

STA-209 Exam #2

Name: _____

Directions

- You **do not** need to fill all the space provided. In many cases, large amounts of white space only exist to correct for general spacing of the exam
- You **do not** need to write in complete sentences: bullet points are completely acceptable and even preferred
- Use your critical value sheet to find critical values appropriate for the distribution at hand. When $df > 40$, use the standard normal approximation
- Have fun

$$\bar{X} \sim N(\mu, \sigma/\sqrt{n})$$

$$\hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

[This page unintentionally left blank]

Question 1

Answer the following short-answer questions:

Part A What is a distributional parameter? What are the distributional parameters for the Normal distribution and the t -distribution?

Part B What is the relationship between standard deviation and standard error

Part C Write out the expression for a confidence interval. Which terms impact the width of the interval? Which terms impact the coverage probability

Part D What is the relationship between confidence and our error rate?

Part E What are bias and variability?

Part F What issue would we run into if we had a sample of size $n = 10$ and we used a *standard normal* distribution to estimate the sampling distribution of the statistic $\frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{n}}$?

What could we do to correct this issue?

Part G Suppose I have two samples with:

- **Sample 1:** $n_1 = 25$ and $\sigma_1 = 10$
- **Sample 2:** $n_2 = 50$ and $\sigma_2 = 15$

Which sample will have the *least* variability in the distribution of \bar{x} ?

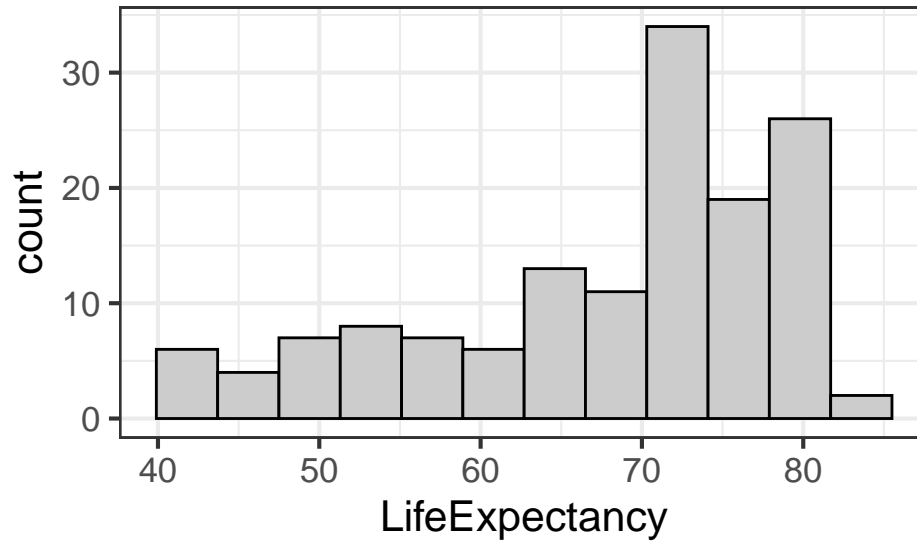
Question 2

Which of these would have a larger impact on the width of a 95% confidence interval centered around $\bar{x} = 50$ when $\hat{\sigma} = 1$ and $n = 20$

- Changing n to 25 (don't forget to change C value, too)
- Changing \bar{x} to 60
- Changing our confidence interval from 95% to 90%

Question 3

The Happy Planet dataset includes a number of indices of well-being across 143 countries. One of these indices, Life Expectancy, is visualized in the histogram below.



We are interested in determining the *proportion* of countries in which the average life expectancy is greater than 70 years. To this end, we collected two different samples:

- **Sample 1:** 13 countries had life expectancy greater than 70, 12 did not
- **Sample 2:** 16 countries had life expectancy greater than 70, 14 did not

Use this information to answer the following questions

Part A Find the estimated proportion and standard error for each of the samples provided. Round your estimates for the standard error to four decimal places.

Part B Which sample provides more evidence against the null hypothesis $H_0 : p_0 = 0.725$?

Question 4

Suppose we are interested in determining if the composition of public and private schools is the same between the Plains and Great Lakes regions of the US

##	Type		
## Region	Private	Public	
## Great Lakes	125	64	
## Plains	84	42	

Part A Find the proportion and associated standard error of private schools in each of the regions.

Part B Construct a 95% confidence interval for the *difference in proportions* (for $df > 40$ use the standard normal approximation)

Part C Construct a t -statistic testing the hypothesis that $p_1 = p_2$. Perform a hypothesis test with the null hypothesis $H_0 : p_1 = p_2$ with 95% confidence. What conclusion do you come to?

[Scratch Paper]