# Math 332 A - Mathematical Statistics

Ethan Jensen

March 6, 2020

HW p.320 #4,9  p.328 #22,24,34,36

## 1 p. 320 #4

Show that the $(1-\alpha)100\%$ confidence interval

$$\overline{x} - z_{\alpha/2} \cdot \frac{\sigma}{n} < \mu < \overline{x} + z_{\alpha/2} \cdot \frac{\sigma}{n}$$

is shorter than the $(1-\alpha)100\%$ confidence interval

$$\overline{x} - z_{2\alpha/3} \cdot \frac{\sigma}{n} < \mu < \overline{x} + z_{\alpha/3} \cdot \frac{\sigma}{n}$$

Let f(x) be the standard normal distribution. We know

$$\int_{z_{2\alpha/3}}^{(z_{\alpha/3}+z_{2\alpha/2})/2} f(x)dx + \int_{(z_{\alpha/3}+z_{2\alpha/2})/2}^{z_{\alpha/3}} f(x)dx = \int_{z_{2\alpha/3}}^{z_{\alpha/2}} f(x)dx + \int_{z_{\alpha/2}}^{z_{\alpha/3}} f(x)dx$$

$$\int_{z_{2\alpha/3}}^{z_{\alpha/2}} f(x)dx = \int_{z_{\alpha/2}}^{z_{\alpha/3}} f(x)dx = \frac{\alpha}{6}$$

Since $z_{2\alpha/3} < z_{\alpha/3}$ and f(x) is decreasing on $(0,\infty)$, we know

$$\int_{z_{2\alpha/3}}^{(z_{\alpha/3}+z_{2\alpha/2})/2} f(x)dx > \int_{(z_{\alpha/3}+z_{2\alpha/2})/2}^{z_{\alpha/3}} f(x)dx$$

Thus, making a substitution into the first equation we have

$$2\int_{z_{2\alpha/3}}^{(z_{\alpha/3}+z_{2\alpha/2})/2} f(x)dx > 2\int_{z_{2\alpha/3}}^{z_{\alpha/2}} f(x)dx$$

Since the integrals have the same lower bounds, and f(x) is always positive,

$$(z_{\alpha/3} + z_{2\alpha/2})/2 > z_{\alpha/2}$$

$$z_{\alpha/2} + z_{\alpha/2} < z_{\alpha/3} + z_{2\alpha/2}$$

$\boxed{\therefore \text{The first confidence interval is shorter than the second.}}$

## 2 p. 320 #9

Show that $S_p^2$ is an unbiased estimator of $\sigma^2$ and find its variance under the conditions of Theorem 11.5.

The pooled estimator of $\sigma^2$ is defined as follows:

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

Because Expected Value is a linear operator,

$$E(S_p^2) = \frac{1}{n_1+n_2-2}[(n_1-1)E(S_1^2) + (n_2-1)E(S_2^2)]$$

Since $S_1^2$ and $S_2^2$ are both unbiased estimators of $\sigma^2$,

$$E(S_p^2) = \frac{1}{n_1+n_2-2}[(n_1-1)\sigma^2 + (n_2-1)\sigma^2] = \frac{n_1-1+n_2-1}{n_1+n_2-2}\sigma^2$$

$$E(S_p^2) = \sigma^2$$

$\boxed{\text{So } E(S_p^2) \text{ is an unbiased estimator of } \sigma^2.}$

By Theorem 8.11, under the conditions of Theorem 11.5

$$\frac{(n_1-1)S_1^2}{\sigma^2} \text{ and } \frac{(n_2-1)S_2^2}{\sigma^2}$$

have chi-square distirbutions with respctively $n_1-1$ and $n_2-1$ degrees of freedom.

Using Theorem 8.9,

$$X = \frac{(n_1+n_2-2)S_p^2}{\sigma^2}$$

has a chi-square distribution with $n_1+n_2-2$ degrees of freedom.

This means $S_p^2$ is a linear combination of a chi-square random variable and Corollary 4.3 applies.

$$\mathrm{var}(S_p^2) = \left(\frac{\sigma^2}{n_1+n_2-2}\right)^2 \mathrm{var}(X)$$

Using Corollary 6.2 we have

$$\mathrm{var}(S_p^2) = \left(\frac{\sigma^2}{n_1+n_2-2}\right)^2 2(n_1+n_2-2)$$

$$\boxed{\mathrm{var}(S_p^2) = \frac{2\sigma^4}{n_1+n_2-2}}$$

## 3 p. 328 #22

A medical research worker intends to use the mean of a random sample of size $n = 120$ to estimate the mean blood pressure of women in their fifties. If, based on experience, he knows that $\sigma = 10.5$mm of mercury, what can he assert with probability 0.99 about the maximum error?

By Theorem 11.1, the maximum error is less than

$$z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

With Table III, we find that $z_{\alpha/2} = 2.575$.
Plugging in $n = 120, \sigma = 10.5$, we have that our maximum error is equal to $2.575 \cdot \frac{10.5}{\sqrt{120}} \approx 2.468$.

$\boxed{\text{with probability 99\%, the maximum error is about 2.468mm of mercury.}}$

# 4 p. 328 #24

A study of the annual growth of certain cacti showed that 64 of them, selected at random in a desert region, grew on average 52.80mm with a standard deviation of 4.5mm. Construct a 99% confidence interval for the true average annual growth of the given kind of cactus.

With Theorem 11.1, the maximum error is less than

$$z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

With Table III, we find that $z_{\alpha/2} = 2.575$.
Plugging in $\mu = 52.80, n = 64, \sigma = 4.5$, we can construct the confidnece interval. $52.80 \pm 2.575 \cdot \frac{4.5}{\sqrt{64}} = 51.35$ and 54.25 respectively.

$\boxed{\text{The 99\% confidence interval of the mean is } (51.35, 54.25) \text{ mm growth.}}$

## 5 p. 328 #34

A study of two kinds of photocopying equipment shows that 61 failures of the first kind of equipment took on the average 80.7 minutes to repair with a standard deviation of 19.4 minutes, whereas 61 failures of the second kind of equipment took on the average 88.1 minutes to repair with a standard deviation of 18.8 minutes. Find a 99% confidence interval for the difference between the true average amounts of time it takes to repair failures of the two kinds of photocopying equipment.

With Theorem 11.4 we can deduce the 99% confidence interval.
Plugging in $\mu_1 = 80.7, \mu_2 = 88.1, n_1 = 61, n_2 = 61, \sigma_1 = 19.4, \sigma_2 = 18.8$
And using our spicy value $z_{\alpha/2} = 2.575$ we have

$$E = 2.575 \cdot \sqrt{\frac{19.4^2}{61} + \frac{18.8^2}{61}} = 8.907$$

$$\mu_1 - \mu_2 = 7.4$$

$\boxed{\text{The 99\% confidence interval for the true difference of the means is } (0, 16.307)}$

*Note that I'm forcing the difference to be a positive value.*

# 6 p. 328 #36

The following are the heat-producing capacities of coal from two mines (in millions of calories per ton):

*Mine A:* 8,500, 8,330, 8,480, 7,960, 8,030

*Mine B:* 7,710, 7,890, 7,920, 8,270, 7,860

Assuming that the data constitute independent random samples from normal populations with equal variances, construct a 99% confidence interval for the difference between the true average heat-producing capacities of coal from the two mines.

Since the two populations have equal variances, we can use Theorem 11.5. to construct the confidence interval.

First, we have to calculate $\overline{x}_1$ and $\overline{x}_2$, which are 8260 and 7930 respectively. Next, we calculate $S_1^2$ and $S_2^2$, which are 63450 and 42650 respectively. Then, we calculate the pooled variance $S_p$ which is 205.03.

Then, we look up the value of $t_{0.005,8}$ in Table 4 which is 3.355.

Using Theorem 11.5 we calculate the confidence interval which is $(0, 816.40)$.

$\boxed{\text{The 99\% confidence interval for the true difference of the means is } (0, 816.40)}$

*Note that I'm forcing the difference to be a positive value.*