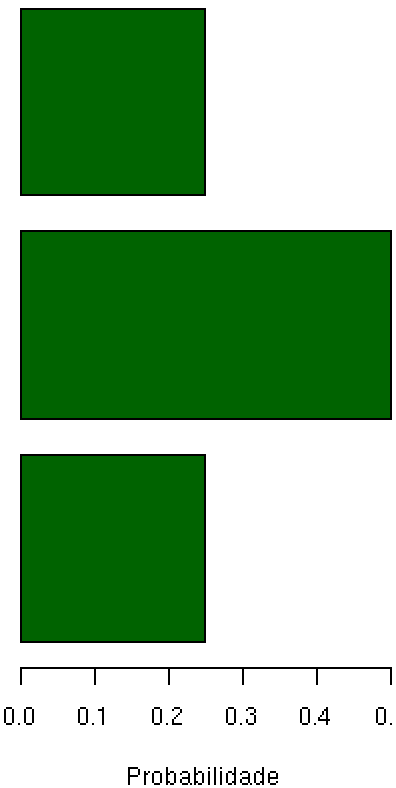
BIE5781

Aula 2

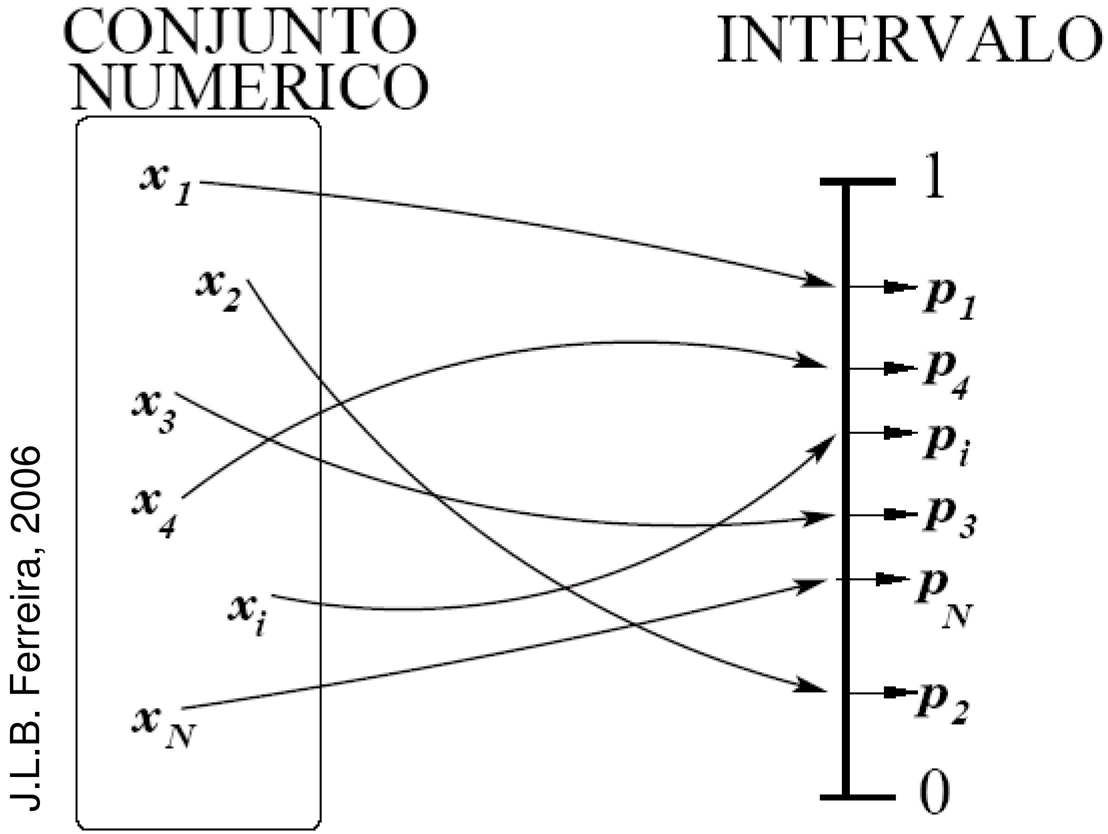
Variáveis Aleatórias Discretas

Variável Aleatória

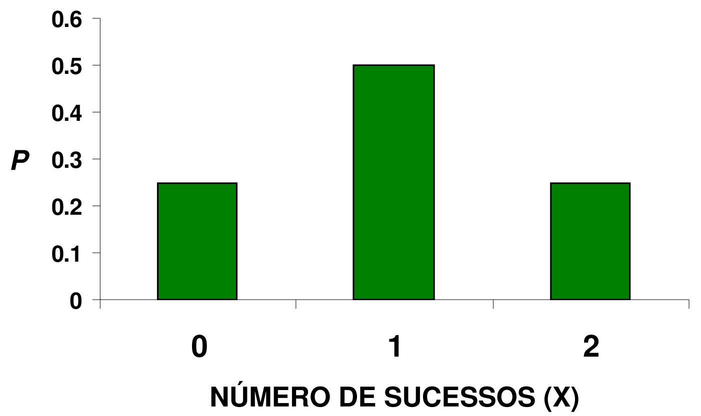
SUCESSOS	COMBINAÇÕES POSSÍVEIS	P
0		1/4
1	 ou 	2/4
2		1/4



Funções de Probabilidade



$$f(x) = P(X = x_i)$$



Função de Probabilidade

$$f(x) = \binom{\boldsymbol{n}}{x} \boldsymbol{p}^x (1 - \boldsymbol{p})^{\boldsymbol{n} - x}$$

Em negrito: parâmetros da distribuição

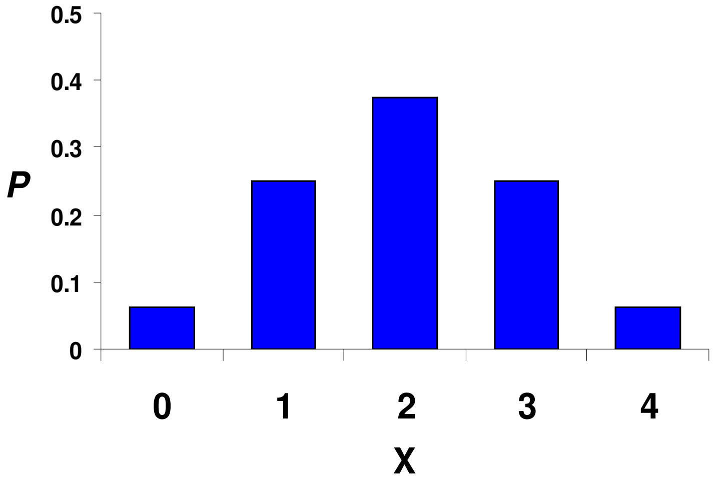
x	**f(x) = P(X = x)**
0	$1 \times 0{,}5^0 \times (1 - 0{,}5)^{2-0} = 1 \times 1 \times 0{,}25 = 0{,}25$
1	$2 \times 0{,}5^1 \times (1 - 0{,}5)^{2-1} = 2 \times 0{,}5 \times 0{,}5 = 0{,}50$
2	$1 \times 0{,}5^2 \times (1 - 0{,}5)^{2-2} = 1 \times 0{,}25 \times 1 = 0{,}25$

Esperança e Variância

X	*P*	X x *P*	$(X - E(X))^2$	$(X - E(X))^2$ x *P*
0	0,0625	0	4	0,25
1	0,250	0,25	1	0,25
2	0,375	0,75	0	0
3	0,250	0,75	1	0,25
4	0,0625	0,25	4	0,25
		2		1

$$E[X] = \sum x_i P_i$$

$$VAR[X] = E[(x - E[X])^2]$$



RESUMINDO ...

Variável aleatória = distribuição de probabilidades = função de probabilidade.

A distribuição de probabilidades é uma função que estabelece uma relação entre um conjunto de valores e o intervalo de probabilidades [0,1].

Cada distribuição têm parâmetros próprios, que a definem.

Média e variância são valores esperados, deduzidos da distribuição.

Variável Bernoulli

Variável discreta com apenas dois valores possíveis, com uma probabilidade p de assumir um dos valores (e portanto probabilidade 1-p de assumir o outro.

Exemplos:

- cara/coroa em um lançamento de moeda
- menino/menina em um nascimento
- sobreviveu/morreu em um período de tempo

Variável Bernoulli

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

$$E[X] = p$$

$$VAR[X] = pq$$

p = probabilidade de sucesso

Variável Binomial

Número de sucessos em **N** experimentos de Bernoulli.

Exemplos:

- Número de caras em 10 lançamentos de moeda
- Número de meninas em famílias com 5 filhos
- Número de sobreviventes entre N peixes em um aquário após aplicação de herbicida (dose-resposta)

Variável Binomial

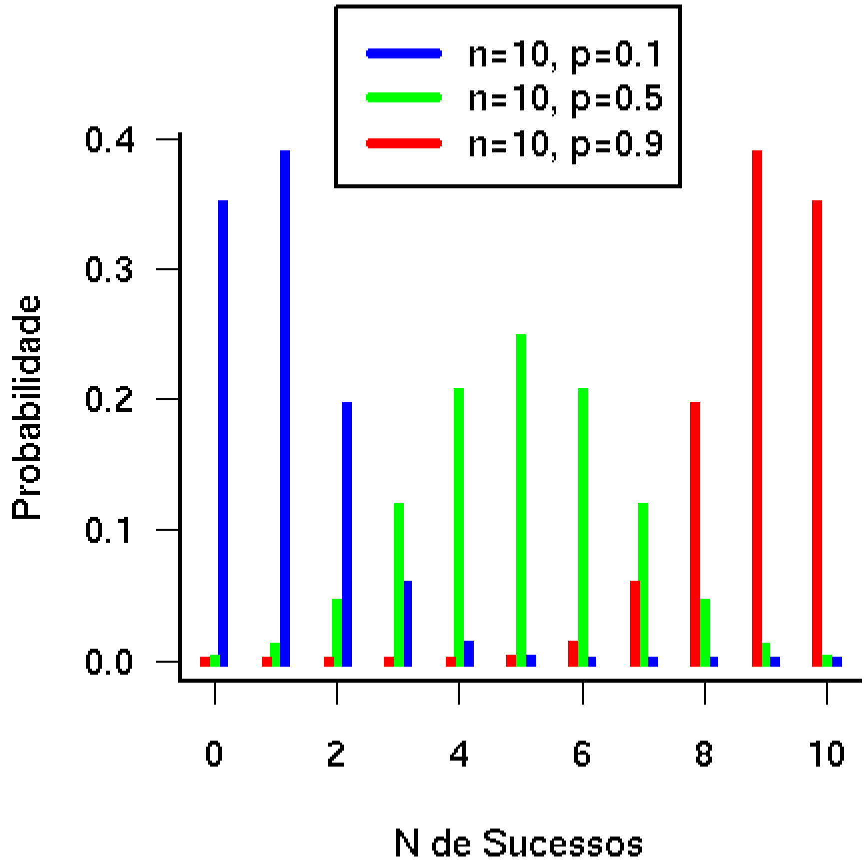
$$f(x) = \frac{N!}{N!(N-x)!} p^x (1-p)^{N-x}$$

N = número de tentativas (inteiro positivo)

p = probabilidade de sucesso por tentativa (0 a 1)

$$E[X] = Np$$

$$VAR[X] = Np(1-p)$$



Variável Poisson

Número de eventos em uma unidade fixa de tempo ou espaço, dada uma taxa de ocorrência por unidade.

Exemplos:

- Árvores por parcela
- Presas capturadas por intervalo de tempo
- Bombas por quarteirão*

* R.D. Clarke, An Application of the Poisson Distribution, Journal of the Institute of Actuaries, vol. 72 (1946), p. 481.

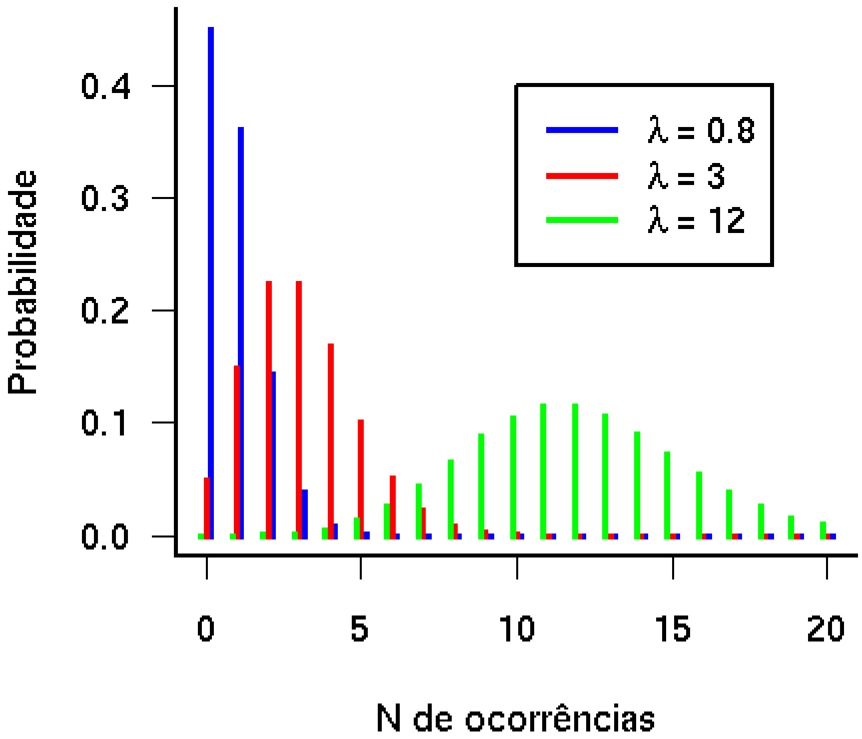
Variável Poisson

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

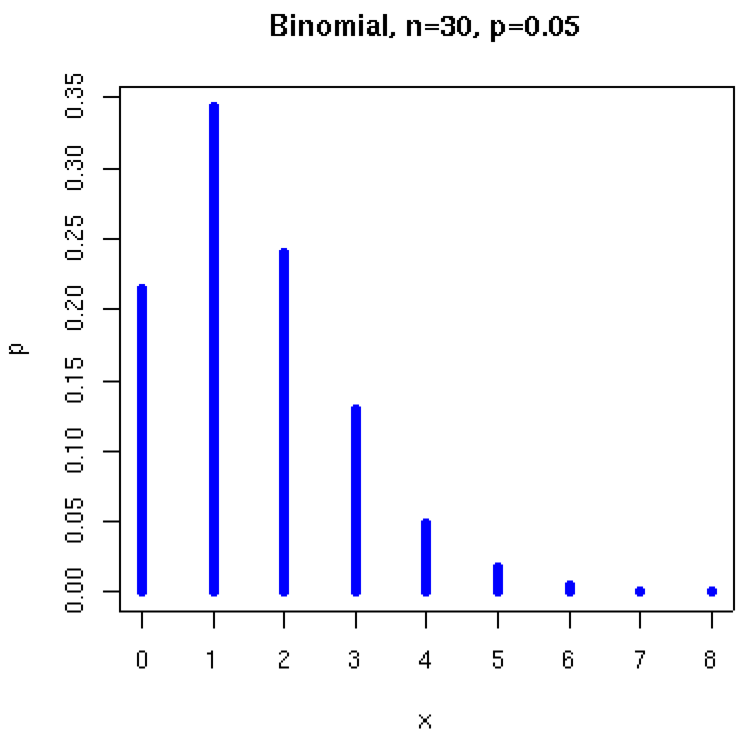
λ = taxa de ocorrência por unidade de espaço ou tempo (real positivo)

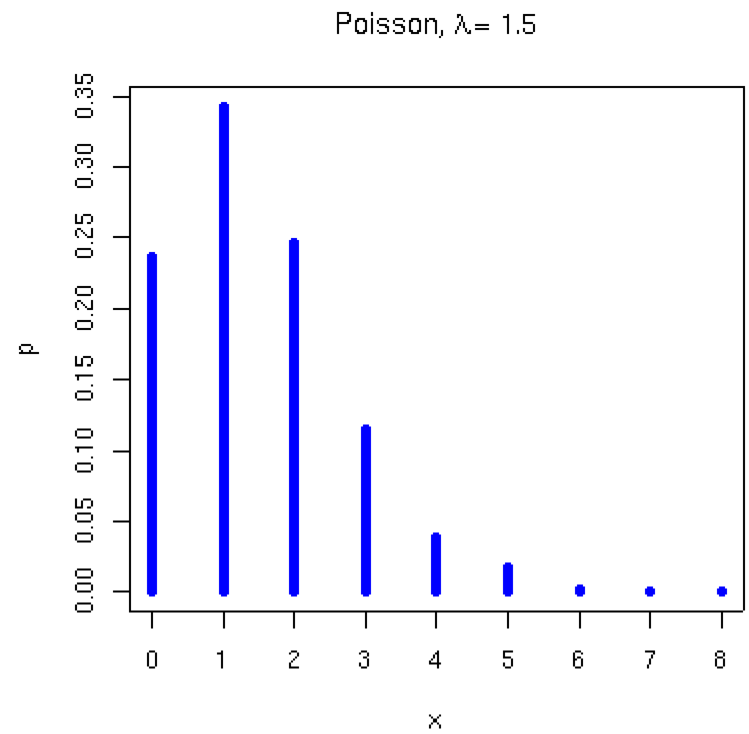
$$E[X] = \lambda$$

$$VAR[X] = \lambda$$



Relações entre Distribuições - 1





Binomial $\xrightarrow{\textbf{n alto, p baixo}}$ Poisson

Variável Geométrica

- Número de fracassos até o primeiro sucesso em uma série de experimentos de Bernoulli
- Exemplos:
  - Sobreviventes de uma coorte após cada estação
  - Passagem para um estádio fenológico ou reprodutivo a cada estação
  - Ocorrência de divórcio a cada carnaval

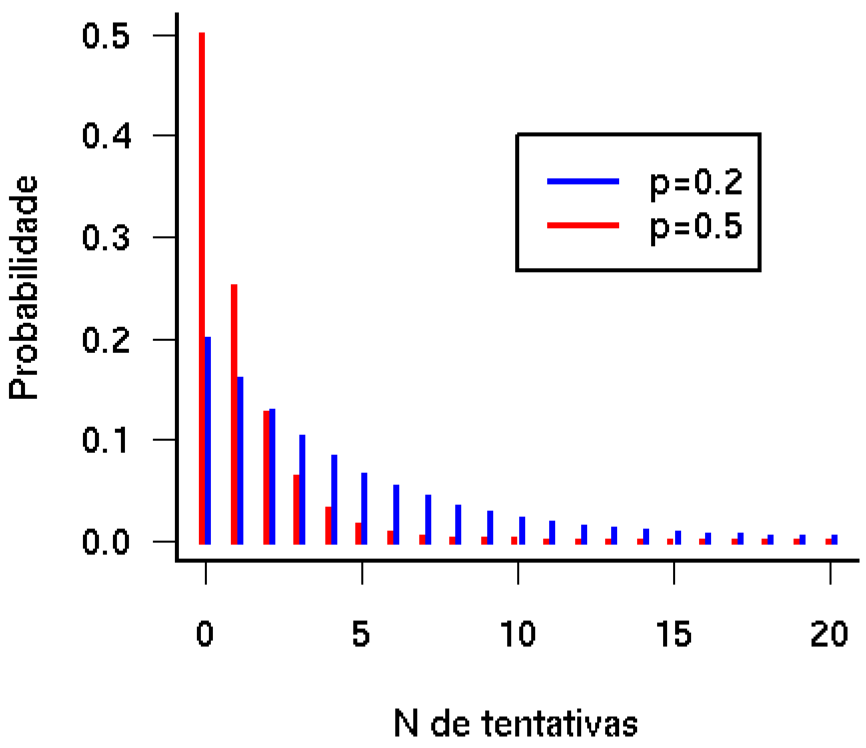
Variável Geométrica

$$f(x) = p(1-p)^x$$

p = probabilidade de sucesso por tentativa (0 a 1)

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

$$VAR[X] = \frac{1-p}{p^2}$$



Variável Binomial Negativa

Número de fracassos até o N sucessos em uma série de experimentos de Bernoulli.

Exemplos:

- Número de tentativas para se obter duas caras
- Em biologia mais usada como modelo heurístico para eventos agregados no tempo ou espaço (segue)

Distribuição Binomial Negativa

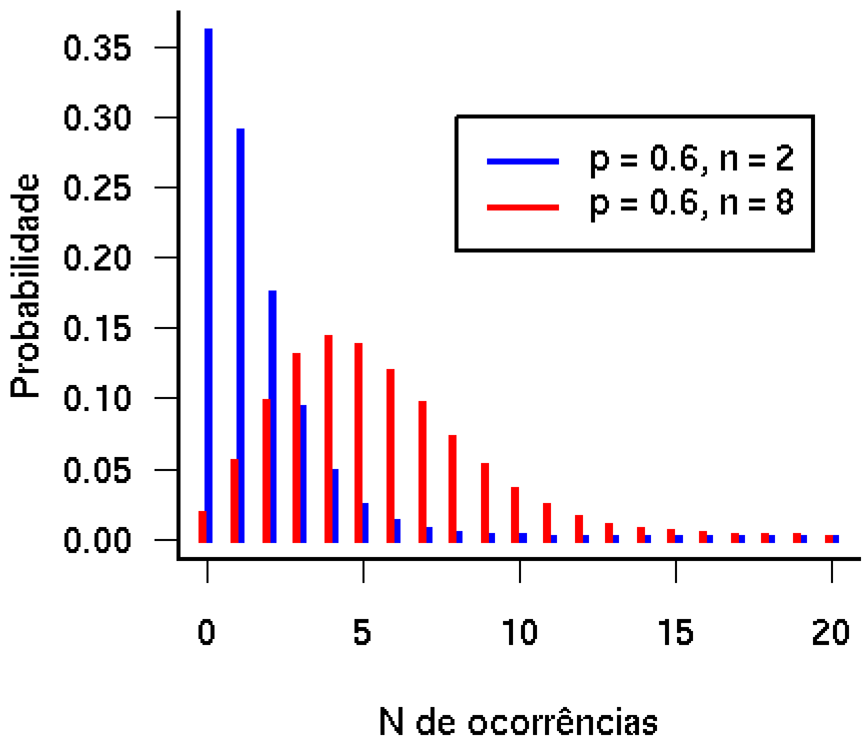
$$f(x) = \frac{(n+x-1)!}{(n-1)!\,x!} p^n (1-p)^x$$

p = probabilidade de sucesso por tentativa (0 a 1)

n = número de sucessos a aguardar (inteiro positivo)

$$E[X] = \frac{n(1-p)}{p}$$

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



Relações entre Distribuições - 2

$$\boxed{f(x) = \frac{(n+x-1)!}{(n-1)!\,x!}\,p^n(1-p)^x}$$

$$n = 1 \;\rightarrow\; \frac{(1+x-1)!}{(1-1)!\,x!}\,p^1(1-p)^x \;=\; p(1-p)^x$$

$$\text{Binomial Negativa} \xrightarrow{\textbf{n =1}} \text{Geométrica}$$

Variável Binomial Negativa

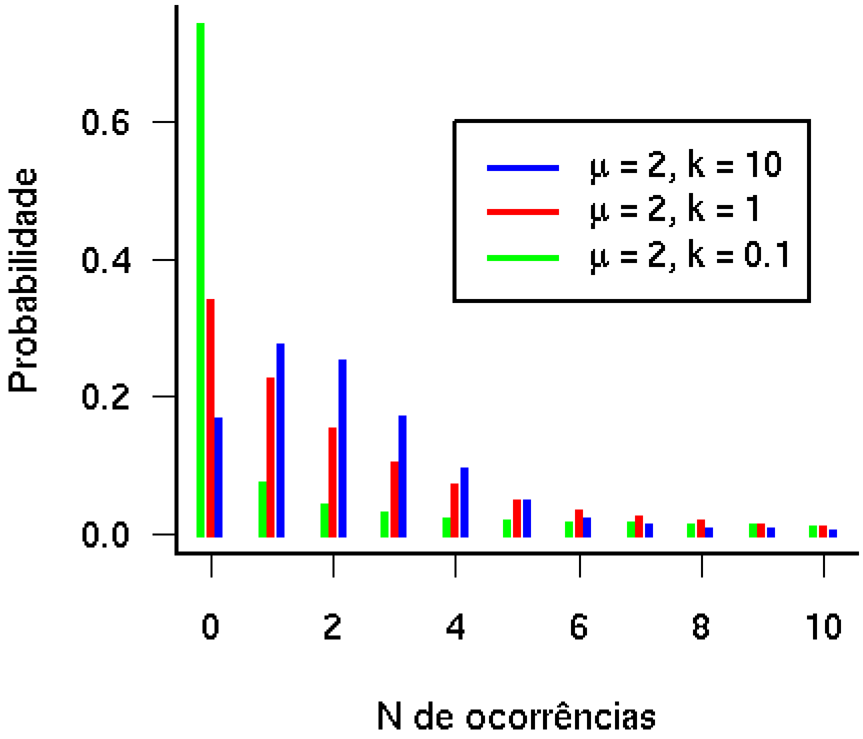
$$f(x) = \frac{\Gamma(k+x)}{\Gamma(k)\,x!} \cdot \left(\frac{k}{k+\mu}\right)^k \cdot \left(\frac{\mu}{k+\mu}\right)^x$$

μ = média de ocorrências (real positivo)

k = índice de dispersão (real positivo)

$$E[X] = \mu$$

$$VAR[X] = \mu + \frac{\mu^2}{k}$$



Variável Binomial Negativa

Número de eventos em uma unidade fixa de tempo ou espaço, dada uma taxa de ocorrência por unidade, e uma variância maior do que a Poisson (agregação)

Exemplos:

- Árvores por parcela
- Capturas por armadilha
- Bombas por continente

RESUMINDO ...

Distribuições de probabilidade discretas associam valores às suas probabilidades (PMF).

Distribuições de probabilidade contínuas associam valores às suas densidades de probabilidade (PDF).

Algumas distribuições são casos especiais de outras.

Algumas distribuições são casos-limite de outras.

Muitas distribuições discretas têm uma correspondente contínua.

RESUMINDO ...

O valor dos parâmetros determina a forma das distribuições.

Muitas distribuições podem ser generalizadas para aplicações diferentes daquelas para quais foram propostas.

Uma distribuição pode ser re-parametrizada.

Média e variância <u>não são</u> parâmetros da distribuição, embora possam ser expressas como funções destes.

Em muitas distribuições, média e variância estão correlacionadas.

Distribuições de Probabilidade no R

d`[distr]` = densidade probabilística (pdf)
p`[distr]` = probabilidade acumulada (cdf)
q`[distr]` = quantil
r`[distr]` = sorteio de valores

DISTRIBUIÇÕES ESTATÍSTICAS NO R

Distribuição	Nome no R	Parâmetros[1)]
beta	beta	shape1, shape2, ncp
binomial	binom	size, prob
Cauchy	cauchy	location, scale
qui-quadrado	chisq	df, ncp
exponential	exp	rate
F	f	df1, df2, ncp
gamma	gamma	shape, scale
geométrica	geom	prob
hypergeométrica	hyper	m, n, k
log-normal	lnorm	meanlog, sdlog
logística	logis	location, scale
binomial negativa	nbinom	size, prob
normal	norm	mean, sd
Poisson	pois	lambda
t de Student	t	df, ncp
uniforme	unif	min, max
Weibull	weibull	shape, scale
Wilcoxon	wilcox	m, n

http://lmq.esalq.usp.br/wiki/doku.php?id=biometria:r-tutor:

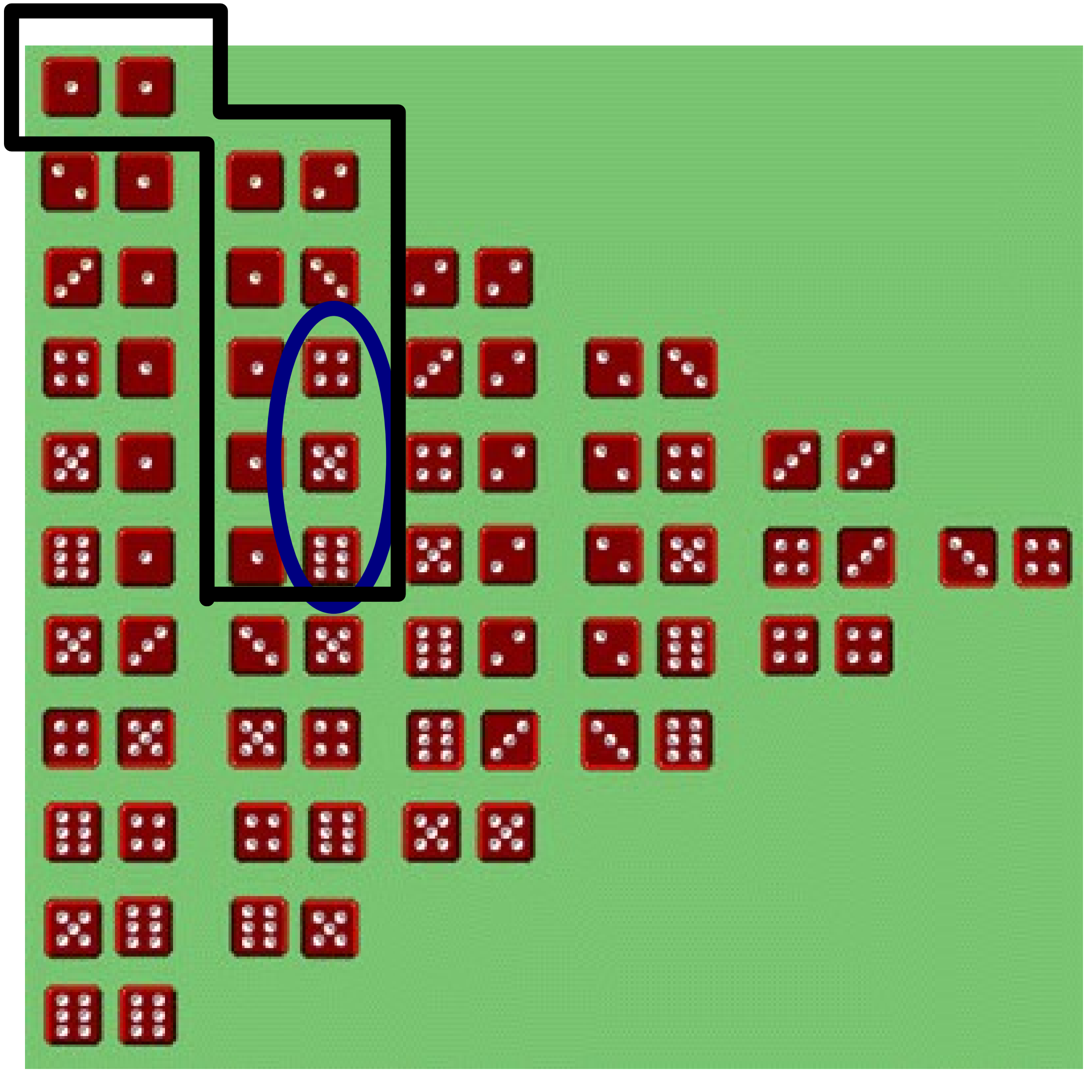
Probabilidades

A probabilidade de eventos excludentes é a soma das probabilidades de cada um.

A soma das probabilidades de todos os eventos possíveis é um (consequência da anterior).

A probabilidade de eventos independentes é o produto de suas probabilidades.

Probabilidade Condicional



$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Qual a probabilidade de um valor maior que 4 se o primeiro lançamento deu 1?

$$\frac{3/35}{6/35} = 0{,}5$$

Teorema de Bayes

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

$$P(H|D) = \frac{P(D|H)P(H)}{\sum P(D|H_i)P(H_i)}$$

Teorema de Bayes
Exemplo

H1: O baralho tem só ases de espadas
H2: O baralho é um baralho comum
D : sorteio de uma carta e ela é um ás de espada

$$P(D|H_1) = 1{,}0$$

$$P(H_1) = 0{,}5$$

$$\sum P(D|H_i)P(H_i) = 1{,}0 \times 0{,}5 + 1/52 \times 0{,}5 = 0{,}5096$$

$$P(H_1|D) = \frac{1{,}0 \times 0{,}5}{0{,}5096} = 0{,}981$$

Teorema de Bayes
Outro Exemplo

- 1% da população de um país tem uma certa doença
- O teste para a doença produz falsos negativos em 0.1% de suas aplicações
- O teste para a doença produz falsos positivos em 0.5% de suas aplicações
- Uma pessoa tomada ao acaso teve resultado positivo
- Dada esta observação, qual a probabilidade dessa pessoa estar doente?

Teorema de Bayes
Outro Exemplo

A pessoa não está doente $P(H_1)=0{,}99$

A pessoa está doente $P(H_2)=0{,}01$

A pessoa não está doente, dado um teste positivo $P(D|H_1)=0{,}001$

A pessoa está doente, dado um teste positivo $P(D|H_2)=0{,}999$

$$\sum P(D|H_i)P(H_i)=0{,}99\times 0{,}001+0{,}01\times 0{,}999=0{,}01098$$

$$P(H_1|D)=\frac{0{,}99\times 0{,}001}{0{,}01098}=0{,}0902$$

Teorema de Bayes

Razão de verossimilhanças

$$P(H|D) = \frac{P(D|H)}{\sum P(D|H_i)} \cdot \frac{P(H)}{\sum P(H_i)}$$

Razão de risco
(odds-ratio)

A razão de verossimilhanças é o que modifica as chances relativas de uma hipótese estar correta, em relação às demais.

Leituras Recomendadas

Otto, S. P. & Day, T. (2007). A biologist's guide to mathematical modelling in ecology and evolution. Princenton, Princenton University Press. (Primer 3)

Bolker, B. (2008). Ecological Models and Data in R. Princeton, Princeton University Press. (Cap.4)