

STATISTICS

Estimation

Warsaw School of Economics
Winter Semester 2018

Exercise 1 – Estimation of mean

A market research company carried out a survey to estimate the acceptable price of a new energy drink. The potential clients who tried a sample were asked about the price. Calculate the confidence intervals (CI) in the following situations:

- 1) Let's assume that the acceptable price of the drink in the population has normal distribution with standard deviation 1,50. The mean from the sample of 15 persons amounted to 3,20. Estimate the 90% CI and interpret the result. What is the sampling error (SE)?
- 2) Let's assume that the acceptable price of the drink in the population has normal distribution but we don't know the parameters in the population. Estimate the 95% CI for the average price offered by potential clients if sample size amounted to 15 persons mean 3,20 and standard deviation amounted to 1,50. Interpret the result.
- 3) Suppose that the distribution in the population is unknown but company has collected data from 100 experiments and statistics from the sample were following: mean: 3,20 and standard deviation 1,50. Estimate and interpret the 99% CI.

Exercise 1 – continued

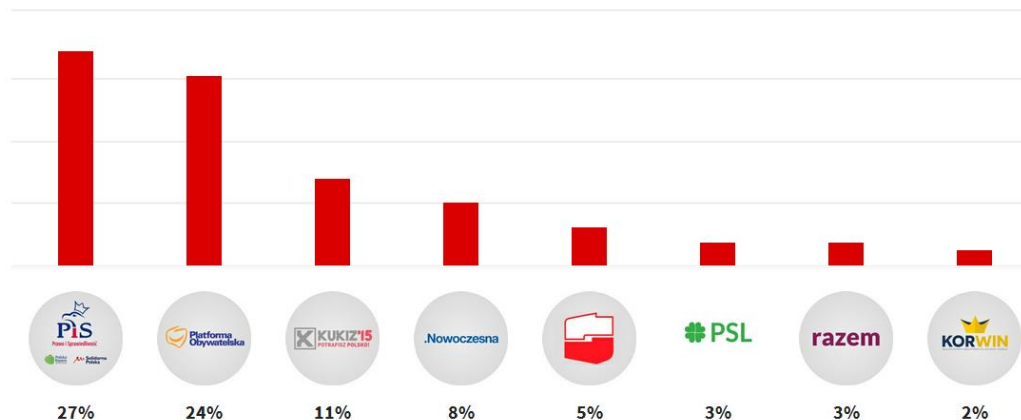
Consider the situation in point 3):

- 4) How big is sampling error (SE) in point 3). How many persons should be asked to increase the precision of estimation twice (to reduce SE by half) ?
- 5) Suppose that the market research company want to compare the results of this survey with the results of the previous survey (assessment of older version of the drink without special ingredient). Let's assume in both cases the accepted price of the drink in the population of young people have normal distributions. The results of previous sample are as follows: $n=120$, average=2,95, standard deviation = 1,10. The results of current sample: $n=100$, average =3,1, standard deviation= 1,50 Please estimate the 95% confidence interval for the difference between mean accepted price of the new drink and mean accepted price of an older version.

Exercise 2 – Estimation of proportion

According to the latest polls (TNS Polska) the „Law & Justice - PiS” party is supported by 27% of respondents. The survey was carried out on the sample of 1025 persons.

- 1) Calculate the standard error of estimation and the sampling error. Estimate 95% CI for proportion of PiS in the population
- 2) In the same polls the second party (Civic Platform) was supported by 24% of respondents. Estimate the 95% CI for this proportion. What is the conclusion for the comparison of the support of two biggest parties?
- 3) How many persons should be asked to reduce the sampling error of 95% CI for PiS to 1 percentage point. Assuming that the unit cost of one interview amounts to 20PLN what is the additional cost of improving precision of the estimation.



Exercise 3

Tax-exempt charities. Donations to tax-exempt organizations such as the Red Cross, the Salvation Army, the YMCA, and the American Cancer Society not only go to the stated charitable purpose, but are also used to cover fund-raising expenses and overhead. For a sample of 30 charities, the next table lists their *charitable commitment*, the percentage of their expenses that go toward the stated charitable purpose.

- a. Give a point estimate for the mean charitable commitment of tax-exempt organizations.
- b. Locate a 95% confidence interval for the true mean charitable commitment of tax-exempt organizations on the printout. Interpret the result.
- c. Why is the confidence interval of part b a better estimator of the mean charitable commitment than the point estimator of part a? Explain.

⇒ see file CHARITY.xls

Exercise 4

Problem Unoccupied seats on flights cause airlines to lose revenue. Suppose a large airline wants to estimate its average number of unoccupied seats per flight over the past year. To accomplish this, the records of 225 flights are randomly selected, and the number of unoccupied seats is noted for each of the sampled flights. Descriptive Statistics for the data are displayed in the MINITAB Printout.

Estimate μ , the mean number of unoccupied seats per flight during the past year, using a 90% confidence interval.

Variable	N	Mean	StDev	SE Mean
NOSHOWS	225	11.5956	4.1026	0.2735

Exercise 5

Hospital length of stay. Health insurers and the federal government are both putting pressure on hospitals to shorten the average length of stay (LOS) of their patients. The average LOS for men in the United States is 5.3 days and the average for women is 4.6 days (*Statistical Abstract of the United States: 2005*). A random sample of 20 hospitals in one state had a mean LOS for women of 3.8 days and a standard deviation of 1.2 days.

- a. Use a 90% confidence interval to estimate the population mean LOS for women for the state's hospitals.
- b. Interpret the interval in terms of this application.
- c. What is meant by the phrase “90% confidence interval”?

Exercise 6

Interviewing candidates for a job. The costs associated with conducting interviews for a job opening have skyrocketed over the years. According to a Harris Interactive survey, 211 of 502 senior human resources executives at U.S. companies believe that their hiring managers are interviewing too many people to find qualified candidates for the job (*Business Wire*, June 8, 2006).

- a. Describe the population of interest in this study.
- b. Identify the population parameter of interest, p .
- c. Find and interpret an interval estimate for the true proportion of senior human resources executives who believe that their hiring managers interview too many candidates during a job search. Use a confidence level of 98%.
- d. If you had constructed a 90% confidence interval, would it be wider or narrower?

Exercise 7

Accuracy of price scanners at Wal-Mart. The National Institute for Standards and Technology (NIST) mandates that for every 100 items scanned through the electronic checkout scanner at a retail store, no more than 2 should have an inaccurate price. A recent study of the accuracy of checkout scanners at Wal-Mart stores in California was conducted (*Tampa Tribune*, Nov. 22, 2005). At each of 60 randomly selected Wal-Mart stores, 100 random items were scanned. The researchers found that 52 of the 60 stores had more than 2 items that were inaccurately priced.

- a. Give an estimate of p , the proportion of Wal-Mart stores in California that have more than 2 inaccurately priced items per 100 items scanned.
- b. Construct a 95% confidence interval for p .
- c. Give a practical interpretation of the interval, part b.
- d. Suppose a Wal-Mart spokesperson claims that 99% of California Wal-Mart stores are in compliance with the NIST mandate on accuracy of price scanners. Comment on the believability of this claim.

Exercise 8

Security of information submitted over the Internet. As Internet usage proliferates, so do questions of security and confidentiality of personal information, including such things as social security and credit card numbers. NCR Corporation surveyed 1,000 U.S. adults and asked them under what circumstances they would give personal information to a company. Twenty-nine percent said they would never give personal data to a company, while 51% said they would if the company had strict privacy guidelines in place (*Precision Marketing*, Oct. 4, 1999).

Construct a 95% confidence interval for p and interpret your result in the context of the problem.

Exercise 9

IRS answers to gift tax questions. According to estimates made by the General Accounting Office, the Internal Revenue Service (IRS) answered 18.3 million telephone inquiries during a recent tax season, and 17% of the IRS offices provided answers that were wrong. These estimates were based on data collected from sample calls to numerous IRS offices. How many IRS offices should be randomly selected and contacted in order to estimate the proportion of IRS offices that fail to correctly answer questions about gift taxes with a 90% confidence interval of width .06?

Exercise 10

Sick leave taken by employees. A company is interested in estimating μ , the mean number of days of sick leave taken by all its employees. The firm's statistician selects at random 100 personnel files and notes the number of sick days taken by each employee. The following sample statistics are computed: $\bar{x} = 12.2$ days, $s = 10$ days.

- a. Estimate μ using a 90% confidence interval. Interpret the result.
- b. How many personnel files would the statistician have to select in order to estimate μ to within 2 days with a 99% confidence interval?

Exercise 11

Preventing production of defective items. It costs more to produce defective items—since they must be scrapped or reworked—than it does to produce nondefective items. This simple fact suggests that manufacturers should ensure the quality of their products by perfecting their production processes rather than through inspection of finished products (Deming, 1986). In order to better understand a particular metal-stamping process, a manufacturer wishes to estimate the mean length of items produced by the process during the past 24 hours.

- a. How many parts should be sampled in order to estimate the population mean to within .1 millimeter (mm) with 90% confidence? Previous studies of this machine have indicated that the standard deviation of lengths produced by the stamping operation is about 2 mm.

Exercise 11, continued

- b. Time permits the use of a sample size no larger than 100. If a 90% confidence interval for μ is constructed using $n = 100$, will it be wider or narrower than would have been obtained using the sample size determined in part a? Explain.
- c. If management requires that μ be estimated to within .1 mm and that a sample size of no more than 100 be used, what is (approximately) the maximum confidence level that could be attained for a confidence interval that meets management's specifications?