SMT 2 Solutions

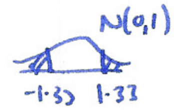
Q1 sample size $n = 200$ $\bar{x} = 1.5, \sigma = 3$ $\bar{x} \sim N(\mu, 3/\sqrt{200})$

a) $C = 0.95$ $z^* = 1.96$ $1.5 \pm 1.96 \times 3/\sqrt{200}$ $(1.08, 1.92)$

b) $C = 0.88$ $z^* = 1.55$ $1.5 \pm 1.55 \times 3/\sqrt{200}$ $(1.17, 1.82)$

Q2 sample size $n = 64$ a) $H_0: \mu = 0$, $H_a: \mu \neq 0$

b) $\sigma = 12$ $\bar{x} \sim N(\mu, \underbrace{12/\sqrt{64}}_{=1.5})$ ~~p-value = P(|z| ≥~~ test statistic $z = 2/1.5 = 1.33$

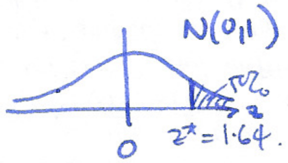


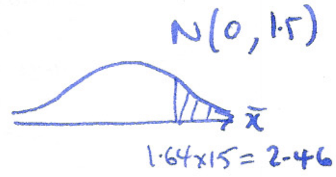
p-value $= \mathbb{P}(|z| \geq 1.33) = 2\,\text{pnorm}(-1.33) = 0.184 \approx 18\%$.

c) Fail to reject null hypothesis at 95% significance level.

Q3 a) $H_0: \mu = 0$, $H_a: \mu > 0$

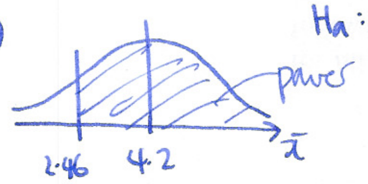
b) sample size $n = 400$, $\sigma = 30$ $\bar{x} \sim N(\overset{0}{\mu}, \overset{1.5}{30/\sqrt{400}})$ (under H_0 true)





reject null hypothesis if $\bar{x} > 2.46$

c) $H_a: \mu = 4.2$, then $\bar{x} \sim N(4.2, 1.5)$



power $= 1 - \text{pnorm}\left(\frac{2.46 - 4.2}{1.5}\right) = 0.877 \approx 88\%$.

Q4 a) assume null hypothesis true, reject null hypothesis.

b) type 1 errors less common, type 2 errors more common.

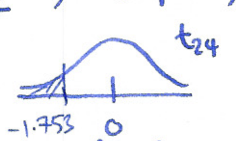
c) s.d. of the sampling dist of the sample mean, pop. std. dev.

Q5 a) $\mu = 74{,}341$ b) $\bar{x} = 69{,}421$ c) 2nd statement. d) smaller e) smaller

Q6 a) larger b) $2(1 - \text{pnorm}(1.82))$ ~~≈ 7%~~. $= 0.0688$ c) $1 - \text{pnorm}(1.82) = 0.0344$ d) $\text{pnorm}(1.82) = 0.966$.

Q7 a) $H_0: \mu = 8$, $H_a: \mu < 8$ b) t-statistic is $\frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{7.73 - 8}{0.77/\sqrt{25}} = -1.753$ $df = 24$

c)



p-value in R: $pt(-1.753, 24) = 0.0462$

d) significant at the 5% level. e) yes. f) $7.73 \pm \overset{\cancel{2.06}\ \cancel{1.712}\ 1.712}{qt(\cancel{0.05}\ \cancel{0.025}\ 0.05, 24)} * 0.77/5 = \cancel{(7.41, 8.05)}$ $(\overset{7.46}{\cancel{8.00}}, 8.00)$

confidence interval needs to be symmetric

Q8 $\bar{x}_1 - \bar{x}_2 = 27.88 - 22.92 = 4.96 \pm t^* \sqrt{s_1^2/n_1 + s_2^2/n_2}$ ← $\sqrt{\frac{5.29^2}{26} + \frac{5.01^2}{26}} = \cancel{3.258}\ 1.429$

t^* from t-dist with $df = 25$: $t^* = 2.485$

$4.99 \pm 2.485 * \overset{1.429}{\cancel{3.258}} =$ ~~(-3.11, 13.09)~~ $(1.44, 8.54)$

Q9 a) p = proportion of graduating class who got jobs. $\hat{p} = 348/400 = 0.87$

b) yes, number of success 348, number of failures 52 >> 20.

c) $\hat{p} \pm z_* \sqrt{\hat{p}(1-\hat{p})/n}$ 95% $\Rightarrow z_* = 1.96$

$0.87 \pm 1.96\sqrt{0.87 \times (1-0.87)/400} = (0.84, 0.90)$

d) 95% of all samples of size 400 from the graduating class will give a confidence interval which contains the population proportion p.

e) 99% $\Rightarrow z_* = 2.~~83~~$ 58 $\quad 0.87 \pm 2.58 * 0.01682 = (0.83, 0.91)$

99% of all samples of size 400 will give a confidence interval containing p.

f) 99% confidence interval is larger.

<u>Q10</u>

grp	n	p
1	11,545	0.080
2	4,691	0.088

a) $\hat{p}_1 - \hat{p}_2 \pm \underbrace{z_*}_{1.96} \underbrace{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}_{0.0047587}$ $\quad \hat{p} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1+n_2} = 0.082311$

$0.008 \pm 1.96 \times 0.0047587 = (-0.00133, 0.017327)$

b) $H_0 : \hat{p}_1 - \hat{p}_2 = 0$ $\quad$ test statistic: $\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.008}{0.0047587} = 1.6811.$

$H_1 : \hat{p}_1 - \hat{p}_2 > 0$

z_* critical value for 1-sided 95% test is $z_* = 1.644$. $1.6811 > 1.644$, so significant at ~~≥~~95% for 1-sided test. This is why you shouldn't use 1-sided tests!