## Chapters covered: Chapter 10

## Show your work to receive full credit.

- **Problem 1** For each part below answer
  (i) state the name of relevant hypothesis test:
  - One sample $z$-test for proportion
  - One sample $t$-test for mean
  - Matched pairs $t$-test for mean of difference
  - Independent two sample $t$-test for difference of means
  - Independent two sample $z$-test for difference of proportions

  (ii) define parameter(s) of interest and state the null and alternative hypotheses.

**Example :** To test the claim that Startbucks Coffee's daily sale in Twin Cities is higher than that in Des Moines, IA, you randomly select 30 stores from each location and calculate the sample daily sales mean for each state.
**Answer :** Independent two sample $t$-test for difference of means.
$\mu_1$: mean daily sale of Startbucks coffee's in Twin Cities
$\mu_2$: mean daily sale of Startbucks coffee's in Des Moines, IA
$H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 > \mu_2$

a. Past experience indicates that the time required for high school seniors to complete a standardized test is a normal random variable with a mean of 35 minutes. A researcher wants to test the claim that high school seniors from 2019 requires less amount of time than before. A random sample of 20 high school seniors took 33.1 minutes to complete this test.

b. A marketing expert for a pasta-making company believes that more than 40% of pasta lovers prefer lasagna. To test this claim, the researcher collected a random sample of 20 pasta lovers. 9 out of 20 choose lasagna over other pastas.

c. Forty-four sixth graders were randomly selected from a school district. Then, they were divided into 22 matched pairs, each pair having equal IQ's. One member of each pair was randomly selected to receive special training. Then all of the students were given an IQ test. The researcher wants to test if the special training has any effect.

d. In a winter of an epidemic flu, the parents of 2000 babies were surveyed by researchers at a well-known pharmaceutical company to determine if the company's new medicine was better than their old medicine after two days. Among 120 babies who had the flu and were given the new medicine, 29 were cured within two days. Among 280 babies who had the flu and were given the old medicine, 56 recovered within two days. Test the company's claim.

e. An experiment investigated whether cell phone use impairs drivers' reaction times, using an sample of 64 students from the University of Utah. Students were randomly assigned to a cell phone group or to a control group, 32 to each. On a simulation of driving situations, a target flashed red or green at irregular periods. Participants pressed a brake button as soon as they detected a red light. The control group listened to radio or audio book while they performed the simulated driving. The experiment measured each group's mean response time over many trials.

- **Problem 2** Textbook Exercise 10.4 Smoking and lung obstruction (page 458)
- **Problem 3** Textbook Exercise 10.6 Aspirin and heart attacks in Sweden. Use the provided MINITAB output on page 459.
- **Problem 4** Textbook Exercise 10.7
  Comments on c. Physician Health Study (Examples 2,3,4 from page 446 - 451) used over 10,000 subjects in each group (placebo and Aspirin) and recorded whether each subject died of cancer or not. The study finds the difference between two sample proportion $\hat{p}_{placebo} - \hat{p}_{aspirin} = 0.03 - 0.027 = 0.007$ with $p$-value 0.026

  Additional problems :
  For b. and interpret the $p$-value and draw a conclusion in the context of the problem. Use $\alpha = 0.1$
  For d. and use $\alpha = 0.1$ to draw a conclusion for the new alternative hypothesis.
  e. Report the $P$-value for the one-sided alternative hypothesis that the chance of death due to heart attack is lower for the placebo group.
  f. Verify that $z$-test statistic is 1.46 from MINITAB output on page 459. Due to 'continuity correction', your test statistic value might be slightly different.

- **Problem 5** Textbook Problem 10.29 (Study time, page 472)
  Use the following R command to input data.

```
grad<-c(15, 7, 15, 10, 5, 5, 2, 3, 12, 16, 15, 37, 8, 14, 10, 18, 3, 25, 15, 5, 5)
noGrad<-c(6, 8, 15, 6, 5, 14, 10, 10, 12, 5)
```

  For part b. Hint; Standard error is estimated standard deviation of statistic.
  For part c. Do not use `t.test()`. Select the correct value of multiplier ($z_{\alpha/2}$ or $t_{\alpha/2,df}$) to construct a 95% confidence interval for difference in means. Construct the confidence interval and interpret.

```
> qt(0.95, df=9)
[1] 1.83
> qt(0.975, df=9)
[1] 2.26
> qt(0.95, df=20)
```

```
[1] 1.72
> qt(0.975, df=20)
[1] 2.09
> qnorm(0.95)
[1] 1.645
> qnorm(0.975)
[1] 1.96
```

- **Problem 6** Refer to Problem 5 (Study time).
  a. Show all five steps to test if mean study times are equal or not at $\alpha = 0.1$. Let $\mu_1$ be the mean study time for students who plan to go to graduate school and $\mu_2$ be the mean study time for the other students. Do not use `t.test()`
  Use the R commands below to find the range of $p$-value such as $p$-value$> 0.1$ or $p$-value$< 0.1$

```
> qt(0.05, df=9)
[1] -1.833113
> qt(0.025, df=9)
[1] -2.262157
```

  b. Do you think samples in Problem 4 were random or convenience samples? Why?

R Problem.

- **Problem 7** Refer Problem 5 (Study time). Use `t.test()` to construct 90% CI for the difference of mean study times. Copy and paste R output. Is it wider or narrower than 95% CI from Problem 5 part c.?

```
t.test(x=_______, y=_______, conf.level = ______, alternative ="_____")
```

  Note that R uses 'Welch's degree of freedom' rather than $df$=min$(n_1 - 1, n_2 - 1)$.