Notas de Aulas do prof. João Maurício Araújo Mota da disciplina CC0282- Probabilidade I ministrada em 2021 .1.

1 Distribuição Binomial

1.1 Introdução

Fato 1. Considere que um experimento de Bernoulli é ensaiado n vezes independentemente e que a probabilidade de sucesso em cada repetição é sempre constante e igual a p, $0 < p < 1$. Seja X a variável aleatória correspondente ao número total de sucessos da amostra. A f.p. de X é dada por

$$f(x) = \binom{n}{x} p^x \, q^{n-x} I_{\{0,1,\ldots,n\}}(x) \tag{1}$$

Prova: O suporte da distribuição é o conjunto $A = \{0, 1, \ldots, n\}$ e seja $x \in A$. Vai-se calcular a $f(x) = P(X = x)$, isto é, a probabilidade de que sejam obtidos x sucessos. A probabilidade de que as x primeiras repetições sejam sucessos e as restantes fracassos é dada por:

$$P(SS\ldots SF\ldots F) = p^x \, q^{n-x},$$

pois as repetições são independentes.

Qualquer permutação da sequencia $(SS\ldots SF\ldots F)$ gera x sucessos e $n - x$ fracassos, e como existem:

$$\frac{n!}{x!(n-x)!} = \binom{n}{x}$$

permutações com repetição desse tipo a probabilidade de aparecer x sucessos é dada por:

$$f(x) = P(X = x) = \binom{n}{x} p^x \, q^{n-x} I_{\{0,1,\ldots,n\}}(x).$$

Esta distribuição é chamada na literatura de distribuição binomial e é uma das distribuições mais importantes da Estatística. Quem primeiro estudou essa distribuição foi Jacob Bernoulli (1713) que publicou um trabalho sem, contudo, chamá-la de distribuição binomial. Somente em 1911, Yule chamou-a de distribuição binomial. Para empregá-la , em situações práticas, precisa assegurar que as seguintes pressuposições:

i. tem-se n repetições de um ensaio de Bernoulli;

ii. as repetições são independentes; e

iii. a probabilidade de sucesso em cada ensaio é constante.

1.2 Definição

Uma variável aleatória discreta X é dita possuir distribuição *Binomial* de parâmetros n e p, onde $p = 1-q$, $0 \leq p \leq 1$, se sua função de probabilidade ($f.p.$) é da forma:

$$f(x) = \binom{n}{x} p^x \, q^{n-x} I_{\{0,1,\ldots,n\}}(x) \tag{2}$$

Notação: $X \sim B(n,p)$.

Observação 1: Lê-se a notação acima do seguinte modo: X segue distribuição *Binomial* de parâmetros n e p.

Observação 2: O suporte da distribuição binomial é:

$$A = \{0, 1, 2, \ldots, n\}.$$

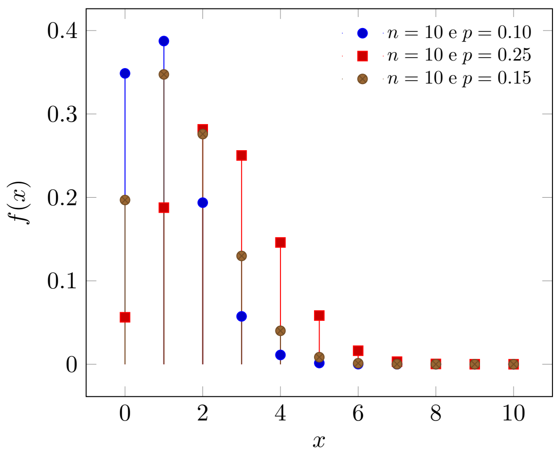

Figura 1: Gráfico da Função de Probabilidade Binomial

A Figura 1 apresenta o gráfico da distribuição binomial para certos valores dos parâmetros n e p.

1.3 Propriedades da função de probabilidade

Fato 2. A expressão (2) é realmente uma função de probabilidade.

Prova: deve-se verificar que

i. $f(x) > 0, \ x \in A.$; e

ii. $\sum_{A} f(x) = 1,$

sendo $A = \{x \in R | f(x) > 0\}$ o suporte da distribuição de X. Como $A = \{0, 1, \ldots, n\}$ e para $0 < p < 1$ tem-se $f(x) = \binom{n}{x} p^x \, q^{n-x} > 0$ para qualquer ponto do suporte . A segunda propriedade nos diz que a soma dos valores das probabilidades para os pontos do suporte é 1. Assim

$$\sum_{x=0}^{n} \binom{n}{x} p^x \, q^{n-x} = (p+q)^n = 1 \qquad \blacksquare$$

1.4 Função geradora de probabilidades

Fato 3. Se $X \sim B(n, p)$, então $G(t) = (pt + q)^n$.

Prova:

$$G(t) = E(t^X) = \sum_{x=0}^{n} t^x \binom{n}{x} p^x \, q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (pt)^x \, q^{n-x} = (pt+q)^n \qquad \blacksquare$$

1.5 Função geradora de momentos

Fato 4. Se $X \sim B(n, p)$, então $M(t) = (pe^t + q)^n$.

Prova:

$$M(t) = E(e^{tX}) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} \, p^x \, q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} \, (pe^t)^x \, q^{n-x} = (pe^t+q)^n \qquad \blacksquare$$

1.6 Função geradora de Cumulantes

Fato 5. Se $X \sim B(n, p)$, então $K(t) = n \log(pe^t + q)$.

Prova:

$$K(t) = \log(M(t)) = \log\left((pe^t+q)^n\right) = n \log(pe^t + q) \qquad \blacksquare$$

1.7 Momentos fatoriais

Fato 6. Se $X \sim B(n,p)$, então os quatro primeiros momentos fatoriais são dados por:

$$E\left(X_{[r]}\right) = \begin{cases} np & \text{, se } r=1 \\ n(n-1)p^2 & \text{, se } r=2 \\ n(n-1)(n-2)p^3 & \text{, se } r=3 \\ n(n-1)(n-2)(n-3)p^4 & \text{, se } r=4 \end{cases}$$

Prova: como $E\left(X_{[1]}\right) = E(X) =$ tem-se:

$$E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x\ q^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x\ q^{n-x},$$

mas

$$x\binom{n}{x} = x\frac{n!}{x!(n-x)!} = x\frac{n(n-1)!}{x(x-1)!(n-x)!} = n\frac{(n-1)!}{(x-1)!(n-x)!)} = n\binom{n-1}{x-1},$$

assim,

$$E(X) = \sum_{x=1}^{n} n\binom{n-1}{x-1} p^x\ q^{n-x} = np\sum_{x=1}^{n}\binom{n-1}{x-1} p^{x-1}\ q^{(n-1)-(x-1)} = np,$$

e fazendo a mudança $y = x-1$ no somatório tem-se que para $x=1, y=0$ e para $x=n, y=n-1$. Portanto

$$\sum_{x=1}^{n}\binom{n-1}{x-1} p^{x-1}\ q^{(n-1)-(x-1)} = \sum_{y=0}^{n-1}\binom{n-1}{y} p^{y}\ q^{n-1-y} = (p+q)^{n-1} = 1.$$

Vai-se calcular agora o segundo momento fatorial, assim

$$E\left(X_{[2]}\right) = E\left[X(X-1)\right] = \sum_{x=0}^{n} x(x-1)\binom{n}{x} p^x\ q^{n-x} = \sum_{x=2}^{n} x(x-1)\binom{n}{x} p^x\ q^{n-x},$$

mas,

$$x(x-1)\binom{n}{x} = x(x-1)\frac{n!}{x!(n-x)!} = x\frac{n(n-1)!}{x(x-1)(x-2)!(n-x)!} = n(n-1)\frac{(n-2)!}{(x-2)!(n-x)!},$$

e portanto,

$$x(x-1)\binom{n}{x} = n(n-1)\binom{n-2}{x-2},$$

e por fim,

$$E\left(X_{[2]}\right) = \sum_{x=2}^{n} n(n-1)\binom{n-2}{x-2} p^x\ q^{n-x},$$

colocando-se p^2 e $n(n-1)$ para fora do somatório tem-se:

$$E\left(X_{[2]}\right) = n(n-1)p^2\sum_{x=2}^{n}\binom{n-2}{x-2} p^{x-2}\ q^{n-x} = n(n-1)p^2,$$

e fazendo a mudança $y = x - 2$ no somatório tem-se que para $x = 2, y = 0$ e para $x = n, y = n - 2$, e portanto

$$\sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2}\, q^{(n-2)-(x-2)} = \sum_{y=0}^{n-2} \binom{n-2}{y} p^{y}\, q^{n-2-y} = (p+q)^{n-2} = 1.$$

Vai-se calcular agora o terceiro momento fatorial, assim

$$E\left(X_{[3]}\right) = E\left[X(X-1)(X-2)\right] = \sum_{x=3}^{n} x(x-1)(x-2)\binom{n}{x}\, p^{x}\, q^{n-x},$$

mas,

$$\begin{aligned} x(x-1)(x-2)\binom{n}{x} &= x(x-1)(x-2)\frac{n!}{x!(n-x)!} = x\frac{n(n-1)(n-2)!}{x(x-1)(x-2)(x-3)!(n-x)!} \\ &= n(n-1)(n-2)\frac{(n-3)!}{(x-3)!(n-x)!}, \end{aligned}$$

e portanto,

$$x(x-1)(x-2)\binom{n}{x} = n(n-1)(n-2)\binom{n-3}{x-3},$$

e por fim,

$$E\left(X_{[3]}\right) = \sum_{x=3}^{n} n(n-1)(n-2)\binom{n-3}{x-3} p^{x}\, q^{n-x},$$

colocando-se p^3 e $n(n-1)(n-2)$ para fora do somatório tem-se:

$$E\left(X_{[3]}\right) = n(n-1)(n-2)p^{3} \sum_{x=3}^{n} \binom{n-3}{x-3} p^{x-3}\, q^{n-x} = n(n-1)(n-2)p^{3},$$

pois, fazendo a mudança $y = x - 3$ no somatório tem-se que, para $x = 3, y = 0$ e para $x = n, y = n - 3$, e portanto

$$\sum_{x=3}^{n} \binom{n-3}{x-3} p^{x-3}\, q^{(n-3)-(x-3)} = \sum_{y=0}^{n-3} \binom{n-3}{y} p^{y}\, q^{n-3-y} = (p+q)^{n-3} = 1.$$

Vai-se calcular agora o quarto momento fatorial, assim

$$E\left(X_{[4]}\right) = E\left[X(X-1)(X-2)(X-3)\right] = \sum_{x=4}^{n} x(x-1)(x-2)(x-3)\binom{n}{x}\, p^{x}\, q^{n-x},$$

mas,

$$\begin{aligned} x(x-1)(x-2)(x-3)\binom{n}{x} &= x(x-1)(x-2)(x-3)\frac{n!}{x!(n-x)!} \\ &= x\frac{n(n-1)(n-2)(n-3)!}{x(x-1)(x-2)(x-3)!(n-x)!} \\ &= n(n-1)(n-2)(n-3)!\frac{(n-4)!}{(x-4)!(n-x)!}, \end{aligned}$$

e portanto,

$$x(x-1)(x-2)(x-3)\binom{n}{x} = n(n-1)(n-2)(n-3)\binom{n-4}{x-4},$$

e por fim,

$$E\left(X_{[4]}\right) = \sum_{x=4}^{n} n(n-1)(n-2)(n-3)\binom{n-4}{x-4} p^{x}\, q^{n-x},$$

colocando-se p^4 e $n(n-1)(n-2)(n-3)$ para fora do somatório tem-se:

$$E\left(X_{[4]}\right) = n(n-1)(n-2)(n-3)p^4 \sum_{x=4}^{n} \binom{n-4}{x-4} p^{x-4}\, q^{n-x} = n(n-1)(n-2)(n-3)p^4,$$

pois, fazendo a mudança $y = x-4$ no somatório tem-se que, para $x = 4, y = 0$ e para $x = n, y = n-4$, e portanto

$$\sum_{x=4}^{n} \binom{n-4}{x-4} p^{x-4}\, q^{(n-4)-(x-4)} = \sum_{y=0}^{n-4} \binom{n-4}{y} p^{y}\, q^{n-4-y} = (p+q)^{n-4} = 1 \qquad \blacksquare$$

1.8 Momentos em relação à origem

Fato 7. Se $X \sim B(n,p)$, então os quatro primeiros momentos em relação à origem são dados por

$$E\left(X^r\right) = \begin{cases} np & \text{, se } r = 1 \\ np(np+q) & \text{, se } r = 2 \\ np\left[(n-1)(n-2)p^2 + 3(n-1)p + 1\right] & \text{, se } r = 3 \\ np\left[n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2 + np\right] & \text{, se } r = 4 \end{cases}$$

Prova: O primeiro momento em relação à origem é igual ao primeiro momento fatorial e portanto $E(X) = \mu = np$. O segundo momento em relação à origem, $E(X^2)$, é calculado por:

$$E(X^2) = E(X(X-1)) + E(X) = n(n-1)p^2 + np = np(np-p+1) = np(np+q) = np(np+1-p) = np((n-1)p+1).$$

O terceiro momento em relação à origem, $E(X^3)$, é calculado por:

$$E(X^3) = E\left[X(X-1)(X-2)\right] + 3E(X^2) - 2E(X).$$

Logo

$$E(X^3) = n(n-1)(n-2)p^3 + 3np(np+q) - 2np = np\left[(n-1)(n-2)p^2 + 3(np+q) - 2\right],$$

como $3(np+q)-2 = 3(np+1-p)-2 = 3((n-1)p+1)-2 = 3(n-1)p+3-2 = 3(n-1)p+1$, tem-se:

$$E(X^3) = np\left[(n-1)(n-2)p^2 + 3(n-1)p + 1\right].$$

O quarto momento em relação à origem, $E(X^4)$, é calculado por:

$$E(X^4) = E\left[X(X-1)(X-2)(X-3)\right] + 6E(X^3) - 11E(X^2) + 6E(X).$$

Logo,

$$\begin{aligned}
E(X^4) &= E\left[X(X-1)(X-2)(X-3)\right] + 6E(X^3) - 11E(X^2) + 6E(X) \\
&= n(n-1)(n-2)(n-3)p^4 + 6np[(n-1)(n-2)p^2 + 3(n-1)p + 1] + 11np(np+q) + 6np \\
&= np[(n-1)(n-2)(n-3)p^3 + 6[(n-1)(n-2)p^2 + 3(n-1)p + 1] - 11(np + 1 - p) + 6] \\
&= np[(n-1)(n-2)(n-3)p^3 + 6(n-1)(n-2)p^2 + 18(n-1)p + 6 - 11(n-1)p - 11 + 6] \\
&= np[(n-1)(n-2)(n-3)p^3 + 6(n-1)(n-2)p^2 + 7(n-1)p + 1].
\end{aligned}$$

1.9 Momentos centrais

Fato 8. Se $X \sim B(n,p)$, então$\mu_2 = E\left[(X-\mu)^2\right] = E(X^2) - E^2(X) = npq$.

Prova:

$$var(X) = np(np+q) - n^2p^2 = (n^2p^2 + npq - n^2p^2 = npq \qquad \blacksquare$$

Assim, a variância de $X \sim B(n,p)$, é dada por

$$Var(X) = npq. \tag{3}$$

Uma propriedade importante da distribuição binomial é que a variância é sempre menor que a média. Assim

$$V(X) = npq < np.1 = E(X), \quad 0 < q < 1.$$

O terceiro momento central de $X \sim B(n,p)$

$$\mu_3 = npq(q-p). \tag{4}$$

Prova:

$$
\begin{aligned}
\mu_3 &= E[(X-\mu)^3] \\
&= E(X^3) - 3E(X^2)E(X) + 2E^3(X) \\
&= np\left[(n-1)(n-2)p^2 + 3(n-1)p + 1\right] - 3np\left[(n-1)p+1\right]np + 2(np)^3 \\
&= np\left[(n-1)(n-2)p^2 + 3(n-1)p + 1 - 3n(n-1)p^2 - 3np + 2n^2p^2\right] \\
&= np\left[(n^2-3n+2)p^2 + 3np - 3p + 1 - 3n(n-1)p^2 - 3np + 2n^2p^2\right] \\
&= np\left[(n^2-3n+2-3n^2-3n+2n^2)p^2 + (3n-3-3n)p + 1\right] \\
&= np\left[2p^2 - 3p + 1\right] \\
&= np\left[2p^2 - 2p - p + 1\right] \\
&= np\left[2p(p-1) + 1 - p\right] \\
&= np\left[-2p(1-p) + (1-p)\right] \\
&= np\left[-2pq + q\right] \\
&= npq\left[-2p + 1\right] \\
&= npq\left[-p + 1 - p\right] \\
&= npq\left[q - p\right].
\end{aligned}
$$

O quarto momento central de $X \sim B(n,p)$

$$\mu_4 = npq(1 + 3(n-2)pq) = 3n^2p^2q^2 + npq(1 - 6pq). \tag{5}$$

Prova:

$$
\begin{aligned}
\mu_4 &= E[(X-\mu)^4] \\
&= E(X^4) - 4E(X^3)E(X) + 6E(X^2)E^2(X) - 3E^4(X) \\
&= P_1 - P_2 + P_3 - P4.
\end{aligned}
$$

Vamos analisar cada parcela separadamente:

$$
\begin{aligned}
E(X^4) &= np[(n-1)(n-2)(n-3)p^3 + 6(n-1)(n-2)p^2 + 7(n-1)p + 1] \\
&= np[(n^3 - 6n^2 + 11n - 6)p^3 + (6n^2 - 18n + 12)p^2 + (7n-7)p + 1]
\end{aligned}
$$

Vamos para a segunda parcela:

$$
\begin{aligned}
4E(X^3)E(X) &= 4np[n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np] \\
&= 4np[(n^3 - 3n^2 + 2n)p^3 + (3n^2 - 3n)p^2 + np] \\
&= np[(4n^3 - 12n^2 + 8n)p^3 + (12n^2 - 12n)p^2 + 4np]
\end{aligned}
$$

Vamos para a terceira parcela:

$$
\begin{aligned}
6E(X^2)E^2(X) &= 6E(X)E(X^2)E(X) \\
&= np6[n(n-1)p^2 + np]np \\
&= np6[(n^2 - n)p^2 + np]np \\
&= np[(6n^3 - 6n^2)p^3 + 6n^2p^2].
\end{aligned}
$$

A quarta parcela é dada por:

$$
\begin{aligned}
E^4(X) &= E(X)E^3(X) \\
&= np[n^3p^3].
\end{aligned}
$$

Finalmente temos:

$$
\begin{aligned}
\mu_4 &= np[(n^3 - 6n^2 + 11n - 6)p^3 + (6n^2 - 18n + 12)p^2 + (7n-7)p + 1] \\
&- np[(4n^3 - 12n^2 + 8n)p^3 + (12n^2 - 12n)p^2 + 4np] \\
&+ np[(6n^3 - 6n^2)p^3 + 6n^2p^2] \\
&- np[3n^3p^3] \\
&= np\left[(n^3 - 6n^2 + 11n - 6)p^3 + (6n^2 - 18n + 12)p^2 + (7n-7)p + 1 - (4n^3 - 12n^2 + 8n)p^3\right. \\
&- \left.(12n^2 - 12n)p^2 - 4np + (6n^3 - 6n^2)p^3 + 6n^2p^2 - 3n^3p^3\right] \\
&= np\left[(n^3 - 6n^2 + 11n - 6 - 4n^3 + 12n^2 - 8n + 6n^3 - 6n^2 - 3n^3)p^3\right. \\
&+ \left.(6n^2 - 18n + 12 - 12n^2 + 12n + 6n^2)p^2 + (7n - 7 - 4n)p + 1\right] \\
&= np\left[(3n-6)p^3 + (-6n+12)p^2 + (3n-7)p + 1\right] \\
&= np\left[3(n-2)p^3 - 6(n-2)p^2 + (3(n-2)-1)p + 1\right] \\
&= np\left[3(n-2)p^3 - 6(n-2)p^2 + 3(n-2)p + 1 - p\right] \\
&= np\left[3(n-2)p(p^2 - 2p + 1) + q\right] \\
&= np\left[3(n-2)p(1-p)^2 + q\right] \\
&= np\left[3(n-2)pq^2 + q\right] \\
&= npq\left[1 + 3(n-2)pq\right].
\end{aligned}
$$

1.10 Função de distribuição

Fato 9. A função de distribuição de $X \sim B(n,p)$

$$F(x) = \sum_{i=0}^{[x]} \binom{n}{i} p^i \, q^{n-i} I_{[\ 0,\ n)}(x) + I_{[\ n,\ \infty)}(x),$$

em que $[x]$ é o maior inteiro que não ultrapassa x.

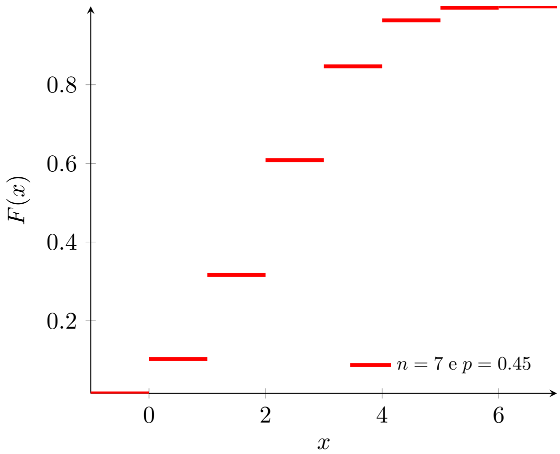

Figura 2: Gráfico da Função de Distribuição Binomial

A Figura 2 mostra a função de distribuição binomial para $n = 7$ e $p = 0,45$.

1.11 Função de sobrevivência

Fato 10. A função de distribuição de $X \sim B(n,p)$

$$S(x) = I_{(-\ \infty,\ 0)}(x) + \sum_{i=[x]+1}^{n} \binom{n}{i} p^i \, q^{n-i} I_{[\ 0,\ n)}(x).$$

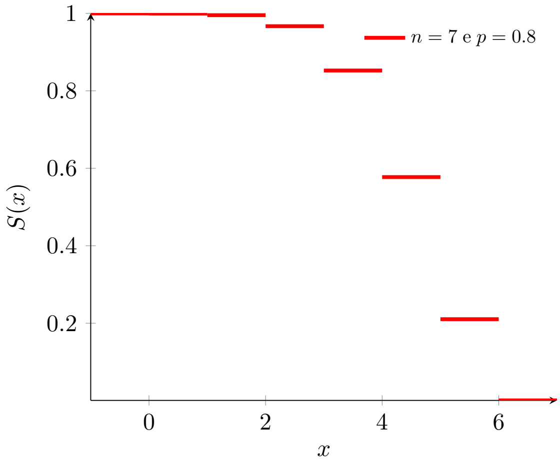

Figura 3: Gráfico da Função de Sobrevivência Binomial

A Figura 3 mostra a função de sobrevivência binomial com parâmetros $n = 7$ e $p = 0,8$.

1.12 Mediana

Fato 11. Não há uma forma fechada para a mediana da distribuição de $X \sim Bin(n, p)$ mas algumas informações especiais devem ser dadas:

a. Se np é inteiro então a média, mediana e a moda são iguais a np

```
> n=4;p=1/2;EX=n*p
> EX##### 2 é a média !!!.
[1] 2
> x=0:n
> fx=dbinom(x,n,p)
> tab1=cbind(x,fx);tab1
     x    fx
[1,] 0 0.0625
[2,] 1 0.2500
[3,] 2 0.3750
[4,] 3 0.2500
[5,] 4 0.0625
> max(fx)
[1] 0.375
> Mo=2####A moda é 2!!!!
> Fx=pbinom(x,n,p)
> tab2=cbind(x,fx,Fx);tab2
```

```
     x   fx     Fx
[1,] 0 0.0625 0.0625
[2,] 1 0.2500 0.3125
[3,] 2 0.3750 0.6875
[4,] 3 0.2500 0.9375
[5,] 4 0.0625 1.0000
    >  ##2 é uma mediana!!!
>
```

b. Uma moda da binomial é um valor inteiro pertencente ao intervalo $\big[\lfloor np \rfloor\,,\, \lceil np \rceil\big]$. ou $(n+1)p-1 \leq x \leq (n+1)p$.

```
> ####X ~Bin(n=7,p=1/4)
> n=7;p=1/4;EX=n*p;EX
[1] 1.75
> x_f=floor(EX);x_f
[1] 1
> x_t=ceiling(EX);x_t
[1] 2
> ####Assim 1  e 2 são modas. Vamos olhar na f.p.:
> x=0:n
> fx=dbinom(x,n,p)
> tab1=cbind(x,fx);round(tab1,5)
     x    fx
[1,] 0 0.13348
[2,] 1 0.31146
[3,] 2 0.31146
[4,] 3 0.17303
[5,] 4 0.05768
[6,] 5 0.01154
[7,] 6 0.00128
[8,] 7 0.00006
> max(fx)
[1] 0.3114624
> ###Confirma as modas!!!
```

c. Quando $p = 1/2$ e n é ímpar qualquer m pertencente ao intervalo

$$\frac{n-1}{2} \leq m \leq \frac{n+1}{2},$$

é mediana. Quando $p = 1/2$ e n é par então $m = \dfrac{n}{2}$ é a única mediana.

1.13 Moda

Fato 12. A moda da distribuição de $X \sim Bin(n,p)$

$$moda = \begin{cases} [\,(n+1)p\,] & (n+1)p-1 \text{ não for inteiro} \\ (n+1)p-1 \text{ e } (n+1)p & \text{se } (n+1)p-1 \text{ for inteiro} \end{cases},$$

onde $[a]$ é o maior inteiro que não ultrapassa a.

Prova: Para $x = 0, 1, \ldots, n-1$

Seja

$$g(x) = \frac{\binom{n}{x+1} p^{x+1}\, q^{n-x-1}}{\binom{n}{x} p^x\, q^{n-x}} = \frac{\binom{n}{x+1}}{\binom{n}{x}} \frac{p}{1-p}.$$

Mas,

$$\frac{\binom{n}{x+1}}{\binom{n}{x}} = \frac{n!}{(x+1)!(n-x-1)!} \frac{x!(n-x)!}{n!} = \frac{x+1}{n-x}.$$

e

$$g(x) = \frac{n-x}{x+1} \frac{p}{1-p}.$$

Vamos analisar o crescimento e o decrescimento de $g(x)$

Assim $g(x) = 1$ quando

$$\frac{n-x}{x+1} = \frac{1-p}{p},$$

e

$$\frac{x+1}{n-x} = \frac{p}{1-p},$$

$$\frac{x+1}{n+1} = \frac{p}{1} = p.$$

Logo,

$$x + 1 = (n+1)p,$$

e

$$x = (n+1)p - 1.$$

De maneira similar $g(x) > 1$ implica

$$x < (n+1)p - 1.$$

Além disso $g(x) > 1$ implica

$$x > (n+1)p - 1.$$

Precisamos analisar dois casos.

Caso 1: $(n+1)p-1$ for um inteiro.

Assim

$$g(0) < g(1) < \ldots < g((n+1)p-1) = g((n+1)p) > g((n+1)p+1) > \ldots > g(n-1) > g(n)$$

e assim a distribuição é bimodal. As modas são $Mo_1 = (n+1)p - 1$ e $Mo_1 = (n+1)p$.

Caso 2: $(n+1)p-1$ não for um inteiro.

Seja $a = [(n+1)p-1]$. Assim a função é estritamente crescente para $x = 0, 1, \ldots, a$ e a partir daí decresce. Assim o ponto de máximo é:

$$Mo = a + 1 = [\ (n+1)p\ ].$$

Assim

$$g(0) < g(1) < \ldots < g((n+1)p-1) = g((n+1)p) > g((n+1)p+1) > \ldots > g(n-1) > g(n)$$

1.14 Coeficiente de Assimetria

Fato 13. O coeficiente de assimetria de $X \sim B(n,p)$

$$\alpha_3 = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{np(1-p)}}.$$

Prova:

$$\alpha_3 = \frac{E\,(X-\mu)^3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{npq(q-p)}{(npq)^{3/2}} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{np(1-p)}} \qquad \blacksquare$$

Assim pode-se classificar a distribuição Binomial quanto à assimetria como: Se $p < 1/2$ a distribuição é assimétrica positiva, se $p > 1/2$ a distribuição é assimétrica negativa . Se $p = 1/2$ então $\alpha_3 = 0$

e nada se pode afirmar sobre a simetria da distribuição usando esse coeficiente. Vai-se usar a definição para para discutir a simetria de $X \sim Bin(n, 1/2)$.

Sua f.p., no suporte, é dada por:

$$f(x) = \frac{\binom{n}{x}}{2^n} = \frac{\binom{n}{n-x}}{2^n} = f(n-x),$$

assim se n é par então $f(x)$ é simétrica em torno do ponto $c = \frac{n}{2}$, e se n é ímpar então $f(x)$ é simétrica em torno do ponto

$$c = \frac{\left[\frac{n}{2}\right] + \frac{n+1}{2}}{2}.$$

1.15 Coeficiente de Curtose

Fato 14. O coeficiente de curtose de $X \sim B(n, p)$

$$\alpha_4 = 3 + \frac{1-6pq}{npq}.$$

Prova:

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{3n^2p^2q^2 + npq(1-6pq)}{(npq)^2} = 3 + \frac{1-6pq}{npq}.$$

Assim pode-se classificar a distribuição binomial quanto à curtose como:

$$\text{leptocurtica} \quad \text{se } pq < \frac{1}{6}.$$

$$\text{mesocurtica} \quad \text{se } pq = \frac{1}{6}. \quad .$$

$$\text{platicurtica} \quad \text{se } pq > \frac{1}{6}$$

Prova:

Como

$$\alpha_4 = 3 + \frac{1-6pq}{npq}.$$

tem-se que :

$$\text{Se } pq < \frac{1}{6}, \quad \alpha_4 > 3.$$

$$\text{Se } pq = \frac{1}{6}, \quad \alpha_4 = 3.$$

$$\text{Se } pq > \frac{1}{6}, \quad \alpha_4 < 3.$$

1.16 Coeficiente de Variação

Fato 15. O coeficiente de variação de $X \sim B(n,p)$

$$CV = \frac{\sigma}{\mu} = \sqrt{\frac{q}{np}}.$$

Prova:

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{npq}}{np} = \sqrt{\frac{q}{np}} \qquad \blacksquare$$

1.17 Aproximação Normal à Binomial

As probabilidades associadas aos experimentos binomiais são facilmente calculadas quando o parâmetro n é pequeno. Alguns livros trazem tabelas da binomial para $n \leq 25$ e $p = 0,10;\ 0,20;\ \ldots;\ 0,9$. Hoje em dia o pacote `R` resolve com bastante propriedade essas questões. Por exemplo, seja $X \sim B(10, p = 0.3)$, calcule $P(X = 5)$ e $P(X \leq 5)$. A probabilidade de $P(X = 5)$ é calculada através do comando:

```
> dbinom(5,10,0.3)
[1] 0.1029193
```

A probabilidade de $P(X \leq 5)$ é calculada através do comando

```
> pbinom(5,10,0.3)
[1] 0.952651
```

A função de probabilidade de $X \sim B(10, p = 0.3)$ é gerada através dos comandos

```
> x=0:10
> fx=dbinom(x,10,0.3)
```

A função de distribuição acumulada de $X \sim B(10, p = 0.3)$ é gerada através dos comandos:

```
> Fx=pbinom(x,10,0.3)
```

A função de sobrevivência de $X \sim B(10, p = 0.3)$ é gerada através dos comandos:

```
> Sx=1-Fx
```

A saída das três funções é dada a seguir:

```
> tabela=cbind(x,fx,Fx,Sx)
> round(tabela,4)
```

```
       x     fx     Fx     Sx
 [1,]  0 0.0282 0.0282 0.9718
 [2,]  1 0.1211 0.1493 0.8507
 [3,]  2 0.2335 0.3828 0.6172
 [4,]  3 0.2668 0.6496 0.3504
 [5,]  4 0.2001 0.8497 0.1503
 [6,]  5 0.1029 0.9527 0.0473
 [7,]  6 0.0368 0.9894 0.0106
 [8,]  7 0.0090 0.9984 0.0016
 [9,]  8 0.0014 0.9999 0.0001
[10,]  9 0.0001 1.0000 0.0000
[11,] 10 0.0000 1.0000 0.0000
```

Há uma aproximação bastante boa para a binomial, quando n é grande e p próximo de $\frac{1}{2}$. Esta aproximação funciona até para n pequeno desde que p não esteja muito próximo de 0 ou 1. Esta aproximação é dada por:

$$P(X \leq r) \approx \Phi(z), z = \frac{r + 1/2 - np}{\sqrt{np(1-p)}},$$

onde Φ é a função de distribuição acumulada da normal padrão e o fator $\frac{1}{2}$ é chamado de correção de continuidade.A tabela a seguir apresenta os valores exatos e aproximados de $X \sim B(10, 0.3)$. Note que mesmo para $n = 10$ a aproximação funciona.

```
> m=10;p=0.3
> Faprox=pnorm((x +1/2-m*p)/sqrt(m*p*(1-p)))
>
> tabela=cbind(x,Fx,Faprox);round(tabela,4)
       x     Fx Faprox
 [1,]  0 0.0282 0.0422
 [2,]  1 0.1493 0.1503
 [3,]  2 0.3828 0.3650
 [4,]  3 0.6496 0.6350
 [5,]  4 0.8497 0.8497
 [6,]  5 0.9527 0.9578
 [7,]  6 0.9894 0.9921
 [8,]  7 0.9984 0.9990
 [9,]  8 0.9999 0.9999
[10,]  9 1.0000 1.0000
 [11,] 10 1.0000 1.0000
```

Seja $X \sim B(50, 1/2)$. Calcule $P(X \leq 30)$ exatamente e através da aproximação normal sem e com a correção de continuidade.

```
> pexata=pbinom(30,50,1/2);pexata
[1] 0.9405398
> paprox=pnorm(30,50*1/2,sqrt(50*1/2*1/2));paprox
[1] 0.9213504
>
> paproxcc=pnorm(30+1/2,50*1/2,sqrt(50*1/2*1/2));paproxcc
[1] 0.9401025
>
```

Percebe-se que a aproximação é realmente interessante.

1.18 Transformações Importantes

Fato 16. Se $X \sim B(m,p)$ e $Y \sim B(n,p)$ são variáveis aleatórias independentes, então $S = X + Y \sim B(m+n,p)$.

Prova: A função geradora de S é dada por

$$\begin{aligned} G_S(t) &= G_{X+Y}(t) = G_X(t)G_Y(t) \\ &= (pt+q)^m(pt+q)^n = (pt+q)^{m+n}, \end{aligned}$$

que é a f.g.p. de uma binomial de parâmetros $(m+n)$ e p e cuja f.p. é dada por

$$f(s) = \binom{m+n}{s} p^s q^{m+n-s} \; I_{\{0,1,\ldots,m+n\}}(s) \qquad \blacksquare$$

Fato 17. Se $X \sim B(m,p)$ e $Y \sim B(n,p)$, então e $X|S = s \sim Hiper(m+n,m,n)$.

Prova:

$$\begin{aligned} P(X|S=s) &= \frac{P(X=x,\, X+Y=s)}{P(S=s)} = \frac{P(X=x,\, Y=s-x)}{P(S=s)} \quad \text{(por independência)} \\ &= \frac{P(X=x)P(Y=s-x)}{P(S=s)} = \frac{\binom{m}{x}p^x q^{m-x}\binom{n}{s-x}p^{s-x}q^{n-s+x}}{\binom{m+n}{s}p^s q^{m+n-p}} = \frac{\binom{m}{x}\binom{n}{s-x}}{\binom{m+n}{s}} I_{\{0,1,\ldots,s\}}, \end{aligned}$$

que é a função de probabilidade da hipergeométrica de parâmetros $N = m+n$, $A = m$ e $n = s$ $\blacksquare$

Fato 18. Sejam $X_1, X_2, \ldots, X_n \overset{iid}{\sim} B(p)$, então $S = \sum_{i=1}^{n} X_i \sim B(n,p)$.

Prova: Sabemos que $G_X(t) = pt + q$ e que $G_S(t) = \prod_{i=1}^{n} G_{X_i}(t) = (pt+q)^n$. Assim $S \sim B(n,p)$ $\blacksquare$

Vamos obter a esperança e variância de uma variável aleatória que segue binomial de outra forma:

$$E[S] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np, \text{ e por independência}$$

$$Var[S] = \sum_{i=1}^{n} Var[X_i] = \sum_{i=1}^{n} pq = npq.$$

1.19 Aproximação Poisson à Binomial

Para grandes valores de n e pequenos valores de p usa-se uma aproximação Poisson para a binomial da seguinte maneira:

$$P(X = r) \approx= \frac{(np)^r \ e^{-np}}{r!}.$$

Seja $X \sim B(300, p = 1/300)$. Calcular $P(X > 4)$. Olhe com cuidado a saída do `R` e comente.

```
>
> n=300;p=1/300
> mu=n*p;mu
[1] 1
> vx=n*p*(1-p);vx
[1] 0.9966667
> pexata= 1-pbinom(4,n,p);pexata
[1] 0.003583469
>
> paproxpoi=1-ppois(4,mu);paproxpoi
[1] 0.003659847
> paproxnorm=1- pnorm(4+1/2,mu,sqrt(vx));paproxnorm
[1] 0.0002275776
```

Note que a média e a variância da aproximação são quase iguais. Isto é um indicativo que a distribuição de Poisson pode ser usada, pois pode ser comprovada comparando as duas probabilidades obtidas pela binomial diretamente com a obtida pela Poisson. Por outro lado, mesmo com $n = 300$, a aproximação pela normal não é tão eficiente.

1.20 Estimação do parâmetro p

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória tamanho m de $X \sim B(m, p)$. O número total de sucessos da amostra, $S = \sum_{i=1}^{n} X_i$, é usado para se estimar p. Assim a frequência amostral de sucessos é usada como estimador de p. Ele é definido por:

$$\hat{p} = \frac{S}{mn} = \frac{\sum_{i=1}^{n} X_i}{mn}$$

Pode-se demonstrar que $E(\hat{p}) = p$ e $Var(\hat{p}) = \dfrac{pq}{mn}$. Além disso, S tem uma distribuição binomial de parâmetros nm e p.

1.21 Simulação

Vai-se discutir agora a simulação de uma amostra aleatória de tamanho n de $X \sim B(m,p)$ usando o pacote estatístico `R`. Como exemplo vai-se simular uma amostra de tamanho 100 de $X \sim B(5, 0.4)$ Isto é feito através dos comandos:

0. set.seed(32) : garante a distinta reprodução da amostra;
1. Amostra=rbinom(100,5,0.4) : Simula a amostra;
2. table(Amostra) : Tabela de frequências absolutas da amostra;
3. mean(Amostra) : Estimativa da probabilidade de sucesso; e
4. var(Amostra) : Variância amostral.

```
> set.seed(32)
> m=5;p=0.4
> muX=m*p;muX #média populacional
[1] 2
> varX=m*p*(1-p);varX #variância populacional
[1] 1.2
> Amostra=rbinom(100,m,p);Amostra
  [1] 2 2 3 3 1 4 3 3 2 2 2 1 2 3 1 3 3 2 2 3 3 3 2 1 2 4 2
 [28] 2 2 2 2 3 3 3 2 3 2 1 3 4 2 3 1 2 3 1 2 4 3 0 3 4 3 2
 [55] 0 0 3 3 2 4 3 1 3 1 1 5 3 1 4 2 3 1 2 0 1 1 3 1 2 2 2
 [82] 2 2 0 1 2 2 3 1 3 0 1 1 1 1 0 1 2 1 1
> table(Amostra)
Amostra
 0  1  2  3  4  5
 7 24 32 29  7  1
> mean(Amostra);var(Amostra)
[1] 2.08 [1] 1.185455
>
```

1.22 Exercícios Resolvidos

1. Seja $X \sim B(20; 0,3)$. Calcule usando o R:

a $P(X = 10)$ e $P(X > 12)$

```
> ## X~B(20;0,3); P(X=10), P(X >12)=1-F(12)
> dbinom(10,20,0.3)
[1] 0.03081708
> 1-pbinom(12,20,0.3)
[1] 0.001278880
> pbinom(12,20,0.3,lower.tail=F)
[1] 0.001278880
```

b. Qual a média e a variância de X?

```
> x=0:20
> fx=dbinom(x,20,0.3)
> EX=sum(x*fx);EX
[1] 6
> VX=sum((x-6)^2*fx);VX
[1] 4.2
```

1.23 Exercícios propostos

1. (Bussab & Morettin-pgs 151 e 152- Questão 20) Para os exercícios (a) a (e), considere o enunciado:

   Das variáveis descritas, a seguir, assinale quais são binomiais, e para essas dê os respectivos campos de definição e função de probabilidade. Quando julgar que a variável não é binomial, aponte as razões de sua conclusão.

   a. De uma urna com dez bolas brancas e 20 pretas, vamos extrair, com reposição, cinco bolas. X é o número de bolas brancas nas cinco extrações.

   b. Refaça o exercício anterior, mas dessa vez as cinco extrações são sem reposição.

   c. Temos cinco urnas com bolas pretas e brancas e vamos extrair uma bola de cada urna. Suponha que X seja o número de bolas brancas obtidas no final.

   d. Vamos realizar uma pesquisa em dez cidades brasileiras, escolhendo ao caso um habitante de cada uma delas e classificando-o como pró ou contra um projeto federal. Suponha que X seja o número de indivíduos contra o projeto no final da pesquisa.

   e. Em uma indústria existem 100 máquinas que fabricam determinada peça. Cada peça é classificada como boa e defeituosa. Escolhemos ao acaso um instante de tempo e verificamos um peça de cada uma das máquinas. Suponha que X seja o número de peças defeituosas.

2. (Bussab & Morettin-pg 152- Questão 21) Seja $X \sim B(n, p)$ sabendo-se que $E(X) = 12$ e $\sigma^2 = 3$, determinar:

   a. n.

b. p.

c. $P(X < 12)$.

d. $P(X \geq 14)$.

e. $E(Z)$ e $Var(Z)$, onde $Z = \dfrac{X - 12}{\sqrt{3}}$.

f. $P(Y \geq \dfrac{14}{16})$, onde $Y = \dfrac{X}{n}$.

g. $(Y \geq \dfrac{12}{16})$, onde $Y = \dfrac{X}{n}$.

3. (Bussab & Morettin-pgs 152 e 153- Questão 25) Examinaram-se 2000 ninhadas de cinco porcos cada uma, segundo o número de machos. Os dados estão representados na tabela a seguir:

Número de Machos	Número de Ninhadas
0	20
1	360
2	700
3	680
4	200
5	40
Total	2000

a. Calcule a proporção média de machos.

b. Calcule, para cada valor de X, o número de ninhadas que você deve esperar se $X \sim Bin(5, p)$, onde p é a proporção média de machos calculada em a.

4. (Bussab & Morettin-pg 153- Questão 26) Se X tem distribuição binomial com parâmetros $n = 5$ e $p = 1/2$, faça os gráficos da função de probabilidade de X e da função de distribuição acumulada de X. Faça também usando o R. Comente!!!!!!

5. (Bussab & Morettin-pg 157- Questão 31) Na manufatura de certo artigo, é sabido que um entre dez artigos é defeituoso. Qual a probabilidade de que uma amostra casual de tamanho quatro contenha:

a. nenhum defeituoso?

b. exatamente um defeituoso?

c. exatamente dois defeituosos?

d. não mais do que dois defeituosos?

e. Refaça usando o R.

6. (Bussab & Morettin-pg 158- Questão 33) Um curso de treinamento aumenta a produtividade de uma certa população de funcionários em 80% dos casos. Se dez funcionários quaisquer participam desse curso, encontre a probabilidade de que:

a. exatamente sete funcionários aumentarem a produtividade.

b. não mais que oito funcionários aumentarem a produtividade.

c. pelo menos três funcionários não aumentarem a produtividade.

e. Refaça usando o R.

7. (Bussab & Morettin-pg 160- Questão 40) Um industrial fabrica peças, das quais 1/5 são defeituosas. Dois compradores A e B, classificaram as partidas adquiridas em categorias I e II, pagando 1,20 u.m. e 0,80 respectivamente so 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 maior lucro?

8. (Bussab & Morettin-pg 160- Questão 48) Num teste tipo certo/errado, com 50 questões, qual é a probabilidade de que um aluno acerte 80% das questões, supondo que ele as responda ao acaso? Indique sua resposta e calcule-a usando o R.

9. (Bussab & Morettin-pg 160- Questão 50) Em um experimento binomial com três provas, a probabilidade de exatamente dois sucessos é 12 vezes a probabilidade de três sucessos. Encontre p.

10. (Bussab & Morettin-pg 161- Questão 52) Se $X \sim Bin(n, p)$ prove que:

$$P(X = k + 1) = \frac{(n - k)}{(k + 1)(1 - p)} P(X = k).$$

11. (Montgomery & Runger -pg 61 -Questão 4.52) Um produto eletrônico contém 40 circuitos integrados. A probabilidade de que qualquer circuito integrado seja defeituoso é de 0,01. Os circuitos integrados são independentes. O produto opera somente se não houver circuitos integrados defeituosos. Qual é a probabilidade de que o produto opere?

12. (Montgomery & Runger -pg 61 -Questão 4.54) As linhas telefônicas em um sistema de reservas de uma companhia aérea estão ocupadas 40% do tempo. Suponha que os eventos em que as linhas estejam ocupadas em sucessivas chamadas sejam independentes. Considere que 10 chamadas aconteçam

    a. Qual é a probabilidade de que, para exatamente três chamadas as linhas estejam ocupadas?

    b. Qual é a probabilidade de que, para no mínimo uma chamada, linhas não estejam ocupadas?

    c. Qual é o número esperado de chamadas em que todas as linhas estejam ocupadas?

13. (Montgomery & Runger -pg 61 -Questão 4.56) Um exemplo de gráfico de controle estatístico de processo.

Amostras de 20 peças de um processo de corte metálico são selecionadas a cada hora. tipicamente, 1% das peças requer conserto. Seja X o número de peças na amostra de 20 que requerem conserto. Suspeita-se de um problema no processo se X exceder sua média em mais de três desvios padrões.

a. Se a percentagem de peças que requerem conserto permanecer em 1%, qual será a probabilidade de X exceder sua média em mais de três desvios padrões ?

b. Se a percentagem de conserto aumentar para 4%, qual é a probabilidade de qual será a probabilidade de X exceder um?

c. Se a percentagem de conserto aumentar para 4%, qual é a probabilidade de qual será a probabilidade de X exceder um em, no mínimo, uma das próximas cinco horas de amostras?

14. (Montgomery & Runger -pg 61 -Questão 4.57 Porque nem todos os passageiros de aviões aparecem na hora do embarque, uma companhia aérea vende 125 bilhetes para um vôo que suporta 120 passageiros. A probabilidade de que um passageiro não apareça é 0,10 e os passageiros se comportam independentemente. Qual é a probabilidade de que:

    a. cada passageiro que apareça possa embarcar?

    b. o vôo decole com assentos vazios?

15. (Davidi-pg 128-questão 6) Uma grande empresa industrial permite um desconto de 30% em qualquer fatura paga dentro de 30 dias. De todas as faturas, 10% recebem o desconto. Em uma auditoria na empresa, 12 faturas são amostradas aleatoriamente.

    a. Qual é a probabilidade de exatamente 4 delas receberem o desconto?

    b. Qual é a probabilidade de menos do que três delas receberem o desconto?

    c. Qual é a probabilidade de nenhuma delas receberem o desconto?

    d. Determine o número médio que recebe o desconto?

    e. Determine o desvio padrão do número que recebe o desconto?

16. (Davidi-pg 129-questão 11) Um sistema k de n é aquele no qual existe n componentes e o sistema funciona se pelo menos k dos componentes funcionarem. Considere que os componentes funcionam independentemente um do outro.

    a. Em um sistema 3 de 5, cada componente tem probabilidade 0,9 de funcionar. Qual é a probabilidade do sistema funcionar?

    b. Em um sistema 3 de n, cada componente tem probabilidade 0,9 de funcionar, qual é o menor valor de n necessário de modo que a probabilidade do sistema funcionar seja pelo menos 0,90?

17. (Davidi-pg 129-questão 12) Para uma determinado sistema 4 de 6, considere que, em um dia chuvoso, cada componente de funcionar, e que, em um dia não chuvoso, cada componente tem probabilidade 0,9 de funcionar.

a. Qual é a probabilidade do sistema funcionar em um dia chuvoso?

b. Qual é a probabilidade do sistema funcionar em um dia não chuvoso?

c. Considere que a probabilidade de chover amanhã é 0,20. Qual é a probabilidade do sistema funcionar amanhã?

18. (Davidi-pg 129-questão 13) Um grande carregamento vem com uma garantia de que não há mais do que 15% de itens defeituosos. Se a proporção de itens defeituosos for maior do que 15%, o carregamento pode ser devolvido. Você retira uma amostra de 10 itens. Seja X o número de itens defeituosos na amostra.

    a. Se de fato 15% dos itens no carregamento estiverem com defeitos(de modo que o carregamento é bom, mas por pouco), qual é o valor de $P(X \geq 7)$?

    b. Baseado na resposta em (a), se 5% dos itens no carregamento estiverem com defeitos, sete itens defeituosos em uma amostra de 10 seria um número excepcionalmente grande?

    c. Se você determinar que sete dos 10 itens da amostra estavam com defeitos, isso seria uma evidencia convincente de que o carregamento deveria ser devolvido?

    d. Se de fato 15% dos itens no carregamento estiverem com defeitos(de modo que o carregamento é bom, mas por pouco), qual é o valor de $P(X \geq 2)$?

    e. Baseado na resposta em (d), se 5% dos itens no carregamento estiverem com defeitos, dois itens defeituosos em uma amostra de 10 seria um número excepcionalmente grande?

    f. Se você encontrar dois dos 10 itens da amostra estavam com defeitos, isso seria uma evidencia convincente de que o carregamento deveria ser devolvido? Explique.

19. (Davidi-pg 129-questão 14) Uma companhia de seguros oferece um desconto aos proprietários de casas que instalam detectores de fumaça em suas casas. Um representante da companhia afirma que 80% ou mais dos segurados tem detectores de fumaça. Você extrai uma amostra aleatória de oito segurados. Seja X o número de segurados na amostra que tem detectores de fumaça.

    a. Se exatamente 80% dos segurados tem detectores de fumaça(de modo que a afirmação do representante seja verdadeira, mas por pouco), qual é o valor de $P(X \leq 1)$?

    b. Baseado na resposta em (a), se 80% dos segurados tem detectores de fumaça, um deles com detector de fumaça em uma amostra de tamanho 8 em um número excepcionalmente pequeno?

    c. Se você identificar que exatamente um dos oito da amostra de segurados tem um detector de fumaça, seria uma evidência convincente de que a afirmação é falsa?

    d. Se exatamente 80% dos segurados têm detectores de fumaça, qual é o valor de $P(X \leq 6)$?

    e. Baseado na sua resposta em (d), se 80% dos segurados tem detectores de fumaça, seis deles com detectores de fumaça em uma amostra de tamanho 8 seria um número excepcionalmente pequeno?

    f. Se você identificou que exatamente seis na amostra de 8 clientes de seguro tem detectores de fumaça, seria uma evidência convincente de que a afirmação é falsa? Explique.

20. (Davidi-pg 130-questão 15) Uma mensagem consiste em uma sequência de bits (0 ou 1s). Devido ao ruído no canal de comunicação, cada bit tem probabilidade 0,3 de ser invertido (ou seja, um 0 ser trocado por um 1 ou vice-versa). Para melhorar a precisão da comunicação, cada bit é enviado cinco vezes. Por exemplo, 0 é enviado como 00000. O receptor atribui o valor de 0 se três ou mais desses forem decodificados como 0 e atribui 1 se três ou mais bits forem decodificados como 1. Considere que os erros ocorrem independentemente.

   a. Um 0 é enviado (como 00000). Qual é a probabilidade do receptor atribuir o valor correto a 0?

   b. Considere que cada bit seja enviado n vezes, em que n é um número ímpar e que o receptor atribui o valor decodificado na maioria dos bits. Qual é o valor mínimo de n necessário de modo que a probabilidade do valor correto atribuído ser pelo menos 0,90?

21. (Davidi-pg 130-questão 16) Um projeto de um sistema requer a instalação de dois componentes idênticos. O sistema funcionará se pelo menos um dos componentes funcionarem. Um projeto alternativo requer quatro desses componentes e o sistema funcionará se pelo menos dois dos quatro componentes funcionarem. Se a probabilidade de um componente funcionar é de 0,9 e se os componentes funcionam independentemente, qual projeto tem maior probabilidade de funcionar?

22. (LISPCHUTZ-pg 201 Questão 6.32) Três cartas são selecionadas, com reposição, de um baralho comum de 52 cartas. Encontre a probabilidade de

   (i) duas copas serem retiradas;

   (ii) três serem retiradas; e

   (iii) pelo menos um copa ser retirada.

23. (LISPCHUTZ-pg 201 Questão 6.33) A média de rebatidas de um jogador de beisebol é 0,3. Se ele bate 4 vezes, qual a probabilidade dele obter

   (i) dois acertos; e

   (ii) pelo menos um acerto?

24. (LISPCHUTZ-pg 201 Questão 6.34) Um caixa contém 3 bolas vermelhas e 2 brancas. Três bolas são retiradas, com reposição, da caixa. Encontre a probabilidade de

   (i) 1 bola vermelha ser retirada;

   (ii) 2 bolas vermelhas serem retiradas; e

   (iii) pelo menos uma bola vermelha ser retirada.

25. (LISPCHUTZ-pg 201 Questão 6.35) O time A tem $\frac{2}{5}$ de probabilidade de vitória sempre que joga. Se disputa 4 partidas, encontre a probabilidade de vencer duas partidas, pelo menos uma partida e mais que a metade das partidas.

26. (LISPCHUTZ-pg 201 Questão 6.36) Uma carta é retirada e recolocada em um baralho comum. Quantas vezes devemos repetir esse procedimento, de mo que

    (i) a probabilidade de retirar uma copa seja maior que $\frac{1}{2}$; e

    (ii) a probabilidade de retirar uma copa seja maior que $\frac{3}{4}$?

27. (LISPCHUTZ-pg 202 Questão 6.37) A probabilidade de um homem acertar um alvo é $\frac{1}{3}$. Então

    (i) e se ele atirar 5 vezes, qual é a probabilidade de acertar o alvo pelo menos duas vezes?; e

    (ii) quantas vezes ele deve atirar de modo que a probabilidade de acertar, ao menos uma vez, seja maior que 90%?

28. (LISPCHUTZ-pg 202 Questão 6.38) O Departamento de Matemática tem 8 professores graduados ocupando o mesmo gabinete. Cada um tanto estuda em casa como no gabinete. Quantas escrivaninhas deve haver no gabinete de modo que cada professor tenha uma pelo menos 90% do tempo?

29. (LISPCHUTZ-pg 202 Questão 6.39) Dos parafusos produzidos por uma fábrica, 2% são defeituosos. Em cada depósito de 3600 parafusos da fábrica, encontre o n^o esperado de parafusos com defeitos e o respectivo desvio-padrão.

30. (LISPCHUTZ-pg 202 Questão 6.40) Um dado não-viciado é atirado 1620 vezes. Encontre o número esperado de vezes em que o número 6 ocorre e o desvio-padrão.

31. (LISPCHUTZ-pg 202 Questão 6.41) Seja X v.a. com distribuição binomial com $E(X) = 2$ e $Var(X) = \frac{4}{3}$. Encontre a distribuição de X.

32. (Dalton & Ogliari-pg 206 Questão 2) Suponhamos que a porcentagem de germinação de sementes de feijão seja de 70%. Vão ser semeadas quatro sementes por cova, as quais serão espaçadas de 0,4m entre as linhas e 0,2m entre covas. Supondo-se que cada canteiro a ser semeado conste de seis linhas de 5m de comprimento, qual o número médio esperado de covas falhadas (nem uma semente germinou, das quatros semeadas) por canteiro?

33. (Dalton & Ogliari-pg 207 Questão 3) Sabendo-se qua a probabilidade de nascer uma bezerra (bezerro fêmea) é $0,6$. Calcular o valor esperado (esperança) e desvio-padrão da variável número de bezerras nascidas em 30 partos.

34. (Dalton & Ogliari-pg 207 Questão 6) Um produtor de sementes vende pacotes com 20 sementes cada. Os pacotes que apresentarem mais de uma semente sem germinar serão indenizados. A probabilidade de uma semente germinar é $0,98$. Portanto,

    a. calcule a média e a variância da variável aleatória “número de sementes que não germinam por pacote”;

    b. qual a probabilidade de um pacote não ser indenizado?;

c. se o produtor vende mil pacotes, qual o número esperado de pacotes indenizados?; e

d. quando o pacote é indenizado, o produtor tem um prejuízo de 1,2u.m. (unidades monetárias) e, quando o pacote não for indenizado, ele tem um lucro de 2,5u.m. Qual o lucro líquido esperado por pacote?

35. (Dalton & Ogliari-pg 207 Questão 7) Sabendo que a probabilidade de um animal submetido a um certo tratamento não sobreviver é 0,2 e que esse tratamento foi aplicado em 20 animais, calcule, para X sendo o número de animais não-sobreviventes:

   a. a esperança, variância e desvio-padrão;

   b. $P(2 < X \leq 4)$; e

   c. $P(X \geq 2)$.

36. (Dalton & Ogliari-pg 209–210 Questão 14) Avaliaram-se 30 parcelas (área de terra de 20m^2) de 27 plantas de mandioca cada uma, de acordo com o número de plantas doentes colhidas. Os dados estão apresentados na tabela a seguir:

Número de plantas doentes colhidas (X)	Número de plantas com X plantas doentes
0	14
1	8
2	4
3	3
4	1
TOTAL	30

   a. calcule a proporção média de plantas doentes;

   b. calcule para cada X, o número de parcelas que você deve esperar se $X \sim B(27, \pi)$, onde π é a proporção média calculada em (a);

   c. Existe uma boa aproximação entre as proporções observadas e as estimadas pelo modelo binomial?; e

   d. No R, faça o gráfico, onde o eixo das ordenadas tem-se as probabilidades (de encontrar 0 plantas doentes por parcela, 1 planta doente por parcela, 2 plantas doentes por parcela, 3 plantas doentes por parcela e 4 plantas doentes por parcela) no eixo das abscissas têm-se o número de plantas doentes colhidas (0,1,2,4).

37. (Dalton & Ogliari-pg 212 Questão 24) De um lote de sementes cuja proporção de germinação é 0,7, foi retirada uma amostra de 10 sementes e colocada em um germinador. Se X é o número de sementes germinadas, então

a. apresente a função de probabilidade de X e, no R, faça o gráfico da f.p. de X;

b. determine a probabilidade de germinarem duas ou mais sementes; e

c. determine a média, a variância e o desvio-padrão da distribuição de X.

38. (Morgado et al.-pg 168 Questão 1) Sacam-se, com reposição, 4 bolas de uma urna que contém 7 bolas brancas e 3 bolas pretas. Qual a probabilidade de serem sacadas 2 bolas de cada cor? Qual seria a resposta no caso sem reposição?

39. (Morgado et al.-pg 168 Questão 2) Lança-se um dado não-viciado até a obtenção do terceiro 6. Seja X o número do lançamento em que isso ocorre. Calcule

    a. $P(X = 10)$;

    b. $P(X > 10)$; e

    c. $P(X < 10)$.

40. (Morgado et al.-pg 168 Questão 3) Dois adversários A e B disputam uma série de 10 partidas. A probabilidade de A ganhar uma partida é 0,6 e não há empates. Qual é a probabilidade de A ganhar?

41. (Morgado et al.-pg 168 Questão 4) Dois adversários A e B disputam uma série de partidas. o primeiro que obtiver 12 vitórias ganha a série. No momento, o resultado é 6×4 a favor de A. Qual a probabilidade de A ganhar a série sabendo que em cada partida as probabilidades de A e B vencerem são respectivamente 0,4 e 0,6?

42. (Morgado et al.-pg 169 Questão 5) Motores de avião funcionam independentemente e cada motor tem uma probabilidade p de falhar durante um voo. Um avião voa com segurança se a maioria de seus motores funciona. Para que valores de p um avião com 3 motores é preferível a um avião com 5 motores?

43. (Morgado et al.-pg 169 Questão 6) Suponha que uma característica (como a cor dos olhos, por exemplo) dependa de um par de genes. Representaremos por A um gene dominante e por a um gene recessivo. Assim um indivíduo com genes AA é dominante puro, um com genes aa é recessivo puro e com genes Aa é híbrido. Dominantes e híbridos são semelhantes em relação à característica. Filhos recebem um gene do pai e um da mãe. Suponha que os pais sejam híbridos e tiveram 4 filhos. Pergunta-se

    a. qual é a probabilidade do primeiro filho ser um recessivo puro?; e

    b. qual a probabilidade de exatamente um dos 4 filhos ser um recessivo puro?

44. (Morgado et al.-pg 169 Questão 7) (*O problema das caixas de fósforos de Banach*) Um matemático sai de casa todos os dias com duas caixas de fósforos, cada uma com n palitos. Toda vez que ele quer acender um cigarro, ele pega (ao acaso) uma das caixas e retira daí um palito. O matemático é meio distraído, de modo que quando ele retira o último palito de uma caixa, ele não percebe que

a caixa fica vazia. Como ele fuma muito, em certa hora ele pega uma caixa e constata que ela está vazia. Qual a probabilidade de nesse momento a outra caixa conte exatamente k $(0 \leq k \leq n)$ palitos?

45. (Morgado et al.-pg 169 Questão 8) Lança-se repetidamente um par de dados não-tendenciosos. Qual a probabilidade de obtermos duas somas iguais a 7 antes de obtermos três somas iguais a 3?

46. (Morgado et al.-pg 169 Questão 9) Uma moeda tem probabilidade 0,4 de dar cara. Lançando-a 12 vezes qual o mais provável valor do número de caras obtidas?

47. (Pagano & Gauvreau-pg. 173 Questão 10) Considere um grupo de sete indivíduos selecionados de uma população de 65 a 74 anos nos Estados Unidos. O número de pessoas que sofrem de diabetes nessa amostra é uma variável aleatória binomial com parâmetros $n = 7$ e $p = 0,125$. Portanto,

    a. Se você deseja fazer uma lista de sete pessoas escolhidas, de quantas maneiras elas podem ser ordenadas?;

    b. Sem se preocupar com a ordem, de quantas maneiras diferentes você pode selecionar quatro indivíduos desse grupo de sete?;

    c. Qual a probabilidade de que quatro deles tenham diabetes?; e

    d. Qual a probabilidade de que quatro deles tenham diabetes?