1)

```
## Hipótese Nula (H0): p =  0.25 ; Hipótese Alternativa (H1) p =!=  0.25
## Regra de decisão:
## Se p-valor < alpha ( 0.05 ) : Rejeita-se H0
## Se p_valor >= alpha ( 0.05 ): Não rejeita-se H0
## Nível de significância:  98.8941 %
## Valor de p: 0.011059
```

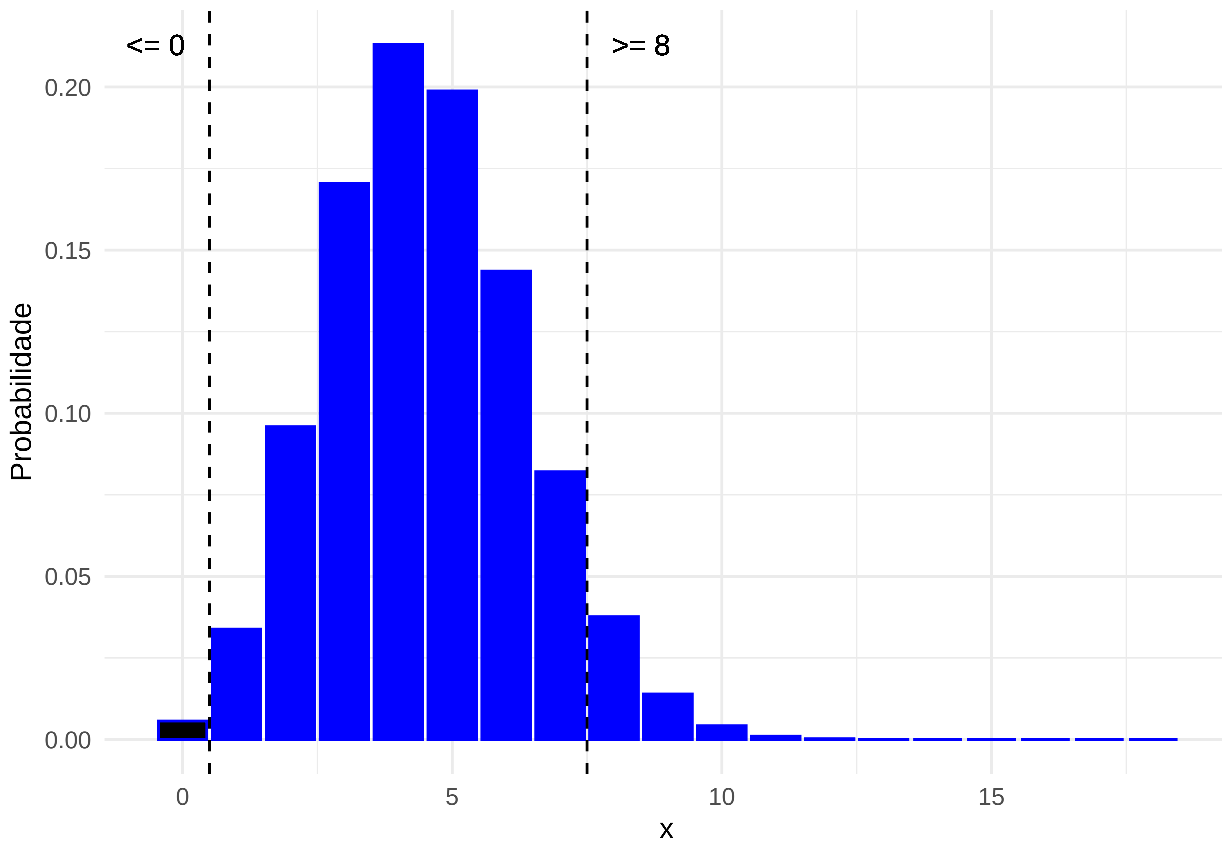


Tabela 1: Binomial (18 , 0.25)

X	Pontual	Acumulada
0	0.005638	0.005638
1	0.033826	0.039464
2	0.095841	0.135305
3	0.170384	0.305689
4	0.212980	0.518669
5	0.198781	0.717451
6	0.143564	0.861015
7	0.082037	0.943052
8	0.037600	0.980652
9	0.013926	0.994578
10	0.004178	0.998756
11	0.001013	0.999769
12	0.000197	0.999966

X	Pontual	Acumulada
13	0.000030	0.999996
14	0.000004	1.000000
15	0.000000	1.000000
16	0.000000	1.000000
17	0.000000	1.000000
18	0.000000	1.000000

Como mostrado acima, o valor observado na amostra (0) ocorreu na região de rejeição (X $\leq$ 0 ou X $\geq$ 8), e portanto rejeitamos a hipótese $H_0 : p = 0.25$. Temos assim relevância estatística para acreditar que a amostra observada não têm a mesma probabilidade de exibir a característica A apresentada pela espécie em geral.

2)

```
## Hipótese Nula (H0): p =  0.4 ; Hipótese Alternativa (H1) p >  0.4
## Regra de decisão:
## Se p-valor < alpha ( 0.05 ) : Rejeita-se H0
## Se p_valor >= alpha ( 0.05 ): Não rejeita-se H0
## Nível de significância:  71.6063 %
## Valor de p: 0.283937
```

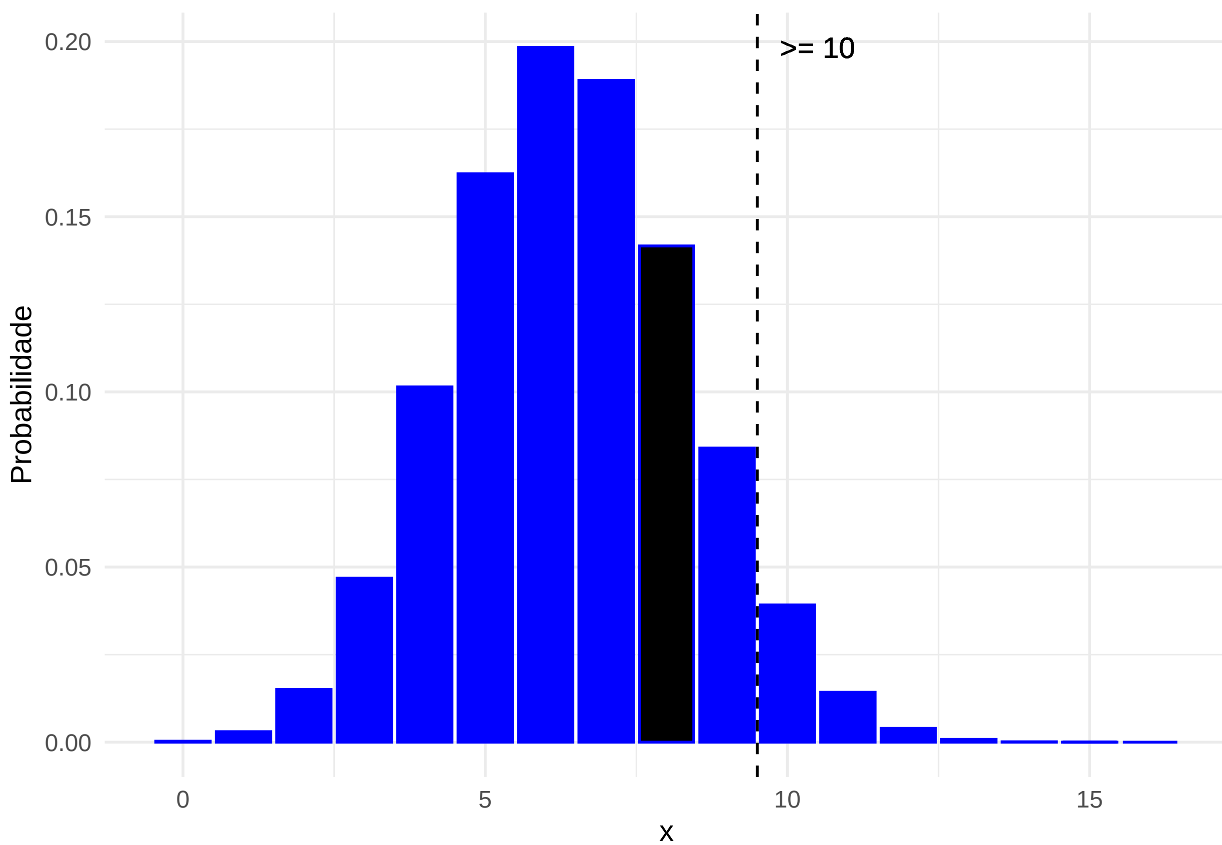


Tabela 2: Binomial (16 , 0.4)

X	Pontual	Acumulada
0	0.000282	0.000282
1	0.003009	0.003291
2	0.015046	0.018337
3	0.046810	0.065147
4	0.101421	0.166567
5	0.162273	0.328840
6	0.198334	0.527174
7	0.188889	0.716063
8	0.141667	0.857730
9	0.083951	0.941681
10	0.039177	0.980858
11	0.014246	0.995104
12	0.003957	0.999062
13	0.000812	0.999873
14	0.000116	0.999989
15	0.000010	1.000000
16	0.000000	1.000000

Como mostrado acima, o valor observado na amostra (8) não ocorreu na região de rejeição (X $\geq$ 10), e portanto não rejeitamos a hipótese $H_0 : p = 0.40$.

3)

```
## Hipótese Nula (H0): p =  0.4 ; Hipótese Alternativa (H1) p =!=  0.4
## Regra de decisão:
## Se p-valor < alpha ( 0.05 ) : Rejeita-se H0
## Se p_valor >= alpha ( 0.05 ): Não rejeita-se H0
## Nível de significância:  81.4451 %
## Valor de p: 0.185549
```

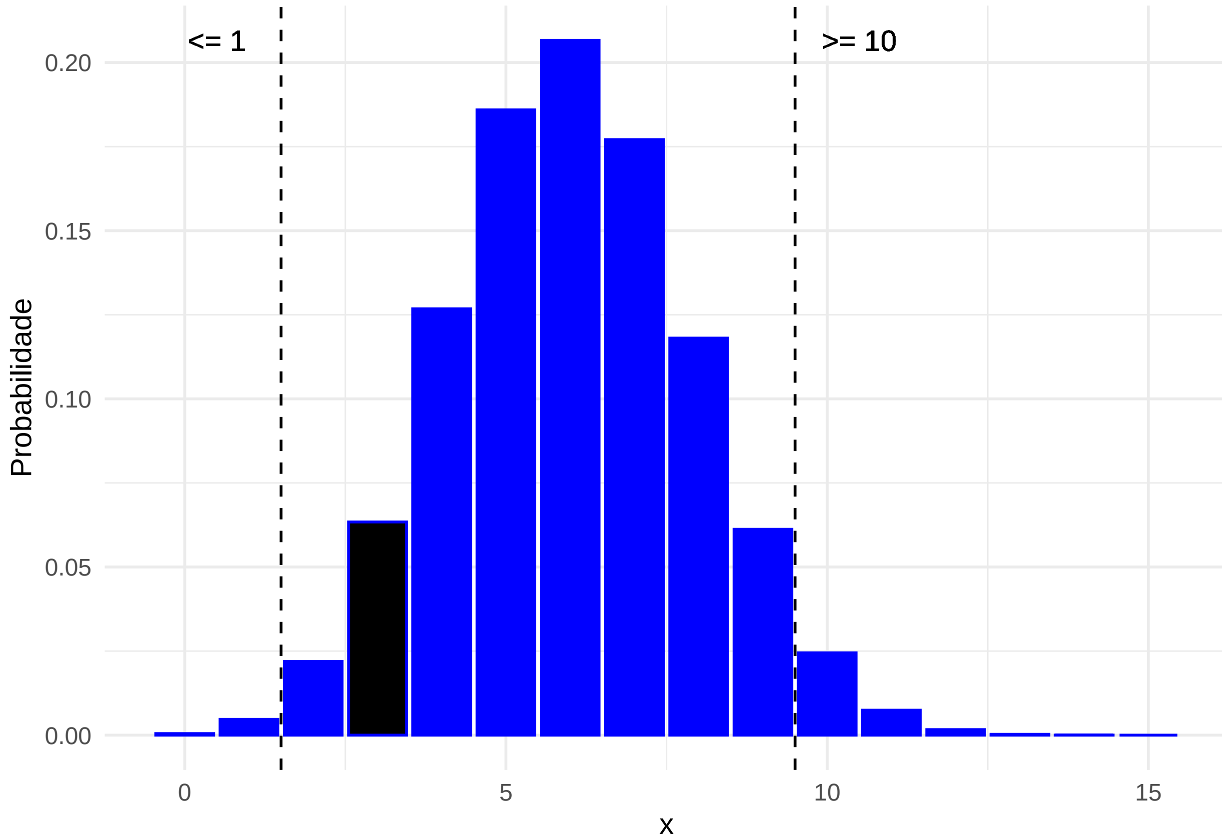


Tabela 3: Binomial (15 , 0.4)

X	Pontual	Acumulada
0	0.000470	0.000470
1	0.004702	0.005172
2	0.021942	0.027114
3	0.063388	0.090502
4	0.126776	0.217278
5	0.185938	0.403216
6	0.206598	0.609813
7	0.177084	0.786897
8	0.118056	0.904953
9	0.061214	0.966167
10	0.024486	0.990652
11	0.007420	0.998072
12	0.001649	0.999721
13	0.000254	0.999975
14	0.000024	0.999999
15	0.000001	1.000000

Como mostrado acima, o valor observado na amostra (3) não ocorreu na região de rejeição (X $\leq$ 1 ou X $\geq$ 10), e portanto não rejeitamos a hipótese $H_0 : p = 0.40$.

4)

```
## Hipótese Nula (H0): p =  0.3 ; Hipótese Alternativa (H1) p >  0.3
## Regra de decisão:
## Se p-valor < alpha ( 0.05 ) : Rejeita-se H0
## Se p_valor >= alpha ( 0.05 ): Não rejeita-se H0
## Nível de significância:  97.1204 %
## Valor de p: 0.028795
```

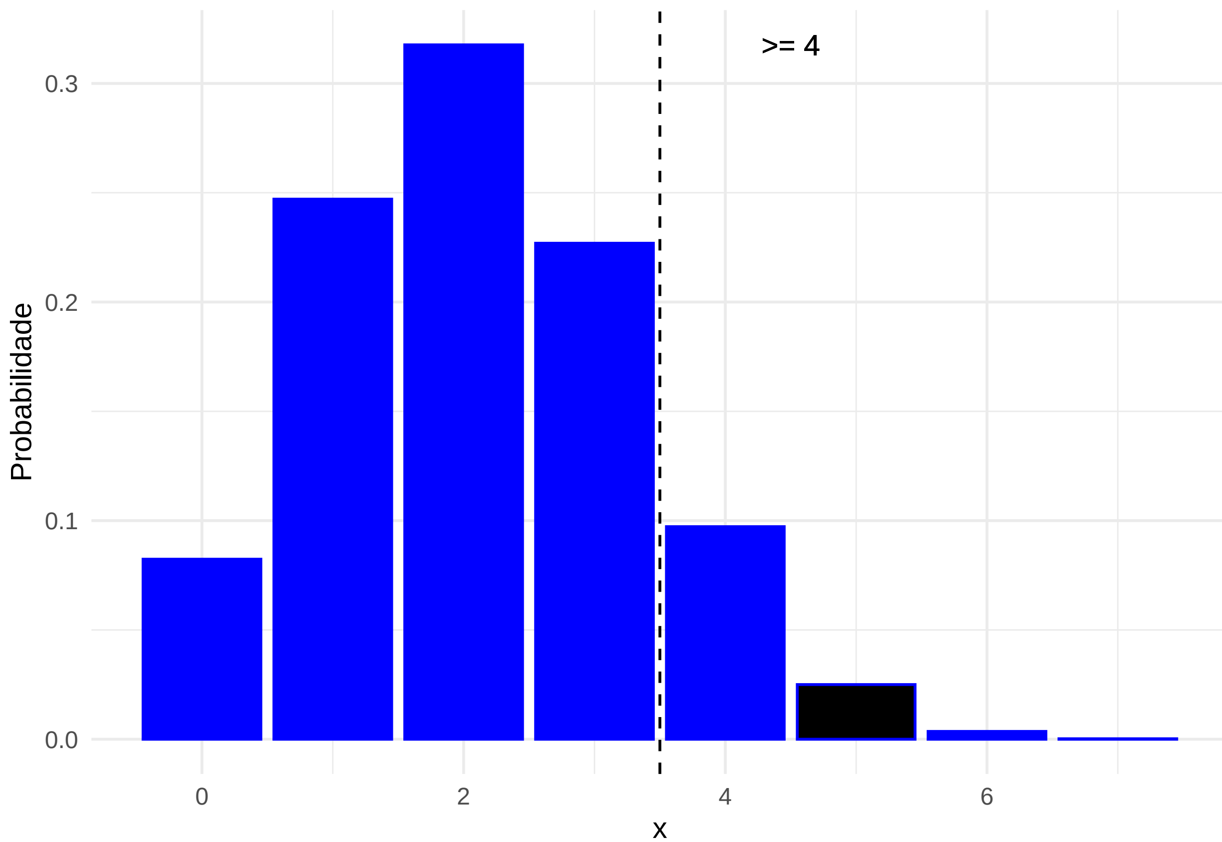


Tabela 4: Binomial (7 , 0.3)

X	Pontual	Acumulada
0	0.082354	0.082354
1	0.247063	0.329417
2	0.317652	0.647070
3	0.226895	0.873964
4	0.097241	0.971205
5	0.025005	0.996209
6	0.003572	0.999781
7	0.000219	1.000000

Como mostrado acima, o valor observado na amostra (5) ocorreu na região de rejeição (X $\geq$ 4), e portanto rejeitamos a hipótese $H_0 : p = 0.30$. Temos assim relevância estatística para assumir $H_1 : p > 0.30$, ou seja, que mais de 30% dos eleitores aprovam a proposição.

5)

Para encontrar os valores do intervalo de confiança para o terceiro quartil, deve-se encontrar os valores r e s, para serem utilizados como as posições da amostra. Como a amostra tem mais de 20 dados, podemos aplicar a aproximação pelo Teorema Central do Limite.

```
## r: 19 ; s: 26
```

Utilizando os valores posicionais de r e s calculados, obtemos que o Intervalo de Confiança 80% para o terceiro quartil é:

$$X^{(19)} = 57 \qquad X^{(26)} = 72 \qquad IC_{80\%}^{Q_3} = [57, 72]$$

6)

Hipótese do pesquisador: A altura das árvores segue uma distribuição uniforme entre 10m e 120m.

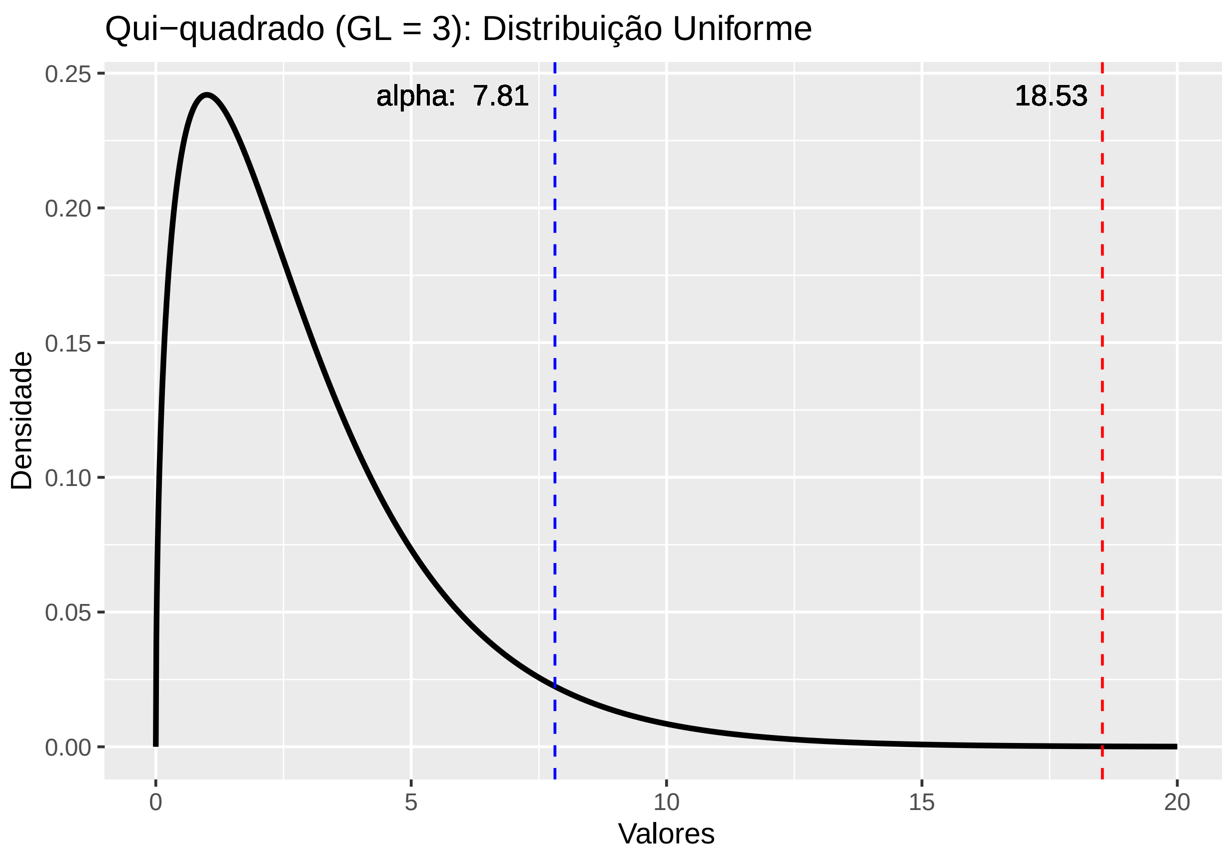
$$H_0 : Y \sim U(10, 120); \qquad H_1 : Y \nsim U(10, 120)$$

A distribuição de referência para teste dessa hipótese será a Qui-Quadrado com GL = número de categorias - 1 = 4 - 1 = 3.

Quartis	(10,37.5]	(37.5,65]	(65,92.5]	(92.5,120]
Observado	7.0	17.0	5.0	1.0
Esperado	7.5	7.5	7.5	7.5

```
## Valor de p: 0.00034138
```

$$\chi_3^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(7-7.5)^2}{7.5} + \frac{(17-7.5)^2}{7.5} + \frac{(5-7.5)^2}{7.5} + \frac{(1-7.5)^2}{7.5} = 18.5333333$$



Conclusão: Como o valor de p (3.4138×10^{-4}) é menor que o $\alpha = 0.05$ estabelecido, rejeitamos a hipótese nula de que a altura das árvores segue uma distribuição uniforme entre 10m e 120m.

7)

Hipótese do pesquisador: A altura das árvores segue uma distribuição normal.

$$H_0 : Y \sim N(\mu, \sigma^2); \qquad H_1 : Y \nsim N(\mu, \sigma^2)$$

Como não foram especificados os valores de μ e σ, teremos de os estimar pela amostra e realizar as seguintes substituições $\mu = \bar{x}$ e $\sigma = s$. Por conta dessa estimativa, a distribuição de referência utilizada será a Qui-Quadrado com GL = número de categorias - número de parâmetros estimados - 1 = 4 - 2 - 1 = 1. Os quartis serão construídos da forma:

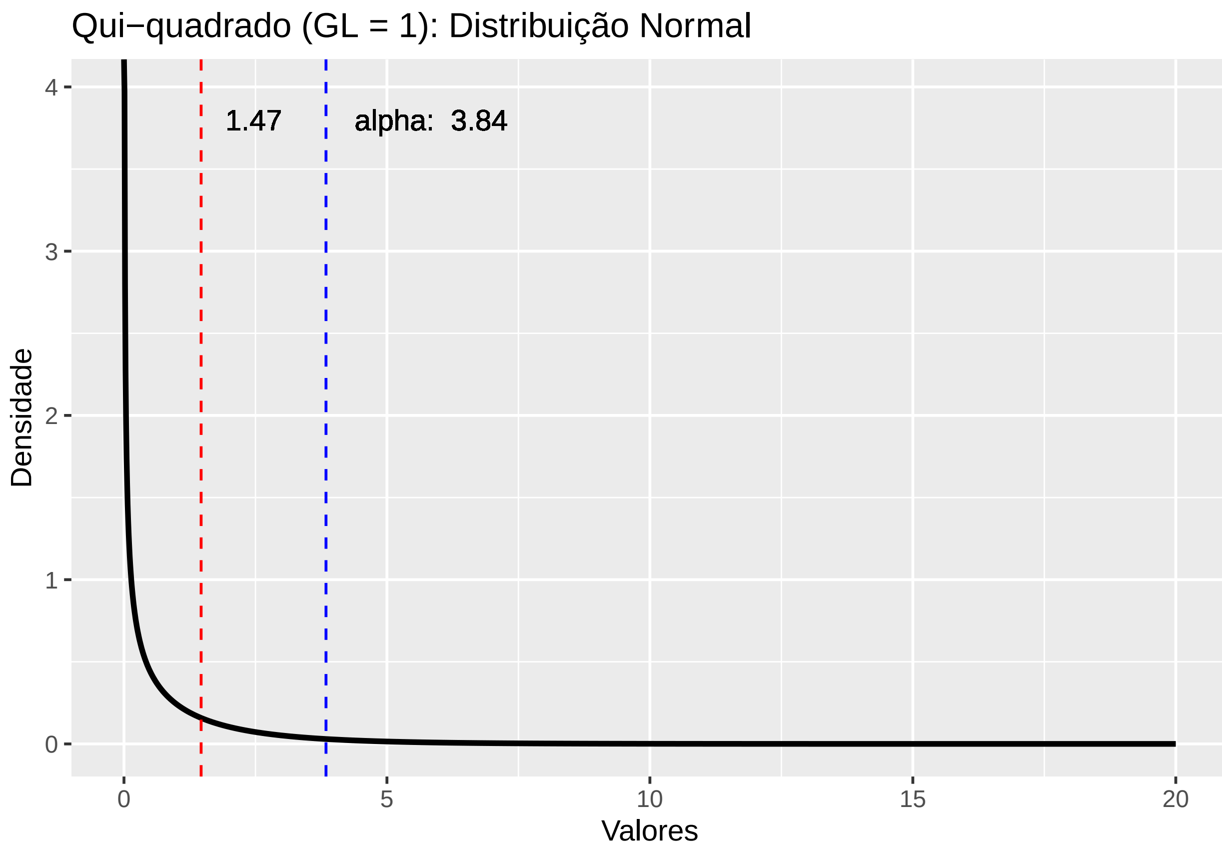
$$Q_1 = \bar{x} + s * Z_{0.25}; \qquad Q_2 = \bar{x}; \qquad Q_3 = \bar{x} + s * Z_{0.75}.$$

```
## [1] 37.66 51.85 66.03
```

Quartis	(-Inf, 37.66]	(37.66 , 51.85]	(51.85 , 66.03]	(66.03 , Inf)
Observado	7.0	9.0	9.0	5.0
Esperado	7.5	7.5	7.5	7.5

```
## Valor de p: 0.68998473
```

$$\chi^2_1 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(7-7.5)^2}{7.5} + \frac{(9-7.5)^2}{7.5} + \frac{(9-7.5)^2}{7.5} + \frac{(5-7.5)^2}{7.5} = 1.4666667$$



Conclusão: Como o valor de p (0.6899847) é maior que o $\alpha = 0.05$ estabelecido, não rejeitamos a hipótese nula de que a altura das árvores segue uma distribuição normal.

8)

Hipótese do pesquisador: A altura das árvores segue uma distribuição exponencial.

$$H_0 : Y \sim Exp(\theta); \qquad H_1 : Y \nsim Exp(\theta)$$

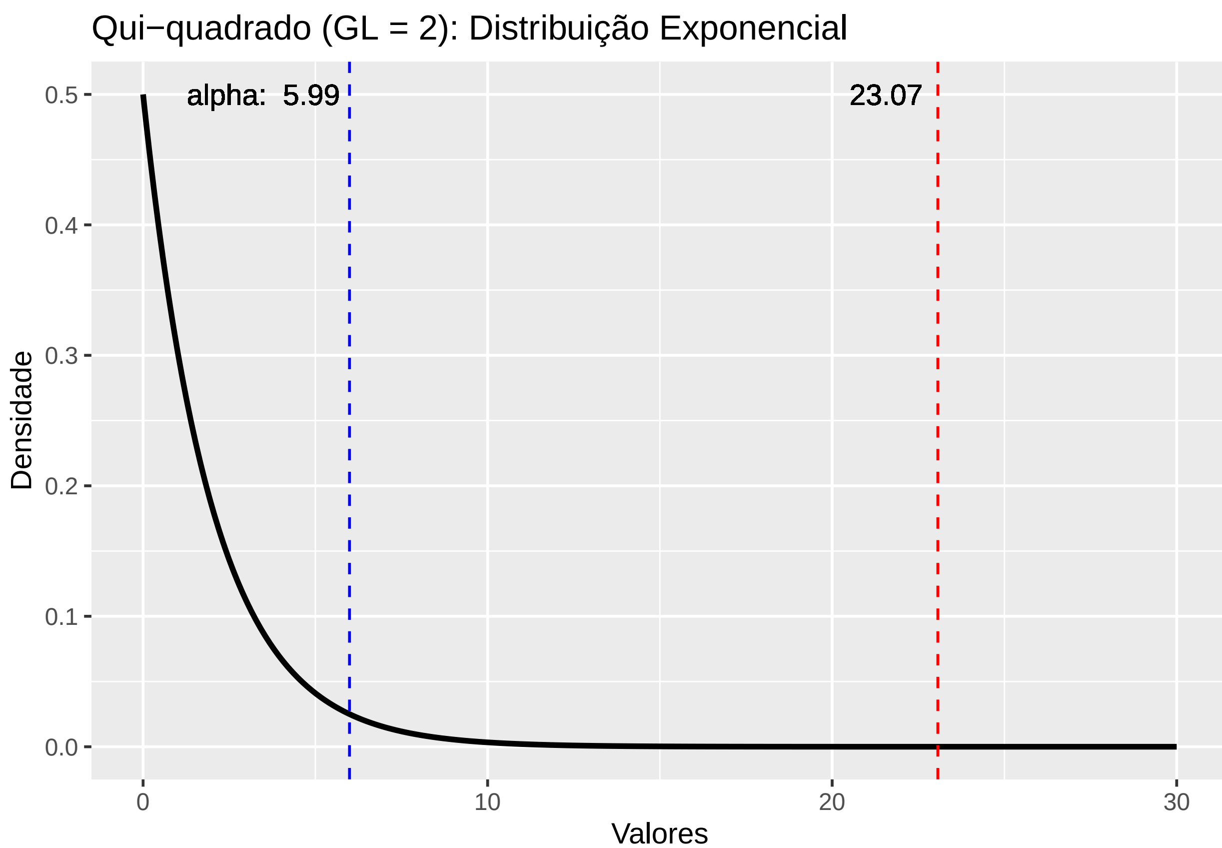
Como não foi especificado o valor de θ, será necessário o estimar pela amostra e realizar as seguinte substituição $\theta = \frac{1}{\bar{x}}$. Por conta dessa estimativa, a distribuição de referência utilizada será a Qui-Quadrado com GL = número de categorias - número de parâmetros estimados - 1 = 4 - 1 - 1 = 2. Os quartis serão construídos da forma:

$$Q_1 = \frac{-\ln(1-0.25)}{\hat{\theta}}; \qquad Q_2 = \frac{-\ln(1-0.5)}{\hat{\theta}}; \qquad Q_3 = \frac{-\ln(1-0.75)}{\hat{\theta}}.$$

Quartis	(0, 14.9]	(14.9 , 35.9]	(35.9 , 71.9]	(71.9 , Inf)
Observado	0.0	7.0	18.0	5.0
Esperado	7.5	7.5	7.5	7.5

```
## Valor de p: 9.8e-06
```

$$\chi_2^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(0-7.5)^2}{7.5} + \frac{(7-7.5)^2}{7.5} + \frac{(18-7.5)^2}{7.5} + \frac{(5-7.5)^2}{7.5} = 23.0666667$$



Conclusão: Como o valor de p (9.8×10^{-6}) é menor que o $\alpha = 0.05$ estabelecido, rejeitamos a hipótese nula de que a altura das árvores segue uma distribuição exponencial.

9)

Hipótese do pesquisador: O tempo de reação segue uma distribuição exponencial.

$$H_0 : Y \sim Exp(\theta); \qquad H_1 : Y \nsim Exp(\theta)$$

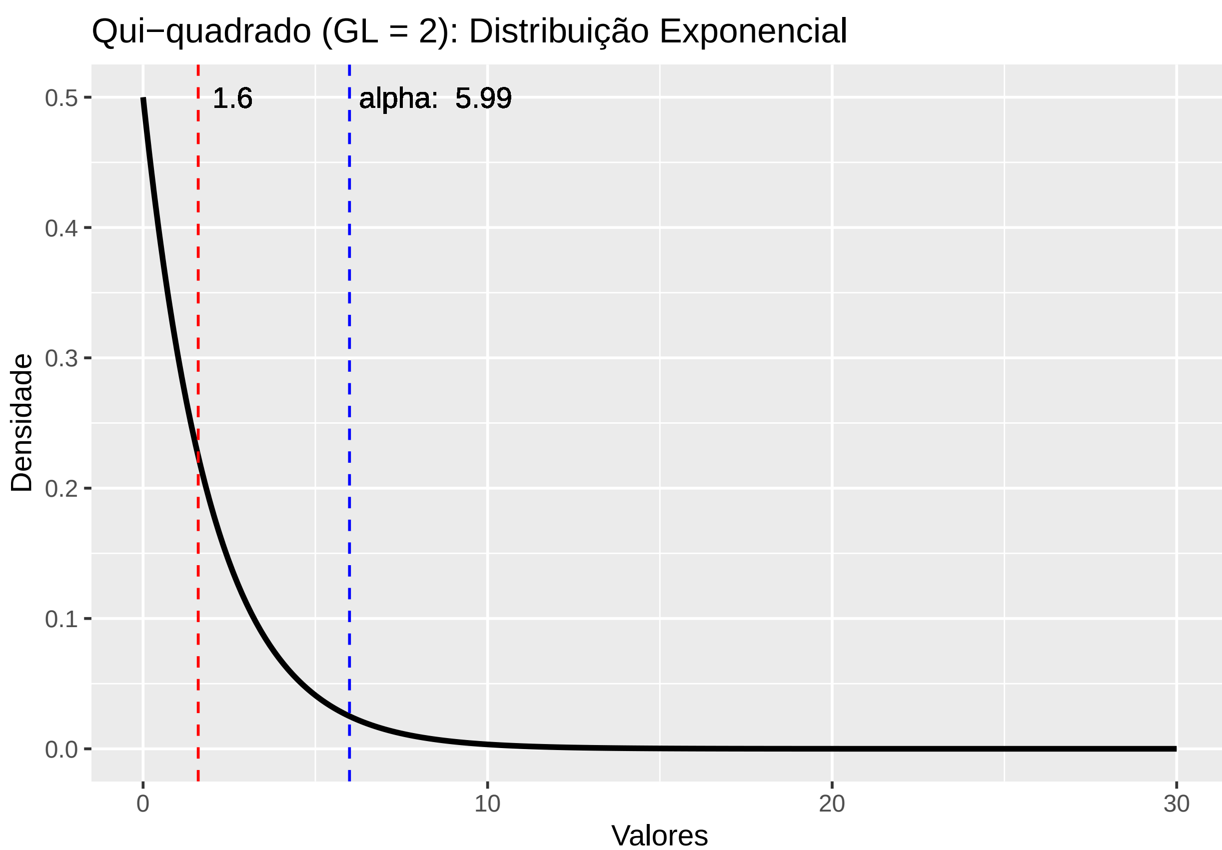
Como não foi especificado o valor de θ, será necessário o estimar pela amostra e realizar as seguinte substituição $\theta = \frac{1}{\bar{x}}$, que é o estimador de máxima verossimilhança para θ. Por conta dessa estimativa, a distribuição de referência utilizada será a Qui-Quadrado com GL = número de categorias - número de parâmetros estimados - 1 = 4 - 1 - 1 = 2. Os quartis serão construídos da forma:

$$Q_1 = \frac{-\ln(1-0.25)}{\hat{\theta}}; \qquad Q_2 = \frac{-\ln(1-0.5)}{\hat{\theta}}; \qquad Q_3 = \frac{-\ln(1-0.75)}{\hat{\theta}}.$$

Quartis	(0, 2.9]	(2.9 , 7]	(7 , 14.1]	(14.1 , Inf)
Observado	5	7	3	5
Esperado	5	5	5	5

```
## Valor de p: 0.44932896
```

$$\chi^2_2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(5-5)^2}{5} + \frac{(7-5)^2}{5} + \frac{(3-5)^2}{5} + \frac{(5-5)^2}{5} = 1.6$$



Conclusão: Como o valor de p (0.449329) é maior que o $\alpha = 0.05$ estabelecido, não rejeitamos a hipótese nula de que o tempo de reação segue uma distribuição exponencial.

10) IC 95% de confiança para a média

a) Levando em consideração a distribuição da variável como sendo exponencial

Para encontrar os valores do intervalo de confiança para a média, devemos lembrar do referencial teórico que nos diz que a soma de variável aleatórias independentes seguindo uma distribuição exponencial em comum segue uma distribuição Gama.

$$X_i \sim Exp(\theta) \quad \rightarrow \quad \sum_{i=1}^{n} X_i \sim G(n, \theta)$$

No caso da amostra acima,

$$G = \sum_{i=1}^{20} X_i = X_1 + \cdots + X_{20} = S \sim G(20, \theta)$$

Sabe-se também a relação direta entre a distribuição Gama e a distribuição Qui-Quadrado: $\chi^2 = 2\theta G \sim \chi^2_{2n}$.

$$\chi^2 = 2\theta S \sim \chi^2_{2n}$$

Da definição da média amostral, podemos encontrar um substituto para a variável S.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \rightarrow \quad n\bar{X} = \sum_{i=1}^{n} X_i \quad \rightarrow \quad n\bar{X} = S$$

Obtendo assim,

$$\chi^2 = 2 * \theta n\bar{x} \sim \chi^2_{2n} \quad \rightarrow \quad 40 * \theta\bar{x} \sim \chi^2_{40}$$

Consultando os quantis de $\frac{\alpha}{2}$ e $1 - \frac{\alpha}{2}$ da distribuição Qui-Quadrado com 40 graus de liberdade, obtemos:

$$P(24.43 < \chi^2_{40} < 59.34) = 1 - \alpha = 95\%$$

Lembremos que $40 * \theta\bar{x} \sim \chi^2_{40}$, portanto:

$$P(24.43 < 40 * \theta\bar{x} < 59.34) = P(\frac{24.43}{40 * \theta} < \bar{x} < \frac{59.34}{40 * \theta}) = 95\%$$

$$IC^{\mu}_{1-\alpha} = (\frac{24.43}{40 * \theta} \; ; \; \frac{59.34}{40 * \theta})$$

```
## [1]  6.19 15.04
```

$$IC^{\mu}_{1-\alpha} = (6.19 \; ; \; 15.04)$$

10) IC 95% de confiança para a média

b) Usando a aproximação pelo Teorema Central do Limite, temos que

$$IC^{\mu}_{1-\alpha} = (\bar{x} + s * Z_{\frac{\alpha}{2}} \; ; \; \bar{x} + s * Z_{1-\frac{\alpha}{2}})$$

```
## [1]  5.82 14.46
```

$$IC^{\mu}_{1-\alpha} = (5.82 \; ; \; 14.46)$$

11)

Para encontrar os valores do intervalo de confiança para a mediana, deve-se encontrar os valores r e s, para serem utilizados como as posições da amostra. Onde r é

```
## r: 6 ; s: 15
```

Utilizando os valores posicionais de r e s calculados, obtemos que o Intervalo de Confiança 95% para o terceiro quartil é:

$$X^{(6)} = 3 \qquad X^{(15)} = 12.3 \qquad IC_{95\%}^{Q_3} = [3, 12.3]$$

O grau de confiança realmente obtido foi: 93.64%

12)

Hipótese do pesquisador: A meditação aumenta o tempo de sono.

$$H_0 : P[+] = P[-] = 0.5 \qquad H_1 : P[+] > P[-]$$

```
## Maior:  16 Menor: 4
```

```
## Hipótese Nula (H0): p =  0.5 ; Hipótese Alternativa (H1) p >  0.5
## Regra de decisão:
## Se p-valor < alpha ( 0.05 ) : Rejeita-se H0
## Se p_valor >= alpha ( 0.05 ): Não rejeita-se H0
## Nível de significância:  99.4091 %
## Valor de p: 0.005909
```

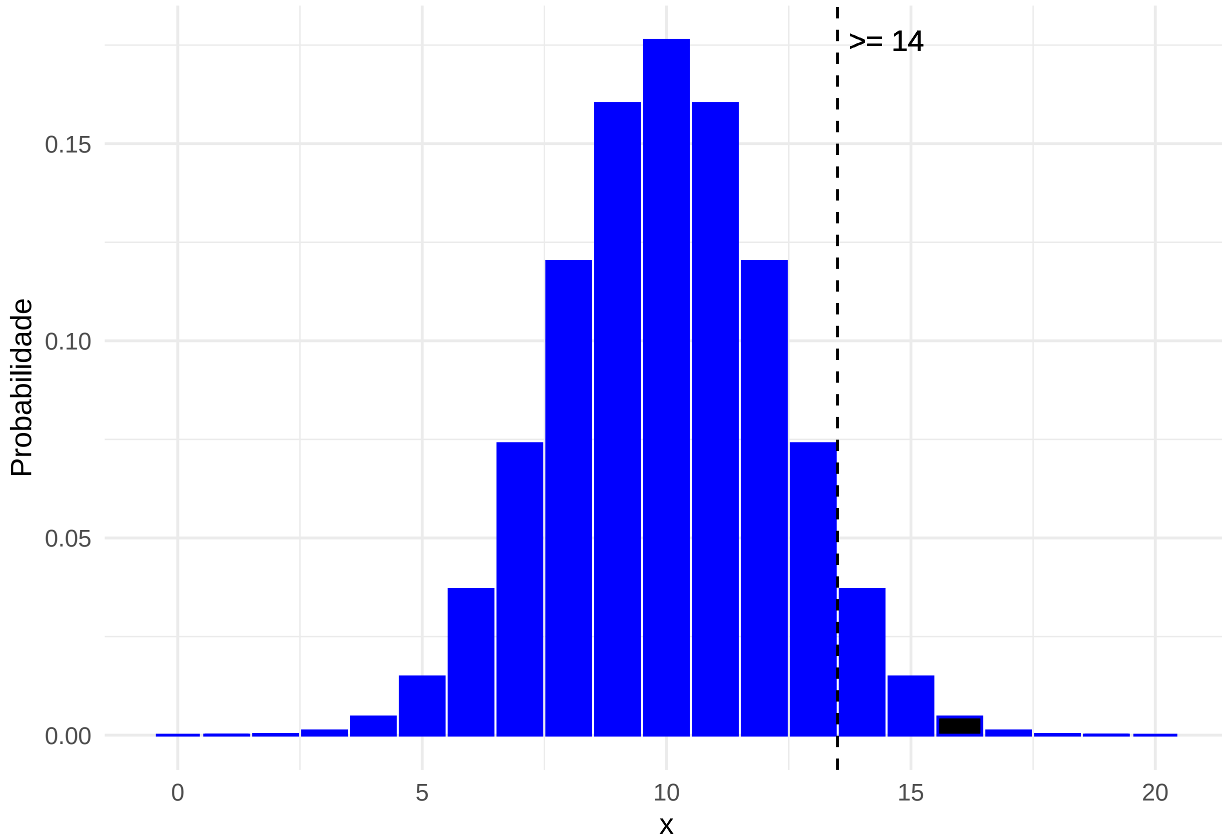


Tabela 9: Binomial (20 , 0.5)

X	Pontual	Acumulada
0	0.000001	0.000001
1	0.000019	0.000020
2	0.000181	0.000201
3	0.001087	0.001288
4	0.004621	0.005909
5	0.014786	0.020695
6	0.036964	0.057659
7	0.073929	0.131588
8	0.120134	0.251722
9	0.160179	0.411901
10	0.176197	0.588099
11	0.160179	0.748278
12	0.120134	0.868412
13	0.073929	0.942341
14	0.036964	0.979305
15	0.014786	0.994091
16	0.004621	0.998712
17	0.001087	0.999799
18	0.000181	0.999980
19	0.000019	0.999999
20	0.000001	1.000000

Como mostrado na tabela acima, o valor observado na amostra (16) ocorreu na região de rejeição (X $\geq$ 14), e portanto rejeitamos a hipótese $H_0 : p = 0.5$. Temos assim relevância estatística para acreditar que a amostra observada realmente teve um aumento do tempo de sono ocasionado pela prática da meditação.

13)

Hipótese do pesquisador: A marca A é melhor do que a marca B.

$$H_0 : P[A] = P[B] = 0.5 \qquad H_1 : P[A] \neq P[B] \neq 0.5$$

```
## Hipótese Nula (H0): p =  0.5 ; Hipótese Alternativa (H1) p =!=  0.5
## Regra de decisão:
## Se p-valor < alpha ( 0.05 ) : Rejeita-se H0
## Se p_valor >= alpha ( 0.05 ): Não rejeita-se H0
## Nível de significância:  92.4481 %
## Valor de p: 0.075519
```

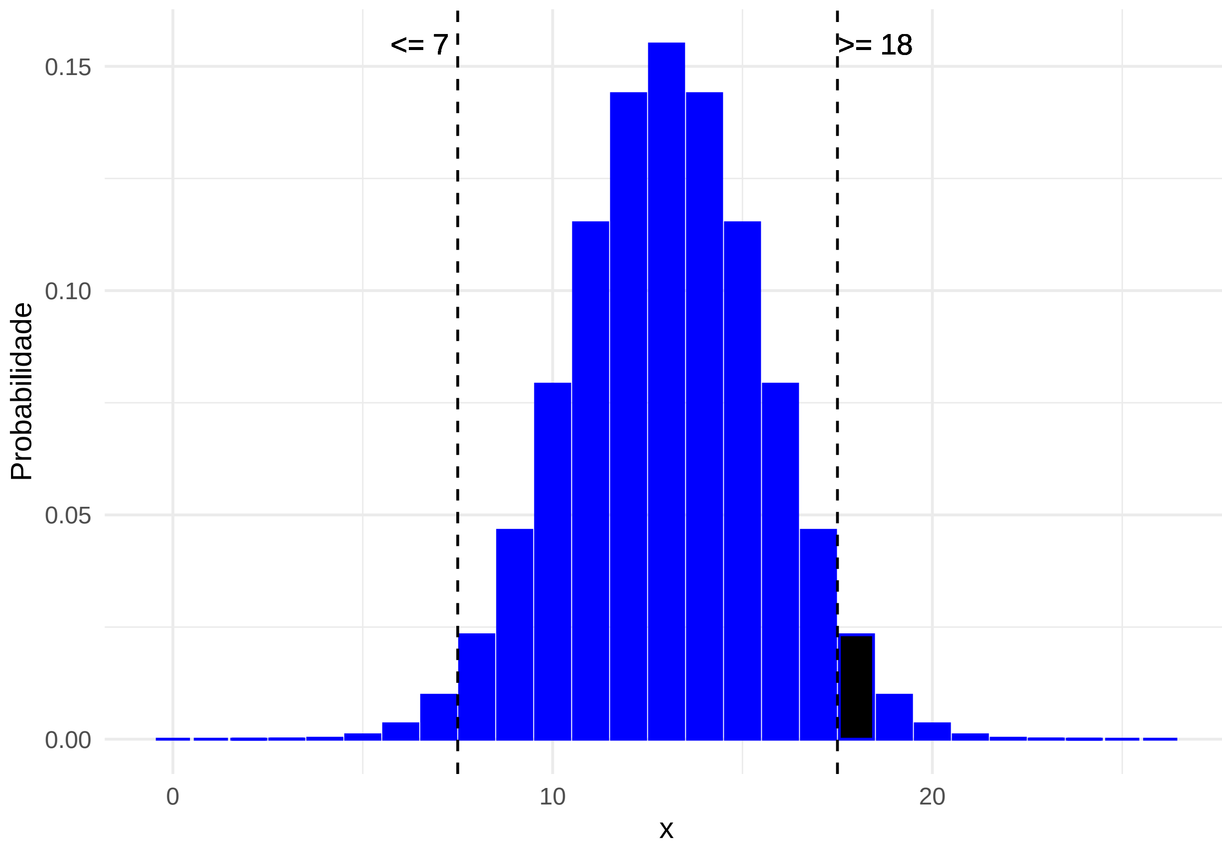


Tabela 10: Binomial (26 , 0.5)

X	Pontual	Acumulada
0	0.000000	0.000000
1	0.000000	0.000000
2	0.000005	0.000005
3	0.000039	0.000044

X	Pontual	Acumulada
4	0.000223	0.000267
5	0.000980	0.001247
6	0.003431	0.004678
7	0.009802	0.014480
8	0.023280	0.037759
9	0.046559	0.084319
10	0.079151	0.163470
11	0.115129	0.278599
12	0.143911	0.422509
13	0.154981	0.577491
14	0.143911	0.721401
15	0.115129	0.836530
16	0.079151	0.915681
17	0.046559	0.962241
18	0.023280	0.985520
19	0.009802	0.995322
20	0.003431	0.998753
21	0.000980	0.999733
22	0.000223	0.999956
23	0.000039	0.999995
24	0.000005	1.000000
25	0.000000	1.000000
26	0.000000	1.000000

Como mostrado na tabela acima, o valor observado na amostra (18) ocorreu na região de rejeição (X $\leq$ 7 ou X $\geq$ 18), e portanto rejeitamos a hipótese $H_0 : P[A] = P[B]$, em favor de $H_1 : P[A] \neq P[B] \neq 0.5$\$. Temos assim relevância estatística para acreditar que há uma diferença na escolha entre as marcas A e B.

14)

a) 1ª Abordagem: Para esse caso, podemos utilizar o teste não paramétrico do sinal. Realizam-se n medições pareadas independentes entre μ e μ^* e a atribuímos a Y a contagem de quantas vezes $\mu > \mu^*$. Ao final, consultamos numa tabela da distribuição Binominal com paramêtros n (número de pares medidos) e p = 0.5. Como o teste é unilateral a direita, escolhemos um Lim para x , o qual contenha $\alpha\%$ de massa de probabilidade. Se $Y > Lim$, rejeita-se $H_0 : \mu = \mu^*$ em favor de $H_1 : \mu > \mu^*$.

b) 2ª Abordagem: Como desenvolvido para a distribuição exponencial na questão 10)a), sabemos que $\chi^2 = 2\mu n\bar{x} \sim \chi^2_{2n}$. Desenvolvendo:

$$P(\chi^2_{\frac{\alpha}{2}}(2n) < 2\mu n\bar{x} < \chi^2_{1-\frac{\alpha}{2}}(2n)) = 1 - \alpha$$

$$P(\frac{\chi^2_{\frac{\alpha}{2}}(2n)}{2n\bar{x}} < \mu < \frac{\chi^2_{1-\frac{\alpha}{2}}(2n)}{2n\bar{x}}) = 1 - \alpha$$

Como o teste de hipótese é unilateral a direita:

$$P(\mu < \frac{\chi^2_{1-\alpha}(2n)}{2n\bar{x}}) = 1 - \alpha$$

$$P(\mu < \frac{\chi^2_{1-\alpha}(2n)}{2n\bar{x}}) = 95\% \Longleftrightarrow P(\mu > \frac{\chi^2_{1-\alpha}(2n)}{2n\bar{x}}) = 5\%$$

Podemos generalizar ainda mais ao lembrar da relação entre a Qui-Quadrado e a distribuição F de Fisher-Snedecor: $\frac{\chi^2(n)}{n} = F(n, \infty)$.

$$P(\mu\bar{x} < F(n, \infty)) = 95\% \Longleftrightarrow P(\mu\bar{x} > F(n, \infty)) = 5\%$$

Podemos definir as hipóteses e a regra de decisão:

$H_0 : \mu = \mu^*$ $\qquad$ $H_1 : \mu > \mu^*$.

Se p-valor $< \alpha = 0.05$: Rejeita-se H_0.

Se p-valor $\geq \alpha = 0.05$: Não rejeita-se H_0.