STUDY SET 4

Confidence Interval Estimation : One Population

1. The number of bolts produced each hour from a particular machine is normally distributed with a standard deviation of 7.4. For a random sample of 15 hours, the average number of bolts produced was 587.3. Find the upper and lower confidence limits of a 98% confidence interval for the population mean number of bolts produced per hour.

   ANSWER:
   UCL = 591.75 and LCL = 582.85

2. In a recent survey of 600 adults, 16.4% indicated that they had fallen asleep in front of the television in the past month. Develop a 95% confidence interval for the population proportion.

   ANSWER:
   UCL = 0.1936 and LCL = 0.1344.

3. In a recent survey of 574 employees, 15.5% indicated that they were not in favor of a particular plan. What is the level of confidence is associated with the interval of 12.85% to 19.89%?

   ANSWER:
   The confidence level is 98%.

4. A student records the time (in minutes) it takes to commute to school for seven days. Those results are: 21, 15, 13, 16, 10, 13, and 18. Assuming the population is normally distributed, develop a 95% confidence interval for the population mean.

   ANSWER:
   UCL=18.496 and LCL=11.79

5. You are told that a 95% confidence interval for the population mean is 17.3 to 24.5. If the population standard deviation is 18.2, how large was the sample?

   ANSWER:
   99

6. The amount of material used in making a custom sail for a sailboat is normally distributed. For a random sample of 15 sails, you find that the mean amount of material is 912 square feet, with a standard deviation of 64 square feet. Develop a 99% confidence interval for the population mean amount of material used in a custom sail.

ANSWER:
912 ± 49.2 or $862.8 < \mu < 961.2$

7. You are interested in determining the amount of time (in minutes) you spend each day on the Internet. For seven days, these values are: 51, 24, 16, 88, 63, 28, and 59. Assume that the amount of time you spend on the Internet each day is normally distributed, and develop a 90% confidence interval for the population average amount of time.

ANSWER:
$28.16 < \mu < 65.84$

8. A mother who is interested in the true proportion of R-rated movies shown on pay TV by a cable system randomly selects 98 listings and finds 14 of them are R-rated movies. In her report to the subcommittee she wants to be 98% confident that the true proportion will be in an interval which she states. She has asked you to assist her by preparing a 98% confidence interval based on the data she collected. What should she report?

ANSWER:
$0.0605 < P < 0.2253.$

THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
A Monte Carlo study involves 10,000 random samples of size 20 from a normal population with mean $\mu = 120$ and standard deviation $\sigma = 20.$ For each sample, the mean and the median are calculated, with the following results:

Estimator	*Mean*	*Median*
Average	120.2	119.96
Variance	25.52	67.89

9. What does the study suggest about the bias of the estimators in this situation?

ANSWER:
The two estimators—the mean and the median—appear to be unbiased. The averages are all very close to 120. This result should hold because the simulation uses a symmetric (normal) population.

10. Which of the two estimators appears most efficient?

ANSWER:

The mean has the smallest variance, and therefore will have the smallest standard error. The sample mean appears most efficient. This result follows a normal population. It may not hold for other populations.

THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
Let X_1, X_2, X_3, and X_4 be a random sample of observations from a population with mean μ and variance σ^2. Consider the following two point estimators of μ:
$\hat{\theta}_1$ = 0.10 X_1 + 0.20 X_2 + 0.40 X_3 + 0.30 X_4, and
$\hat{\theta}_2$ = 0.25 X_1 + 0.25 X_2 + 0.30 X_3 + 0.20 X_4

11. Show that both the estimators are unbiased.

ANSWER:
Since $E\left(\hat{\theta}_1\right) = E\left(\hat{\theta}_2\right) = \mu$, $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased estimators of μ.

12. Which estimator is more efficient, $\hat{\theta}_1$ or $\hat{\theta}_2$? Explain in detail.

ANSWER:
Since Var($\hat{\theta}_2$) < Var($\hat{\theta}_1$), then $\hat{\theta}_2$ is more efficient than $\hat{\theta}_1$.

13. Find the relative efficiency of $\hat{\theta}_2$ with respect to $\hat{\theta}_1$.

ANSWER:
1.176

14. Raising in-state and out-of-state tuition is supposed to reduce the number of students in state supported universities. The registrar of a university wants to estimate the proportion *P* of students who are paying for out-of-state tuition on the installment plan (to later be compared with in-state installment plan payers). A random sample of 80 students who live out-of-state is taken and 50 of them pay tuition on the installment plan. Find a 99-percent confidence interval for *P*, based on these data.

ANSWER:
$0.4854 < P < 0.7646.$

THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
The Daytona Beach Tourism Commission is interested in the average amount of money a typical college student spends per day during spring break. They survey 35 students and find that the mean spending is $63.57 with a standard deviation of $17.32.

15. Develop a 95% confidence interval for the population mean daily spending.

ANSWER:

UCL = 69.548 and LCL = 57.592.

16. Interpret the confidence level in the previous question.

ANSWER:
If independent random samples of size 35 are repeatedly selected from the population and 95% confidence intervals for each of these samples are determined, then over a very large number of repeated trials, 955 of these intervals will contain the value of the true average amount of money a typical college student spends per day during spring break.

17. What level of confidence is associated with an interval of \$58.62 to \$68.52 for the population mean daily spending?

ANSWER:
90%.

THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
A researcher is interested in determining the percentage of all households in the U.S. that have more than one home computer. In a survey of 492 households, 27% indicated that they own more than one home computer.

18. Develop a 90% confidence interval for the proportion of all households in the U.S. with more than one computer.

ANSWER:
UCL= 0.303 and LCL = 0.237

19. The researcher reports a confidence interval of 0.2312 to 0.3096 but neglects to tell you the confidence level. What is the confidence level associated with this interval?

ANSWER:
95%.

THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
A computer is programmed to draw 1000 samples, each of size 40 from a normally distributed population having mean 50 and standard deviation 10. For each sample, the mean and the median are computed. The average value and standard deviation of each set of estimates for the 1000 samples are as follows.

Statistic	Average Value	Standard Deviation
Mean	50.1278	2.2132
Median	50.2109	2.8263

20. Do the two statistics appear to be unbiased?

ANSWER:
The average value of each estimator is very close to the population mean 50, so both estimators appear to be unbiased.

21. Which statistic appears to be more efficient?

ANSWER:
The sample mean has the smallest standard error, so it appears to be most efficient in this situation.

<u>THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
Suppose that the amount of time teenagers spend on the internet is normally distributed with a standard deviation of 1.5 hours. A sample of 100 teenagers is selected at random, and the sample mean is computed as 6.5 hours.

22. Determine the 95% confidence interval estimate of the population mean.

ANSWER:
$6.206 < \mu < 6.794$

23. Interpret what the interval estimate in the previous question tells you.

ANSWER:
If we repeatedly draw independent random samples of size 100 from the population of teenagers and confidence intervals for each of these samples are determined, then over a very large number of repeated trials, 95% of these confidence intervals will contain the value of the true population mean amount of time teenagers spend on the internet.

<u>THE NEXT FOUR QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
The data shown below specify how much a sample of 20 executives paid in federal income taxes, as a percentage of gross income, are reproduced below.

18.0	20.1	20.6	22.2	23.7	24.4	24.4	25.1	25.2	25.5
26.1	26.3	26.7	27.2	27.9	28.3	29.9	30.0	32.4	35.7

Assume that the standard deviation for the underlying population is equal to 4.0.

24. Calculate a 95% confidence interval for the population mean.

ANSWER:
$24.232 < \mu < 27.738$.

25. Calculate a 99% confidence interval for the population mean.

ANSWER:
$23.681 < \mu < 28.289$.

26. Give a careful verbal interpretation of the confidence interval in the previous question.

ANSWER:
We would expect that 99% of all confidence intervals calculated in this manner would include the true value of the parameter μ. This means that if we sampled again and again (20 persons each time) and calculated a sample mean and a 99% confidence interval for the true mean each time, we would expect that 99% of them would include the population mean μ.

27. Closed caption movies allow the hearing impaired to enjoy the dialogue as well as the acting. A local organization for the hearing impaired people of the community takes a random sample of 100 movie listings offered by the cable television company in order to estimate the proportion of closed caption movies offered. They observed that 14 movies were closed caption. The cable television company says at least 5% of the movies shown are closed captioned. Prepare a 90% confidence interval for the true proportion *P* and comment on the cable television company's claim.

ANSWER:
$0.083 \leq P \leq 0.197$.
The organization for hearing impaired people can be 90% confident that the proportion of closed caption movies offered is somewhere between 0.083 (8.3%) and 0.197 (19.7%). The cable television company could therefore conclude that at least 5% of the movies it shows are closed captioned.

<u>THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
The sales manager for a hardware wholesaler finds that 229 of the previous 500 calls to hardware store owners resulted in new product placements. Assume that the 500 calls represent a random sample.

28. Find a 95% confidence interval for the long–run proportion of new product placements.

ANSWER:
$0.414 < P < 0.502$.

29. Give a careful verbal interpretation of the confidence interval in the previous question.

ANSWER:

We are 95% confident that the true long run proportion of new product placements is in the interval (0.414, 0.502). By 95% confident we mean that if this experiment is conducted several times and a confidence interval is calculated for each trial, we expect 95% of them to include the true long run proportion.

30. Find the confidence interval for estimating the population proportion for 90% confidence level; sample size n = 675; and sample proportion $\hat{p}$ =0.10.

ANSWER:
$0.081 < P < 0.119$

THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWINGINFORMATION:
A furniture mover calculates the actual weight as a proportion of estimated weight for a sample of 31 recent jobs. The sample mean is 1.13 and the sample standard deviation is 0.16.

31. Calculate a 95% confidence interval for the population mean using t tables.

ANSWER:
$1.071 < \mu < 1.189.$

32. Assume that the population standard deviation is known to be 0.16. Calculate a 95% confidence interval for the population mean using the z- table.

ANSWER:
$1.074 < \mu < 1.186.$

33. Are the intervals calculated in the previous two questions of roughly similar size? Explain.

ANSWER:
The two intervals are almost the same. The artificial (and probably incorrect) assumption that σ is equal to s allows us to quote a slightly narrower interval.

THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
A regional CPA firm conducted an audit for a discount chain. One part of the audit involved developing an estimate for the mean dollar error in total charges that occur during the checkout process. They wish to develop a 90% confidence interval estimate for the population mean. A simple random sample of n = 20 is selected, with the following data (in dollars):

0.00	1.20	0.43	1.00	1.47	0.83	0.50	3.34	1.58	1.46
-0.36	-1.10	2.60	0.00	0.00	-1.70	0.83	1.99	0.00	1.34

34. Calculate the sample mean.

ANSWER:
$\bar{x} = \sum x / n = \$15.41 / 20 = \$0.77$

35. Calculate the sample standard deviation.

ANSWER:
$s = \sqrt{\sum (x - \bar{x})^2 / (n-1)} = \sqrt{26.906 / 19} = \$1.19$

36. Develop a 90% confidence interval estimate for the population mean.

ANSWER:
$\$0.31 < \mu < \$1.23$

37. In a recent survey of personnel directors, 71% thought that they should hire new personnel over the next three months. The researcher conducting the survey reported that the 99% confidence interval for the proportion of all personnel directors planning to hire personnel over the next three months was from 0.68 to 0.74. What is the sample size taken by the researcher?

ANSWER:
1523

38. The number of television sets coming off a production line each day is known to have a standard deviation of 118.5 sets per day. The production manager tells you that the 90% confidence interval for the population mean was 552.3 to 621.9. How large a sample was this confidence interval based on?

ANSWER:
32

39. The number of television sets coming off a production line each day is known to have a standard deviation of 17.4 sets per day. The production line averaged 852.3 sets per day for 20 randomly selected days. What is the level of confidence associated with the interval 844.75 to 860.0?

ANSWER:
95%.

40. There is concern about the speed of automobiles traveling over US 131. For a random sample of seven automobiles radar indicated the following speeds, in miles per hour: 80, 74, 69, 78, 87, 72, and 70. Assuming a normal population distribution, find the margin of error of a 95% confidence interval for the mean speed of all automobiles traveling over this stretch of highway.

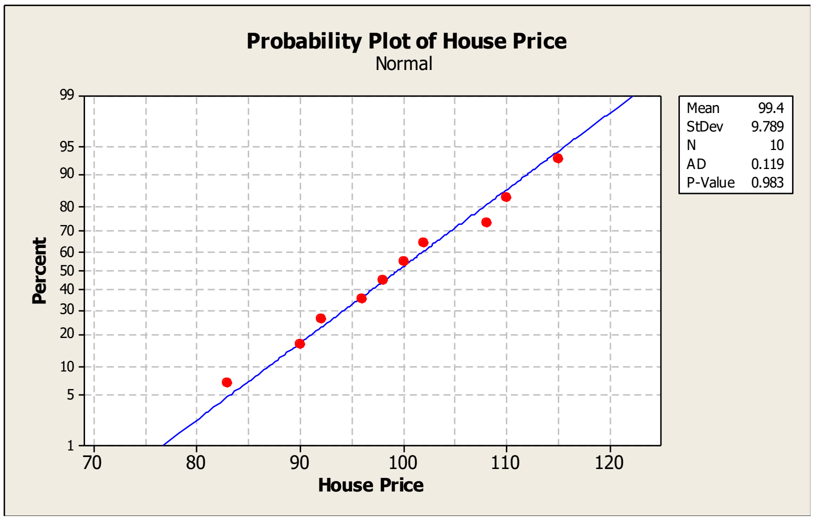
ANSWER:

Margin of error: ± 5.9152 miles

THE NEXT FOUR QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
A random sample of ten homes in a particular suburb had the following selling prices (in thousands of dollars): 92, 83, 110, 115, 108, 96, 102, 90, 100, and 98.

41. Check for evidence of nonnormality.

ANSWER:



There is no evidence of nonnormality.

42. Find a point estimate of the population mean that is unbiased and efficient.

ANSWER:
The minimum variance unbiased point estimator of the population mean is the sample mean, where the value of the sample mean is $\bar{x} = \sum x_i / n = 994 / 10 = 99.4$

43. Use an unbiased estimation procedure to find a point estimate of the variance of the sample mean.

ANSWER:
9.582

44. Use an unbiased estimator to estimate the proportion of homes in this suburb selling for less than $95,000.

ANSWER:
0.30

THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:

Suppose that x_1 and x_2 are random samples of observations from a population with mean μ and variance σ^2. Consider the following three point estimators, *X*, *Y*, and *Z*, of μ: $X = (x_1 + x_2)/2$, $Y = (x_1 + 3x_2)/4$, and $Z = (x_1 + 2x_2)/3$.

45. Show that all three estimators *X*, *Y*, and *Z* are unbiased.

ANSWER:
All three estimators are unbiased.

46. Which of the estimators *X*, *Y*, and *Z* is the most efficient?

ANSWER:
X is the most efficient estimator since $Var(X) < Var(Z) < Var(Y)$.
Note that $X = (X_1 + X_2)/2 = \bar{X}$, then $Var(X) = Var(\bar{X}) = \sigma^2/n = \sigma^2/2$.

47. Find the relative efficiency of *X* with respect to each of the other two estimators *Y* and *Z*.

ANSWER:
Relative efficiency of *X* with respect to *Y* = 1.25

Relative efficiency of *X* with respect to *Z* = 1.11

THE NEXT FOUR QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:
A process producing bricks is known produce bricks whose weights are normally distributed with a standard deviation of 0.11 lb. A random sample of 16 bricks from today's output had a mean weight of 4.08 lb.

48. Find a 99% confidence interval for the mean weight of all bricks produced this day.

ANSWER:
$4.0091 < \mu < 4.1509$

49. Without doing the calculations, explain whether a 95% confidence interval for the population mean would be wider than, narrower than, or the same width as that found in the first question.

ANSWER:
A 95% confidence interval for the population mean would be narrower than that found in the first question since the z score for a 95% confidence interval (1.96) is smaller than the z score for the 99% confidence interval (2.576).

50. It is decided that tomorrow a sample of 20 bricks will be taken. Without doing the calculations, explain whether a correctly calculated 95% confidence interval for

the mean weight of tomorrow's output would be wider than, narrower than, or the same width as that found in the first question.

ANSWER:
A correctly calculated 95% confidence interval for the population mean would be narrower than that found in the first question due to the smaller standard error.

51. Suppose that the population standard deviation for today's output is 0.14 lb (not 0.11 lb). Without doing the calculations, explain whether a correctly calculated 99% confidence interval for the mean weight of today's output would be wider than, narrower than, or the same width as that found in the first question.

ANSWER:
A correctly calculated 99% confidence interval for the population mean would be wider than that found in the first question due to the larger standard error.

<u>THE NEXT TWO QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
From a random sample of 500 registered voters in Los Angeles, 400 indicated that they would vote in favor of a proposed policy in an upcoming election.

52. Calculate a 98% confidence interval estimate for the population proportion in favor of this policy.

ANSWER:
$0.7583 < P < 0.8417$

53. Calculate the width of the 90% confidence interval estimate for the population proportion in favor of this policy.

ANSWER: 0.0589

<u>THE NEXT THREE QUESTIONS ARE BASED ON THE FOLLOWING INFORMATION:</u>
A clinic offers a weight-reduction program. A review of its records revealed the following weight losses, in pounds, for a random sample of 10 of its clients at the conclusion of the program: 20, 15, 21, 28, 22, 19, 9, 14, 18, and 23.

54. Find a point estimate of the population mean and population standard deviation of the weight losses.

ANSWER:
5.301

55. Find a 95% confidence interval for the population mean.

ANSWER:
$15.11 < \mu < 22.69$

56. Without doing any calculations, explain whether a 90% confidence interval for the population mean would be wider than, narrower than, or the same as that found in the second question.

ANSWER:
A 90% confidence interval for the population mean would be narrower than that found in the second question since the t - score for a 90% confidence interval with v = 9 (1.833) is smaller than the t -score for the 95% confidence interval in the previous question (2.262).

57. To calculate a required sample size to estimate a population proportion, given a desired confidence interval and margin of error, the sample proportion is required but often unknown before the sample is collected. How is this predicament resolved?

ANSWER:
Without any additional information, the most conservative approach is to assume that the sample proportion equals 0.50. Using this value, you are assured that the confidence interval and margin of error requirements are met, since this will provide the largest possible sample size. This is because $\hat{p}(1-\hat{p})$ cannot be larger than 0.25, which is its value when the sample proportion is 0.50.

58. How large the sample is needed to estimate the population proportion if ME = 0.09, α = 0.05?

ANSWER:
119