Identifying problems when obtaining population parameters

We estimate population parameters, such as the mean, based on the sample statistics. It is difficult to get a precise value or point estimation of these figures. A more practical and informative approach is to find a range of values in which we expect the population parameters will fall. Such a range of values is called a confidence interval.

1. CONFIDENCE INTERVAL

Definition

The confidence interval is a range of values constructed from sample data so that the population parameter is likely to occur within that range at a specified probability. The specified probability is called the level of confidence.

The shape of the probability distribution of the sample mean allows us to specify an interval of specific probability that the population mean, μ, will fall into.

1. 1 Large Sample Or Standard Deviation Is Known

Case 1:

- The standard deviation Ïƒ is known; or
- It is a large sample (i. e. at least 30 observations).
- The Central Limit Theorem states that the sampling distribution of the sample means is approximately normal. We can use the tables in the Appendix to find the appropriate Z value.

Key Points

The standard normal distribution allows us to draw the following conclusions:

- 68% of the sample means will be within 1 standard deviations of the population mean, μ.
- 95% of the sample means will be within 1. 96 standard deviations of the population mean, μ.
- 99% of the sample means will lie within 2. 58 standard deviations of the population mean.

These intervals are called the confidence interval.

The standard deviation above (i. e. the “ standard error”) is referring to the standard deviation of the sampling distribution of the sample mean.

Locating 0. 475 in the body of the table, read the corresponding row and column values, the value is 1. 96. Thus, the probability of finding a Z value between 0 and 1. 96 is 0. 475. Likewise, the probability of being in the interval between -1. 96 and 0 is also 0. 475. When we combine these two, the probability of being in the interval of -1. 96 to 1. 96 is therefore 0. 95.

1. 1. 1 How do you compute a 95% confidence interval?

Assume our research involves the annual starting salary of business graduates in a local university. The sample mean is $39, 000, while the standard deviation of the sample mean is $250. Assume our sample contains more than 30 observations. The 95% confidence interval is between $38, 510 and $39, 490. Found by $39, 000 +/- 1. 96($250)

In most situations, the population standard deviation is not available, so we estimate it as follows: (Standard Error)

Conclusions:

- 95% confidence interval
- 99% confidence interval
- Confidence interval for the population mean (n > 30)
- Z depends on confidence level

Example 1

The Hong Kong Tourist Association wishes to have information on the mean annual income of tour guides. A random sample of 150 tour guides reveals a sample mean of $45, 420. The standard deviation of this sample is $2, 050. The association would like answers to the following questions:

(a) What is the population mean?

The best estimate of the unknown population value is the corresponding sample statistic. The sample mean of $45, 420 is a point estimate of the unknown population mean.

(b) What is a reasonable range of values for population mean?

The Association decides to use the 95% level of confidence. To determine the corresponding confidence interval, we use the formula:

The endpoints would be $45, 169 and $45, 671 and they are called confidence limits. We could expect about 95% of these confidence intervals contain the population mean. About 5% of the intervals would not contain the population mean annual income, i. e. the μ.

Figure 2 Probability distribution of population mean

1. 2 Small Sample Or Standard Deviation Is Unknown

Case 2:

- The sample is small (i. e. less than 30 observations) or,
- the population standard deviation is not known.

The correct statistical procedure is to replace the standard normal distribution with the t distribution. The t distribution is a continuous distribution with many similarities to the standard normal distribution.

1. 2. 1 Standard normal distribution versus t distribution

Figure 3 Z distribution versus t distribution

- The t distribution is flatter and more spread out than the standard normal distribution.
- The standard deviation of the t distribution is larger than the normal distribution.
- Confidence interval for a sample with unknown population mean, Ïƒ. The confidence interval is
- Assume the sample is from a normal population.
- Estimate the population standard deviation (Ïƒ) with the sample standard deviation (s).
- Use t distribution rather than the Z distribution.

Example 2

A shoe maker wants to investigate the useful life of his products. A sample of 10 pairs of shoes that had been walked for 50, 000 km showed a sample mean of 0. 32 inch of sole remaining with a standard deviation of 0. 09 cm. Constructing a 95% confidence interval for the population mean, would it be reasonable for the manufacturer to conclude that after 50, 000 km the population mean amount of sole remaining is 0. 3 cm?

Assume the population distribution is normal. The sample standard deviation is 0. 09 cm.

There are only 10 observations and hence, we use t distribution

Estimation:

= 0. 32, s = 0. 09, and n = 10.

Step 1: Locate t by moving across the row for the level of confidence required (i. e. 95%).

Step 2: The column on the left margin is identified as “ df”. This refers to the number of degrees of freedom. The number of degree of freedom is the number of observations in the sample minus the number of samples, written n-1.(i. e. 10-1= 9).

Step 3: Confidence Interval =

The endpoints of the confidence interval are 0. 256 and 0. 384.

Step 4: Interpretation – the manufacturer can be reasonably sure (95% confident) that the mean remaining tread depth is between 0. 256 and 0. 384 cm. Because 0. 3 is in this interval, it is possible that the mean of the population is 0. 3.

2. CHOOSING AN APPROPRIATE SAMPLE SIZE

The necessary sample size depends on three factors:

- Level of confidence wanted: To increase level of confidence, increase n.
- Margin of error the researcher will tolerate: To reduce allowable error, increase n.
- Variability in the population being studied: For a more widely dispersed sample, increase n.

We can express the interaction among these three factors and the sample size in the following formula:

Sample size for estimating the population mean,

Note:

- n: Sample size
- Z: Standard normal value
- S: Estimate of population standard deviation
- E: Maximum allowable error

Example 3

An accounting student wants to know the mean amount that independent directors of small companies earn per month as remuneration for being a director. The error in estimating the mean is to be less than $100 with a 95% level of confidence. The student found a report by the government that estimated the standard deviation to be $1000. What is the required sample size?

Maximum allowable error, E, is $100.

Value of Z for a 95% level of confidence is 1. 96, and the estimate of the standard deviation is $1000.

Substitute into , we get

n = [

(1. 96) (1000)

] 2 = 19. 62 = 384. 16

100

The sample of 385 is required to meet the requirements. If the students want to increase the level of confidence, e. g. 99%, this requires a larger sample.

Z = 2. 58, so

n = [

(2. 58) (1000)

] 2 = 25. 82 = 665. 64

100

Sample = 666

3. WHAT IS A HYPOTHESIS?

Definitions

Hypothesis is a statement about a population parameter developed for the purpose of testing.

Hypothesis testing is a procedure based on sample evidence and probability theory to determine whether the hypothesis is a reasonable statement.

In statistical analysis, we always make a claim about the population parameters, i. e. a hypothesis. We collect data and then use the data to test the assertion.

4. 1 Five-Step Procedure For Testing A Hypothesis

Figure 4 How to test a hypothesis

4. 1. 1 Step 1: State null hypothesis (H0) and alternative hypothesis (H1)

The first step is to state the hypothesis being tested. It is called the null hypothesis. We either reject or fail to reject the null hypothesis. Failing to reject the null hypothesis does not prove that H0 is true.

The null hypothesis is a statement that is not rejected unless our sample data provide convincing evidence that it is false.

The alternative hypothesis is a statement that is accepted if the sample data provide sufficient evidence that the null hypothesis is false.

Example 4

A journal has disclosed that the mean age of commercial helicopters is 15 years. A statistical test of this statement would first need to determine the null and the alternate hypotheses.

The null hypothesis represents the current or reported condition. It is written H0: μ = 15.

The alternate hypothesis is that the statement is not true, i. e. H1: μ â‰ 15.

4. 1. 2 Step 2: Select a level of significance

The level of significance is the probability of rejecting the null hypothesis when it is true.

A decision is made to use the 5% level, 1% level, 10% level or any other level between 0 and 1. We must decide on the level of significance before formulating a decision rule and collecting sample data.

- Type I error: Rejecting the null hypothesis, H0, when it is true.
- Type II error: Accepting the null hypothesis when it is false.

Example 5

Suppose AA Watch Ltd has informed bracelet suppliers to bid for contract on the supply of a large amount of bracelets. Suppliers with the lowest bid will be awarded a sizable contract.

Suppose the contract specifies that the watch producer’s quality-assurance department will take samples of the shipment.

H0: The shipment of bracelet contains 6% or less substandard bracelets.

H1: More than 6% of the boards are defective.

A sample of 50 bracelets received August 2 from BB Metals Ltd revealed that four bracelets, or 8%, were substandard. The shipment was rejected because it exceeded the maximum of 6% substandard bracelets. If the shipment was actually substandard, the decision to return the bracelets to the supplier was correct.

However, suppose the four substandard bracelets selected in the sample of 50 were the only substandard bracelets in the shipment of 4, 000 bracelets. Then only 1/10 of 1% were defective (4/4000 = 0. 001). In that case, less than 6% of the entire shipment was substandard and rejecting the shipment was an error.

We may have rejected the null hypothesis that the shipment was not substandard when we should have accepted the null hypothesis.

By rejecting a true null hypothesis, we committed a Type I error.

AA Watch Ltd would commit a Type II error if, unknown to the company an incoming shipment of bracelet from BB Metals Ltd contained 15% substandard bracelets, yet the shipment was accepted. How could this happen?

Suppose two out of the 50 bracelets in the sample (4%) tested were substandard, and 48 out of the 50 were good bracelets. As the sample contained less than 6% substandard bracelets, the shipment was accepted but it could be purely by chance that the 48 good bracelets selected in the sample were the only acceptable ones in the entire shipment.

In conclusion:

Null Hypothesis

Accepts H0

Rejects H0

H0 is true

Correct decision

Type I error

H0 is false

Type II error

Correct decision

4. 1. 3 Step 3: Select the test statistics

There are many test statistics. In this chapter, we use both Z and t as the test statistic.

Definition

A test statistic is a value, determined from sample information, used to determine whether to reject the null hypothesis.

In hypothesis testing for the mean (µ) when Ïƒ is known or the sample size is large, the test statistic Z is computed by:

The Z value is based on the sampling distribution of , which follows the normal distribution when the sample is reasonably large with a mean () equal to µ, and a standard deviation , which is equal to . We can thus determine whether the difference between and μ is statistically significant by finding the number of standard deviations is from μ, using the formula above.

4. 1. 4 Step 4: Formulate the decision rule

Definition

A decision rule is a statement of the specific conditions under which the null hypothesis is rejected and the conditions under which it is not rejected.

The region or area of rejection defines the location of all those values that are so large or so small that the probability of their occurrence under a true null hypothesis is rather remote.

- The area where the null hypothesis is not rejected is to the left of 1. 65.
- The area of rejection is to the right of 1. 65.
- A one-tailed test is being applied.
- The 0. 05 level of significance was chosen.
- The sampling distribution of the statistic Z is normally distributed.
- The value 1. 65 separates the regions where the null hypothesis is rejected and where it is not rejected.
- The value 1. 65 is the critical value.
- The critical value is the dividing point between the region where the null hypothesis is rejected and the region where it is not rejected.

Figure 5 Area of rejection for the null hypothesis

4. 1. 5 Step 5: Make a decision

The final step in hypothesis testing is computing the test statistic, comparing it to the critical value, and making a decision to reject or not to reject the null hypothesis.

Based on the information, Z is computed to be 2. 34, the null hypothesis is rejected at the 0. 05 level of significance. The decision to reject H0 was made because 2. 34 lies in the region of rejection, i. e. beyond 1. 65.

We would reject the null hypothesis, reasoning that it is highly improbable that a computed Z value this large is due to sampling variation.

Had the computed value been 1. 65 or less, say 0. 71, the null hypothesis would not be rejected. It would be reasoned that such a small computed value could be attributed to chance.

Example 6

A large car leasing company wants to buy tires that average about 60, 000 km of wear under normal usage. The company will, therefore, reject a shipment of tires if tests reveal that the life of the tires is significantly below 60, 000 km on the average.

The company would be glad to accept a shipment if the mean life is greater than 60, 000 km. However, it is more concerned that it will have sample evidence to conclude that the tires will average less than 60, 000 km of useful life. Thus, the test is set up to satisfy the concern of the car leasers that the mean life of the tires is less than 60, 000 km.

The null and alternate hypotheses in this case are written H0: µ â‰¥ 60, 000 and H1: µ < 60, 000.

In this problem, the rejection region is pointing to the left, and is therefore in the left tail.

Summary:

- If H1 states a direction, we use a one-tailed test.
- If no direction is specified in the alternate hypothesis, we use a two-tailed test.

Figure 6 One-tailed test

5. TESTING FOR POPULATION MEAN WITH KNOWN POPULATION STANDARD DEVIATION

5. 1 Two-tailed Test

ABC Watch Ltd manufactures luxury watches at several plants in Europe. The weekly output of the Model A33 watch at the Swiss Plant is normally distributed, with a mean of 200 and a standard deviation of 16. Recently, because of market expansion, mechanisation has been introduced and employees laid off. The CEO would like to investigate whether there has been a change in the weekly production of the Model A33 watch. To put it another way, is the mean output at Swiss Plant different from 200 at the 0. 01 significant levels?

5. 1. 1 Step 1: State null hypothesis and alternate hypothesis

The null hypothesis is “ The population mean is 200.” H0: µ = 200.

The alternate hypothesis is “ The mean is different from 200.” H1: µ â‰ 200.

5. 1. 2 Step 2: Select the level of significance

The 0. 01 level of significance is used. This is Î±, the probability of committing a Type I error, and it is the probability of rejecting a true null hypothesis.

5. 1. 3 Step 3: Select the test statistic

The test statistic for the mean of a large sample is Z.

Figure 7 Normalise the standard deviation

5. 1. 4 Step 4: Formulate the decision rule

The decision rule is formulated by finding the critical values of Z from Appendix D.

Since this is a two-tailed test, half of 0. 01, or 0. 005, is placed in each tail. The area where H0 is not rejected, i. e. area between the two tails, is 0. 99.

Appendix D is based on half of the area under the curve, or 0. 5. Then 0. 5 – 0. 005 is 0. 495, so 0. 495 is the area between 0 and the critical value.

The value nearest to 0. 495 is 0. 4951. Then read the critical value in the row and column corresponding to 0. 4951. It is 2. 58.

Decision rule:

- Reject H0 if the computed Z value is not between -2. 58 and +2. 58.

- Do not reject H0 if Z falls between -2. 58 and +2. 58.

Figure 8 Two-tailed test

5. 1. 5 Make a decision and interpret the result

Compute Z and apply the decision rule to decide whether to reject H0.

The mean number of watches produced weekly for last year is 203. 5. The standard deviation of the population is 16 watches.

Because 1. 55 does not fall in the rejection region, H0 is not rejected. We conclude that the population mean is not different from 200.

So we would report to the CEO that the sample evidence does not show that the production rate at the Swiss plant has changed from 200 per week. The difference of 3. 5 units between the historical weekly production rate and the mean number of watches produced weekly for last year can reasonably be attributed to sampling error.

Figure 9 Rejection regions for the two-tailed test

So did we prove that production rate is still 200 per week?

No! Failing to disprove the hypothesis that " the population mean is 200" is not the same thing as proving it to be true.

5. 2 P-value In Hypothesis Testing

Definition

P-value is the probability of observing a sample value as extreme as, or more extreme than, the value observed, given that the null hypothesis is true.

How confident are we in rejecting the null hypothesis?

This approach reports the probability of getting a value of the test statistic at least as extreme as the value actually obtained. This process compares the probability called the P-value, with the significant level.

- If the P-value < significant level, H0 is rejected.
- If the P-value > significant level, H0 is not rejected.

A very small P-value, such as 0. 0001, indicates that there is little likelihood the H0 is true. If a P-value of 0. 2033 means that H0 is not rejected, there is little likelihood that it is false.

Figure 10 P-value

P-value

Interpretation

Less than 0. 1

Some evidence that H0 is not true

Less than 0. 05

Strong evidence that H0 is not true

Less than 0. 01

Very strong evidence that H0 is not true

Less than 0. 001

Extremely strong evidence that H0 is not true

The probability of finding a Z value of 1. 55 or more is 0. 0606, found by 0. 5 – 0. 4394.

The probability of obtaining an greater than 203. 5 if μ = 200 is 0. 0606.

To compute the P-value, we need to be concerned with the region less than -1. 55 as well as the values greater than 1. 55. The two-tailed P-value is 0. 1212, found by 2(0. 0606). The P-value of 0. 1212 is greater than the significance level of 0. 01, so H0 is not rejected.

Chapter Review

- The Central Limit Theorem states that the sampling distribution of the sample means is approximately normal.
- The standard error refers to the standard deviation of the sampling distribution of the sample mean.
- We use t distribution when the sample is less than 30 observations and the population standard deviation is not known.
- The necessary sample size depends on 1) level of confidence wanted ; 2) margin of error the researcher will tolerate; 3)variability in the population.
- By rejecting a true null hypothesis, we committed a Type I error.
- We would reject the null hypothesis when it is highly improbable that a computed Z value this large is due to sampling variation.

What You Need To Know

- Confidence interval: A range of values constructed from sample data so that the population parameter is likely to occur within that range at a specified probability.
- Hypothesis: A statement about a population parameter developed for the purpose of testing.
- Hypothesis testing: A procedure based on sample evidence and probability theory to determine whether the hypothesis is a reasonable statement.
- Critical value: The dividing point between the region where the null hypothesis is rejected and the region where it is not rejected.
- P-value: The probability of observing a sample value as extreme as, or more extreme than, the value observed, given that the null hypothesis is true.

Work Them Out

1. The average number of days in outdoors assignments per year for salespeople employed by an electronic wholesaler needs to be estimated with a 0. 90 degree of confidence. In a small sample, the mean was 150 days and the standard deviation was 14 days. If the population mean is estimated within two days, how many salespeople should be interviewed?

A 134

B 152

C 111

D 120

2. A random sample of 85 staff of managerial grade revealed that a person spent an average of 6. 5 years on the job before being promoted. The standard deviation of the sample was 1. 7 years. Using the 0. 95 degree of confidence, what is the confidence interval for the population mean?

A 6. 19 and 6. 99

B 6. 15 and 7. 15

C 6. 14 and 6. 86

D 6. 19 and 7. 19

3. The mean weight of lorries travelling on a particular highway is not known. A state highway authority needs an estimate of the mean. A random sample of 49 lorries was selected and finds the mean is 15. 8 tons, with a standard deviation of 3. 8 tons. What is the 95 per cent interval for the population mean?

A 14. 7 and 16. 9

B 14. 2 and 16. 6

C 14. 0 and 18. 0

D 16. 1 and 18. 1

4. A bank wants to estimate the mean balances owed by platinum Visa card holders. The population standard deviation is estimated to be $300. If a 98%

confidence interval is used and an interval of $75 is desired, how many platinum cardholders should be taken into sample?

A 84

B 82

C 62

D 87

5. A sample of 20 is selected from the population. To determine the appropriate critical t-value, what number of degrees of freedom should be used?

A 20

B 19

C 23

D 27

6. If the null hypothesis that two means are equal is true, where will 97% of the computed z-values lie between?

A ± 2. 58

B ± 2. 38

C ± 2. 17

D ± 1. 68

7. Suppose we are testing the difference between two proportions at the 0. 05 level of significance. If the computed z is -1. 57, what is our decision?

A Reject the null hypothesis

B Do not reject the null hypothesis

C Review the sample

D Own judgment

8. The net weights of a sample of bottles filled by a machine manufactured by Dame, and the net weights of a sample filled by a similar machine manufactured by Putne Inc, are (in grams):

Dame: 5, 8, 7, 6, 9 and 7

Putne: 8, 10, 7, 11, 9, 12, 14 and 9

Testing the claim at the 0. 05 level that the mean weight of the bottles filled by the Putne machine is greater than the mean weight of the bottles filled by the Dame machine, what is the critical value?

A 2. 215

B 2. 175

C 1. 782

D 1. 682

9. Which of the following conditions must be met to conduct a test for the difference in two sample means?

A Data must be of interval scale

B Normal distribution for the two populations

C Same variances in the two populations

D All the above are correct

10. Take two independent samples from two populations in order to determine if a statistical difference on the mean exists. The number for the first sample and the number in the second sample are 15 and 12 respectively. What is the degree of freedom associated with the critical value?

A 24

B 25

C 26

D 27

SHORT QUESTIONS

A consumer group would like to estimate the mean monthly water charge for a single family house in June within $5 using a 99% level of confidence. Similar research has found that the standard deviation is estimated to be $25. 00.

What would be the sample size?

The manager of the Kingsway Mall wants to estimate the mean amount spent per shopping visit by customers. A sample of 20 customers reveals the following amounts spent.

$48 $42 $46 $51 $23 $41 $54 $37 $52 $48

$50 $46 $61 $61 $49 $61 $51 $52 $58 $43

What is the best estimate of the population mean?

Determine a 99 per cent confidence interval. Interpret the result.

Would it be reasonable to conclude that the population mean is $50? What about $60?

ESSAY QUESTION

1. ABC Film Ltd knows that a certain favourite movie ran an average of 84 days, and the corresponding standard deviation was 10 days. The manager of New Westminster district was interested in comparing the movie's popularity in his region with that in all of Canada's other theatres. He randomly selected 70 theatres in his region and found that they showed the movie for an average of 82 days.

(a) State appropriate hypotheses for testing whether there was a significant difference in the length of the picture's run between theatres in the New Westminster district and all of Canada's other theatres.

(b) Test these hypotheses at a 1% significance level.