CHAPTER 5

SAMPLING AND ESTIMATION

1. (B) unbiased and efficient.

Explanation

The estimator is unbiased because the expected value of the sample mean is equal to the population mean. The estimator is efficient because the variance of the sampling distribution is smaller than that for other estimators of the parameter. The estimator is not consistent. To be consistent, as the sample size increases, the standard error of the sample mean must decrease.

Related Material

SchweserNotes - Book 1

2. (A) approaches a normal distribution.

Explanation

As *n* gets larger, the variance of the distribution of sample means is reduced, and the distribution of sample means approximates a normal distribution.

(Study Session 2, module 5.1, LOS 5.d)

Related Material

SchweserNotes - Book 1

3. (A) data snooping bias.

Explanation

Data snooping bias can result when the same data is used with different methods until the desired results are obtained.

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

4. (C) look-ahead bias.

Explanation

Look-ahead bias occurs when a study examines an effect based on information that was not yet available at the time being tested. In this case, year-end book values per share are not known until well into the first quarter of the following year.

Time-period bias is present when a study covers either too short a period (the proposed relationship may only hold during that time frame) or too long a period (the proposed relationship may have changed during that span). Sample selection bias refers to taking a sample that is not representative of the population being studied.

Related Material

SchweserNotes - Book 1

5. (A) **Test the strategy on a different data set than the one used to develop the rules.**

Explanation

Data snooping bias occurs when the analyst repeatedly "drills" the dataset until something statistically significant is found.

The *best way* to avoid data snooping is to test a potentially profitable trading rule on a data set different than the one you used to develop the rule (out-of-sample data). Neither a larger sample size nor a data set free of survivorship bias will prevent data snooping.

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

6. (A) **0.85 to 5.15.**

Explanation

The standard error of the sample is the standard deviation divided by the square root of n, the sample size. 6% / $30^{1/2}$ = 1.0954%.

The confidence interval = point estimate +/- (reliability factor x standard error) confidence interval = 3 +/- (1.96 x 1.0954) = 0.85 to 5.15

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

7. (B) **the difference between a sample statistic and its corresponding population parameter.**

Explanation

This is the definition.

(Study Session 2, module 5.1, LOS 5.b)

Related Material

SchweserNotes - Book 1

8. (A) **expands as the probability that a point estimate falls within the interval decreases.**

Explanation

A confidence interval contracts as the probability that a point estimate falls within the interval decreases.

(Study Session 2, module 5.2, LOS 5.g)

Related Material

SchweserNotes - Book 1

9. (A) 0.4.

Explanation

The standard error of the sample mean is estimated by dividing the standard deviation of the sample by the square root of the sample size. The standard deviation of the sample is calculated by taking the positive square root of the sample variance $4^{1/2} = 2$. Applying the formula: $s_X = s/n^{1/2} = 2 / (25)^{1/2} = 2 / 5 = 0.4$.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

10. (C) $60 \pm 1.645(2)$.

Explanation

Because we can compute the population standard deviation, we use the z-statistic. A 90% confidence level is constructed by taking the population mean and adding and subtracting the product of the z-statistic reliability ($z_{a/2}$) factor times the known standard deviation of the population divided by the square root of the sample size (note that the population variance is given and its positive square root is the standard deviation of the population): $x \pm z_{\alpha/2} \times (\sigma / n^{1/2}) = 60 +/- 1.645 \times (100^{1/2}/25^{1/2}) = 60 +/- 1.645 \times (10/5) = 60 +/- 1.645 \times 2$.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

11. (C) data snooping.

Explanation

Data snooping occurs when the analyst continually uses the same database to search for patterns or trading rules until he finds one that *works.* If you are reading research that suggests a profitable trading strategy, make sure you heed the following warning signs of data snooping:

Evidence that the author used many variables (most unreported) until he found ones that were significant.

The lack of any economic theory that is consistent with the empirical results. (Study Session 2, Module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

12. (A) 73.63 to 78.37 hours.

Explanation

The confidence interval is equal to 76 + or - $(2.056)(6/\sqrt{27})$ = 73.63 to 78.37 hours. Because the sample size is small, we use the t-distribution with (27 - 1) degrees of freedom.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

13. (C) confidence interval.

Explanation

A confidence interval is a range of values within which the actual value of a parameter will lie, given a specified probability level. A point estimate is a single value used to estimate a population parameter. An example of a point estimate is a sample mean. The degree of confidence is the confidence level associated with a confidence interval and is computed as 1 - α.

(Study Session 2, module 5.2, LOS 5.g)

Related Material

SchweserNotes - Book 1

14. (C) 16.62.

Explanation

According to the central limit theorem, the mean of the distribution of sample means will be equal to the population mean. n > 30 is only required for distributions of sample means to approach normal distribution.

(Study Session 2, module 5.1, LOS 5.d)

Related Material

SchweserNotes - Book 1

15. (A) -1.584% to 15.584%.

Explanation

The standard error for the mean = s / $(n)^{0.5}$ = 25% / $(60)^{0.5}$ = 3.227%. The critical value from the t-table should be based on 60 - 1 = 59 df. Since the standard tables do not provide the critical value for 59 df the closest available value is for 60 df. This leaves us with an approximate confidence interval. Based on 99% confidence and df = 60, the critical t-value is 2.660. Therefore the 99% confidence interval is approximately: 7% ± 2.660(3.227) or 7% ± 8.584% or -1.584% to 15.584%.

If you use a z-statistic, the confidence interval is 7% ± 2.58(3.227) = -1.326% to 15.326%, which is closest to the correct choice.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

16. (A) overstate returns for the fund sector.

Explanation

A sample suffers from survivorship bias if only surviving funds are captured in the data. Funds that cease to exist (due to poor performance) are excluded. This results in an average that overstates the annual return an investor in the sector could actually have expected to earn over the period.

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

17. (C) **$27,159 to $28,842.**

Explanation

Confidence interval = mean ± t_c(S / √n}

= 28,000 ± (1.98) (4,250 / √100) or 27,159 to 28,842

If you use a z-statistic because of the large sample size, you get 28,000 ± (1.96) (4,250 / √100) = 27,167 to 28,833, which is closest to the correct answer.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

18. (C) **convenience sampling.**

Explanation

Convenience sampling refers to sampling an element of a population based on ease of access.

(Study Session 2, module 5.1, LOS 5.c)

Related Material

SchweserNotes - Book 1

19. (C) **any probability distribution.**

Explanation

The central limit theorem tells us that for a population with a mean p and a finite variance σ^2, the sampling distribution of the sample means of all possible samples of size *n* will be approximately normally distributed with a mean equal to μ and a variance equal to σ^2/n, *no matter the distribution of the population,* assuming a large sample size.

(Study Session 2, module 5.1, LOS 5.d)

Related Material

SchweserNotes - Book 1

20. (C) **-0.794% to 8.794%**

Explanation

The standard error for the mean = $s/(n)^{0.5}$ = $15\%/(40)^{0.5}$ = 2.372%. The critical value from the t-table should be based on 40 - 1 = 39 df. Since the standard tables do not provide the critical value for 39 df the closest available value is for 40 df. This leaves us with an approximate confidence interval. Based on 95% confidence and df = 40, the critical t-value is 2.021. Therefore the 95% confidence interval is approximately: 4% ± 2.021(2.372) or 4% ± 4.794% or -0.794% to 8.794%.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

21. (A) 2.313 to 2.687.

Explanation

The reliability factor for a 95% confidence level for the Student's t-distribution with (20 - 1) degrees of freedom is 2.093. The confidence interval is equal to: 2.5 $\pm$ 2.093(0.4 /$\sqrt{20}$) = 2.313 to 2.687. (We must use the Student's t-distribution and reliability factors because of the small sample size.)

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

22. (A) 34 to 82.

Explanation

This question has a bit of a trick. To answer this question, remember that the mean is at the midpoint of the confidence interval. The correct confidence interval will have a midpoint of 58. (34 + 82) / 2 = 58.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

23. (C) survivorship bias.

Explanation

When constructing samples, researchers must be careful not to include just survivors (e.g., surviving companies, mutual funds, or investment newsletters). Since survivors tend to be those that have done well (by skill or chance), funds that have 10-year track records will exhibit performance histories with upward bias-mutual fund companies regularly discontinue funds with poor performance histories or roll their assets into better performing funds. Time period bias occurs when the period chosen is so short that it shows relationships that are unlikely to recur, or so long that it includes fundamental changes in the relationship being observed. A 10-year period typically includes a full economic cycle and is likely to be appropriate for this test. Look-ahead bias is present if the test relates a variable to data that were not available at the points in time when that variable's outcomes were observed.

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

24. (A) time-period bias.

Explanation

Time-period bias can result if the time period over which the data is gathered is either too short because the results may reflect phenomenon specific to that time period, or if a change occurred during the time frame that would result in two different return distributions. In this case the time period sampled is probably not large enough to draw any conclusions about the long-term relative performance of value and growth stocks, even if the sample size within that time period is large.

Look-ahead bias occurs when the analyst uses historical data that was not publicly available at the time being studied. Survivorship bias is a form of sample selection bias in which the observations in the sample are biased because the elements of the sample that *survived* until the sample was taken are different than the elements that dropped out of the population.

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

25. (B) 0.0200.

Explanation

The standard error of the mean = $\sigma/\sqrt{n}$ or = $s/\sqrt{n}$ if the population variance is unknown.

The standard error is, therefore, $0.16/\sqrt{64} = 0.02$.

A mean calculated from a sample selected at random from the population will deviate from the true population mean. This deviation is referred to as the standard error.

The smaller the standard error, the smaller the deviation; hence, the closer the sample mean is likely to be to the true population mean.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

26. (A) 25 to 29.

Explanation

The standard error of the sample mean = $5 / \sqrt{25} = 1$

Degrees of freedom = 25 - 1 = 24

From Student's t-table, $t_{S/2} = 2.064$

The confidence interval is: $27 \pm 2.064(1) = 24.94$ to 29.06 or 25 to 29.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

27. (C) data snooping bias.

Explanation

The practice of data snooping involves repeatedly analyzing the same data in the hope of detecting a pattern. Data snooping bias may result because some number of apparently significant relationships are likely to appear by chance.

Look-ahead bias occurs by using information that is not available or known in the analysis period. An example is back testing a trading strategy that uses information in year-end accounts to 31st December 2019 to assist trading in the first week on January. This would not be possible as the financial statements are released several weeks after the year-end.

Sample selection bias occurs when certain information is excluded from a sample due to lack of availability. The sample will, therefore, not be a random one (i.e., where each member of the population has an equal likelihood of being selected).

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

28. (A) For a sample size of 30, using a t-statistic will result in a wider confidence interval for a population mean than using a z-statistic.

Explanation

Although the t-distribution begins to approach the shape of a normal distribution for large sample sizes, at a sample size of 30 a t-statistic produces a wider confidence interval than a *z*-statistic. A confidence interval for the population mean is the sample mean plus-or-minus the appropriate critical value times the *standard error,* which is the standard deviation divided by the square root of the sample size. If a population is normally distributed, we can use a *t*-statistic to construct a confidence interval for the population mean from a small sample, even if the population variance is unknown.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

29. (A) 0.04.

Explanation

The standard error = sample standard deviation / √n.

If the sample variance in 0.16, then the sample standard deviation is √10.16.

The standard error is, therefore, √0.16 / √1100 = 0.04.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

30. (C) The standard error of the sample means when the standard deviation of the population is unknown equals s / √n, where s = sample standard deviation.

Explanation

The probability that a parameter lies within a range of estimated values is given by 1 - α. The standard error of the sample means when the standard deviation of the population is known equals α / √n, where a = *population* standard deviation.

(Study Session 2, module 5.2, LOS 5.g)

Related Material

SchweserNotes - Book 1

31. (A) **has a sampling distribution with a smaller variance than that of all other unbiased estimators of the parameter.**

Explanation

An estimator is *efficient* if the variance of its sampling distribution is smaller than that of all other unbiased estimators of the parameter. An *unbiased* estimator has an expected value equal to the parameter it is estimating. A *consistent* estimator becomes more accurate as the sample size increases.

(Study Session 2, module 5.1, LOS 5.f)

Related Material

SchweserNotes - Book 1

32. (C) **higher degree of confidence.**

Explanation

A higher degree of confidence (e.g. 99% instead of 95%) would require a higher reliability factor (2.575 instead of 1.96 assuming a normal distribution). A wider confidence interval corresponds to a lower alpha significance level and the point estimate does not affect the width of the confidence interval.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

33. (B) **sample mean.**

Explanation

The central limit theorem tells us that for a population with a mean m and a finite variance σ^2, the sampling distribution of the *sample means* of all possible samples of size n will approach a normal distribution with a mean equal to μ and a variance equal to σ^2 / n as *n* gets large.

(Study Session 2, module 5.1, LOS 5.d)

Related Material

SchweserNotes - Book 1

34. (A) **The standard error decreases and the confidence interval narrows.**

Explanation

Increasing the sample size will improve the accuracy of the estimated mean. As the sample size increases, both the standard error of the sample mean and the width of the confidence interval decrease.

Related Material

SchweserNotes - Book 1

35. (C) 0.64.

Explanation

The standard error of the sample mean when the standard deviation of the population is known is equal to the standard deviation of the population divided by the square root of the sample size. In this case, $3.2 / \sqrt{125} = 0.64$.

It is a measure of how much the sample mean is likely to deviate from the population mean. The larger the sample selected, the lower the standard error, and so the less the sample mean will deviate from the true population mean.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

36. (B) **The size of each sub-sample is selected to be the same across strata.**

Explanation

In stratified random sampling we first divide the population into subgroups, called strata, based on some classification scheme. Then we randomly select a sample from each stratum and pool the results. *The size of the samples from each strata is based on the relative size of the strata relative to the population and are not necessarily the same across strata.*

(Study Session 2, module 5.2, LOS 5.c)

Related Material

SchweserNotes - Book 1

37. (C) **stratified random sampling.**

Explanation

In stratified random sampling, a researcher classifies a population into smaller groups based on one or more characteristics, takes a simple random sample from each subgroup, and pools the results.

A random sample is one where each member of the population has an equal chance of being selected.

Systematic sampling is where every *n*th member of the population is selected, also known as nonrandom sampling.

(Study Session 2, module 5.1, LOS 5.c)

Related Material

SchweserNotes - Book 1

38. (A) $0.7713 to $1.2287.

Explanation

The population *standard deviation* is the square root of the variance ($\sqrt{0.49} = 0.7$). Because we know the population standard deviation, we use the z-statistic. The z-statistic reliability factor for a 95% confidence interval is 1.960. The confidence interval is $\$1.00 \pm 1.960(\$0.7 / \sqrt{36})$ or $\$1.00 \pm \$0.2287$.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

39. (A) $667.

Explanation

The sample standard deviation is the square root of the variance: $(36{,}000{,}000)^{1/2}$ = \$6,000. The standard error of the sample mean is estimated by dividing the standard deviation of the sample by the square root of the sample size:

$\sigma_{mean} = s / (n)^{1/2} = 6{,}000 / (81)^{1/2} = \$667$.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

40. (B) 3.00.

Explanation

The standard error of the sample mean equals the standard deviation of the population divided by the square root of the sample size $\sigma / \sqrt{n} = 15 / \sqrt{125} = 3$.

The standard error measures how much the sample mean deviates from the true population mean. The smaller the standard error, the closer the sample mean is likely to lie to the true population mean.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

41. (C) **The distribution is nonnormal, the population variance is unknown, and the sample size is at least 30.**

Explanation

The t-distribution is the theoretically correct distribution to use when constructing a confidence interval for the mean when the distribution is nonnormal and the population variance is unknown but the sample size is at least 30.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

42. (A) **expected value of the sample mean is equal to the population mean.**

Explanation

An unbiased estimator is one for which the expected value of the estimator is equal to the parameter you are trying to estimate.

(Study Session 2, module 5.1, LOS 5.f)

Related Material

SchweserNotes - Book 1

43. (A) **A point estimate is a single estimate of an unknown population parameter calculated as a sample mean.**

Explanation

Time-series data are observations taken at specific and equally-spaced points.

A confidence interval estimate consists of a range of values that bracket the parameter with a specified level of probability, $1 - \alpha$.

(Study Session 2, module 5.2, LOS 5.g)

Related Material

SchweserNotes - Book 1

44. (A) **95.706 to 96.294.**

Explanation

Because we can compute the population standard deviation, we use the z-statistic. A 95% confidence level is constructed by taking the population mean and adding and subtracting the product of the z-statistic reliability ($z_{\alpha/2}$) factor times the known standard deviation of the population divided by the square root of the sample size (note that the population variance is given and its positive square root is the standard deviation of the population): $x \pm z_{\alpha/2} \times (\sigma / n^{1/2}) = 96 \pm 1.96 \times (9^{1/2} / 400^{1/2}) = 96 \pm 1.96 \times (0.15) = 96 \pm 0.294 = 95.706$ to 96.294.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

45. (C) **1.0% ± 1.9%.**

Explanation

If the distribution of the population is *nonnormal,* but we *don't know the* population variance, we can use the Student's t-distribution to construct a confidence interval. The sample standard deviation is the square root of the variance, or 6%. Because there are 41 observations, the degrees of freedom are 40. From the Student's *t* distribution, we can determine that the reliability factor for $t_{0.025}$, is 2.021. Then the 95% confidence interval is 1.0% ± 2.021(6 / √14) or 1.0% ± 1.9%.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

46. (B) **-0.4% to 3.4%.**

Explanation

Because we know the population standard deviation, we use the z-statistic. The z-statistic reliability factor for a 99% confidence interval is 2.575. The confidence interval is 1.5% ± 2.575[(8.0%) / √121] or 1.5% ± 1.9%.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

47. (C) the expected value of the estimator is equal to the population parameter.

Explanation

Desirable properties of an estimator are unbiasedness, efficiency, and consistency. An estimator is unbiased if its expected value is equal to the population parameter it is estimating. An estimator is efficient if the variance of its sampling distribution is smaller than that of all other unbiased estimators. An estimator is consistent if an increase in sample size decreases the standard error.

(Study Session 2, module 5.1, LOS 5.f)

Related Material

SchweserNotes - Book 1

48. (B) stratified random sampling.

Explanation

In stratified random sampling we first divide the population into subgroups, called strata, based on some classification scheme. Then we randomly select a sample from each stratum and pool the results. The size of the samples from each strata is based on the relative size of the strata relative to the population. Simple random sampling is a method of selecting a sample in such a way that each item or person in the population being studied has the same (non-zero) likelihood of being included in the sample.

(Study Session 2, module 5.1, LOS 5.c)

Related Material

SchweserNotes - Book 1

49. (B) Point estimate +/- (Reliability factor x Standard error).

Explanation

We can construct a confidence interval by adding and subtracting some amount from the point estimate. In general, confidence intervals have the following form:

Point estimate +/- Reliability factor x Standard error

Point estimate = the value of a sample statistic of the population parameter

Reliability factor = a number that depends on the sampling distribution of the point estimate and the probability the point estimate falls in the confidence interval $(1 - \alpha)$

Standard error = the standard error of the point estimate

(Study Session 2, module 5.2, LOS 5.g)

Related Material

SchweserNotes - Book 1

50. (A) stratified random sampling.

Explanation

Stratified random sampling is used to preserve characteristics of an underlying dataset.

Related Material

SchweserNotes - Book 1

51. (B) 25.

Explanation

Given the population standard deviation and the standard error of the sample mean, you can solve for the sample size. Because the standard error of the sample mean equals the standard deviation of the population divided by the square root of the sample size, $4 = 20 / n^{1/2}$, so $n^{1/2} = 5$, so $n = 25$.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

52. (B) The lower the significance level, the wider the confidence interval.

Explanation

A higher degree of confidence requires a wider confidence interval. The degree of confidence is equal to one minus the significance level, and so the wider the confidence interval, the higher the degree of confidence and the lower the significance level.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

53. (B) corporate bonds.

Explanation

Stratified sampling is most often used for bond portfolios.

Related Material

SchweserNotes - Book 1

54. (C) Stratified random sampling.

Explanation

In stratified random sampling, we first divide the population into subgroups based on some relevant characteristic(s) and then make random draws from each group.

(Study Session 2, module 5.1, LOS 5.c)

Related Material

SchweserNotes - Book 1

55. (C) Sampling errors are errors due to the wrong sample being selected from the population.

Explanation

Sampling error is the difference between a sample statistic (the mean, variance, or standard deviation of the sample) and its corresponding population parameter (the mean, variance, or standard deviation of the population).

(Study Session 2, module 5.1, LOS 5.b)

Related Material

SchweserNotes - Book 1

56. (A) **judgmental sampling.**

Explanation

judgmental sampling refers to using expert or professional judgement to select observations from a population.

(Study Session 2, module 5.1, LOS 5.c)

Related Material

SchweserNotes - Book 1

57. (A) **5.00.**

Explanation

The standard error of the sample mean = population standard deviation / $\sqrt{n}$.

If the population variance is 2500, then the standard deviation is $\sqrt{2500}$. Therefore, the standard error is $\sqrt{2500} / \sqrt{100} = 5$.

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

58. (C) **sample mean provides a more accurate estimate of the population mean as the sample size increases.**

Explanation

A consistent estimator provides a more accurate estimate of the parameter as the sample size increases.

(Study Session 2, module 5.1, LOS 5.f)

Related Material

SchweserNotes - Book 1

59. (C) **Involves a trade-off between the cost of increasing the sample size and the value of increasing the precision of the estimates.**

Explanation

A larger sample reduces the sampling error and the standard deviation of the sample statistic around its population value. However, this does not imply that the sample should be as large as possible, or that the sampling error must be as small as can be achieved. Larger samples might contain observations that come from a different population, in which case they would not necessarily improve the estimates of the population parameters. Cost also increases with the sample size. When the cost of increasing the sample size is greater than the value of the extra precision gained, increasing the sample size is not appropriate.

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

60. (C) -48.7% to 84.1 %.

Explanation

A 95% confidence interval is ± 1.96 standard deviations from the mean, so 0.177 ± 1.96(0.339) = (-48.7%, 84.1%).

(Study Session 2, Module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

61. (A) the sample size n > 30.

Explanation

The Central Limit Theorem states that if the sample size is sufficiently large (i.e. greater than 30) the sampling distribution of the sample means will be approximately normal.

(Study Session 2, module 5.1, LOS 5.d)

Related Material

SchweserNotes - Book 1

62. (A) 60 ± 1.96(1.85).

Explanation

Because the sample size is sufficiently large, we can use the z-statistic. A 95% confidence level is constructed by taking the sample mean and adding and subtracting the product of the *z*-statistic reliability factor ($z_{\alpha/2}$) times the standard error of the sample mean: $x \pm z_{\alpha/2}$ x (s /$n^{1/2}$) = 60 ± (1.96) x (16 / $75^{1/2}$) = 60 ± (1.96) x (16 / 8.6603) = 60 ± (1.96) x (1.85).

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

63. (C) $45,000 to $55,000.

Explanation

Because the sample size is large and the population is normally distributed, we can acceptably use a z-statistic. A 90% confidence level for the population mean is constructed by taking the sample mean and adding and subtracting the product of the z-statistic reliability ($z_{\alpha/2}$) factor times the sample standard deviation divided by the square root of the sample size: $x \pm z_{a/2}$ x (s /$n^{1/2}$) = 50,000 ± 1.645 x (30,000 / $100^{1/2}$) = 50,000 ± 4,935 = $45,065 to $54,935.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

64. (C) (1 - α) percent confidence interval.

Explanation

A 95% confidence interval for the population mean (α = 5%), for example, is a range of estimates within which the actual value of the population mean will lie with a probability of 95%. Point estimates, on the other hand, are *single* sample) values used to estimate population parameters. There is no such thing as a α percent *point estimate* or a (1 - α) percent *cross-sectional point estimate.*

(Study Session 2, module 5.1, LOS 5.e)

Related Material

SchweserNotes - Book 1

65. (C) no test statistic is available.

Explanation

No test statistic is available when sampling from a *non-normally distributed* population. If a *normally distributed* population has a known variance, a z-test is appropriate, even with a small sample, and if its population variance is unknown, a t-statistic is appropriate with a small sample.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

66. (A) 15 ± 1.96(0.4).

Explanation

Because we can compute the population standard deviation, we use the z-statistic. A 95% confidence level is constructed by taking the population mean and adding and subtracting the product of the z-statistic reliability ($z_{\alpha/2}$) factor times the known standard deviation of the population divided by the square root of the sample size (note that the population variance is given and its positive square root is the standard deviation of the population) : $x \pm z_{\alpha/2} \times (\sigma / n^{1/2}) = 15 \pm 1.96 \times (4^{1/2} / 25^{1/2}) = 15 \pm 1.96 \times (0.4)$.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

67. (C) data snooping.

Explanation

Data snooping refers to the extensive review of the same database searching for patterns.

(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

68. (B) **The standard error of the sample mean will increase as the sample size increases.**

Explanation

The standard error of the sample mean is equal to the sample standard deviation divided by the square root of the sample size. As the sample size increases, this ratio decreases. The other two choices are predictions of the central limit theorem.
(Study Session 2, module 5.1, LOS 5.d)

Related Material

SchweserNotes - Book 1

69. (B) **$70.27 to $79.73.**

Explanation

If the distribution of the population is *normal,* but we *don't know the* population variance, we can use the Student's t-distribution to construct a confidence interval. Because there are 41 observations, the degrees of freedom are 40. From Student's *t* table, we can determine that the reliability factor for $t_{\alpha/2}$, or $t_{0.05}$, is 1.684. Then the 90% confidence interval is $75.00 ± 1.684($18.00 / √41), or $75.00 ± 1.684 x $2.81 or $75.00 ± $4.73
(Study Session 2, module 5.1, LOS 5.h)

Related Material

SchweserNotes - Book 1

70. (B) **Survivorship bias.**

Explanation

Survivorship bias is not likely to significantly influence the results of this study because the authors looked at the stocks in the S&P 500 at the beginning of the year and measured performance over the following three months. Look-ahead bias could be a problem because earnings-price ratios are calculated and the trading strategy implemented at a time before earnings are actually reported. Finally, the study is conducted over a relatively short time period during the long bull market of the 1990s. This suggests the results may be time-specific and the result of time-period bias.
(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

71. (B) **The study suffers from look-ahead bias.**

Explanation

The study suffers from look-ahead bias because traders at the beginning of the year would not be able to know the book value changes. Financial statements usually take 60 to 90 days to be completed and released.
(Study Session 2, module 5.2, LOS 5.j)

Related Material

SchweserNotes - Book 1

72. (C) **2.475 to 2.525.**

Explanation

The Z-score corresponding with a 5% significance level (95% confidence level) is 1.96. The confidence interval is equal to: $2.5 \pm 1.96(0.4 / \sqrt{1,000}) = 2.475$ to 2.525. (We can use Z-scores because the size of the sample is so large.)

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1

73. (B) **states that for a population with mean p and variance ci², the sampling distribution of the sample means for any sample of size n will be approximately normally distributed.**

Explanation

This question is asking you to select the inaccurate statement. The CLT states that for a population with mean μ and a finite variance σ^2, the sampling distribution of the sample means becomes approximately normally distributed *as the sample size becomes large.* The other statements are accurate.

(Study Session 2, module 5.1, LOS 5.d)

Related Material

SchweserNotes - Book 1

74. (C) **difference between a sample statistic and its corresponding population parameter.**

Explanation

Sampling error is the difference between any sample statistic (the mean, variance, or standard deviation of the sample) and its corresponding population parameter (the mean, variance or standard deviation of the population). For example, the sampling error for the mean is equal to the sample mean minus the population mean.

(Study Session 2, module 5.1, LOS 5.b)

Related Material

SchweserNotes - Book 1

75. (A) **$172,754 to $227,246.**

Explanation

If the distribution of the population is *normal,* but we *don't know the* population variance, we can use the Student's t-distribution to construct a confidence interval. Because there are 61 observations, the degrees of freedom are 60. From the student's t table, we can determine that the reliability factor for $t_{\alpha}n$, or $t_{0.005}$, is 2.660. Then the 99% confidence interval is $200,000 ± 2.660($80,000 / √61) or $200,000 ± 2.660 x $10,243, or $200,000 ± $27,246.

(Study Session 2, module 5.2, LOS 5.h)

Related Material

SchweserNotes - Book 1