Student Number:_____________________ **Quiz 4**
March 1, 2004 **Confidence Intervals and Sample Size (Ch 7)**

Ground rules:

- DO NOT WRITE YOUR NAME ON THE QUIZ. If you're worried about losing the first page, write your student number on each page.
- You can use your notes or books, but you may not communicate with other people about this exam nor the material covered by it. You may use a calculator.
- In order to receive as much credit as possible, please show all of your work. Showing that you understand the question and know how to set up the solution correctly is more important than arriving at the exact answer.
- Read each question carefully and answer all parts of each question. Make sure to interpret your answer for a non-technical audience.
- Good luck!

---

1. A university dean is interested in determining the proportion of students who receive some sort of financial aid. Rather than examine the records for all students, the dean randomly selects 200 students and finds that 118 are receiving financial aid. Use a 90% confidence interval to estimate the true proportion of students on financial aid. Make sure to state your assumptions and to interpret your results in non-technical language for the dean. (10 points)

We assume that the sample is large enough to use the sample variance to estimate the standard error and that a normal approximation can be used (that is, the sample size is large enough for the sampling distribution of $\hat{p}$ to be normal and that the mean of the sampling distribution is an unbiased estimator of the true population proportion).

First we need to figure out the point estimate for the proportion of students receiving financial aid: $\hat{p} = \frac{118}{200} = .59$

Need a 90% confidence interval, leaving 10% in the tail. So, α=.10, α/2=.05, and z=1.645.

We'll use this equation to estimate the confidence interval:

$$\hat{p} \pm z_{a/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

To form the confidence interval we need to estimate the standard error: $\sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(.59)(1-.59)}{200}} = \sqrt{\frac{(.59)(.41)}{200}} = \sqrt{\frac{.2419}{200}} = \sqrt{0.00121} = .034778$

So, to form the 90% confidence interval:

[.59 ±1.645(.034778)] or [.53279, .64721].

We are 90% confident that the true proportion of students receiving financial aid is between 53% and 65%.

2. The dean would like to know what would cause the width of the range estimate to become more narrow. What 3 factors could change the width of the estimate? (10 points)

The dispersion of the data—if the sample were less dispersed (that is, the standard deviation were smaller), the width of the estimate would be smaller.

A larger sample size would produce a narrower estimate.

The level of confidence would also change the estimate. The more confident, the wider the interval. That is, to be more sure, the interval would have to include a wider range. So, to produce a narrower interval, the level of confidence would have to decrease. We'd be less sure of our estimate.

3. The dean would like to know if the sample of 200 students she selected is large enough to estimate the true proportion of all students to within plus or minus 3% with 99% reliability. If the sample is not large enough, tell the dean the sample size she would need (don't worry about correcting the size for without replacement sampling). Make sure to show how you arrived at your answer and interpret your result in nontechnical language for the dean. (10 points)

No, the sample of 200 is not large enough to make an estimate within plus or minus 3% with 99% confidence.

To figure out what sample size she would need, I used the formula for estimating a sample size using the dean's new parameters. Here's the formula.

$$n = \frac{z_{a/2}^2 pq}{B^2}$$

Here are the pieces of the equation:

z?/2: Since the level of confidence is 99%, ?=.01, ?/2=.005, and so z?/2=2.576.
B: the half-bound is 3% or .03.
Variance Estimate (pq): You could interpret this problem in two ways. To produce the largest and most conservative sample, one

could say that p=.5, so q=.5. OR I could also have used the estimate from the sample, where $\hat{p} = \frac{118}{200} = .59$. I've solved the problem using either estimate.

Where we assume p=.5:

$$n = \frac{z_{\alpha/2}^2 pq}{B^2} = \frac{2.576^2(.5)(.5)}{.03^2} = \frac{6.635776(0.25)}{.0009} = 1843.271$$

So, we could tell the dean she would need to have an effective sample size of 1,844 students.

Where we assume $\hat{p} = .59$:

$$n = \frac{z_{\alpha/2}^2 pq}{B^2} = \frac{2.576^2(.59)(.41)}{.03^2} = \frac{6.635776(0.2419)}{.0009} = 1783.549$$

So, we could tell the dean she would need to have an effective sample size of 1784 students.

4. Reprise last quiz: Please state whether the following statements are true or false. If the statement is false, please write a correction below the statement.

4a. The Central Limit Theorem guarantees that the population that you are sampling from is approximately normal whenever you have selected a sufficiently large sample. (5 points)

False. The Central Limit Theorem guarantees that the **sampling distribution** is approximately normal whenever you have selected a sufficiently large sample, **regardless of the distribution of the population.**

4b. As the sample size taken gets larger, the standard error of the sampling distribution of the sample mean gets larger as well. (5 points)

False. As the sample size taken gets larger, the standard error of the sampling distribution of the sample mean gets **smaller.**

4c. The standard error of the sampling distribution of the sample mean is equal to s, the standard deviation of the population. (5 points)

False. The standard error of the sampling distribution of the sample mean is equal to s/**n**, the standard deviation of the population **divided by the square root of the sample size.**

5. Extra Credit (OPTIONAL): Correct your estimate of the sample size the dean would need in 3 for without replacement sampling. Explain why the sample size is different from the original sample size calculated. (10 points)

Where we assume p=.5:

$$n_{WOR} = \frac{n}{1+\frac{n}{N}} = \frac{1784}{1+\frac{1784}{20{,}000}} = 1638$$

So, we would need an effective sample size of 1,638 to estimate the true proportion within plus or minus 3% with 99% condifidence.

Where we assume p=.59:

$n_{WOR} = \frac{n}{1 + \frac{n}{N}} = \frac{1844}{1 + \frac{1844}{20{,}000}} = 1688$ So, we would need an effective sample size of 1,688 to estimate the true proportion within plus or minus 3% with 99% confidence.

In both cases, the number we would need to sample is smaller than the original estimate because we are sampling without replacement. That is, each population element included in the sample, included unique information and is in the sample only once, unlike in the original sample that assumed with replacement sampling.