Sec. 8.4 Ex. 11

Let $Y_1 \sim b(n_1, p_1)$ and $Y_2 \sim b(n_2, p_2)$ with $n_1 = n_2 = 1000$

We want to test $H_0 : p_1 = p_2$ vs $H_1 : p_1 \neq p_2$

Let $\hat{p}_1 = Y_1/n_1$ and $\hat{p}_2 = Y_2/n_2$ and note that since $n_1 = n_2 = 1000$ are both large, the Central Limit Theorem guarantees that:

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \quad \text{is approximately } N(0,1)$$

Under $H_0 : p_1 = p_2$, we may use the following pooled estimate for p_1 and p_2: $\hat{p} = \dfrac{y_1 + y_2}{n_1 + n_2}$

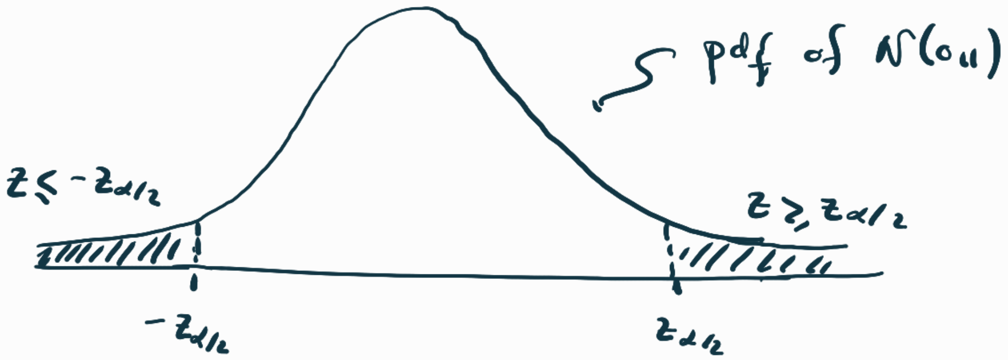
so that the test statistic is:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{approximately } N(0,1)$$

recall that under H_0 $p_1 - p_2 = 0$

We reject H_0 at significance level α if the observed value z of the test statistic Z is such that

$$|z| \geq z_{\alpha/2}, \text{ that is, } z \geq z_{\alpha/2} \text{ or } z \leq -z_{\alpha/2}$$



Critical region corresponds to shaded intervals

b) Since we are given that $y_1 = 37$ and $y_2 = 53$, we have that

$$\hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{37 + 53}{1000 + 1000} = \frac{90}{2000}$$

and

$$z = \frac{\frac{37}{1000} - \frac{53}{1000}}{\sqrt{\frac{90}{2000} \cdot \frac{1910}{2000} \cdot \left(\frac{1}{1000} + \frac{1}{1000}\right)}} \approx -1.73$$

At significance level $\alpha = 0.05$, we have

$$z_{\alpha/2} = z_{0.025} = 1.96$$

and so $|z| = 1.73 < 1.96 = z_{0.025}$

Thus, we do not reject H_0.