

Show your work to receive full credit.

Problem 1 Part 1

In a clinical study, 1000 subjects were vaccinated with a vaccine manufactured with a new ingredient. Over a period of roughly 28 weeks, 6 of these subjects developed the flu.

- Find the point estimate of the population proportion that were vaccinated but still developed the flu. Specify its value and statistical notation for this estimate.
- Show set up (formula with numbers plugged in) and verify that the standard error of this estimate is 0.0024. Briefly explain what this number means.
- What is the value of z-multiplier $z_{\alpha/2}$ for a 98% confidence interval? Find the margin of error for a 98% confidence interval. Round your margin of error to the nearest four decimal places.

Problem 1 Part 2 (e) updated

- Use your answers from Problem 1 Part 1 and construct a 98% confidence interval. **Interpret.**
- Based on your confidence interval, ~~is it plausible~~ **can we conclude** that less than 1% of all people vaccinated with the vaccine will develop the flu? Explain.

Problem 2

Refer to the Problem 1.

Does the information provided in Problem 1 satisfy conditions (assumptions) of the large number confidence interval for p ? State two assumptions and determine whether each is met or not. (Hint: See lecture notes on page 106. There are two assumptions.)

Problem 3

An election is expected to be close. Pollsters planning an exit poll want their estimates to be to within ± 0.02 with 90% confidence. How large should be the sample size if they have no prior information?

Problem 4 Part 1

The 2012 General Social Survey asked, “What do you think is the ideal number of children for a family to have?”. The 590 females who gave a numeric response from 0 to 6 had a median of 2, a mean of 2.56, and a standard deviation of 0.84.

- What is the point estimate of the population mean? Specify its value and statistical notation for this estimate.
- Find the standard error of this estimate. Round your answer to the nearest four decimal places. Explain what this number means.
- What is the value of t-multiplier $t_{\alpha/2, n-1}$ for a 99% confidence interval? Find the margin of error for a 99% confidence interval. Round your margin of error to the nearest two decimal places.

Problem 4 Part 2 (e) updated)

- 99% confidence interval is (2.47, 2.65). Interpret the interval in context.
- ~~Is it plausible~~ **Can we conclude** that the population mean of females’ ideal number of children is greater than 2?

Problem 5

Refer to Problem 4. State assumptions of the confidence interval for the population mean. Determine whether each assumption is met based on Problem 4 description. Explain.

Problem 6: Multiple Choice Question (Option D removed)

The following interprets the 99% confidence interval for population mean of females’ ideal number of children from Problem 4 part d) (2.47, 2.65). Which of the following is correct?

- In the entire population, 99% of females’ ideal number of children fall within 2.47 and 2.65.
- We are 99% confident the sample mean of females’ ideal number of children falls between 2.47 and 2.65.
- If we repeatedly sample 590 females randomly and construct 99% confidence intervals, then in the long run, 99% of those intervals contain the population mean of females’ ideal number of children.
- ~~There is 99% probability population mean is between 2.47 and 2.65.~~

R Problem

In this problem, we will use the class survey data set. We will assume that the population of interest is *all students at the University of Minnesota* and the variable of interest is average weekday sleep hours. Assume that *the survey data represents a random sample of 391 students at the University of Minnesota*.

Use the following R code to import data from the survey data.

```
survey_f22<-read.csv("http://users.stat.umn.edu/~parky/Fall2022Survey.csv", header=TRUE)
```

- a) Construct a histogram and Q-Q plot of students' weekday sleep hours. Include both plots in your submission. Describe the shape of the distribution.
- b) (Multiple Choice Question) The histogram from the previous part is :
 - (i) a population distribution
 - (ii) a data distribution (sample distribution)
 - (iii) a sampling distributionBriefly explain.
- c) State the assumptions of confidence interval for the mean. Determine whether our sample satisfies this assumption.
- d) Use `t.test()` command to construct a 90% confidence interval to estimate the population mean weekday sleep hours. Submit R command and output. Interpret the interval.