17. Em uma amostra de 500 famílias da cidade A, constatou-se que 298 haviam comprado, durante os últimos 30 dias, o refrigerante Meca-Mela em sua nova versão incolor. Na cidade B esse número foi de 147 em 300 famílias entrevistadas. Na cidade A foi feita uma campanha publicitária através da rádio local, e não na cidade B. Os resultados trazem evidências de que as campanhas locais aumentam as vendas?

Solução:

$$X = \begin{cases} 1 & \text{se a família da cIdade A compra a nova versão ;} \\ 0 & \text{caso contrário .} \end{cases}$$

.

$$p_A = P(X = 1).$$

$$X \sim Ber(p_A).$$

$$Y = \begin{cases} 1 & \text{se a família da cIdade B compra a nova versão ;} \\ 0 & \text{caso contrário .} \end{cases}$$

.

$$p_B = P(Y = 1).$$

$$Y \sim Ber(p_B),$$

X e Y independentes.

Queremos testar:

$$H_0 : p_A - p_B = 0 \text{ contra } H_1 : p_A - p_B > 0.$$

Foi tirada uma amostra de tamanho $n = 500$ de X e observado $S_A = 298$ sucessos e uma amostra de tamanho $m = 300$ de X e observado $S_B = 147$ sucessos. Vamos construir uma intervalo de confiança para $p_A - p_B$ com 95% de confiança.

As estimativas pontuais de p_A e p_B são dadas por:

$$\hat{p}_A = \frac{S_A}{n} = \frac{298}{500} = 0,596.$$

$$\hat{p}_B = \frac{S_B}{m} = \frac{147}{300} = 0,49.$$

Logo

$$\hat{p}_A - \hat{p}_b = 0,576 - 0,49 = 0,106.$$

$$Var\,(\hat{p}_A - \hat{p}_B) = \frac{\hat{p}_A(1 - \hat{p}_A)}{n} + \frac{\hat{p}_B(1 - \hat{p}_B)}{m}.$$

$$Var\,(\hat{p}_A - \hat{p}_B) = \frac{0,596 \times 0,404}{500} + \frac{0,49 \times 0,51}{300} = 0,00131457.$$

O erro padrão desse estimador é dado por:

$$epm = 0,0363$$

O erro amostral é dado por:

$$e = 1,96 \times 0,0363 = 0,07106.$$

Nosso intervalo de confiança é dado por:

$$0,11 \pm 0,07.$$

$$IC\left[p_A - p_B \ \ ; \ \ 95\%\right] = [0,04 \ 0,18].$$

Note que o ponto 0 não pertence ao IC.

Vamos explicar a saída do **R**:

$$H_0 : p_A - p_B = 0 \ \text{ contra } \ H_1 : p_A - p_B > 0.$$

```
>
> n=500;m=300
> SA=298;SB=147
> p_est_A=SA/n;p_est_A
[1] 0.596
> p_est_B=SB/m;p_est_B
[1] 0.49
>
> aux1=p_est_A*(1- p_est_A)/n;aux1
[1] 0.000481568
> aux2=p_est_B*(1- p_est_B)/m;aux2
[1] 0.000833
>
> aux1+aux2
[1] 0.001314568
>
> epm=sqrt(aux1+aux2);epm
[1] 0.03625697
>
> e=1.96*epm;e
[1] 0.07106366
>
> delta_est=p_est_A-p_est_B;delta_est
[1] 0.106
>
> IC95=delta_est+c(-1,1)*e;IC95
[1] 0.03493634 0.17706366
> round(IC95,2)
[1] 0.03 0.18
>
>
```

```
> prop.test(c(298,147),c(500,300))

2-sample test for equality of proportions with continuity correction

data:  c(298, 147) out of c(500, 300)
X-squared = 8.111, df = 1, p-value = 0.0044
alternative hypothesis: two.sided
95 percent confidence interval:
0.03227097 0.17972903
sample estimates:
prop 1 prop 2
0.596  0.490

>
> prop.test(c(298,147),c(500,300),alternative="greater")

2-sample test for equality of proportions with continuity correction

data:  c(298, 147) out of c(500, 300)
X-squared = 8.111, df = 1, p-value = 0.0022
alternative hypothesis: greater
95 percent confidence interval:
0.04369592 1.00000000
sample estimates:
prop 1 prop 2
0.596  0.490
```

Vamos explicar a saída do **R**:

Para testar:

$$H_0 : p_1 - p_2 = 0 \text{ contra } H_1 : p_1 - p_2 > 0.$$

Vamos utilizar a estatística:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} \sim N(0,1)$$

Se H_0 é verdade temos:

$$H_0 : p_1 = p_2 = p$$

Vamos estimar p por

$$\hat{p}_c = \frac{S_1 + S_2}{n + m}.$$

Logo

$$\hat{p}_c = \frac{298 + 147}{500 + 300} = 0,55625.$$

A estatística fica:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}_c(1 - \hat{p}_c)\left(\frac{1}{n} + \frac{1}{m}\right)}} \sim N(0,1)$$

Este é o valor observado>

$$z_{cal} = \frac{0,596 - 0.490}{0,013} = 2,921476.$$

Este valor $2,9215$ é um valor significativo?

Vamos calcular a probabilidade de ocorres na amostra um valor igual ou mais desfavorável a hipótese nula do que $z_{cal} = 2,9215$.

Esta probabilidade é conhecida como nível mínimo de significância.

$$nd = P(Z \geq z_{cal}) = P(Z \geq 2,92) = 0,001742.$$

Se a hipótese alternativa for $H_1 : p_1 - P_2 a < 0$:

o nível descritivo é:

$$nd = P(Z \leq z_{cal}).$$

Se a hipótese alternativa for $H_1 : p_1 - P_2 a \neq 0$:

o nível descritivo é:

$$nd = P(|Z| \geq |z_{cal}|).$$

Vamos estudar as saídas:

```
    2-sample test for equality of proportions without continuity correction

data:  c(298, 147) out of c(500, 300)
X-squared = 8.535, df = 1, p-value = 0.001742
alternative hypothesis: greater
95 percent confidence interval:
0.04636259 1.00000000
sample estimates:
prop 1 prop 2
0.596  0.490
```

Para obter o valor do qui-quadrado basta elevar ao quadrado:

$$z_2 = 2,921476^2 = 8,535021.$$

Acontece geralmente o oposto. Temos o valor do qui-quadrado e queremos o valor de z. Basta extrair a raiz quadrada e colocando o sinal da diferença $(\hat{p}_1 - \hat{p}_2)$.

```
n=500;m=300
> SA=298;SB=147
> pc_est=(SA+SB)/(n+m);pc_est
[1] 0.55625
> aux=pc_est*(1-pc_est)*(1/n+1/m)
> aux
[1] 0.001316458
> sqrt(aux)
[1] 0.03628303
> z=(p_est_A-p_est_B)/sqrt(aux);z
[1] 2.921476
> z^2
[1] 8.535021
>
```

```
      2-sample test for equality of proportions without continuity correction

data:  c(298, 147) out of c(500, 300)
X-squared = 8.535, df = 1, p-value = 0.001742
alternative hypothesis: greater
95 percent confidence interval:
0.04636259 1.00000000
sample estimates:
prop 1 prop 2
0.596  0.490
```

```
sign(-7)
sign(7)
sign(0)
```

```
z_1=sign(p_est_A-p_est_B)*sqrt(z^2);z_1
```

```
prop.test(c(147,298),c(300,500),alternative="greater",correct=F)
```

```
z=(p_est_B-p_est_A)/sqrt(aux);z

z^2

z_2=sign(p_est_B-p_est_A)*sqrt(z^2);z_2
```

Ficaram evidenciadas o aumento das vendas deviso a propaganda.