Taller 07

Universidad Externado de Colombia
Departamento de Matemáticas
Estadística 2

Juan Sosa, Ph. D.

June 25, 2018

Tamaños de muestra

1. De Berenson et al. (2014), leer y sintetizar la Sección 8.5 (*Determining Sample Size*, p. 297).

2. *(Anderson et al. 2011, Chap. 8, problem 60)* Although airline schedules and cost are important factors for business travelers when choosing an airline carrier, a USA Today survey found that business travelers list an airline's frequent flyer program as the most important factor. From a sample of $n = 1993$ business travelers who responded to the survey, 618 listed a frequent flyer program as the most important factor.

   (a) What is the point estimate of the proportion of the population of business travelers who believe a frequent flyer program is the most important factor when choosing an airline carrier?

   (b) Develop a 95% confidence interval estimate of the population proportion.

   (c) How large a sample would be required to report the margin of error of 0.01 at 95% confidence? Would you recommend that USA Today attempt to provide this degree of precision? Why or why not?

3. *(Berenson et al. 2012, problem 8.38)* A survey is planned to determine the mean annual family medical expenses of employees of a large company. The management of the company wishes to be 95% confident that the sample mean is correct to within of the population mean annual family medical expenses. A previous study indicates that the standard deviation is approximately $400.

   (a) How large a sample is necessary?

   (b) If management wants to be correct to within how many employees need to be selected?

Pruebas de hipótesis

1. De Thijseen (2016), leer y sintetizar la Sección 4.3 (*Hypothesis Tests*, p. 51).

2. *(Keller et al. 2011, problem 12.42)* Companies that sell groceries over the Internet are called e-grocers. Customers enter their orders, pay by credit card, and receive delivery by truck. A potential e-grocer analyzed the market and determined that the average order would have to exceed $85 if the e-grocer were to be profitable. To determine whether an e-grocery would be profitable in one large city, she offered the service and recorded the size of the order for a random sample of customers. Can we infer from these data that an e-grocery will be profitable in this city? The data set is avaible in `Xr12-42.xlsx`.

3. *(Anderson et al. 2011, Chap. 9, problem 32)* According to the National Automobile Dealers Association, the mean price for used cars is $10,192. A manager of a Kansas City used car dealership reviewed a sample of 50 recent used car sales at the dealership in an attempt to determine whether the population mean price for used cars at this particular dealership differed from the national mean. The prices for the sample of 50 cars are shown in the file named Used- Cars. The data set is available in `UsedCars.csv`.

   (a) Formulate the hypotheses that can be used to determine whether a difference exists in the mean price for used cars at the dealership.

   (b) What is the p-value?

    (c) At $\alpha = .05$, what is your conclusion?

4. *(Anderson et al. 2011, Sec. 9.8, problem 53)* Sparr Investments, Inc., specializes in tax-deferred investment opportunities for its clients. Recently Sparr offered a payroll deduction investment program for the employees of a particular company. Sparr estimates that the employees are currently averaging $100 or less per month in tax-deferred investments. A sample of 40 employees will be used to test Sparr's hypothesis about the current level of investment activity among the population of employees. Assume the employee monthly tax-deferred investment amounts have a standard deviation of $75 and that a 0.05 level of significance will be used in the hypothesis test.

    (a) What is the Type II error in this situation?

    (b) What is the probability of the Type II error if the actual mean employee monthly investment is $120?

    (c) What is the probability of the Type II error if the actual mean employee monthly investment is $130?

    (d) Assume a sample size of 80 employees is used and repeat parts (b) and (c).

Quality Associates, Inc.

(Anderson et al. 2011, Chap. 9, Case Problem 1) Quality Associates, Inc., a consulting firm, advises its clients about sampling and statistical procedures that can be used to control their manufacturing processes. In one particular application, a client gave Quality Associates a sample of 800 observations taken during a time in which that client's process was operating satisfactorily. The sample standard deviation for these data was .21; hence, with so much data, the population standard deviation was assumed to be .21. Quality Associates then suggested that random samples of size 30 be taken periodically to monitor the process on an ongoing basis. By analyzing the new samples, the client could quickly learn whether the process was operating satisfactorily. When the process was not operating satisfactorily, corrective action could be taken to eliminate the problem. The design specification indicated the mean for the process should be 12. The hypothesis test suggested by Quality Associates

follows.

$$H_0 : \mu = 12 \qquad \text{vs.} \qquad H_0 : \mu \neq 12$$

Corrective action will be taken any time H_0 is rejected.

The samples given in `Quality.csv` were collected at hourly intervals during the first day of operation of the new statistical process control procedure.

1. Conduct a hypothesis test for each sample at the .01 level of significance and determine what action, if any, should be taken. Provide the test statistic and p-value for each test.

2. Compute the standard deviation for each of the four samples. Does the assumption of .21 for the population standard deviation appear reasonable?

3. Compute limits for the sample mean around $\mu = 12$ such that, as long as a new sample mean is within those limits, the process will be considered to be operating satisfactorily. If exceeds the upper limit or if is below the lower limit, corrective action will be taken. These limits are referred to as upper and lower control limits for quality control purposes.

4. Discuss the implications of changing the level of significance to a larger value. What mistake or error could increase if the level of significance is increased?