ME951 - Estatística e Probabilidade I

Parte 10

Notas de aula de ME414 produzidas pelos professores **Samara Kiihl**, **Tatiana Benaglia** e **Benilton Carvalho** modificadas e alteradas pela Profa. **Larissa Avila Matos**

Principais Modelos Discretos

Distribuição Uniforme Discreta

Uniforme Discreta

Exemplo: Uma rifa tem 100 bilhetes numerados de 1 a 100. Tenho 5 bilhetes consecutivos numerados de 21 a 25, e meu colega tem outros 5 bilhetes, com os números 1, 11, 29, 68 e 93.

Quem tem maior probabilidade de ser sorteado?

- espalhar os números é a melhor forma de ganhar o sorteio?
- assumindo honestidade da rifa, todos os números têm a mesma probabilidade de ocorrência, com $\frac{1}{100}$ para cada um.
- como eu e meu colega temos 5 bilhetes, temos a mesma probabilidade de ganhar a rifa: $\frac{5}{100} = \frac{1}{20}$.
- assim, a probabilidade de ganhar depende somente da quantidade de bilhetes que se tem na mão, independente da numeração.

Uniforme Discreta

- A v.a. discreta X, assumindo valores $x_1, x_2, ..., x_k$, segue uma distribuição uniforme discreta se, e somente se, cada valor possível tem a mesma probabilidade de ocorrer, isto é,

$$P(X = x_i) = p(x_i) = \frac{1}{k}, \quad \forall\, 1 \leq i \leq k.$$

Exemplo: lançamento de um dado honesto de 6 faces

X	1	2	3	4	5	6
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Uniforme Discreta

A média e a variância são dadas por:

$$\mathbb{E}(X) = \sum_{i=1}^{k} x_i P(X = x_i) = \frac{1}{k} \sum_{i=1}^{k} x_i,$$

$$Var(X) = \sum_{i=1}^{k} (x_i - \mathbb{E}(X))^2 P(X = x_i) = \frac{1}{k} \sum_{i=1}^{k} (x_i - \mathbb{E}(X))^2$$

ou

$$Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{1}{k} \left[\sum_{j=1}^{k} x_j^2 - \frac{1}{k} \left(\sum_{j=1}^{k} x_j \right)^2 \right].$$

Uniforme Discreta

Exemplo: lançamento de um dado honesto de 6 faces

X	1	2	3	4	5	6
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Calculando a esperança e a variância de X, temos

$\mathbb{E}(X) = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = 3.5$

$$Var(X) = \frac{1}{6}\left[(1+4+9+16+25+36) - \frac{1}{6}(21)^2\right] = \frac{1}{6} \times \frac{35}{2} = 2.92$$

Uniforme Discreta

Cálculo da função de distribuição acumulada (f.d.a.) de uma variável uniforme discreta:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} \frac{1}{k} = \frac{\#(x_i \leq x)}{k}$$

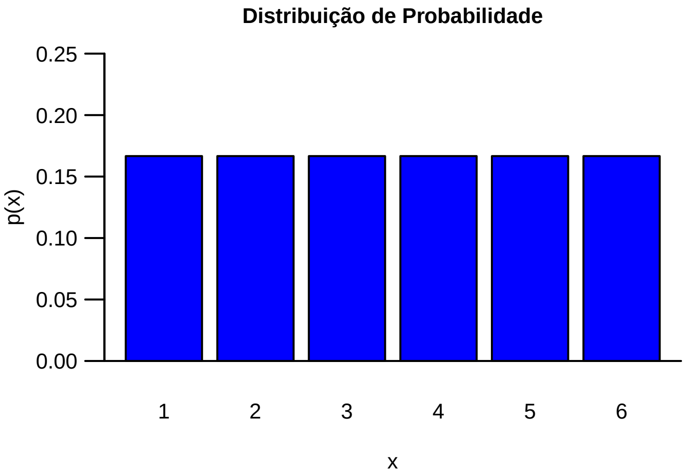
Exemplo: voltando ao exemplo do lançamento de um dado honesto de 6 faces

- $F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = \frac{2}{6}$
- $F(2.5) = P(X \leq 2.5) = P(X = 1) + P(X = 2) = \frac{2}{6}$

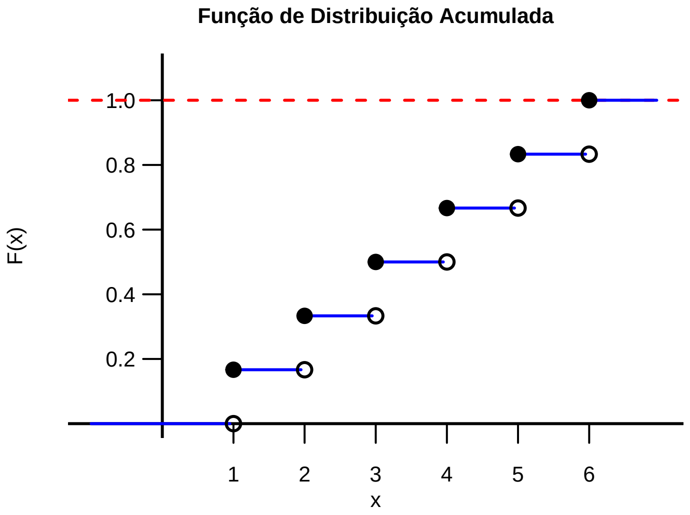
Uniforme Discreta - f.d.a

X	$p(x)$	$F(x)$
1	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$
3	$\frac{1}{6}$	$\frac{3}{6}$
4	$\frac{1}{6}$	$\frac{4}{6}$
5	$\frac{1}{6}$	$\frac{5}{6}$
6	$\frac{1}{6}$	$\frac{6}{6}$

Uniforme Discreta - Gráficos



Uniforme Discreta - Gráficos



Distribuição Bernoulli

Modelo Bernoulli

Em muitas aplicações, cada observação de um experimento aleatório é **binária**: tem apenas dois resultados possíveis.

Por exemplo, uma pessoa pode:

- aceitar ou recusar uma oferta de cartão de crédito de seu banco.
- votar sim ou não em uma assembléia.

Modelo Bernoulli

Considere um experimento aleatório com dois resultados possíveis: **sucesso** e **fracasso**.

Seja p a probabilidade de sucesso.

Exemplo: lançar uma moeda e verificar se cai cara ou coroa. Consideramos como sucesso, a obtenção de cara. Se a moeda é honesta, $p = 1/2$.

Esse tipo de experimento é conhecido como **Ensaio de Bernoulli**.

Modelo Bernoulli

Exemplo: lançar um dado, sendo “sucesso” o obtenção da face 6. Se o dado é honesto, a probabilidade de sucesso é $p = 1/6$.

Exemplo: Uma pessoa é escolhida ao acaso entre os moradores de uma cidade, e é perguntado a ela se concorda com um projeto.

As possíveis respostas são apenas “Sim” ou “Não”.

$$\Omega = \{\text{Sim}, \text{Não}\}$$

$$X = \begin{cases} 1, & \text{se a pessoa respondeu sim } (sucesso) \\ 0, & \text{caso contrário } (fracasso) \end{cases}$$

$$P(X = 1) = P(sucesso) = p \;\; \rightarrow \;\; P(X = 0) = P(fracasso) = 1 - p$$

Modelo Bernoulli

Seja X uma v.a. discreta assumindo apenas valores 0 e 1, onde $X = 1$ corresponde a sucesso e seja p a probabilidade de sucesso.

A distribuição de probabilidade de X é dada por:

$$P(X = x) = \begin{cases} p & \text{se } x = 1 \\ 1 - p & \text{se } x = 0 \end{cases}$$

Ou de forma equivalente, podemos escrever como:

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

Notação: $X \sim b(p)$

Modelo Bernoulli - Esperança e Variância

Se X é um v.a. Bernoulli, $X \sim b(p)$, então:

$$\mathbb{E}(X) = p \quad \text{e} \quad Var(X) = p(1-p)$$

Demonstração: Temos que,

$\mathbb{E}(X) = 0 \times (1-p) + 1 \times p = p$

e

$\mathbb{E}(X^2) = 0^2 \times (1-p) + 1^2 \times p = p$

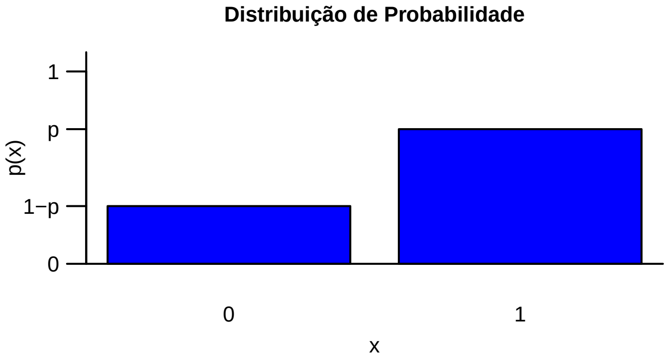
Então,

$$\begin{aligned} Var(X) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\ &= p - p^2 = p(1-p) \end{aligned}$$

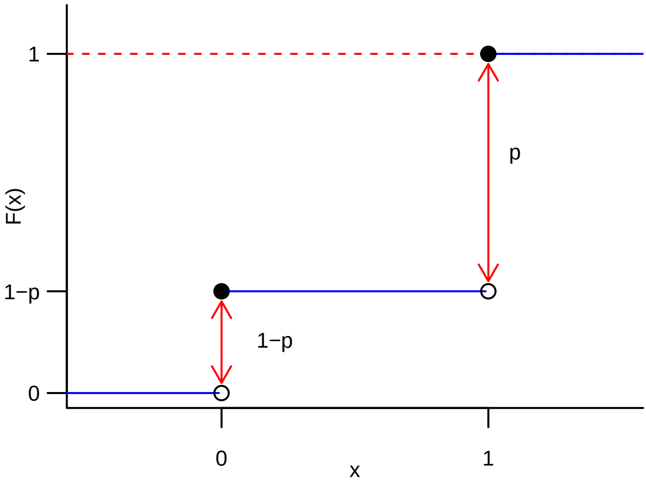
Modelo Bernoulli

A função de distribuição acumulada de uma v.a. Bernoulli é:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$



Função de Distribuição Acumulada



Modelo Bernoulli

Exemplo: lançamos um dado e consideramos sucesso a obtenção da face 5.

Supondo que o dado seja honesto, a probabilidade de sucesso é $p = 1/6$. Então:

$$\begin{aligned} P(X = x) &= \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{1-x} \qquad \text{para } x = 0, 1 \\ &= \begin{cases} 5/6 & \text{se } x = 0 \\ 1/6 & \text{se } x = 1 \end{cases} \end{aligned}$$

Encontre a esperança e variância de X.

$\mathbb{E}(X) = \frac{1}{6}$ $\qquad \mathbb{E}(X^2) = \frac{1}{6}$

$Var(X) = \frac{1}{6} - \frac{1}{36} = \frac{6-1}{36} = \frac{5}{36}$

Distribuição Binomial

Modelo Binomial

Ao obtermos uma amostra do experimento/fenômeno aleatório com observações binárias, podemos resumir os resultados usando o número ou a proporção de observações com o resultado de interesse.

*Sob certas condições, a v.a. X que conta o número de vezes que um resultado específico ocorreu, dentre dois possíveis, tem uma distribuição de probabilidade chamada **Binomial**.*

Considere um experimento aleatório com espaço amostral Ω e o evento A.

Vamos dizer que ocorreu sucesso se o evento A aconteceu. Se A não aconteceu ocorreu fracasso.

Repetimos o experimento n vezes, de forma independente.

Seja X o número de sucessos nos n experimentos.

Exemplo: vacinas

Sabe-se que a eficiência de uma vacina é de 80%.

Um grupo de 3 indivíduos é sorteado, dentre a população vacinada, e cada um é submetido a testes para averiguar se está imunizado.

Nesse caso, consideramos como sucesso a imunização.

$$X_i = \begin{cases} 1, & \text{indivíduo i está imunizado} \\ 0, & \text{caso contrário} \end{cases}$$

Pelo enunciado, sabe-se que $P(X_i = 1) = p = 0.8$.

Exemplo: vacinas

Os indivíduos 1, 2 e 3 são independentes.

As v.a.'s X_1, X_2 e X_3 são Bernoulli.

Se o interesse está em estudar X = número de indivíduos imunizados no grupo, X poderá assumir valores $\{0, 1, 2, 3\}$.

Note que $X = X_1 + X_2 + X_3$.

Exemplo: vacinas

evento	P(evento)	X
$X_1 = 0,\ X_2 = 0,\ X_3 = 0$	$(0.2)^3$	0
$X_1 = 1,\ X_2 = 0,\ X_3 = 0$	$0.8 \times (0.2)^2$	1
$X_1 = 0,\ X_2 = 1,\ X_3 = 0$	$0.8 \times (0.2)^2$	1
$X_1 = 0,\ X_2 = 0,\ X_3 = 1$	$0.8 \times (0.2)^2$	1
$X_1 = 1,\ X_2 = 1,\ X_3 = 0$	$(0.8)^2 \times 0.2$	2
$X_1 = 1,\ X_2 = 0,\ X_3 = 1$	$(0.8)^2 \times 0.2$	2
$X_1 = 0,\ X_2 = 1,\ X_3 = 1$	$(0.8)^2 \times 0.2$	2
$X_1 = 1,\ X_2 = 1,\ X_3 = 1$	$(0.8)^3$	3

Modelo Binomial

Assim, as probabilidades de cada valor possível de X são:

X	0	1	2	3
$P(X = x)$	$(0.2)^3$	$3 \times 0.8 \times (0.2)^2$	$3 \times (0.8)^2 \times 0.2$	$(0.8)^3$

E o comportamento de X é completamente determinado pela função:

$$P(X = x) = \binom{3}{x}(0.8)^x(0.2)^{3-x}, \qquad x = 0, 1, 2, 3$$

Modelo Binomial

Modelo Geral: Considere a repetição de n ensaios X_i Bernoulli independentes e todos com a mesma probabilidade de sucesso p.

A variável aleatória $X = X_1 + ... + X_n$ representa o total de sucessos e corresponde ao modelo Binomial com parâmetros n e p, ou seja, $X \sim Bin(n, p)$.

A probabilidade de se observar x é dada pela expressão geral:

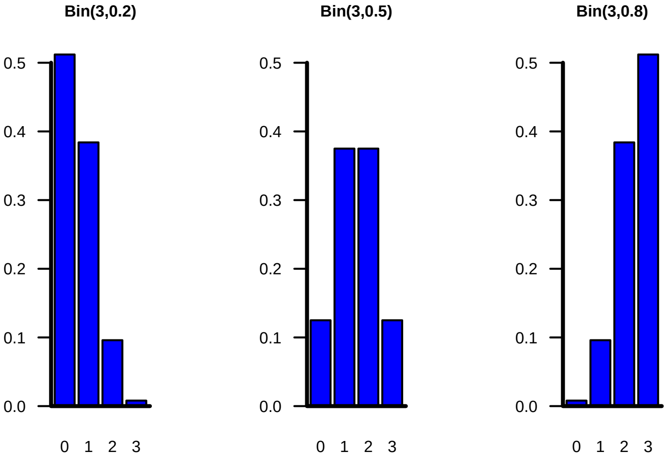
$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, ..., n$$

A esperança e variância de uma v.a. Binomial são dadas por:

$$\mathbb{E}(X) = np \qquad \text{e} \qquad Var(X) = np(1-p)$$

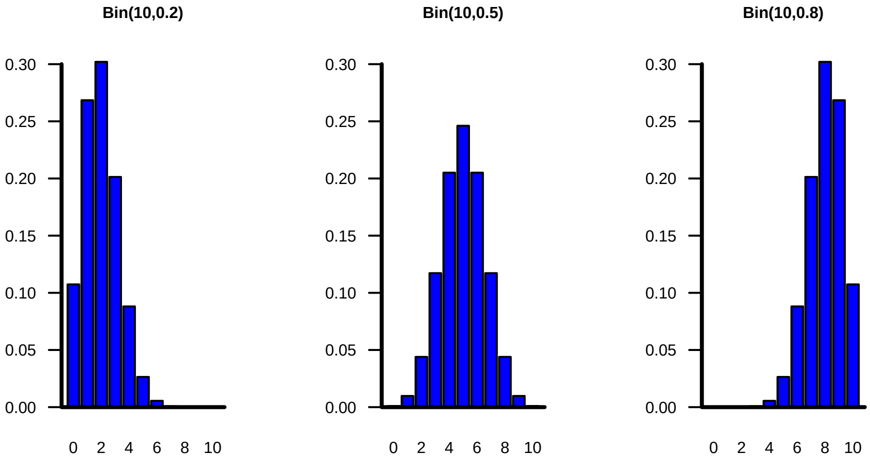
Modelo Binomial

Distribuição de probabilidade de uma $Bin(3, p)$, com $p = 0.2, 0.5$ e 0.8.



Modelo Binomial

Distribuição de probabilidade de uma $Bin(10, p)$, com $p = 0.2, 0.5$ e 0.8.



Exemplo: Vacina

No exemplo da vacina, temos então que o número de indíviduos imunizados segue uma distribuição Binomial com $n = 3$ e $p = 0.8$

$$X \sim Bin(3, 0.8)$$

Qual a probabilidade de que dentre os 3 indíviduos, nenhum tenha sido imunizado?

$$P(X = 0) = \binom{3}{0}(0.8)^0(0.2)^3 = 0.008$$

Encontre a esperança e variância.

$$\mathbb{E}(X) = 3 \times 0.8 = 2.4 \qquad \text{e} \qquad Var(X) = 3 \times 0.8 \times 0.2 = 0.48$$

Exemplo: Inspeção

Um inspetor de qualidade extrai uma amostra aleatória de 10 tubos armazenados num depósito onde, de acordo com os padrões de produção, se espera um total de 20% de tubos defeituosos.

Qual é a probabilidade de que não mais do que 2 tubos extraídos sejam defeituosos?

Se X denotar a variável "número de tubos defeituosos em 10 extrações independentes e aleatórias", qual o seu valor esperado? Qual a variância?

Exemplo: Inspeção

Note que a variável aleatória X = número de tubos defeituosos em 10 extrações tem distribuição binomial, com parâmetros $n = 10$ e $p = 0.2$.

Então, "não mais do que dois tubos defeituosos" é o evento $\{X \leq 2\}$.

Sabemos que, para $X \sim Bin(10, 0.2)$

$$P(X = x) = \binom{10}{x}(0.2)^x(0.8)^{10-x}, \qquad x = 0, 1, \ldots, 10$$

e que

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= (0.8)^{10} + 10(0.2)(0.8)^9 + 45(0.2)^2(0.8)^8 = 0.678 \end{aligned}$$

Exemplo: Inspeção

Se $X \sim Bin(n, p)$, então $\mathbb{E}(X) = np$ e $Var(X) = np(1-p)$

Então:

$\mathbb{E}(X) = 10(0.2) = 2$

$Var(X) = 10(0.2)(0.8) = 1.6$

Quando se encontram quatro ou mais tubos defeituosos, o processo de produção é interrompido para revisão. Qual é a probabilidade que isto aconteça?

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) \\ &= 1 - P(X \leq 3) \\ &= 1 - 0.879 = 0.121 \end{aligned}$$

Exemplo: Comprador A ou B?

Exemplo: Um industrial fabrica peças, das quais 1/5 são defeituosas. Dois compradores **A** e **B** classificaram um grande lote de peças adquiridas em categorias I e II, pagando \$1.20 e \$0.80 por peça, respectivamente, do seguinte modo:

- **Comprador A**: retira uma amostra de cinco peças; se encontrar mais que uma defeituosa, classifica como II.
- **Comprador B**: retira uma amostra de dez peças; se encontrar mais que duas defeituosas, classifica como II.

Em média, qual comprador oferece mais lucro?

Fonte: Morettin & Bussab, Estatística Básica 5^a edição, pág 159.

Exemplo: Comprador A ou B?

Sabemos que 1/5 das peças são defeituosas.

Podemos nos concentrar na probabilidade dos vendedores julgarem um lote como tipo I ou II.

Seja X o número de peças defeituosas em n testes.

O experimento do **comprador A** tem distribuição $X_A \sim Bin(5, 1/5)$ enquanto o experimento do **comprador B** tem distribuição $X_B \sim Bin(10, 1/5)$.

Para o **comprador A**, temos que:

$$\begin{aligned} P(X_A > 1) &= 1 - P(X_A = 0) - P(X_A = 1) \\ &= 1 - \binom{5}{0}\left(1 - \frac{1}{5}\right)^5 - \binom{5}{1}\left(\frac{1}{5}\right)\left(1 - \frac{1}{5}\right)^4 = 0.263 \end{aligned}$$

Exemplo: Comprador A ou B?

De modo similar, para o **comprador B** temos:

$$P(X_B > 2) = 1 - \binom{10}{0}\left(1-\frac{1}{5}\right)^{10} - \binom{10}{1}\frac{1}{5}\left(1-\frac{1}{5}\right)^{9} - \binom{10}{2}\left(\frac{1}{5}\right)^{2}\left(1-\frac{1}{5}\right)^{8} = 0.322$$

Como o segundo comprador irá classificar o lote como II com maior probabilidade que o primeiro, ele é o que oferece menor lucro para o fornecedor.

Mas podemos verificar o lucro esperado do vendedor.

Exemplo: Comprador A ou B?

Preço por peça na categoria I: \$1.20. Preço por peça na categoria II: \$0.80.

Se o industrial decidir vender o lote para o **comprador A**, temos que

$$\mathbb{E}(\text{lucro A}) = 1.20 \times 0.737 + 0.80 \times 0.263 \approx 1.09$$

Ou seja, ele irá lucrar em média \$1.09 por peça.

Já se ele vender para o **comprador B**, temos que

$$\mathbb{E}(\text{lucro B}) = 1.20 \times 0.678 + 0.80 \times 0.322 \approx 1.07$$

que é um lucro dois centavos inferior.

Portanto, é mais interessante ao industrial que o comprador A examine mais peças.

Exemplo: Sherlock

Um inimigo de Sherlock propõe um jogo, que consiste no lançamento de uma moeda honesta várias vezes, em quatro versões:

1. Se a proporção de caras for maior ou igual do que 0.60, Sherlock vence.
2. Se a proporção de caras for maior ou igual do que 0.40, Sherlock vence.
3. Se a proporção de caras estiver entre 0.40 e 0.60 (inclusive), Sherlock vence.
4. Se a proporção de caras for menor ou igual do que 0.30, Sherlock vence.

O inimigo escolhe primeiro qual a versão do jogo e depois Sherlock terá que escolher se quer jogar com 10 ou 100 lançamentos da moeda.

Exemplo: Sherlock

Se o inimigo escolhe a versão 1, Sherlock deve escolher 10 ou 100 lançamentos?

Seja X_i a v.a. que indica o resultado do i-ésimo lançamento da moeda, ou seja,

$$X_i = \begin{cases} 1, & \text{se sair cara} \\ 0, & \text{se sair coroa} \end{cases} \quad \text{e} \quad P(X_i = 0) = P(X_i = 1) = 0.5$$

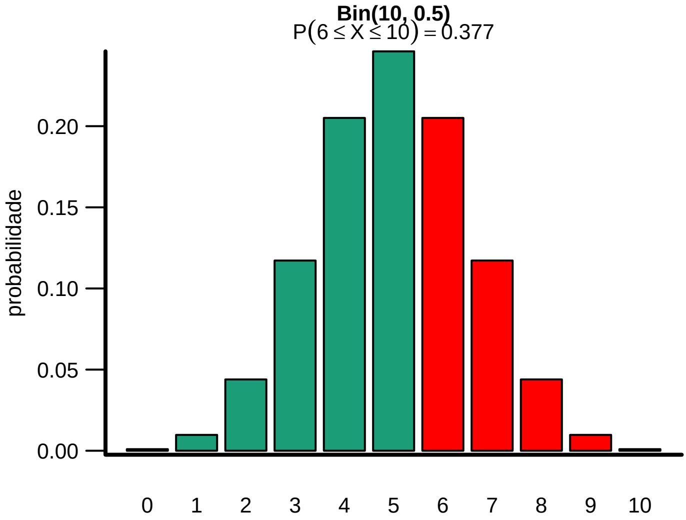
Seja $X = \sum_{i=1}^{n} X_i$ a v.a. que indica o número de caras em n lançamentos da moeda. Então, $X \sim \text{Bin}(n, 0.5)$.

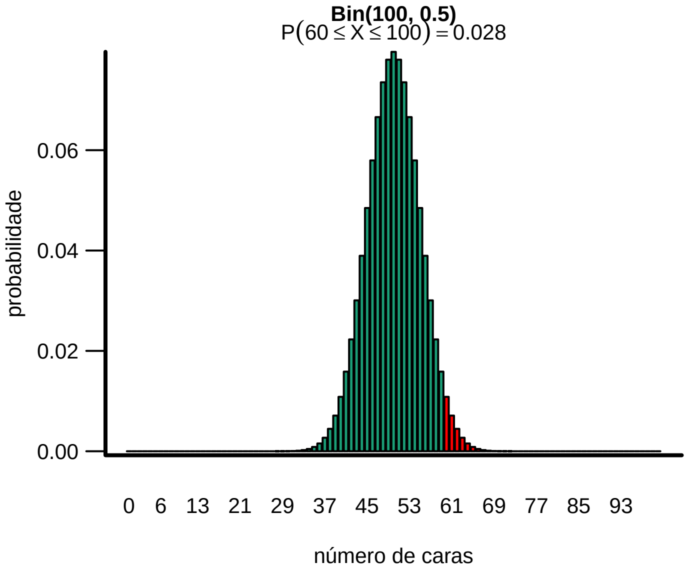
Na versão 1, Sherlock vence se a proporção de caras é maior do que 0.60, ou seja, se $X \geq n \times 0.6$.

Basta então Sherlock comparar $P(X \geq n \times 0.6)$ para $n = 10$ ou $n = 100$ e escolher o que resultar em maior probabilidade.

Exemplo: Sherlock - Versão 1

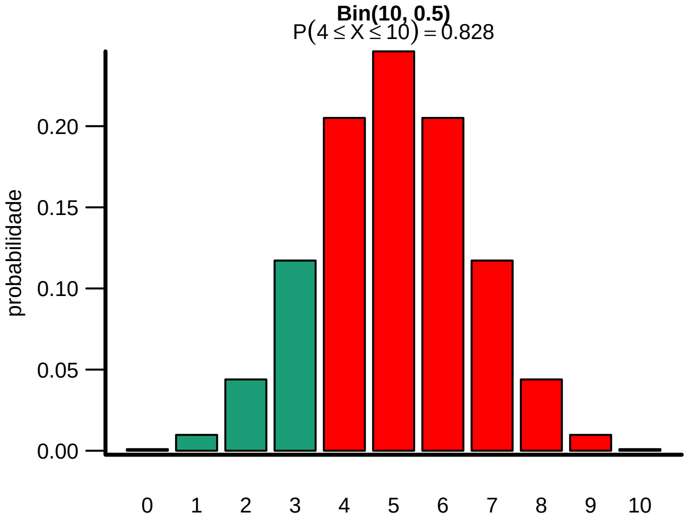
Se a proporção de caras for maior ou igual do que 0.60, Sherlock vence.

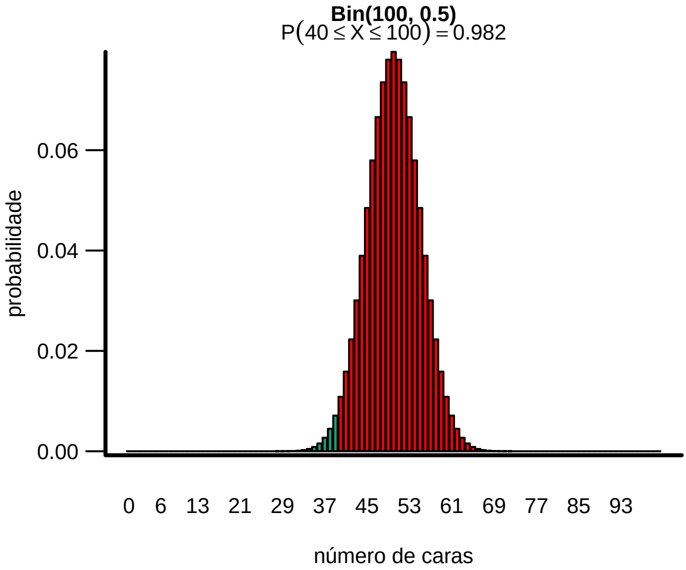



Bin(100, 0.5)
$P(60 \le X \le 100) = 0.028$
probabilidade
0.06
0.04
0.02
0.00
0 6 13 21 29 37 45 53 61 69 77 85 93
número de caras

Exemplo: Sherlock - Versão 2

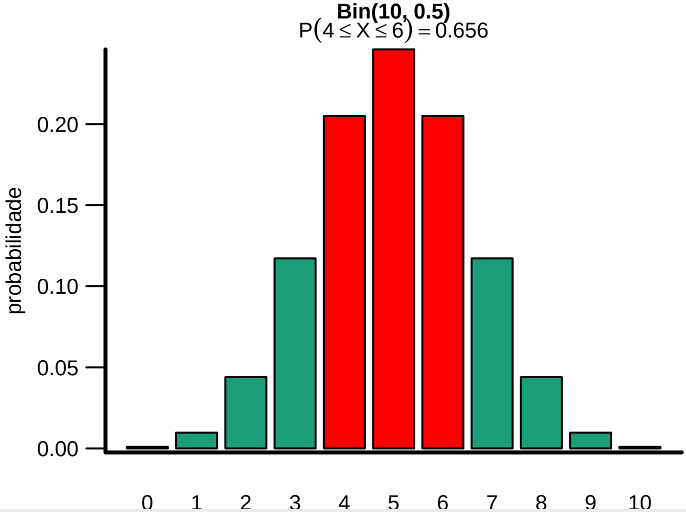
Se a proporção de caras for maior ou igual do que 0.40, Sherlock vence.

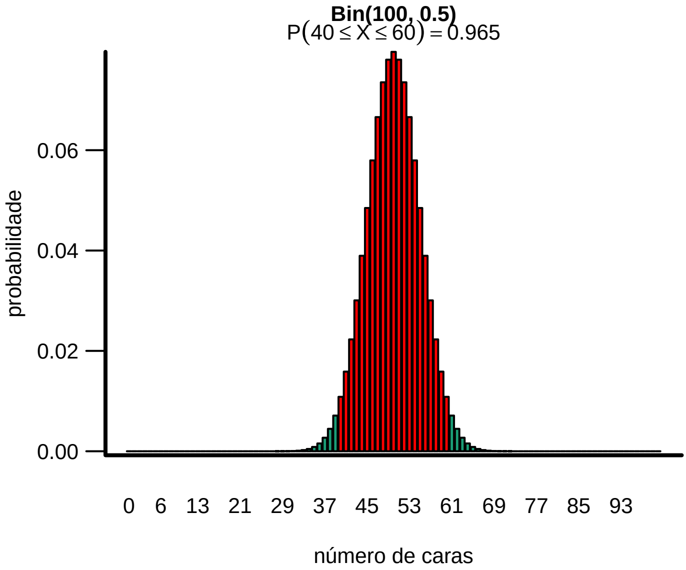



Bin(100, 0.5)
$P(40 \le X \le 100) = 0.982$
probabilidade
0.06
0.04
0.02
0.00
0 6 13 21 29 37 45 53 61 69 77 85 93
número de caras

Exemplo: Sherlock - Versão 3

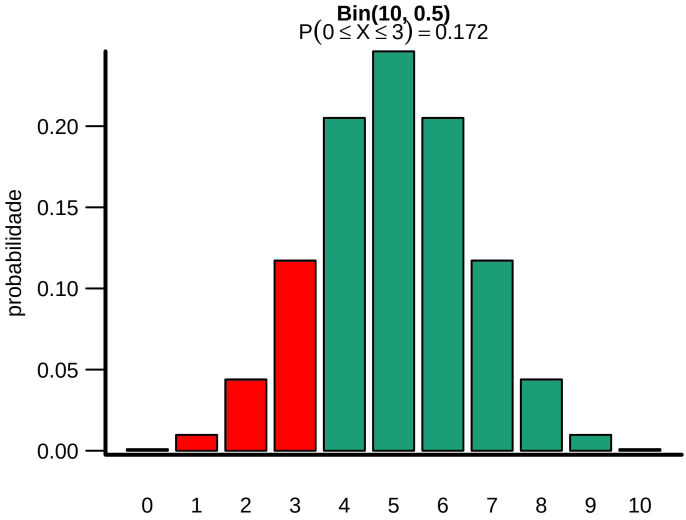
Se a proporção de caras estiver entre 0.40 e 0.60 (inclusive), Sherlock vence.

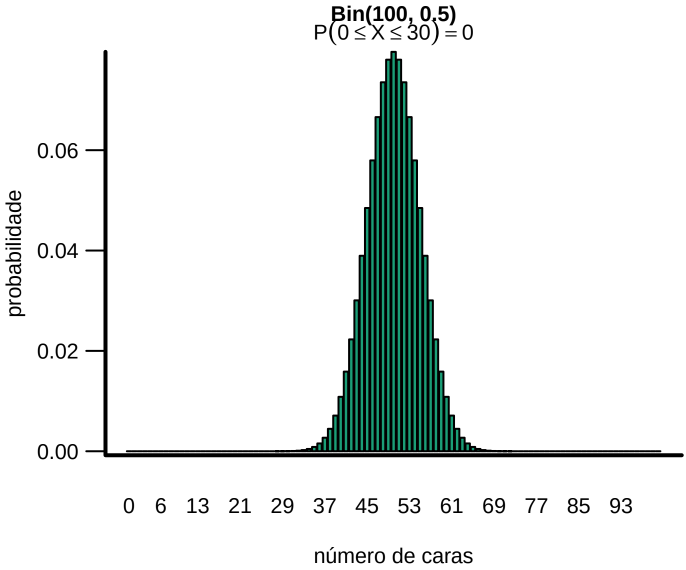



Bin(100, 0.5)
P(40 ≤ X ≤ 60) = 0.965
probabilidade
0.06
0.04
0.02
0.00
0 6 13 21 29 37 45 53 61 69 77 85 93
número de caras

Exemplo: Sherlock - Versão 4

Se a proporção de caras for menor ou igual do que 0.30, Sherlock vence.





Bin(100, 0.5)
$P(0 \le X \le 30) = 0$
0.06
0.04
0.02
0.00
probabilidade
0
6
13
21
29
37
45
53
61
69
77
85
93
número de caras

Exemplo: Sherlock

Na prática, para tomar uma decisão rápida, Sherlock deve considerar que a proporção esperada de caras é sempre 0.5, mas que quando n é menor, existe maior variabilidade em torno desse valor esperado.

Se a proporção de caras for maior do que 0.60, Sherlock vence.

Aqui Sherlock deve escolher $n = 10$, pois a variância de X/n (proporção de caras) é maior com $n = 10$.

Se a proporção de caras for maior do que 0.40, Sherlock vence.

Aqui Sherlock deve escolher $n = 100$, pois a variância de X/n é menor com $n = 100$, portanto chances maiores da proporção observada estar próxima da proporção esperada.

Exemplo: Sherlock

Se a proporção de caras estiver entre 0.40 e 0.60, Sherlock vence.

Mesmo raciocínio do item anterior.

Se a proporção de caras for menor do que 0.30, Sherlock vence.

Mesmo raciocínio do item 1.

Distribuição Geométrica

Distribuição Geométrica

Consideremos novamente um experimento aleatório com espaço de resultados Ω e o evento A.

Vamos dizer que ocorreu sucesso se o evento A aconteceu e $p = P(\text{sucesso})$.

Repetimos o experimento até o primeiro sucesso.

Seja X o número de repetições até o primeiro sucesso.

Exemplo: lançar uma moeda repetidas vezes até a primeira cara e $p = P(cara)$.

Os valores possíveis de X são $\{1, 2, 3, ...\}$.

$$
\begin{aligned}
&P(X = 1) = p && \text{(sucesso logo na primeira tentativa)} \\
&P(X = 2) = (1 - p)p && \text{(1 fracasso seguido de 1 sucesso)} \\
&P(X = k) = (1 - p)^{k-1}p && (k - 1 \text{ fracassos sucessivos e 1 sucesso})
\end{aligned}
$$

Distribuição Geométrica

Modelo Geral: Suponha uma sequência de ensaios de Bernoulli independentes com probabilidade de sucesso p.

Seja X a v.a. que representa o número de ensaios de Bernoulli até a ocorrência do primeiro sucesso. Então dizemos que X segue uma distribuição **Geométrica** com parâmetro p, ou seja, $X \sim G(p)$.

A probabilidade de se observar x é dada por:

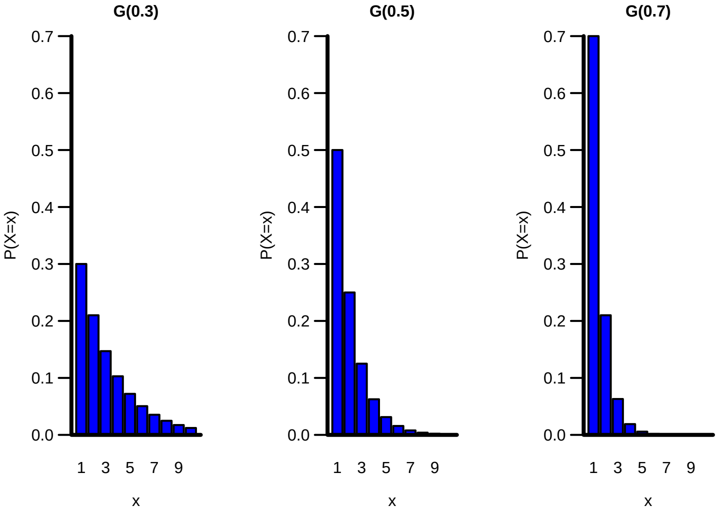
$$P(X = x) = (1 - p)^{x-1}p, \qquad x = 1, 2, \ldots$$

A esperança e variância de uma v.a. Geométrica são dadas por:

$$\mathbb{E}(X) = \frac{1}{p} \qquad \text{e} \qquad Var(X) = \frac{1-p}{p^2}$$

Distribuição Geométrica

Distribuição de probabilidade de uma $G(p)$, com $p = 0.3, 0.5$ e 0.7.



Distribuição Geométrica

A função de distribuição acumulada de uma v.a. $G(p)$ é dada por:

$$F(x) = P(X \leq x) = 1 - (1-p)^x$$

A distribuição geométrica tem uma propriedade que serve para caracterizá-la no conjunto das distribuições discretas: a propriedade de perda de memória!

Propriedade de Perda de Memória

$$P(X > x + m \mid X > m) = P(X > x)$$

Interpretação: O fato de já termos observado m fracassos sucessivos não muda a probabilidade do número de ensaios até o primeiro sucesso ocorrer.

Propriedade de perda de memória

$$P(X > x + m \mid X > m) = P(X > x)$$

Demonstração:

Lembre-se que:

$$F(x) = P(X \leq x) = 1 - (1-p)^x \quad \Longrightarrow \quad P(X > x) = (1-p)^x$$

Então,

$$\begin{aligned} P(X > x + m \mid X > m) &= \frac{P(X > x + m, X > m)}{P(X > m)} \\ &= \frac{P(X > x + m)}{P(X > m)} = \frac{(1-p)^{x+m}}{(1-p)^m} \\ &= (1-p)^x = P(X > x) \end{aligned}$$

Exemplo: Sinal de trânsito

A probabilidade de se encontrar aberto o sinal de trânsito numa esquina é 0.2.

Qual a probabilidade de que seja necessário passar pelo local 5 vezes para encontrar o sinal aberto pela primeira vez?

X = número de vezes necessárias para encontrar o sinal aberto.

$p = P(\text{sinal aberto}) = 0.2$

$$\begin{aligned} P(X = 5) &= (1 - p)^4 p \\ &= 0.8^4 \times 0.2 \\ &= 0.0819 \end{aligned}$$

Exemplo: lançamento de um dado

Qual a probabilidade de que um dado deva ser lançado 15 vezes para que ocorra a face 6 pela primeira vez?

X = número de vezes necessárias para ocorrer o resultado 6.

$p = P(\text{face } 6) = 1/6$

$$\begin{aligned} P(X = 15) &= (1-p)^{15-1} p \\ &= \left(\frac{5}{6}\right)^{14} \frac{1}{6} \\ &= 0.01298 \end{aligned}$$

Exemplo: Roleta Russa

Em sua autobiografia *A Sort of Life*, o autor inglês Graham Greene descreveu um período de grave depressão em que ele jogava roleta russa. Esse "jogo" consiste em colocar uma bala em uma das seis câmaras de um revólver, girar o tambor e disparar a arma contra a própria cabeça.

- Greene jogou seis partidas deste jogo, e teve a sorte da arma nunca ter disparado. Qual a probabilidade desse resultado?
- Suponha que ele continue jogando roleta russa até a arma finalmente disparar. Qual é a probabilidade de Greene morrer na k-ésima jogada?

Fonte: A. Agresti, Categorical Data Analysis.

Exemplo: Roleta Russa

Greene jogou seis partidas deste jogo, e teve a sorte da arma nunca ter disparado. Qual a probabilidade desse resultado?

Ao girar o tambor, a arma disparar ou não é um ensaio de Bernoulli com probabilidade 1/6 de disparar. Como cada uma das jogadas é independente, a probabilidade da arma não ter disparado em nenhuma das seis vezes é

$$\left(\frac{5}{6}\right)^6 = 0.33489$$

Seja X o número de disparos na roleta russa em 6 partidas e $p = 1/6$ a probabilidade de disparo. Então, $X \sim Bin(6, 1/6)$.

$$P(X = 0) = \binom{6}{0} p^0 (1-p)^{6-0} = \left(\frac{5}{6}\right)^6 = 0.33489$$

Exemplo: Roleta Russa

Qual é a probabilidade de Greene morrer na k-ésima jogada?

Ao efetuar a 1ª jogada, o autor pode morrer com probabilidade $1/6$, ou continuar jogando. Se ele sobreviver à primeira, pode jogar pela 2ª vez, e morrer com probabilidade $5/6 \times 1/6$, ou continuar jogando. Repetindo esse raciocínio, a probabilidade de morte na k-ésima jogada é

$$\left(\frac{5}{6}\right)^{k-1} \frac{1}{6}$$

Ou podemos definir X como o número de tentativas até o 1º disparo e $p = 1/6$ a probabilidade de um disparo ocorrer. Então, $X \sim G(1/6)$ e:

$$P(X = k) = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}$$

Exemplo: Banco de Sangue

Um banco de sangue necessita sangue do tipo O negativo. Suponha que a probabilidade de uma pessoa ter este tipo de sangue seja 0.10. Doadores permanentes chegam ao hemocentro para fazer sua doação rotineira. Calcule a probabilidade de que o primeiro doador com sangue do tipo O negativo seja:

- o primeiro a chegar;
- o segundo;
- o sétimo.
- Quantos doadores esperamos passar pelo hospital até encontrarmos um com sangue O negativo?

Fonte: Prof. Mario Gneri, Notas de Aula.

Exemplo: Banco de Sangue

Seja X o número de doadores que chegam no hemocentro até a chegada do primeiro doador com sangue O negativo.

Novamente temos um experimento com distribuição geométrica. Usando a fórmula para a função de probabilidade, send $X \sim G(0.1)$:

$$P(X = x) = 0.9^{x-1} 0.1, \qquad x = 1, 2, \ldots$$

Temos que

- $P(X = 1) = 0.1$
- $P(X = 2) = 0.9 \times 0.1 = 0.09$
- $P(X = 7) = 0.9^6 \times 0.1 = 0.053$
- $\mathbb{E}(X) = 1/0.1 = 10$. Neste caso, esperamos que dez doadores passem pelo hospital, em média, para encontrarmos o primeiro com sangue O negativo.

Hipergeométrica

Distribuição Hipergeométrica

População dividida em duas características.

Extrações casuais sem reposição.

Detalhes:

- N objetos;
- r têm a característica A;
- $N - r$ têm a característica B;
- Um grupo de n elementos é escolhido ao acaso, dentre os N possíveis, sem reposição.

Objetivo: calcular a probabilidade de que este grupo de n elementos contenha x elementos com a característica A.

Distribuição Hipergeométrica

Seja X a v.a. que representa o número de elementos com a característica A dentre os n selecionados.

Então dizemos que X segue uma distribuição **Hipergeométrica** com parâmetros N, n, r, ou seja, $X \sim Hip(N, n, r)$.

A probabilidade de se observar x é dada por:

$$P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \qquad 0 \leq x \leq min\{r, n\}$$

A esperança e variância são, respectivamente:

$$E(X) = \frac{nr}{N} \qquad \text{e} \qquad Var(X) = \frac{nr}{N}\left(1 - \frac{r}{N}\right)\frac{(N-n)}{(N-1)}$$

Exemplo: Urna

Uma urna contém 10 bolas: 6 brancas e 4 pretas.

Qual a probabilidade de obter 3 bolas brancas dentre 4 bolas retiradas?

Seja X o número de bolas brancas dentre as 4 bolas retiradas.

Então, $X \sim Hip(N = 10, n = 4, r = 6)$ e

$$P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \qquad 0 \leq x \leq min\{r, n\}$$

Portanto,

$$P(X = 3) = \frac{\binom{6}{3}\binom{4}{1}}{\binom{10}{4}} = \frac{8}{21}$$

Exemplo: Comissão

Voltando ao exemplo: O Departamento de Estatística é formado por 25 professores, sendo 17 homens e 8 mulheres. Uma comissão será formada por 3 professores. Queremos saber qual é a probabilidade da comissão ser formada por pelo menos duas mulheres?

Seja X o número de mulheres na comissão, então
$X \sim Hip(N = 25, n = 3, r = 8)$

$$P(X = 2) = \frac{\binom{8}{2}\binom{17}{1}}{\binom{25}{3}} = 0.21$$

$$P(X = 3) = \frac{\binom{8}{2}\binom{17}{0}}{\binom{25}{3}} = 0.02$$

$$P(X \geq 2) = P(X = 2) + P(X = 3) = 0.21 + 0.02 = 0.23$$

Exemplo: Loteria

Um jogo de loteria consiste em selecionar 6 dezenas de 00 a 99, com uma bola para cada dezena e sem reposição.

Numa aposta o jogador pode escolher de 6 a 10 dezenas.

Qual a probabilidade de acertar a quina (5 dezenas) marcando-se 10 dezenas na aposta?

- N = 100 (total de dezenas)
- n = 6 (dezenas sorteadas)
- r = 10 (dezenas escolhidas pelo jogador)
- x = 5 (número de sucessos, queremos 5)

$$P(X = 5) = \frac{\binom{10}{5}\binom{100-10}{6-5}}{\binom{100}{6}} = 0.000019$$

Exemplo: Mega-Sena

Qual a probabilidade de um jogador ganhar na Mega-Sena jogando 6 dezenas?

- $N = 60$ (dezenas de 01 a 60)
- $n = 6$ (dezenas sorteadas)
- $r = 6$ (dezenas escolhidas pelo jogador)
- $x = 6$ (número de sucessos, queremos 6)

Então, a probabilidade de ganhar na Mega-Sena é:

$$P(X = 6) = \frac{\binom{6}{6}\binom{54}{0}}{\binom{60}{6}} = \frac{1}{50063860}$$

Aplicação: Controle de Qualidade

Suponha um lote com $N = 100$ elementos a ser analisado.

São escolhidas $n = 5$ peças sem reposição.

Sabendo que neste lote de 100 elementos, $r = 10$ são defeituosos.

Se X é o número de peças defeituosas em 5 escolhidas, então

$$X \sim Hip(N = 100, n = 5, r = 10)$$

A probabilidade de nenhuma peça defeituosa na amostra retirada é:

$$P(X = 0) = \frac{\binom{10}{0}\binom{100-10}{5-0}}{\binom{100}{5}} = \frac{\binom{90}{5}}{\binom{100}{5}} \approx 0.584$$

Aplicação: Controle de Qualidade

A probabilidade de pelo menos uma peça defeituosa é:

$$P(X \geq 1) = \sum_{i=1}^{5} P(X = i) = 1 - P(X = 0) \approx 0.416$$

A média e a variância são:

$$\begin{aligned}
\mathbb{E}(X) &= \frac{nr}{N} = \frac{5 \times 10}{100} = 0.5 \\
Var(X) &= \frac{nr}{N}\left(1 - \frac{r}{N}\right)\frac{(N-n)}{(N-1)} \\
&= \frac{5 \times 10}{100}\left(1 - \frac{10}{100}\right)\frac{(100-5)}{(100-1)} \approx 0.409
\end{aligned}$$

Exemplo

Pequenos motores são guardados em caixas de 50 unidades. Um inspetor de qualidade examina cada caixa, antes da remessa, testando 5 motores. Se nenhum motor for defeituoso, a caixa é aceita. Se pelo menos um for defeituoso, todos os 50 motores são testados. Há 6 motores defeituosos numa caixa. Qual a probabilidade de que seja necessário examinar todos os motores?

X = número de motores defeituosos da amostra.

$N = 50, n = 5$ e $r = 6$. Então, $X \sim Hip(N = 50, n = 5, r = 6)$

Se pelo menos 1 é defeituoso, inspeciona todos os 50.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X < 1) = 1 - P(X = 0) \\ &= 1 - \frac{\binom{6}{0}\binom{44}{5}}{\binom{50}{5}} = 1 - 0.5126 = 0.4874 \end{aligned}$$

Exemplo

Uma firma compra lâmpadas por centenas. Examina sempre uma amostra de 15 lâmpadas para verificar se estão boas. Se uma centena inclui 12 lâmpadas queimadas, qual a probabilidade de se escolher uma amostra com pelo menos uma lâmpada queimada?

X = número de lâmpadas queimadas da amostra.

$N = 100, n = 15$ e $r = 12$. Então, $X \sim Hip(N = 100, n = 15, r = 12)$

Pelo menos uma queimada:

$$
\begin{aligned}
P(X \geq 1) &= 1 - P(X < 1) = 1 - P(X = 0) \\
&= 1 - \frac{\binom{12}{0}\binom{88}{15}}{\binom{100}{15}} = 0.8747
\end{aligned}
$$

Aproximação da Binomial pela Poisson

Distribuição de Poisson

Muitas vezes, em problemas em que seria natural usar a distribuição binomial, temos n muito grande ($n \to \infty$) e p muito pequeno ($p \to 0$).

Nesses casos, o cálculo fica difícil com calculadoras comuns.

Considerando uma v.a. $X \sim \text{Bin}(n, p)$, quando temos grandes valores para n e p pequeno (mantendo-se o produto $np = \lambda$ constante), podemos usar a seguinte aproximação para a probabilidade:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \approx \frac{e^{-np}(np)^x}{x!}, \quad x = 0, 1, 2, \ldots n$$

Geralmente considera-se o critério $np \leq 7$ para usar essa aproximação.

Demonstração

Exemplo

$X \sim Bin(100, 0.065)$, deseja-se obter $P(X = 10)$

No modelo Binomial:

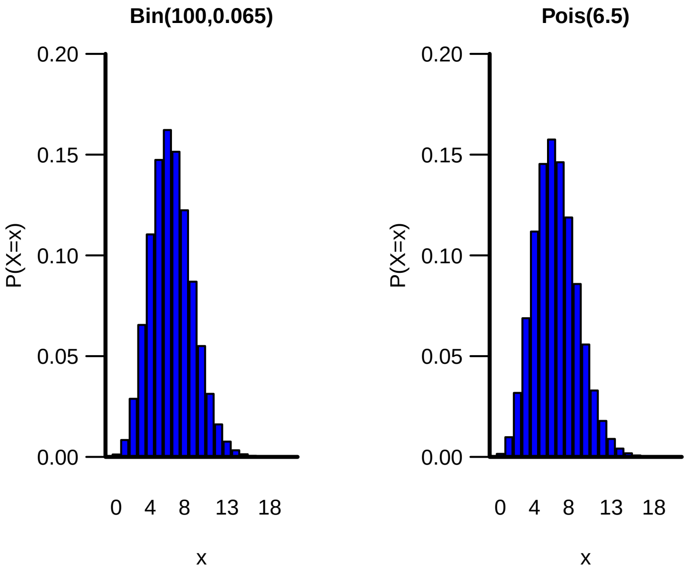
$$P(X = 10) = \binom{100}{10}(0.065)^{10}(0.935)^{100-10} = 0.055$$

$\lambda = np = 100 \times 0.065 = 6.5 \leq 7$

No modelo Poisson:

$$P(X = 10) = \frac{e^{-6.5}(6.5)^{10}}{10!} \approx 0.056$$

Poisson para Aproximar uma Binomial



Exemplo

A probabilidade de uma lâmpada se queimar ao ser ligada é 1/100. Numa instalação com 100 lâmpadas, qual a probabilidade de 2 lâmpadas se queimarem ao serem ligadas?

No modelo Binomial: $X \sim \text{Bin}(100, 0.01)$

$$P(X = 2) = \binom{100}{2}(0.01)^2(0.99)^{100-2} = 0.1849$$

$\lambda = np = 100 \times 0.01 = 1 \leq 7$

No modelo Poisson:

$$P(X = 2) = \frac{e^{-1}(1)^2}{2!} \approx 0.1839$$

Distribuição de Poisson

Outro caso em que a distribuição de Poisson é utilizada:

- Considere a probabilidade de ocorrência de sucessos em um determinado intervalo.
- A probabilidade de ocorrência de um sucesso no intervalo é proporcional ao intervalo.
- A probabilidade de mais de um sucesso nesse intervalo é bastante pequena com relação à probabilidade de apenas um sucesso.

Distribuição de Poisson

Seja X o número de sucessos no intervalo, então:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

onde λ é a esperança.

Notação: $X \sim \text{Poisson}(\lambda)$.

Distribuição de Poisson

A distribuição de Poisson é muito usada na distribuição do número de:

- carros que passam por um cruzamento por minuto, durante uma certa hora do dia;
- erros tipográficos por página, em um livro;
- defeitos por unidade (m^3, m^2, m, etc...) por peça fabricada;
- colônias de bactérias numa dada cultura por $0.01mm^2$, numa plaqueta de microscópio;
- mortes por ataque de coração por ano, em um certo bairro;
- etc...

Distribuição de Poisson

Para uma v.a. quantificando eventos raros, sob algumas suposições, podemos usar a distribuição de Poisson.

Uma variável aleatória X tem distribuição de Poisson com parâmetro $\lambda > 0$, se sua função de probabilidade é dada por:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, ...$$

λ é chamado de taxa de ocorrência

$\mathbb{E}(X) = Var(X) = \lambda$

Notação: $X \sim P(\lambda)$

Exemplo: Erros em um livro

Num livro de 800 páginas, há 800 erros de impressão. Qual a probabilidade de que uma página contenha pelo menos 3 erros?

X = número de erros por página

Taxa de ocorrência: $\lambda = 1$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X < 3) \\ &= 1 - [P(X = 0) + P(X = 1) + P(X = 2)] \\ &= 1 - \left\{ \frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!} + \frac{e^{-1}1^2}{2!} \right\} \\ &= 0.08 \end{aligned}$$

Exemplo: Mensagens no Facebook

Uma firma recebe 720 mensagens em sua página do Facebook durante as 8 horas de horário comercial. Qual a probabilidade de que em 6 minutos no horário comercial a firma receba pelo menos 4 mensagens no Facebook?

$$\begin{aligned} 720 \text{ mensagens} &\rightarrow 480 \text{ min} \\ \lambda &\rightarrow 6 \text{ min} \end{aligned}$$

Então, $\lambda = 9$ e

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) = 1 - P(X \leq 3) \\ &= 1 - \left[\frac{e^{-9}9^0}{0!} + \frac{e^{-9}9^1}{1!} + \frac{e^{-9}9^2}{2!} + \frac{e^{-9}9^3}{3!} \right] \\ &= 0.979 \end{aligned}$$

Exemplo: SAC

Numa central de SAC (serviço de atendimento ao consumidor) chegam 300 telefonemas por hora. Qual a probabilidade de que:

- num minuto não haja nenhuma chamada?
- em 2 minutos haja 2 chamadas?
- em t minutos não haja chamadas?

Exemplo: SAC

X = número de chamadas por minuto.

Taxa de ocorrência por minuto: $\lambda = 300/60 = 5$

$$P(X=0) = \frac{e^{-5}5^0}{0!} = 0.0067$$

X = número de chamadas a cada 2 minutos.

Taxa de ocorrência em 2 minutos: $\lambda = 10$

$$P(X=2) = \frac{e^{-10}10^2}{2!} = 0.00227$$

X = número de chamadas a cada t minutos.

Taxa de ocorrência em t minutos: $\lambda = 5t$

$$P(X=0) = \frac{e^{-5t}(5t)^0}{0!} = e^{-5t}$$

Exemplo: Lâmpadas

A experiência mostra que de cada 400 lâmpadas, 2 se queimam ao serem ligadas. Qual a probabilidade de que numa instalação de:

- 600 lâmpadas, no mínimo 3 se queimem?
- 900 lâmpadas, exatamente 8 se queimem?

Exemplo: Lâmpadas

X = número de lâmpadas que se queimam numa instalação de 600 lâmpadas.

De 400, 2 se queimam, ou seja, de 200, 1 se queima.

Taxa de ocorrência para 600 lâmpadas: $\lambda = 600/200 = 3$

600 lâmpadas, no mínimo 3 se queimem:

$$\begin{aligned} P(X \geq 3) = \sum_{x=3}^{\infty} \frac{e^{-3}3^x}{x!} &= 1 - P(X < 3) = \\ &= 1 - [P(X = 0) + P(X = 1) + P(X = 2)] \\ &= 1 - \left[\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!}\right] = 0.5768 \end{aligned}$$

Exemplo: Lâmpadas

X = número de lâmpadas que se queimam numa instalação de 900 lâmpadas.

Taxa de ocorrência para 900 lâmpadas: $\lambda = 900/200 = 4.5$

900 lâmpadas, 8 se queimem:

$$P(X = 8) = \frac{e^{-4.5}(4.5)^8}{8!} = 0.0463$$

Exemplo: Twitter

O número citações de uma certa conta do Twitter ocorre segundo uma distribuição de Poisson, com a média de oito citações por minuto.

Determinar qual a probabilidade de que num minuto se tenha:

1. dez ou mais citações;
2. menos que nove citações;
3. entre sete (inclusive) e nove (exclusive) citações.

Exemplo: Twitter

Sabemos que se $X \sim \text{Poisson}(\lambda)$, então sua função de probabilidade é

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, \ldots$$

Além disso, $\mathbb{E}(X) = \lambda$.

O enunciado diz *média de oito citações por minuto*, então a variável aleatória $X =$ *número de citações por minuto* tem distribuição Poisson(8).

- A probabilidade de dez ou mais chamadas é dada por:

$$\begin{aligned} P(X \geq 10) &= 1 - P(X < 10) = 1 - P(X \leq 9) \\ &= 1 - \sum_{k=0}^{9} \frac{e^{-8}8^k}{k!} = 1 - e^{-8} - \ldots - \frac{e^{-8}8^9}{9!} = 0.2833 \end{aligned}$$

A probabilidade de termos menos que nove citações em um minuto é dada por:

$$P(X < 9) = P(X \leq 8) = e^{-8} + \ldots + \frac{e^{-8}8^8}{8!} = 0.5926$$

A probabilidade de termos entre sete (inclusive) e nove (exclusive) citações em um minuto é dada por:

$$\begin{aligned} P(7 \leq X < 9) = P(7 \leq X \leq 8) &= P(X = 7) + P(X = 8) \\ &= \frac{e^{-8}8^7}{7!} + \frac{e^{-8}8^8}{8!} = 0.2792 \end{aligned}$$

Leituras

- Ross: capítulo 5
- Magalhães: capítulo 3
- OpenIntro: seções 3.3, 3.4, 3.5.2