

STAR511 HW#7

Questions 1 and 2 (Defective Items): A factory manager wants to estimate the proportion of defective items. A random sample of 65 items was inspected and it was found that 4 of them are defective.

1. Is the sample size large enough for the normal approximation to be valid? Justify your response using the criteria discussed in the notes.
2. Provide an estimate of the proportion of defective items and corresponding **90%** confidence interval using the exact binomial method. Fine to just show the resulting output. **Hint:** use `conf.level = 0.90`.

Questions 3 and 4 (Election): Suppose we are interested in estimating the proportion of registered voters who support candidate Jones for mayor. A random sample of $n = 215$ voters were asked “Do you support candidate Jones for mayor?” From this sample, 124 answered yes and 91 answered no. Use large sample normal approximation via `prop.test()` with default `correct = TRUE`.

3. Provide an estimate of the proportion who support candidate Jones and a corresponding 95% confidence interval. Give an interpretation for this confidence interval.
4. Do we have evidence that more than half of registered voters support candidate Jones? In other words, test $H_0: \pi \leq 0.50$ versus $H_A: \pi > 0.50$. Use $\alpha = 0.05$.
 - A. Conduct an appropriate test and show the output.
 - B. Provide a conclusion in context.

Questions 5 through 7 (Survey Planning): A public opinion polling agency plans to conduct a national survey to determine the π of people who would be willing to pay a higher per kilowatt hour rate for electricity provided that renewable sources were used (solar, wind, etc). How many people must be included in the poll to estimate π within 0.06 using a 95% CI? In other words, find the minimum sample size required to achieve $95\% ME \leq 0.06$.

Notes:

- Use the large sample normal approximation.
 - Report an integer value for the sample size (round up!).
5. Suppose the polling agency has no previous information about π .
 6. Suppose the polling agency conjectures that $\pi = 0.3$.
 7. Just for this question, suppose they want to plan a study to compare the proportion of CO vs WY residents who would be willing to pay a higher rate. Investigators conjecture that 40% of CO residents will agree vs 10% of WY residents. What sample size (per group/state) is required to achieve 90% power?

Questions 8 and 9 (Political Debate): The data from O&L problem concerns the effect of a political debate between two candidates (A and B). Using a sample of $n = 75$ registered voters, a political scientist records each voter's preference Before and After the debate.

Notes:

- Full data is given below.
- “After A” indicates support for candidate A after the debate. Other cells are defined similarly.

	After A	After B	Total
Before A	28	13	41
Before B	6	28	34
Total	34	41	75

8. Calculate the proportion that prefer candidate A before the debate. Calculate the proportion that prefer candidate A after the debate.
9. Considering the design, run an appropriate test to see if there is strong evidence that voters who changed their preference after the debate changed in favor of one candidate or the other.

Questions 10 through 14 (Birds): A case-control study in Berlin, reported by Kohlmeier, Armingier, Bartolomeycik, Bellach, Rehm and Thamm (1992) and by Hand et al. (1994) asked 239 lung cancer patients and 429 healthy controls (matched to the cases by age and sex) whether or not they had kept a pet bird during adulthood. The data is summarized here.

	Healthy Controls	Cancer Patients	Total
No Bird	328	141	469
Yes Bird	101	98	199
Total	429	239	668

10. A colleague looks at the data above and says “Wow, an estimated 35% ($239/668$) of the population has lung cancer.” Briefly explain why this is NOT correct.
11. Use `oddsratio()` from the `epitools` package to run an appropriate analysis and show your output. **Note:** Use `method="wald"`.
12. Give an interpretation of the odds ratio from Q11.
13. Do we find evidence of an association between bird ownership and lung cancer? Briefly justify your response using the 95% confidence interval for the odds ratio from Q11.
14. Do we find evidence of an association between bird ownership and lung cancer? Briefly justify your response using the chi-square test p-value from Q11.