Autor: Alberto Rodrigues Ferreira

1 Teste de Hipóteses para Diferença de Proporções e Estimadores

$$H_0 = p_1 - p_2 = 0$$
$$H_1 = p_1 - p_2 < 0$$

Considerando $X_{11}, X_{12}, \ldots, X_{1n} \sim Bernoulli(p_1)$ independentes e $X_{21}, X_{22}, \ldots, X_{2n} \sim Bernoulli(p_2)$ independentes.
Considerando os dois estimadores para os parâmetros p_1 e p_2:

$$\hat{p}_1 = \bar{X}_1$$
$$\hat{p}_2 = \bar{X}_2$$

Pelo Teorema Central do Limite, temos que:

$$\hat{p}_1 \sim A.N\left(p_1, \frac{p_1(1-p_1)}{n}\right)$$

e

$$\hat{p}_2 \sim A.N\left(p_2, \frac{p_2(1-p_2)}{n}\right)$$

Portanto, temos que:

$$\hat{p}_1 - \hat{p}_2 \sim A.N\left(p_1 - p_2, \frac{p_1(1-p_1) + p_2(1-p_2)}{n}\right)$$

As hipóteses de interesse podem ser: $H_0 = p_1 - p_2 = 0$ vs $H_1 = p_1 > p_2$ ou $H_1 = p_1 < p_2$ ou $H_1 = p_1 \neq p_2$
A estatística de teste para todas elas é:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}}$$

onde $\bar{p} = (p_1 + p_2)/2$.
Temos 3 possibilidades:

1. Para $H_1 = p_1 > p_2$, rejeitamos H_0 se $Z > z(1-\alpha)$
2. Para $H_1 = p_1 < p_2$, rejeitamos H_0 se $Z < z(\alpha)$
3. Para $H_1 = p_1 \neq p_2$, rejeitamos H_0 se $|Z| > z(1-\alpha/2)$

Aqui, $z(a)$ denota o quantil a da distribuição Normal Padrão, ou seja, o valor de x onde $\mathbb{P}(Z < x) = a$. Onde Z têm distribuição Normal Padrão sob H_0.

2 Determinação do tamanho amostral

Vamos supor que estamos na situação onde $H_1 = p_1 > p_2$. O erro do tipo 2 é não rejeitar H_0 quando H_0 é falso. Essa probabilidade denotada por β é dada por:

$$\mathbb{P}\left[Z < z(1-\alpha)\right] = \beta$$

Agora precisamos calcular esperança e variância de Z para podermos encontrar o tamanho amostral em função de β, α e proporções.

$$\mathbb{E}\left[Z\right] = \mathbb{E}\left[\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}}\right] = \frac{1}{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}}\mathbb{E}\left[\hat{p}_1 - \hat{p}_2\right] = \frac{p_1 - p_2}{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}}$$

$$Var[Z] = Var\left[\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}}\right] = \frac{1}{\frac{2\bar{p}(1-\bar{p})}{n}}Var[\hat{p}_1 - \hat{p}_2] = \frac{\frac{p_1(1-p_1)+p_2(1-p_2)}{n}}{\frac{2\bar{p}(1-\bar{p})}{n}}$$

Portanto, temos que:

$$\begin{aligned}
\mathbb{P}\left[Z < z(1-\alpha)\right] &= \beta \\
\mathbb{P}\left[\frac{Z - \mathbb{E}[Z]}{\sqrt{Var[Z]}} < \frac{z(1-\alpha) - \mathbb{E}[Z]}{\sqrt{Var[Z]}}\right] &= \beta \\
\Phi\left[\frac{z(1-\alpha) - \mathbb{E}[Z]}{\sqrt{Var[Z]}}\right] &= \beta \\
\Phi\left[\frac{z(1-\alpha) - \frac{p_1-p_2}{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}}}{\sqrt{\frac{\frac{p_1(1-p_1)+p_2(1-p_2)}{n}}{\frac{2\bar{p}(1-\bar{p})}{n}}}}\right] &= \beta \\
\Phi\left[\frac{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}z(1-\alpha) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)+p_2(1-p_2)}{n}}}\right] &= \beta \\
\frac{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}z(1-\alpha) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)+p_2(1-p_2)}{n}}} &= \Phi^{-1}(\beta) = z(\beta) \\
\sqrt{\frac{2\bar{p}(1-\bar{p})}{p_1(1-p_1)+p_2(1-p_2)}}z(1-\alpha) - z(\beta) &= \sqrt{\frac{n}{p_1(1-p_1)+p_2(1-p_2)}}(p_1 - p_2) \\
\sqrt{n} &= \frac{\sqrt{\frac{2\bar{p}(1-\bar{p})}{p_1(1-p_1)+p_2(1-p_2)}}z(1-\alpha) - z(\beta)}{\frac{p_1-p_2}{\sqrt{p_1(1-p_1)+p_2(1-p_2)}}} \\
\sqrt{n} &= \frac{z(1-\alpha)\sqrt{2\bar{p}(1-\bar{p})} - z(\beta)\sqrt{p_1(1-p_1)+p_2(1-p_2)}}{p_1 - p_2} \\
n &= \left[\frac{z(1-\alpha)\sqrt{2\bar{p}(1-\bar{p})} - z(\beta)\sqrt{p_1(1-p_1)+p_2(1-p_2)}}{p_1 - p_2}\right]^2
\end{aligned}$$

Finalmente, substituindo p_1 e p_2 por suas estimativas temos que:

$$n = \left[\frac{z(1-\alpha)\sqrt{2\bar{p}(1-\bar{p})} - z(\beta)\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}}{\hat{p}_1 - \hat{p}_2} \right]^2$$

onde $\bar{p} = (\hat{p}_1 + \hat{p}_2)/2$.
O autor do livro referenciado diz que para $H_1 : p_1 < p_2$, a fórmula é a mesma e para $H_1 : p_1 \neq p_2$, basta trocar $z(1-\alpha)$ por $z(1-\alpha/2)$. Logo, teríamos:

$$n = \left[\frac{z(1-\alpha/2)\sqrt{2\bar{p}(1-\bar{p})} - z(\beta)\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}}{\hat{p}_1 - \hat{p}_2} \right]^2$$