HW3Solutions

Dr. Syring

2/6/2020

4.2.2

An approximate CI is

$$(\bar{X} \pm t_{1-\alpha/2}(n-1)S\sqrt{n}).$$

For the data provided the 95% CI is $(101.15 \pm 2.093 * 105.41/\sqrt{20}) = (51.82, 150.48)$.

4.2.8

The statement

$$0.954 = P(\bar{X} - 1/2 < \mu < \bar{X} + 1/2)$$

implies $1/2 = z_{0.977}\sigma/\sqrt{n} = 1.9954\sqrt{10}/\sqrt{n}$. So, $n = 4 * 10 * 1.9954^2 \approx 160$. Since n is sample size it makes sense to round up to the nearest whole number.

4.2.9

a. $L = 2z_{1-\alpha/2}\sigma/\sqrt{n} = 2 * 1.96 * \sigma/3 = 1.30667\sigma$.

b.
$$E(L) = \frac{2 * 2.306\sigma}{3\sqrt{8}} E\left(\left[\frac{(n-1)S^2}{\sigma^2}\right]^{1/2}\right).$$

Using Thm 3.3.2 from Hogg,

$$E\left(\left[\frac{(n-1)S^2}{\sigma^2}\right]^{1/2}\right) = 2^{1/2}\frac{\Gamma(4.5)}{\Gamma(4)} = 2.741625$$

Then,

$$E(L) = 1.490154\sigma.$$

c. Unsurprisingly, the second CI has larger expected length. Since in the second case the population variance is unknown there is greater uncertainty in the variance of the sample mean. This greater uncertainty manifests in a wider interval estimate for the mean.

4.2.10

Note that $\bar{X} - X_{n+1} \sim N(0, \sigma^2(1 + 1/n))$. Therefore,

$$\frac{\bar{X} - X_{n+1}}{\sigma / \sqrt{\frac{n}{n+1}}} \sim N(0,1).$$

We still have the usual

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

independent of the difference of sample mean and next observation so that

$$T = \frac{\frac{\bar{X}-X_{n+1}}{\sigma/\sqrt{\frac{n}{n+1}}}}{\sqrt{S^2/\sigma^2}} = \frac{\bar{X} - X_{n+1}}{S/\sqrt{\frac{n}{n+1}}} \sim t(n-1).$$

You can think of the next observation X_{n+1} as the unknown "parameter" – in this way prediction is like inference. Then, a CI or PI (prediction interval) for X_{n+1} is given by

$$\left(\bar{X} + t_{\alpha/2}(n-1)S/\sqrt{n/(n+1)},\ \ \bar{X} + t_{1-\alpha/2}(n-1)S/\sqrt{n/(n+1)}\right).$$

In this problem, we have $k = t_{1-\alpha/2}(n-1)\sqrt{(n+1)/n} = t_{0.9}(7)\sqrt{(9/8)} = 1.500753$.

4.2.18

a. From Student's Theorem we know $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ so by definition

$$1-\alpha = P(\chi^2(n-1)_{\alpha/2} < \frac{(n-1)S^2}{\sigma^2} < \chi^2(n-1)_{1-\alpha/2})$$

for quantiles defined $P(Y < \chi^2(k)_\alpha) = \alpha$ when $Y \sim \chi^2(k)$. Then, by a couple algebraic steps

$$1-\alpha = P\left(\frac{(n-1)S^2}{\chi^2(n-1)_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2(n-1)_{\alpha/2}}\right).$$

b. The CI is $(8 * 7.93/17.535,\ 8 * 7.93/2.18) = (3.62,\ 29.10)$.
c. If μ is known then it makes sense to use $S^{*2} = \frac{1}{n}\sum_{i=1}^n (x_i - \mu)^2$ to estimate σ^2. Write this estimator as

$$S^{*2} = \frac{\sigma^2}{n}\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right)^2 = \frac{\sigma^2}{n}\sum_{i=1}^n (Z_i)^2$$

where $Z_i \overset{iid}{\sim} N(0,1)$. Then, $\frac{nS^{*2}}{\sigma^2} \sim \chi^2(n)$. The modified CI for σ^2 is

$$\left(\frac{nS^{*2}}{\chi^2(n)_{1-\alpha/2}},\ \ \frac{nS^{*2}}{\chi^2(n)_{\alpha/2}}\right).$$

4.2.21

Since the population variances are equal we use the pooled version of the CI:

$$\left(\bar{X} - \bar{Y} \pm t_{1-\alpha/2}(n+m-2)\sqrt{S_p^2(1/n+1/m)}\right)$$

where $S_p^2 = \frac{(n-1)S_X^2+(m-1)S_Y^2}{n+m-2}$. For the given data we have

$$(4.8 - 5.6 \pm 2.101\sqrt{8.64/10 + 7.88/10}) = (-3.5,\ 1.9)$$

4.2.22

By the CLT the difference of proportions is approximately normally distributed, i.e.

$$\hat{p}_1 - \hat{p}_2 \dot{\sim} N(p_1 - p_2, p_1(1-p_1)/n + p_2(1-p_2)/m).$$

Then, an approximate CI for $p_1 - p_2$ is given by

$$(\hat{p}_1 - \hat{p}_2 \pm z_{1-\alpha/2}\sqrt{p_1(1-p_1)/n + p_2(1-p_2)/m}).$$

For the given data we compute the CI

$$(0.1 \pm 1.645\sqrt{.25/100 + .24/100})$$

$$= (-0.1514, .21515).$$

4.2.27

Also see notes from 01/29.

$$1 - \alpha = P(F_{\alpha/2}(n-1, m-1) < F < F_{1-\alpha/2}(n-1, m-1))$$

$$= P(F_{\alpha/2}(n-1, m-1) < \frac{S_X^2}{S_Y^2}\frac{\sigma_Y^2}{\sigma_X^2} < F_{1-\alpha/2}(n-1, m-1))$$

$$= P(\frac{S_Y^2}{S_X^2}F_{\alpha/2}(n-1, m-1) < \frac{\sigma_Y^2}{\sigma_X^2} < \frac{S_Y^2}{S_X^2}F_{1-\alpha/2}(n-1, m-1))$$

$$= P(\frac{S_X^2}{S_Y^2}1/F_{1-\alpha/2}(n-1, m-1) < \frac{\sigma_X^2}{\sigma_Y^2} < \frac{S_X^2}{S_Y^2}1/F_{\alpha/2}(n-1, m-1))$$

$$= P(\frac{S_X^2}{S_Y^2}1/F_{1-\alpha/2}(n-1, m-1) < \frac{\sigma_X^2}{\sigma_Y^2} < \frac{S_X^2}{S_Y^2}F_{1-\alpha/2}(m-1, n-1))$$

where the last line is the most common form and uses the fact that $F_\alpha(u, v) = 1/F_{1-\alpha}(v, u)$.