Taller 08

Universidad Externado de Colombia
Departamento de Matemáticas
Estadística 2

Juan Sosa, Ph. D.

June 26, 2018

Pruebas de hipótesis sobre la proporción poblacional

1. De Berenson et al. (2012), leer y sintetizar la Sección 9.4 (*Z Test of Hypothesis for the Proportion*, p. 349).

2. *(Berenson et al. 2012, problem 9.73)* An auditor for a government agency is assigned the task of evaluating reimbursement for office visits to physicians paid by Medicare. The audit was conducted on a sample of 75 of the reimbursements, with the following results:

    - In 12 of the office visits, there was an incorrect amount of reimbursement.
    - The amount of reimbursement was $\bar{x} =$ US 93.7, $s =$ US 34.55.

    (a) At the 0.05 level of significance, is there evidence that the population mean reimbursement was less than US 100?

    (b) At the 0.05 level of significance, is there evidence that the proportion of incorrect reimbursements in the population was greater than 0.10?

(c) Discuss the underlying assumptions of the test used in (a).

(d) What is your answer to (a) if the sample mean equals US 90?

(e) What is your answer to (b) if 15 office visits had incorrect reimbursements?

3. *(Anderson et al. 2011, Chap. 9, problem 39)* Accroding to the Pew Internet & American Life Project, 75% of American adults use the Internet (Pew Internet website, April 19, 2008). The Pew project authors also reported on the percentage of Americans who use the Internet by age group. The data in the file AgeGroup are consistent with their findings. These data were obtained from a sample of 100 Internet users in the 30–49 age group and 200 Internet users in the 50–64 age group. A Yes indicates the survey repondent had used the Internet; a No indicates the survey repondent had not. The data set is available in `AgeGroup.csv`.

   (a) Formulate hypotheses that could be used to determine whether the percentage of Internet users in the two age groups differs from the overall average of 75%

   (b) Estimate the proportion of Internet users in the 30–49 age group. Does this proportion differ significantly from the overall proportion of .75? Use $\alpha = .05$.

   (c) Estimate the proportion of Internet users in the 50–64 age group. Does this proportion differ significantly from the overall proportion of .75? Use $\alpha = .05$.

   (d) Would you expect the proportion of users in the 18–29 age group to be larger or smaller than the proportion for the 30–49 age group? Support you conclusion with the results obtained in parts (b) and (c).

Service Times

(Berenson et al. 2012, problem 9.74) A bank branch located in a commercial district of a city had the business objective of improving the process for serving customers during the noon-to-1:00 P.M. lunch period. The waiting time (defined as the time the customer enters the line until he or she reaches the

teller window) of all customers during this hour is recorded over a period of a week. Data were collected from a random sample of 15 customers, and the results are organized (and stored in `Bank1.xls`) as follows:

1. At the 0.05 level of significance, is there evidence that the population mean waiting time is less than 5 minutes?

2. What assumption about the population distribution is needed in order to conduct the t test in (a)?

3. Construct a boxplot, a normal probability plot (histogram), and a qqplot to evaluate the assumption made in (b).

4. Simulate $M = 1,000$ from a Normal distribution with mean $\mu = \bar{x}$ and variance $\sigma^2 = s^2$, where $\bar{x}$ and s^2 are the sample mean and the sample variance of the waiting times, respectively. Plot the corresponding qqplot for each sample in the same plot (in gray), and on top of that, plot the qqplot for the actual sample (in blue) and the qqline (in red).

5. Do you think that the assumption needed in order to conduct the t test in (a) is valid? Explain.

6. Plot the power curve for the test developed in (a).

7. As a customer walks into the branch office during the lunch hour, she asks the branch manager how long she can expect to wait. The branch manager replies, "Almost certainly not longer than 5 minutes." On the basis of the results of (a), evaluate this statement.

Ethical Behavior of Business Students at Bayview Universtiy

(Anderson et al. 2011, Chap. 9, Case Problem 2) During the global recession of 2008 and 2009, there were many accusations of unethical behavior by Wall Street executives, financial managers, and other corporate officers. At that time, an article appeared that suggested that part of the reason for such unethical business behavior may stem from the fact that cheating has

become more prevalent among business students (Chronicle of Higher Education, February 10, 2009). The article reported that 56 percent of business students admitted to cheating at some time during their academic career as compared to 47 percent of nonbusiness students. Cheating has been a concern of the dean of the College of Business at Bayview University for several years. Some faculty members in the college believe that cheating is more widespread at Bayview than at other universities, while other faculty members think that cheating is not a major problem in the college. To resolve some of these issues, the dean commissioned a study to assess the current ethical behavior of business students at Bayview. As part of this study, an anonymous exit survey was administered to a sample of 90 business students from this year's graduating class. Responses to the following questions were used to obtain data regarding three types of cheating.

- During your time at Bayview, did you ever present work copied off the Internet as your own?

- During your time at Bayview, did you ever copy answers off another student's exam?

- During your time at Bayview, did you ever collaborate with other students on projects that were supposed to be completed individually?

Any student who answered Yes to one or more of these questions was considered to have been involved in some type of cheating. The complete data set is in the file named `BAYVIEW.csv`.

Prepare a report for the dean of the college that summarizes your assessment of the nature of cheating by business students at Bayview University. Be sure to include the following items in your report.

1. Use descriptive statistics to summarize the data and comment on your findings.

2. Develop 95% confidence intervals for the proportion of all students, the proportion of male students, and the proportion of female students who were involved in some type of cheating.

3. Conduct a hypothesis test to determine if the proportion of business students at Bayview University who were involved in some type of cheating

is less than that of business students at other institutions as reported by the Chronicle of Higher Eudcation.

4. Plot the power curve for the previous test.

5. Conduct a hypothesis test to determine if the proportion of business students at Bayview University who were involved in some form of cheating is less than that of nonbusiness students at other institutions as reported by the Chronicle of Higher Education.

6. What advice would you give to the dean based upon your analysis of the data?