CONCORDIA UNIVERSITY
Department of Economics

ECON 221/2 SECTIONS A, B, C and DD
STATISTICAL METHODS I
FALL 2016 – MIDTERM 2 (SOLUTIONS)
Sunday, November 20, 2016, 1pm – 3pm

Name: **I.D:** **Section:**

TOTAL: 50 points

1. (**10 points**) The proportion of adult women in Canada is 51 percent. A marketing survey telephones 400 people at random. Let $\hat{p}$ be the proportion of women in the sample.

   a. (**2 points**) ***Briefly*** describe the sampling distribution of $\hat{p}$.

   Since $np(1-p) = 400 \cdot 0.51 \cdot (1-0.51) = 99.96 > 5$, *the central limit theorem states that the sample size is sufficiently large to assume that the sampling distribution of* $\hat{p}$ *is normal.*

   b. (**2 points**) Calculate the mean and standard deviation of the sampling distribution of $\hat{p}$.

   $$E(\hat{p}) = p = 0.51,\ \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.51 \cdot (1-0.51)}{400}} = 0.0250$$

   c. (**2 points**) Calculate the probability of finding a sample with more than 53 percent women.

   $$\Pr(\hat{p} > 0.53) = \Pr\left(\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} > \frac{0.53-0.51}{\sqrt{\frac{0.51 \cdot (1-0.51)}{400}}} \right) = \Pr(z > 0.80) = 1 - \Pr(z < 0.80)$$
   $$= 1 - 0.7881 = 0.2119$$

   d. (**2 points**) Calculate the probability of finding a sample with less than 47 percent women.

$$\Pr(\hat{p} < 0.47) = \Pr\left(\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} < \frac{0.47-0.51}{\sqrt{\frac{0.51\cdot(1-0.51)}{400}}}\right) = \Pr(z < -1.60) = \Pr(z > 1.60)$$

$$= 1 - \Pr(z < 1.60) = 1 - 0.9452 = 0.0548$$

e. (**2 points**) The probability is 0.15 that the proportion of women in the sample is less than what number?

$$\Pr(\hat{p} < p_0) = 0.15 \Rightarrow \Pr\left(\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} < \frac{p_0-0.51}{\sqrt{\frac{0.51\cdot(1-0.51)}{400}}}\right) = 0.15 \Rightarrow \Pr(z < z_0) = 0.15$$

$$\Rightarrow \Pr(z > -z_0) = 0.15 \Rightarrow 1 - \Pr(z < -z_0) = 0.85 \Rightarrow -z_0 = 1.035 \Rightarrow z_0 = -1.035$$

$$\Rightarrow \frac{p_0-0.51}{\sqrt{\frac{0.51\cdot(1-0.51)}{400}}} = -1.035 \Rightarrow p_0 = 0.4841$$

2. (**10 points**) A marketing researcher for a phone company surveys 100 people and finds that the proportion of clients who plan to switch providers when their contract expires is $\hat{p} = 0.15$.

a. (**2 points**) Calculate the margin of error for a 95 percent confidence interval for the true proportion of clients who plan to switch providers.

$$ME = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{0.15\cdot(1-0.15)}{100}} = 0.0700$$

b. (**2 points**) Construct a 95 percent confidence interval for the true proportion of clients who plan to switch providers.

$$\hat{p} \pm ME = 0.15 \pm 0.07 = (0.08, 0.22)$$

c. (**2 points**) ***Briefly*** interpret the confidence interval you found in part (b).

There is 95 percent confidence that the true proportion of clients who plan to switch providers is between eight and 22 percent.

d. **(2 points)** Suppose you wanted to be 95 percent confident that your estimate is within 0.05 of the true proportion, calculate the required sample size.

$$n = \left(\frac{0.5 z_{\alpha/2}}{ME} \right)^2 = \left(\frac{0.5 \cdot 1.96}{0.05} \right)^2 = 384.16 \approx 385$$

e. (**2 points**) If we only need to be 90 percent confident, ***briefly*** explain whether the margin of error would be larger or smaller than the one found in part (b).

A lower level of confidence would result in a smaller margin of error.

3. (**10 points**) In 2015, the Canadian income per capita was $40,807 with a standard deviation of $12,650. A random sample of 16 Canadians is selected. It is known that income per capita is normally distributed.

a. (**2 points**) Calculate the mean of the sampling distribution.

The mean of the sampling distribution is $40,807.

b. **(2 points)** Calculate the standard deviation of the sampling distribution.

The standard deviation of the sampling distribution is $sd\left(\bar{x}\right) = \frac{\sigma}{\sqrt{n}} = \frac{12650}{\sqrt{16}} = 3162.5$.

c. **(2 points)** Calculate the probability that the mean income in the sample is less than $35,000.

$$\Pr\left(\bar{x} < 35000\right) = \Pr\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < \frac{35000 - 40807}{12650/\sqrt{16}} \right) = \Pr\left(z < -1.84\right) = \Pr\left(z > 1.84\right)$$
$$= 1 - \Pr\left(z < 1.84\right) = 1 - 0.9671 = 0.0329$$

d. (**2 points**) Calculate the probability that the mean income in the sample exceeds $46,000.

$$\Pr\left(\bar{x} > 46000\right) = \Pr\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} > \frac{46000 - 40807}{12650/\sqrt{16}} \right) = \Pr\left(z > 1.64\right) = 1 - \Pr\left(z < 1.64\right) =$$
$$= 1 - 0.9495 = 0.0505$$

e. (**2 points**) ***Briefly*** explain whether it was necessary to know that per capita income is normally distributed.

Yes, it was necessary to know that the underlying population is normally distributed; otherwise, a sample of at least 25 Canadians would be needed to invoke the central limit *theorem. The sampling distribution of $\bar{x}$ is normal only if the underlying population is normal or if the sample size is large enough to invoke the central limit theorem.*

4. (**8 points**) As a result of the decline in the oil and gas industry, home prices in some areas of British Columbia have been falling. In one large community, a sample of 25 homes revealed an average fall in home value of $11,560 with a standard deviation of $1,500.

   a. (**2 points**) Calculate the margin of error for a 90 percent confidence interval for the true mean loss in home values.

   $$ME = t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} = 1.711 \cdot \frac{1500}{\sqrt{25}} = 513.3$$

   b. (**2 points**) Construct a 90 percent confidence interval for the true mean loss in home values.

   $$\bar{x} \pm ME = 11560 \pm 513.3 = (11046.70, 12073.30)$$

   c. (**2 points**) ***Briefly*** explain what assumptions were necessary to construct the confidence interval in part (b).

   It was necessary to assume that the population is distributed normally so that the estimated sample variance could be used in place of the unknown population variance and the t-distribution could be used to construct the confidence interval.

   d. (**2 points**) Suppose that, for the entire province of British Columbia, the average loss in home values is $10,000. ***Briefly*** explain whether we can conclude that the loss in the sampled community differs from the provincial average.

   The provincial mean loss of $10,000 is below the lowest value in the 90 percent confidence interval. Therefore, we can conclude, with 90 percent confidence, that the loss in the sampled community differs from the provincial average.

5. (**8 points**) A trucking company is comparing two different routes for efficiency. Twenty-one truckers randomly assigned to follow route *A* report an average of 40 minutes with a standard deviation of 3 minutes. Twenty-one truckers randomly assigned to follow Route ***B*** report an average of 43 minutes with a standard deviation of 2 minutes. Assume that the random sample observations are from normally distributed populations and that the population variances are assumed to be equal.

   a. (**2 points**) Calculate the pooled sample variance.

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(21-1)\cdot 3^2 + (21-1)\cdot 2^2}{21+21-2} = 6.5$$

b. (**2 points**) Calculate the margin of error for a 99 percent confidence interval for the true difference in the average time for the two routes (ie, route *B* – route *A*).

$$ME = t_{\alpha/2,n_1+n_2-2}\sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}} = 2.704\cdot\sqrt{\frac{6.5}{21}+\frac{6.5}{21}} = 2.1275$$

c. (**2 points**) Construct a 99 percent confidence interval for the true difference in the average time for the two routes.

$$\bar{x}_1 - \bar{x}_2 \pm ME = 43 - 40 \pm 2.1275 = (0.8725, 5.1275)$$

d. (**2 points**) ***Briefly*** explain whether we can conclude whether one route is more efficient than the other.

Since zero is not within the limits of the confidence interval, we can conclude with 99 percent confidence that route A is more efficient on average than route B.

6. (**6 points**) A random sample of 508 men and 508 women in Canada found that 65 percent of men and 45 percent of women supported scrapping the penny. Let p_1 and p_2 represent the proportion of men and women, respectively, who support scrapping the penny.

a. (**2 points**) Calculate the margin of error for a 90% confidence interval for $p_1 - p_2$.

$$n_1 = 508,\ \hat{p}_1 = 0.65,\ n_2 = 508,\ \hat{p}_2 = 0.45$$

$$ME = z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 1.645\sqrt{\frac{0.65\cdot(1-0.65)}{508}+\frac{0.45\cdot(1-0.45)}{508}} = 0.0503$$

b. **(2 points)** Construct a 90 percent confidence interval for $p_1 - p_2$.

$$\hat{p}_1 - \hat{p}_2 \pm ME = 0.65 - 0.45 \pm 0.0503 = (0.1497, 0.2503)$$

c. **(2 points)** Based on the interval in part (c), ***briefly*** explain whether we can conclude that the proportion of men who support scrapping the penny exceeds the proportion of women.

Since zero is not within the limits of the confidence interval, it is not possible that the proportions are the same. Therefore, one must conclude that the proportion of men who support scrapping the penny exceeds the proportion of women.