EXERCISE

TRY TO PREDICT THE WINNER OF NEXT ELECTIONS

$\hat{\pi} = 0.42$ ESTIMATE % OF PEOPLE THAT SUPPORT THE REPUBBLICANS

$n = 1025$ $\begin{cases} \text{- SAMPLE IS RAPPRESENTATIVE} \\ \text{- HONEST ANSWERS} \end{cases}$

$1-\alpha = 0.95 \Rightarrow z_{(1-\alpha)/2} = 1.96 = z_{0.975}$ (QUANTILE)

$$P\left(\hat{\pi} - z_{(1-\alpha)/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} < \pi < \hat{\pi} + z_{(1-\alpha)/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right) = 1-\alpha$$

$$P\left(|\hat{\pi} - \pi| < z_{(1-\alpha)/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right) = 1-\alpha$$

MARGINAL ERROR $= 1.96 \cdot \sqrt{\frac{0.42(1-0.42)}{1025}} \simeq 0.03$ $[0.42 \pm 0.03]$

LET'S SAY WE FIX THE MARGINAL ERROR:

$$ME = 0.01 \geqslant 1.96 \sqrt{\frac{\pi(1-\pi)}{n}}$$

$$\sqrt{n} \geqslant \frac{1.96\sqrt{\pi(1-\pi)}}{0.01}$$

$$n \geqslant \left(\frac{1.96\sqrt{\pi(1-\pi)}}{0.01}\right)^2$$

R's FILE FOR MORE ANALYSIS

LET'S SAY WE OBTAIN 96.2, WE HAVE TO ROUND IT UP TO 97

NOW, WE HAVE TO FIND A GOOD VALUE FOR π:

WE CHOOSE 0.5 BECAUSE THE NUMBER OF OUTCOME IS 2, BUT THERE ARE OTHERS POSSIBILITY.

EXERCISE ESTIMATE DIFFERCE BETWEEN 2 SAMPLE MEANS ASSUMING THA THEY COME FROM T-STUDENT

n SUBJECT $<\begin{matrix} x_{11}, \dots, x_{1n_1} \to \text{PLACEBO} \\ x_{21}, \dots, x_{2n_2} \to \text{DRUG} \end{matrix} \Rightarrow \begin{matrix} 1° \text{ GROUP} \\ 2° \text{ GROUP} \end{matrix}$

1° SAMPLE $\bar{X}_1 \overset{iid}{\sim} N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$; 2° SAMPLE $\bar{X}_2 \overset{iid}{\sim} N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$

W WANT TO ESTIMATE $\mu_1 - \mu_2$

THIS TWO FOLLOW A NORMAL DISTRIBUTON AND SINCE WE SUPPOSE THEM i.i.d :

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

BUT IF WE DON'T KNOW THE TRUE VARIANCE WE HAVE TO COMPUTE THE SAMPLE VARIANCE, THAT IN THIS CASE:

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left((n_1 - 1)S_1^2 + (n_2 - 1)S_2^2\right)$$

WE ASSUME THAT THE TWO GROUP HAS THE SAME VARIANCE

$\sigma_1^2 = \sigma_2^2 = \sigma$

- IF n SMALL, CHECK THIS ASSUMPTION
- IF n LARGE WE USUALLY USE THE ASINTOTIC RESAULT

THE PIVOTAL QUANTITIE THA WE USE:

$$T(\mu_1-\mu_2) = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{2}{n_2}}} \sim t_{n_1+n_2-2}$$

THE CONFIDENCE INTERVALL WE OBTAIN CAN BE COMPUTED AS:

$$\left[\underbrace{\bar{x}_1 - \bar{x}_2}_{\text{DIFFERENCE OF SIMPLE MEAN}} \pm \underbrace{q_{t_{n_1+n_2-2},\,1-\alpha/2}}_{\text{QUANTILE OF STUDENT DISTRIBUTION}} \cdot S_p \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right] = \dots = 8.73 \pm \overbrace{1.375 \cdot 2.74}^{ME}$$

$$= [3.373, 14.2]$$

EXERCISE ON CONFIDENCE INTERVAL FOR A DIFFERENCE MEANS

$\hat{\pi}_1 = 0.00046$

$\hat{\pi}_2 = 0.00057$

WE HAVE TO FIND A CONFIDENCE INTERVAL OF $\delta = \pi_1 - \pi_2$

$\Rightarrow \hat{\delta} = \hat{\pi}_1 - \hat{\pi}_2$

WE NOW CHOOSE A CONFIDENCE INTERVALL

$SE(\hat{\delta}) = ?$

BUT YOU NEED n : $n_1 = 200745$

$n_2 = 201283$

AND NOW, I JUST USE THE i.i.d :

$$\hat{\pi}_1 - \hat{\pi}_2 \sim N\left(\pi_1 - \pi_2, \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}\right)$$

IR FOR THE PROCEDURE