro, ou cinco filhos. Suponha que uma família será escolhida, aleatoriamente, nessa região, e o número de filhos averiguado. Definimos N como sendo a variável aleatória número de filhos (não importa a família, e sim o número de filhos). Qual é a função de probabilidade da variável aleatória discreta N?

Exemplo 1

1) $P(N=3)=P(N=4)=P(N=5)=p$

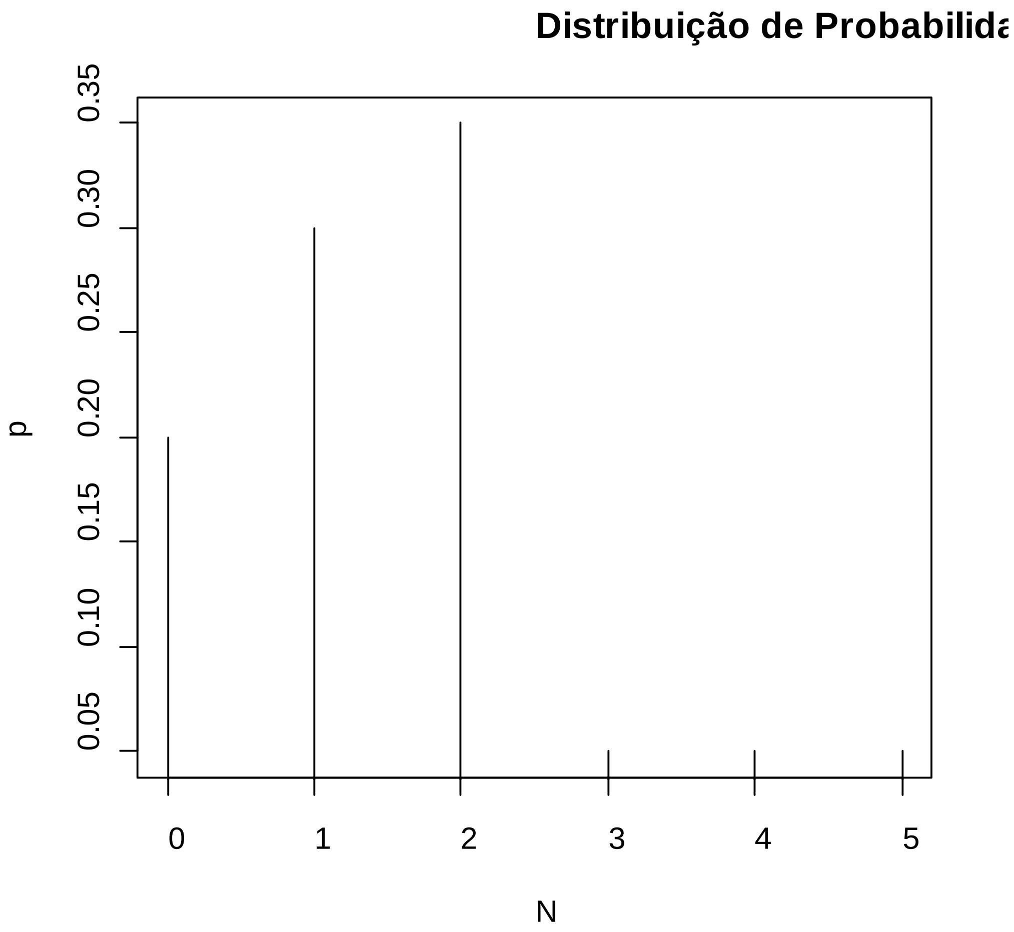
2) $\sum_{i=1}^{5} P(N=i)=1 \therefore 0,20+0,30+0,35+p+p+p=1 \therefore p=0,05$

N	0	1	2	3	4	5
p_i	0,20	0,30	0,35	0,05	0,05	0,05

No R:

```
> N=0:5
> p = c(0.2, 0.3, 0.35, 0.05, 0.05, 0.05)
> plot(N, p, type = "h", main = "Distribuição de Probabilidade de N“)
```

Exemplo 1



Distribuição de Probabilida
p
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
1
2
3
4
5
N

Exemplo 2

- Na construção de um certo prédio, as fundações devem atingir 15 metros de profundidade e, para cada 5 metros de estacas colocadas, o operador anota se houve alteração no ritmo de perfuração previamente estabelecido. Essa alteração é resultado de mudanças, para mais ou para menos, na resistência do subsolo. Nos dois casos, medidas corretivas são necessárias, encarecendo o custo da obra. Com base em avaliações geológicas, admite-se que a probabilidade de ocorrência de alterações é de 0,1 para cada 5 metros.

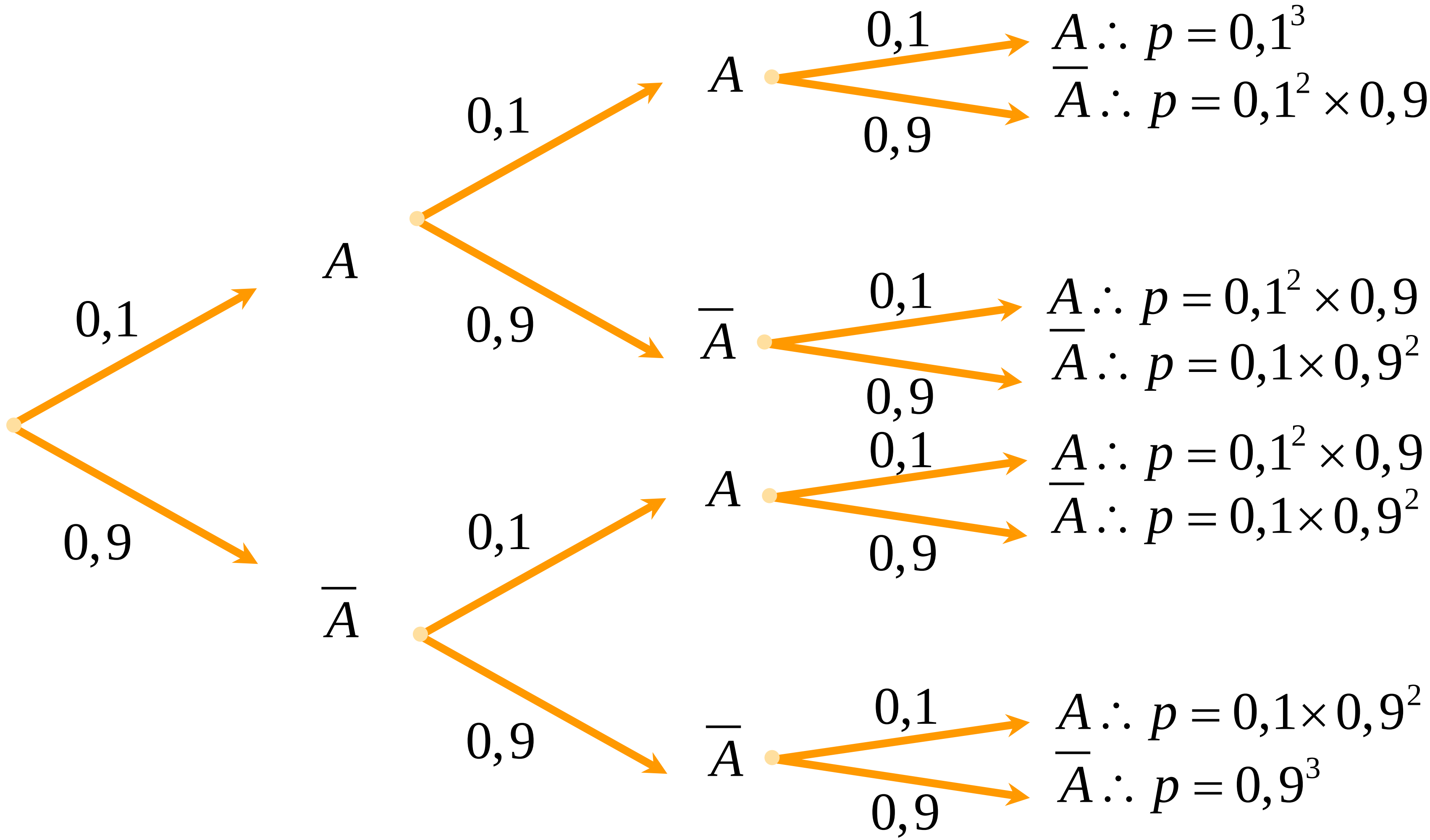
Exemplo 2

- O custo básico inicial é de 100 UPCs (unidade padrão de construção), e será acrescido de $50k$, com k representando o número de alterações observadas. Como se comporta a variável custo das obras de fundação?

- Considerações:

1) *A é o evento que representa a ocorrência de alteração em cada intervalo.*
2) *As alterações ocorrem independentemente entre cada um dos intervalos.*
3) *C é a variável aleatória custo das obras de fundação.*

Exemplo 2



0,1
0,9
A
$\overline{A}$
0,1
0,9
0,1
0,9
A
$\overline{A}$
A
$\overline{A}$
0,1
0,9
0,1
0,9
0,1
0,9
0,1
0,9
$A \therefore p = 0,1^3$
$\overline{A} \therefore p = 0,1^2 \times 0,9$
$A \therefore p = 0,1^2 \times 0,9$
$\overline{A} \therefore p = 0,1 \times 0,9^2$
$A \therefore p = 0,1^2 \times 0,9$
$\overline{A} \therefore p = 0,1 \times 0,9^2$
$A \therefore p = 0,1 \times 0,9^2$
$\overline{A} \therefore p = 0,9^3$

Exemplo 2

Evento	Probabilidade	C (em UPCs)
AAA	$0{,}1^3$	250
$AA\bar{A}$	$0{,}1^2 \times 0{,}9$	200
$A\bar{A}A$	$0{,}1^2 \times 0{,}9$	200
$A\bar{A}\bar{A}$	$0{,}1 \times 0{,}9^2$	150
$\bar{A}AA$	$0{,}1^2 \times 0{,}9$	200
$\bar{A}A\bar{A}$	$0{,}1 \times 0{,}9^2$	150
$\bar{A}\bar{A}A$	$0{,}1 \times 0{,}9^2$	150
$\bar{A}\bar{A}\bar{A}$	$0{,}9^3$	100

C	p_i
100	0,729
150	0,243
200	0,027
250	0,001

Eventos Disjuntos

Função de Distribuição de Probabilidade

- A função de distribuição ou função acumulada de probabilidade de uma variável aleatória discreta X é definida, para qualquer número real x, pela seguinte expressão:

$$F(x) = P(X \leq x)$$

Exemplo 3

- Uma população de 1.000 crianças foi analisada num estudo para determinar a efetividade de uma vacina contra um tipo de alergia. No estudo, as crianças recebiam uma dose de vacina e, após um mês, passavam por um novo teste. Caso ainda tivessem tido alguma reação alérgica, recebiam outra dose da vacina. Ao final de 5 doses todas as crianças foram consideradas imunizadas. Os resultados completos estão apresentados na tabela a seguir.

Exemplo 3

Tabela de Frequências

Doses	1	2	3	4	5
Frequência	245	288	256	145	66

Função de Probabilidade da VA Número de Doses Recebidas

Doses	1	2	3	4	5
p_i	0,245	0,288	0,256	0,145	0,066

Qual é a probabilidade de uma criança escolhida ao acaso ter recebido até duas vacinas?

$$F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = 0{,}533$$

Modelo Uniforme Discreto

- Seja X uma variável aleatória cujos possíveis valores são representados por x_1, x_2, ..., x_k. Dizemos que X segue o Modelo Uniforme Discreto se atribui a mesma probabilidade $1/k$ a cada um desses k valores, isto é, a sua função de probabilidade é dada por

$$P(X = x_j) = 1/k,\ \forall j = 1, 2, \cdots, k.$$

Modelo Bernoulli

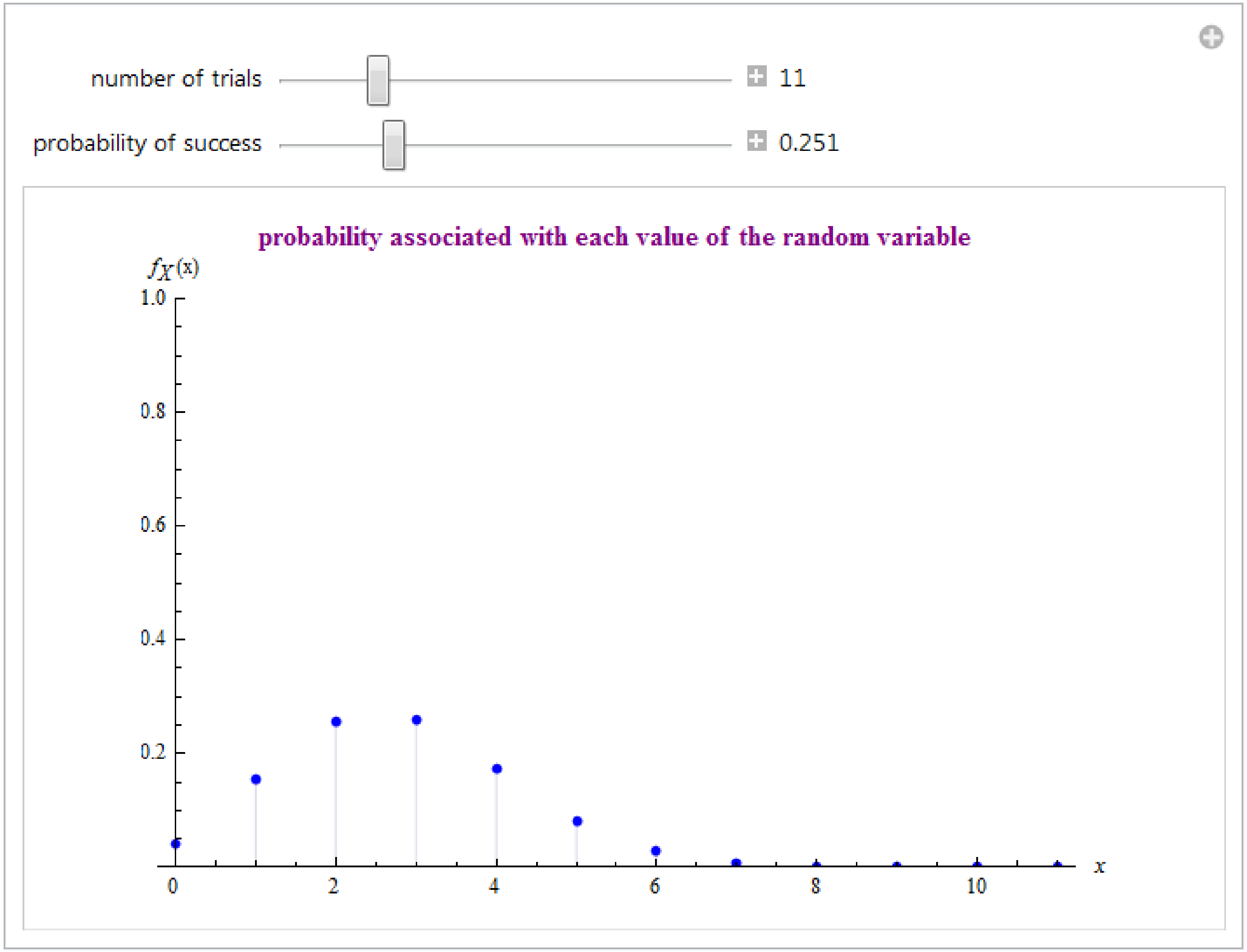
- Dizemos que uma variável X segue o modelo Bernoulli se atribui 0 ou 1 à probabilidade de ocorrência de fracasso ou sucesso, respectivamente (ensaio de Bernoulli). Com p representando a probabilidade de sucesso, $0 \leq p \leq 1$, sua função discreta de probabilidade é dada por

$$P(X = x) = p^x (1 - p)^{1-x},\ x = 0,\ 1.$$

Modelo Binomial

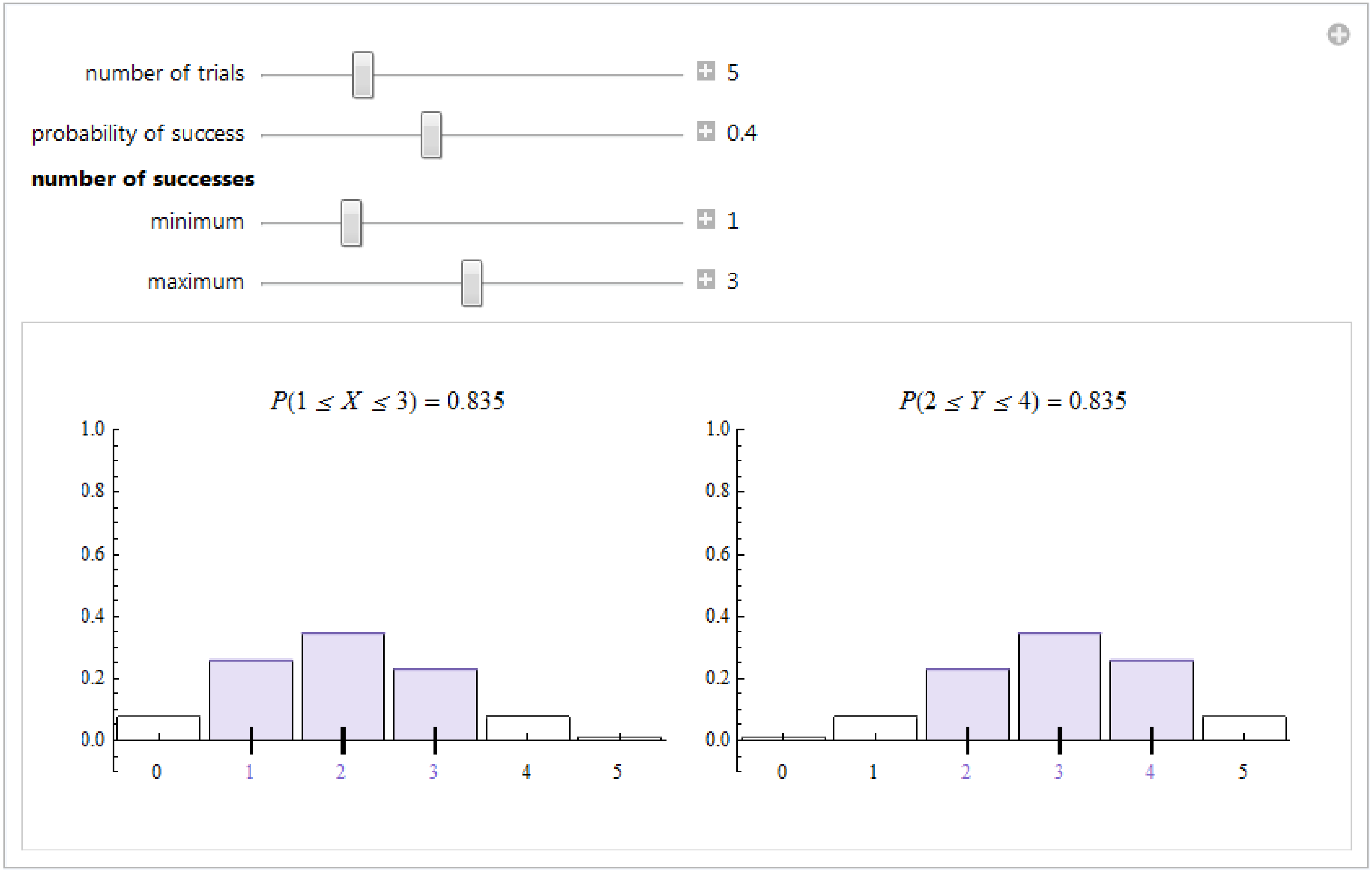
- Considere a repetição de n ensaios de Bernoulli independentes, e todos com a mesma probabilidade de sucesso p. A variável aleatória que conta o número total de sucessos é denominada Binomial com parâmetros n e p e sua função de probabilidade é dada por

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k} \therefore X \sim b(n,p).$$



"Binomial Probability Distribution" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/BinomialProbabilityDistribution/



"Successes and Failures in a Run of Bernoulli Trials" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/SuccessesAndFailuresInARunOfBernoulliTrials/

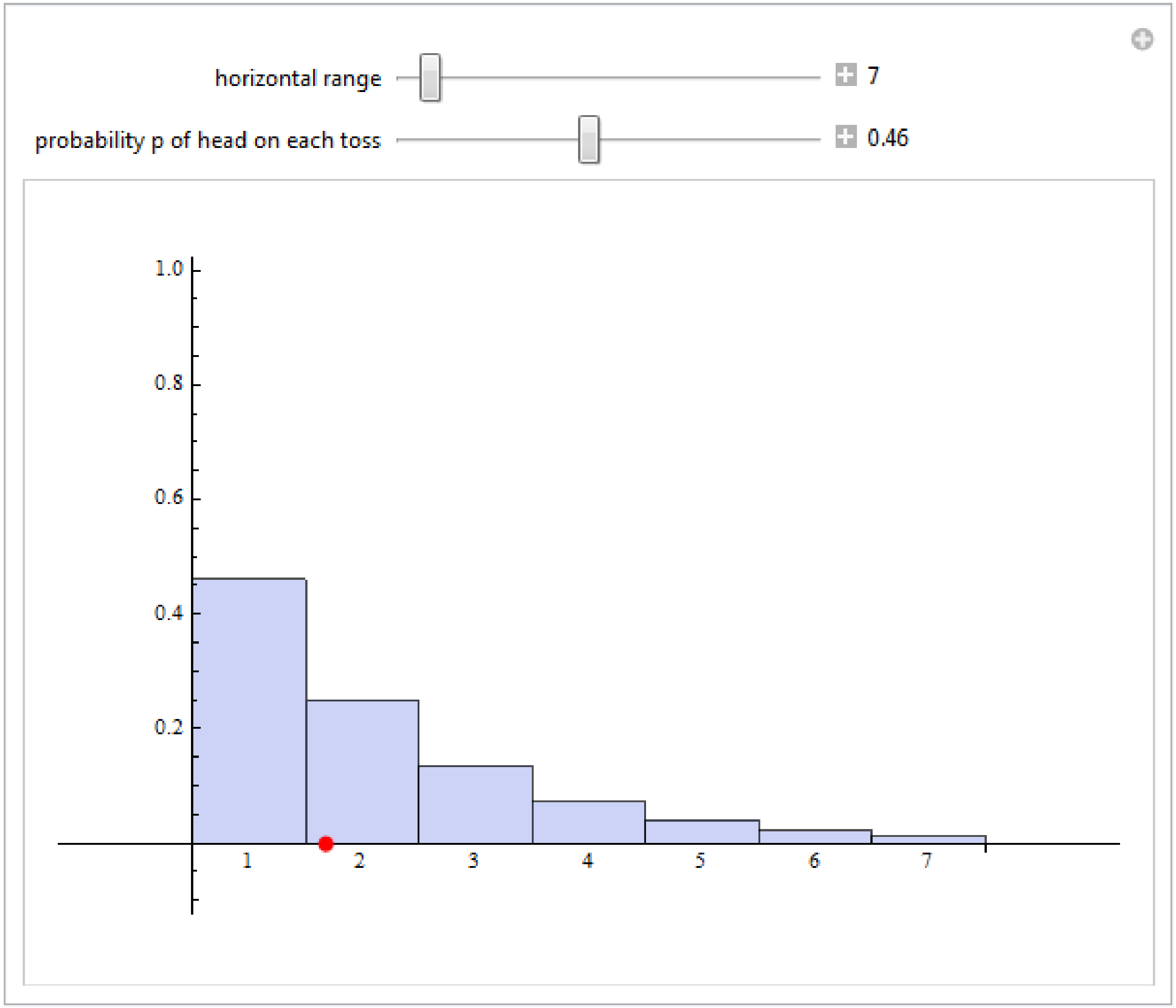
Modelo Geométrico

- Dizemos que uma variável aleatória X tem Distribuição Geométrica de parâmetro p se a sua função de probabilidade tem a forma

$$P(X=k)=p(1-p)^k,\ 0\leq p\leq 1\ e\ k=0,\ 1,\ 2,\cdots.$$

$$\downarrow$$

Interpretando p como a probabilidade de sucesso, a distribuição geométrica pode ser pensada como o número de ensaios de Bernoulli que precedem o primeiro sucesso.



“Geometric Distribution" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/GeometricDistribution/

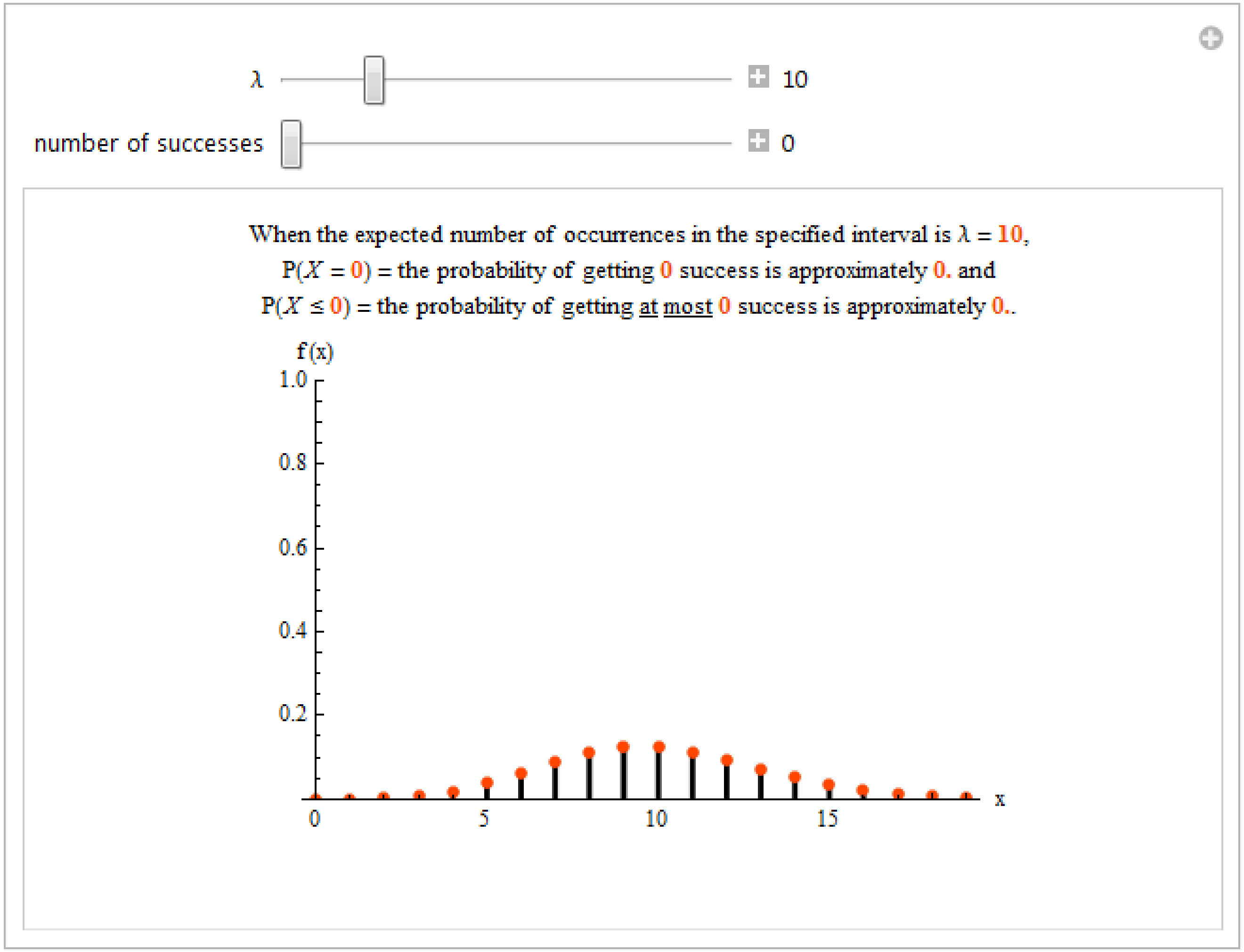
Modelo Poisson

- Uma variável aleatória X tem distribuição de Poisson com parâmetro $\lambda > 0$ (taxa de ocorrência) se a sua função de probabilidade é dada por

$$P(X = k) = \frac{e^{-\lambda}\lambda^{k}}{k!},\ k = 0,\ 1,\ 2,\cdots \therefore X \sim Po(\lambda).$$

$$\downarrow$$

O modelo Poisson tem sido muito utilizado em experimentos físicos e biológicos e, nesses casos, λ é a frequência média ou esperada de ocorrências num determinado intervalo de tempo.



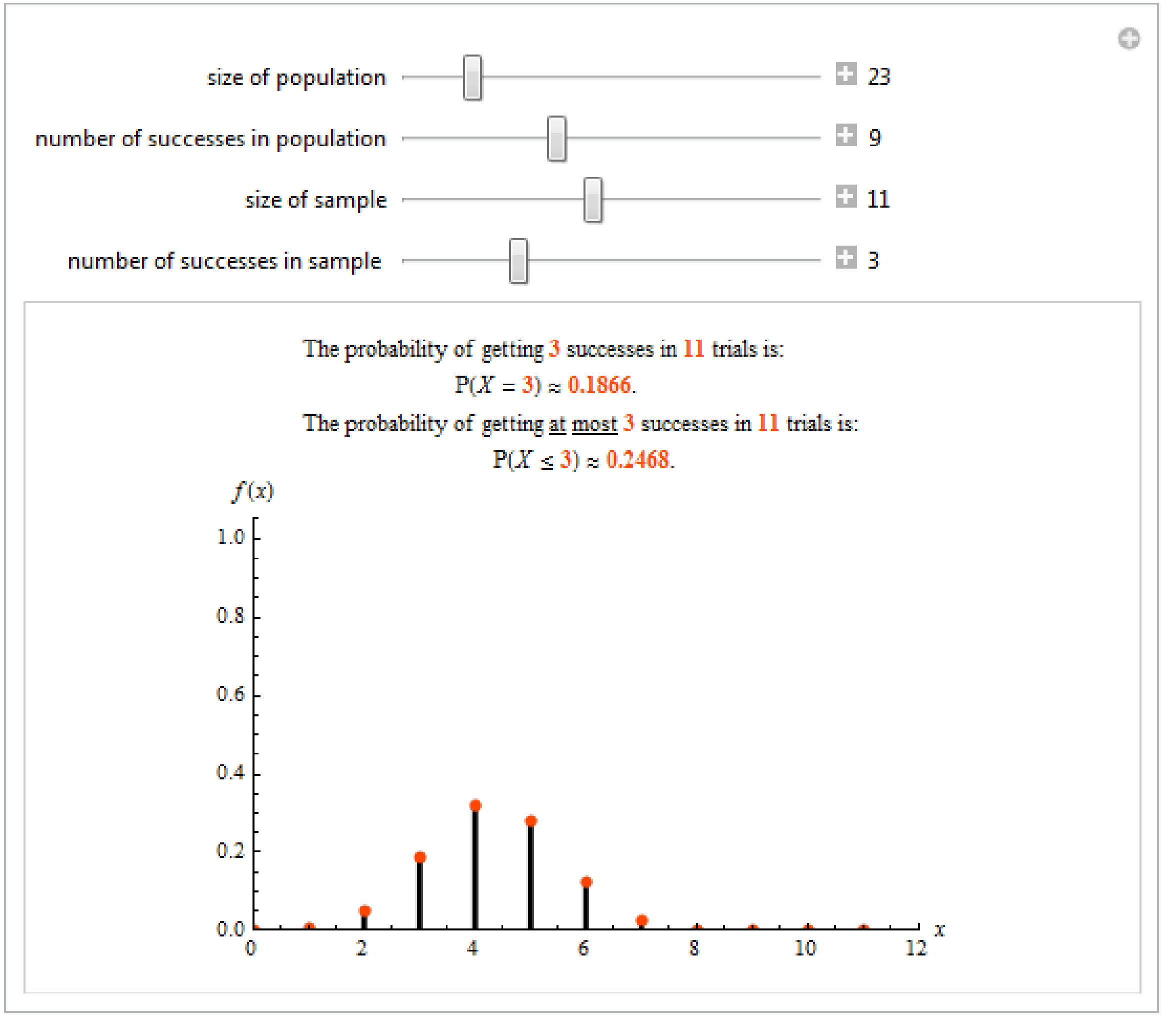
“Individual and Cumulative Poisson Probabilities" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/IndividualAndCumulativePoissonProbabilities/

Modelo Hipergeométrico

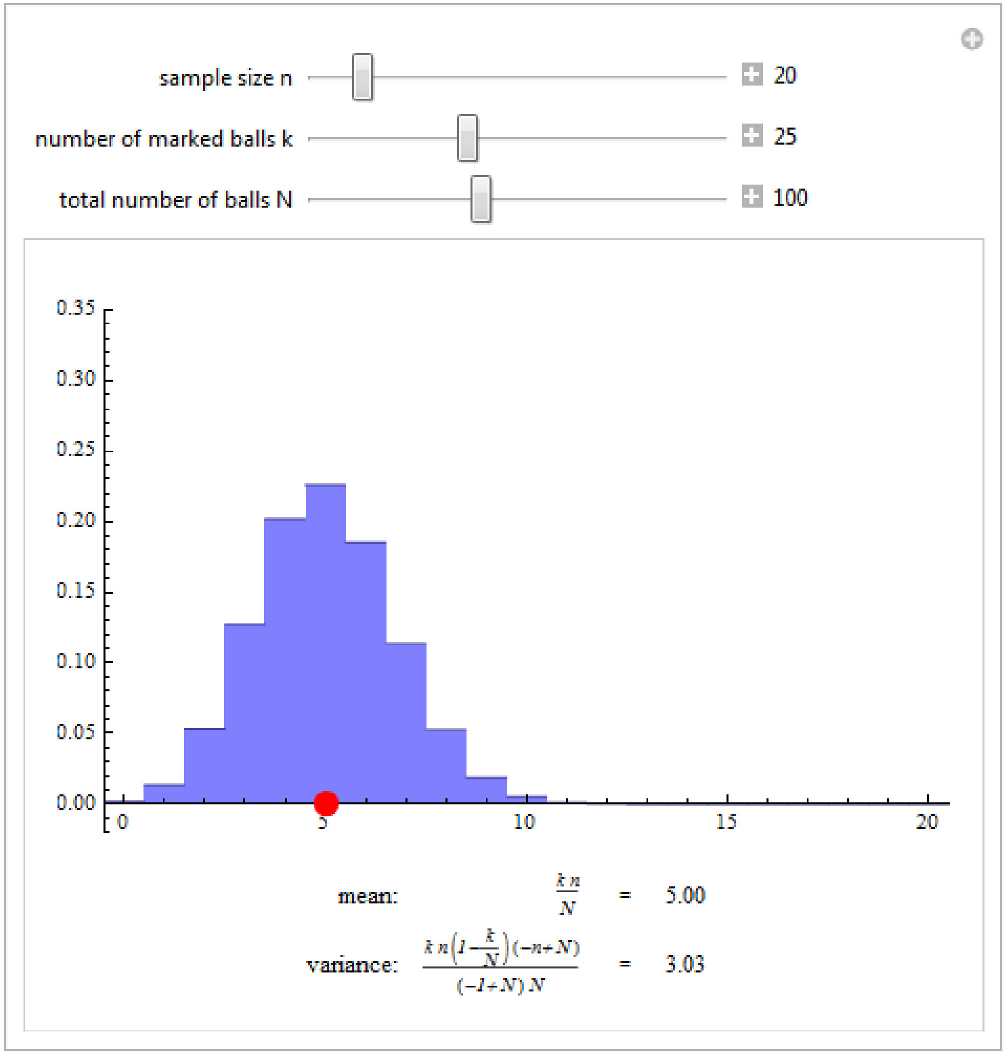
- Considere um conjunto de n objetos dos quais m são do tipo I e $(n - m)$ são do tipo II. Para um sorteio de r objetos $(r < n)$, feito ao acaso e sem reposição, a variável aleatória X que conta o número de objetos de tipo I selecionados segue o modelo Hipergeométrico, e sua função de probabilidade é dada pela expressão

$$P(X=k)=\frac{\binom{m}{k}\binom{n-m}{r-k}}{\binom{n}{r}},\ k=max(0,r-(n-m)),\cdots,min(r,m).$$



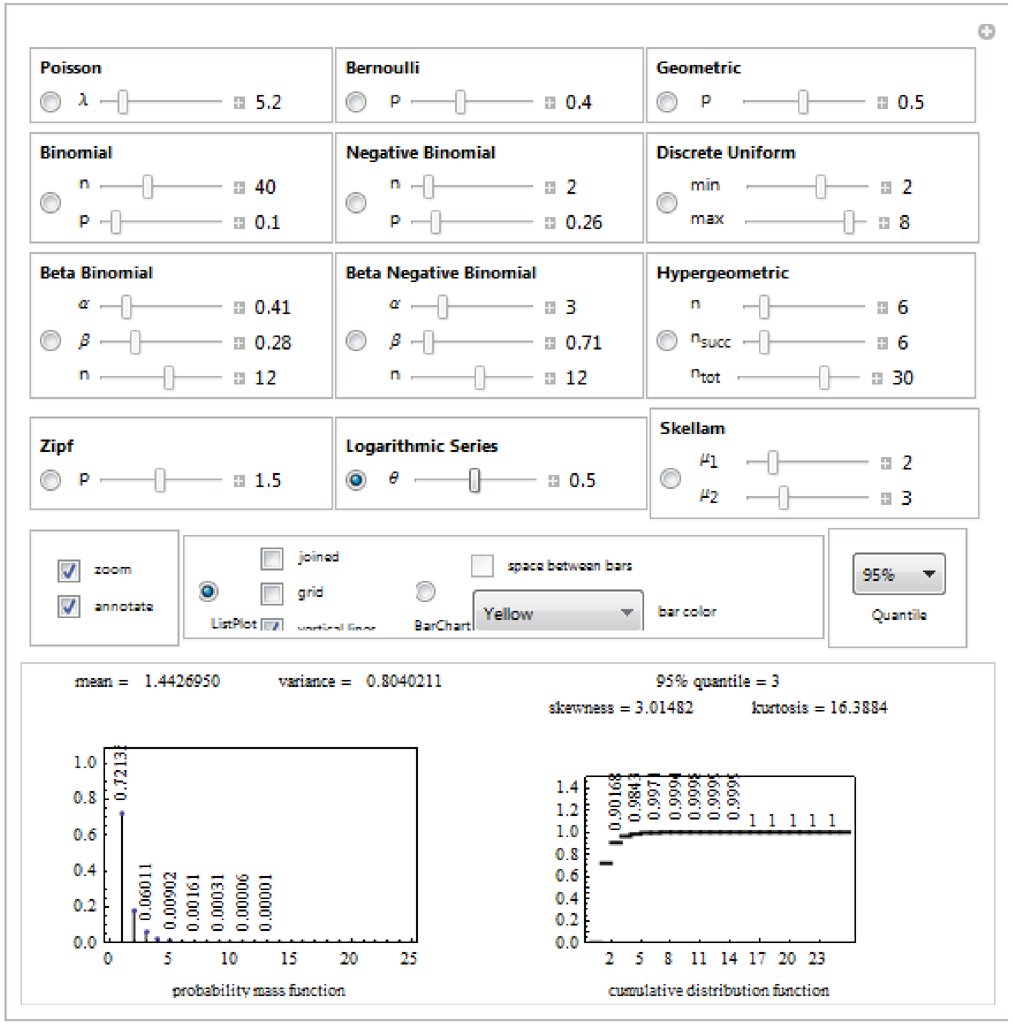
"Computing Individual and Cumulative Hypergeometric Probabilities" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/ComputingIndividualAndCumulativeHypergeometricProbabilities/



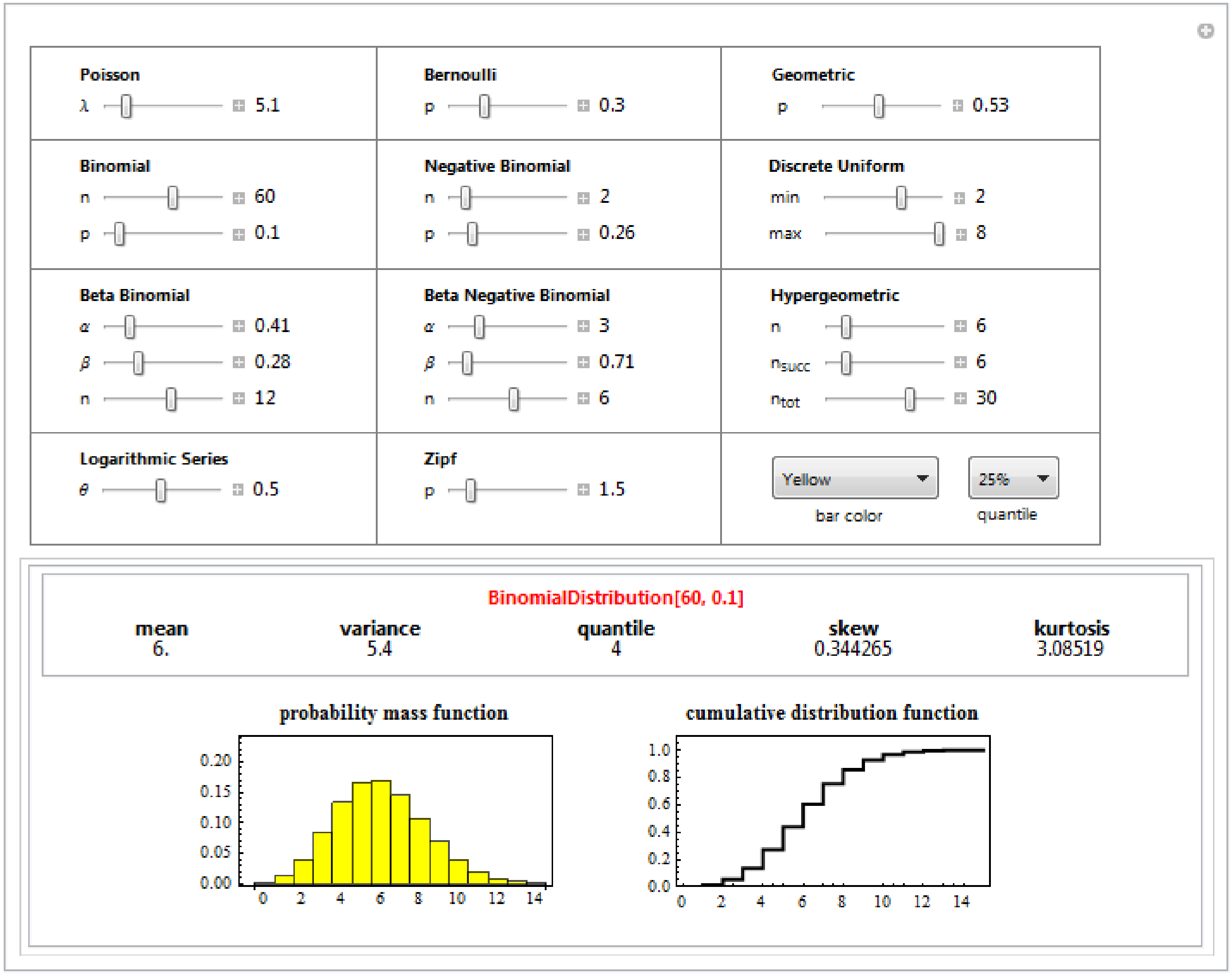
"The Hypergeometric Distribution" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/TheHypergeometricDistribution/



"Illustrating the Use of Discrete Distributions" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/IllustratingTheUseOfDiscreteDistributions/



"Mathematica 7's Discrete Distributions" from the Wolfram Demonstrations Project

http://demonstrations.wolfram.com/Mathematica7sDiscreteDistributions/

Importante

- Encerramos o Capítulo 3 – Variáveis Aleatórias Discretas...
- Agora trabalharemos com mais alguns exemplos do livro...
- Vocês já podem fazer os exercícios das seções 3.1 a 3.4! Trabalharemos com eles na próxima aula!

Ainda há Tempo?

- Vamos brincar um pouco com o R, utilizando o R *Commander*:
  - Gerar gráficos das diferentes distribuições.
  - Determinar probabilidades.
  - E muito mais!