This quiz includes material from section ISI 5.3 and possibly from some preceding sections.

**ISI 5.3.23, PHS ulcers, pp.316-317**

ISI 5.3.23  The Physicians’ Health Study is a very large, randomized study designed to “test the effects of low-dose aspirin... in the prevention of cardiovascular disease (CVD).” The subjects were 22,071 U.S. male physicians (aged 40–84 years, in the year 1982), who were randomly assigned to be in either the low-dose aspirin group or the placebo group. Each participant was required to take the assigned pill every other day for five years. The study was double blind.

ISI 5.3.27 Another outcome of interest in the Physicians’ Health Study was whether the subjects developed ulcers or not. Of the 11,034 physicians who took the placebo, 138 developed ulcers during the study. Of the 11,037 physicians who took aspirin, 169 developed ulcers. The researchers would like to know if taking aspirin is associated with the development of ulcers.


We will conduct a \( z \)-test for the difference of two population proportions and calculate a 95% confidence interval for the appropriate population parameter.

**HT**

To proceed with this exercise, load the appropriate data into the Theory-based Inference applet, as follows:

a. launch the Theory-based Inference applet
   
   select Scenario: two proportions

b. under Group 1, enter the values of \( n \) and \( \text{count} (= \text{number of 'successes') for the placebo group}
   
   under Group 2, enter the values of \( n \) and \( \text{count} (= \text{number of 'successes') for the aspirin group}

   do NOT use commas inside larger numbers for these applets; write 1234 instead of 1,234

   click Calculate

**research question**

1. *State the research question.*
   
   Make clear whether the researchers want to prove that something is \(<, >, \) or \( \neq \) to something else.
   
   This is key for constructing the matching \( H_a.\)

**experiment or observational study?**

2. *What is the key requirement for making this an experiment?*

3. *Is that requirement satisfied for this study?*

4. *If this study is not an experiment, then what is it?*
variables

5. Identify the **observational units**. How many observational units are there all together in this data set? Hint: Check the sample statistics reported by the applet.

6. Identify the **explanatory variable**.
   Is the explanatory variable quantitative or categorical?

7. Identify the **response variable**.
   Is the response variable quantitative or categorical?

hypotheses

8. Define (in words) the **parameters** of interest in this study.

9. Assign symbols to the parameters.
   Parameters refer to populations, so use \( \pi \) or \( \mu \).
   Use subscripts to refer to specific populations \( \pi_{\text{tamoxifen}} \).

10. State the null and alternative **hypotheses in words**.
    \( H_0 \) often states that two variables are independent
    \( H_a \) often states that two variables are associated
    associated means NOT independent

11. State the null and alternative hypotheses using symbols.
    For \( H_0 \) our authors always choose =
    For \( H_a \) we always have three choices: >, <, ≠
    \( H_a \) must be consistent with the research question
    \( H_a \) indicates that we are trying to gather evidence to show that this is the case

test statistic

12. What **test statistic** will you use for this study?
    Express the test statistic for this study in words
13. Express the test statistic for this study in symbols.
   Test statistics are computed from samples, so use \( \hat{p} \) or \( \bar{x} \)
   For proportions, use **conditional proportions**
   Symbols such as \( \hat{p}_{\text{tamoxifen}} \) refer to **conditional sample proportions**
   \( \hat{p}_{\text{tamoxifen}} \) is the observed proportion of participants who _______ given that they were taking

14. *What is the observed value of the test statistic in this study?*
   *Hint: Check the sample statistics reported by the applet.*
   
   observed test statistic = ________

**theory-based inference**

15. *Use the Theory-based Inference applet with Scenario: two proportions*
   *Click on Test of Significance*
   *What is the correct form of the alternative hypothesis for this problem (\(<, >, \neq\)?)?*

16. *Report the value of the standardized test statistic (z or t).*


18. *Include a screenshot of your completed applet displaying the required images and statistics.*

**significance (theory-based inference)**

19. *What is the strength of evidence against the null hypothesis indicated by this theory-based p.value?*
   not much 0.10 moderate 0.05 strong 0.01 very strong

**estimation**

95% CI and its interpretation

\[
CI = \text{point.estimate} \pm \text{multiplier} \times SE
\]
20. Use the Theory-based Inference applet to determine a 95% confidence interval for the appropriate population parameter.

21. Interpret the confidence interval.
   with how much confidence (%), exactly what (in symbols) is where (numerical limits)?

22. What does the phrase “with 95% confidence” refer to in this context?
   hint: it does NOT refer to a specific parameter and interval
   it DOES refer to . . .

23. What is the midpoint of this CI?
   in symbols and numerical value

24. Recall: the entire interval is the interval estimate and the center of the interval is the point estimate
   ... of what?

25. What is the ME (= margin of error) of this interval (numerical value)?

   \[ ME = \text{multiplier} \times SE \]

26. The ME of this interval is an indication of _____ (what?).

27. Report a 99% CI for the same data.

28. What is the relationship (if any) between the center of the 95% CI and the center of the 99% CI?

   \[ CI = \text{point.estimate} \pm ME \]

29. Compared to the 95% CI, is the 99% CI wider, narrower, or the same width?

30. Include screenshots of your completed applet displaying the 95% CI and the 99% CI.