**Business Statistics – Test of Hypothesis**

**Test of Hypothesis (TH)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Test of Hypothesis is a topic in Inferential Statistics.*** | | | | |
| Test of Hypothesis Process… | | | | |
| Hypothesis | 🡪 | Evidence | 🡪 | Inference |
| Make a hypothesis  about a parameter |  | Obtain an estimate  from a random sample |  | State a conclusion  based on TH logic |

**Summary of Test of Hypothesis (TH) Topics**

|  |
| --- |
| 1. TH Concept & Process  2. TH Introduction & Procedure ( Ho, Ha, , CV, RR, TS, Inference)  3. Formulation of Hypotheses  4. p-value or Observed Significance Level  5. Type I & Type II Error |

|  |  |  |
| --- | --- | --- |
|  | **Test of Hypothesis – Concept** |  |
|  |  |  |
| Let X~N(,2). Suppose we know 2 but we do not know .  {Example: We do not know the population mean age of a population but we know the population standard deviation to be =3.7 years.}  2. It is hypothesized that the population mean is 20 years, but we do not know.  3. So, we draw a random sample of size n and obtain a sample mean, X=X/n.  {Example: The random sample of 5 subjects resulted in the ages (24,22,28,17,24) which yields a sample mean, X=X/n = 23 years.}  4. We use the obtained value of the sample mean, X=23, to make an inference about the hypothesis that the population mean is =20.  5. So, we can argue that the closer X is to 20, the stronger the inference that =20. The farther away X is from 20, the weaker the inference that =20.  6. We use the evidence that X=23 to make an inference.  7. To make an inference, use “Test of Hypothesis” procedure. | | |

|  |  |  |
| --- | --- | --- |
|  | **Test of Hypothesis – Introduction & Procedure** |  |

“Test of Hypothesis” is part of inferential statistics that makes inferences from statistical evidence. In this development, the objective is to make an inference about the value of a population parameter, , using the value of a sample estimate, X.

|  |
| --- |
| 1. Suppose a random variable, X, follows a normal distribution with a known population standard deviation of 3.7, or X~N(,2) where =3.7.  2. We want to make an inference about the value of the population mean using a sample mean from a random sample. So we draw a random sample of n=5 that yields the data ( 24, 22, 28, 17, 24 ). From the data, the sample mean isX=23.  3. Consider the question:  “Do the data provide sufficient evidence to infer that the population mean is 20?” |

The “Test of Hypothesis” procedure.

|  |  |
| --- | --- |
| Null Hypothesis 🡪  Alternative Hypothesis 🡪  Level of Significance 🡪  Critical Value(s) 🡪  Rejection Region 🡪  Test Statistic 🡪  Inference 🡪 | Ho:  =   Ha:  / 20    CV: ± Z(0.05) = ± 1.645  RR: < –1.645 and > +1.645  TS: Zt = (23–20)/(3.7/sqrt(5))=+1.813  Reject Ho:  =  |
| Decision Rule: If TS is in the RR defined by CV, then Reject Ho. | |
| Conclusion: Since the test statistic, Zt=1.813, lies in the rejection region, Zt=1.813>+1.645, the data provide sufficient evidence to reject the hypothesis, Ho:=, at a 10% level of significance.  (Or: difference is statistically significant.) | |
| On the X axis: LCV= – Z.05\*/sqrt(n) = 20 – (1.645)\*(3.7)/sqrt(5) = 17.28  On the X axis: UCV= + Z.05\*/sqrt(n) = 20 + (1.645)\*(3.7)/sqrt(5) = 22.72 | |
| On the Z axis: LCV = norm.s.inv(0.05) ≈ –1.645 and UCV = norm.s.inv(0.95) ≈ +1.645. | |

Graphically,

|  |  |  |  |
| --- | --- | --- | --- |
| =0.05 |  |  | =0.05 |
|  |  |  |  |
| LCV  =  UCV X  17.28 22.72  TS = 23 | | | |
| LCV  =  UCV Z  –1.645 +1.645  TS = +1.813  . . . | | | |

|  |
| --- |
| Note the correct inferences:  When the TS (Test Statistic) is in the RR (Rejection Region), then Reject Ho  When the TS (Test Statistic) is NOT in the RR (Rejection Region), then DO NOT Reject Ho |

|  |  |  |
| --- | --- | --- |
|  | **Logic of Test of Hypothesis** |  |

|  |
| --- |
| **Introduction.**  Test of Hypothesis is based on a logical procedure that can be called the “Hypothetical-Deductive Method” of inference. This method establishes a rule based on the hypotheses where a conclusion can then be deductively obtained through the evidence. The method is considered inferential instead of deductive since the level of significance is somewhat arbitrarily chosen. The following discussion describes the procedure that is based on this approach. |
| **Discussion.**   1. Assume the Null Hypothesis is true and search for a contradiction. The “search” for a contradiction is the random sample that results in the sample estimate of the parameter. If the statistical data provide sufficient evidence to contradict the null hypothesis based on the critical value defined by the level of significance, then the inference is to “reject the null hypothesis”. If the statistical data do not provide sufficient evidence to reject the null hypothesis, then we do not reject the null hypothesis. Insufficient evidence does not support the null hypothesis, it merely is not sufficient to reject the null hypothesis. It is improper to ‘accept’ the null hypothesis because the null hypothesis has already been assumed to be true. Thus, if we assume the null hypothesis is true then accept the null hypothesis, that defines “circular reasoning” and a fallacy has been committed. Using fallacious arguments to support hypotheses can be severely attacked along with the statistician. Take care not to fall into that trap. Thus, proper inferences from the test of hypothesis procedure are to “Reject the Null Hypothesis” or “Do Not Reject the Null Hypothesis”. The inference to “Accept the Null Hypothesis” should never exist. However, if the inference is “Reject the Null Hypothesis”, then it is acceptable to state “reject the null hypothesis in favor of the alternative hypothesis” or “accept the alternative hypothesis” since the alterative hypothesis was not originally assumed to be true and now there is support for the alternative hypothesis. 2. Null Hypothesis(Ho) & Alternative Hypothesis(Ha).    1. Hypotheses only test parameters. Hypotheses do not contain statistics. Since parameters are constants, hypotheses to test the value of a parameter are appropriate. Statistics are random variables and a hypothesis to test the value of a random variable does not make sense in this development.    2. Equality is always in the null hypothesis. Since the null hypothesis is assumed to be true and equality is in the null hypothesis, then the procedure has a value for the parameter to test. If equality is not in the null hypothesis, then there is not one value to use in the procedure. 3. Level of Significance: =0.05. The “Level of Significance” is a term that applies to the “Significance” of the test. It is defined as the probability of “Type I Error” which is the probability of rejecting the null hypothesis when it is true. Thus, if the null hypothesis is true and the null hypothesis is rejected, that is Type I Error. The probability of Type I Error is the level of significance and usually expressed as . As the level of significance is decreased, the inference of ‘Reject the Null Hypothesis” becomes stronger since the probability of being incorrect is reduced. 4. Critical Value(CV) & Rejection Region(RR). The CV is defined by  and Ha. The RR is defined by CV and Ha. The CV and RR are used to define the decision rule that is used to make the inference of “Reject Ho” or “Do Not Reject Ho”. 5. Test Statistic(TS). The test statistic is determined from the data in the random sample. The test statistic can either be Xt or Zt. The test statistic is used to make the inference of “Reject Ho” or “Do Not Reject Ho” based on whether the test statistic is in the rejection region or not. |

. . .

|  |  |  |
| --- | --- | --- |
|  | **Formulation of Hypotheses** |  |

|  |  |  |
| --- | --- | --- |
| Rule 1: Only test parameters.  Rule 2: Equality in the null. | | |
| Do these data provide sufficient evidence to support the hypothesis that the population mean is … | Match | “Reject Ho” supports Ha  “Do Not Reject Ho” supports Ho |
| less than 20 ( < 20 )  no less than 20 ( >= 20 )  at least 20 ( >= 20 ) | Ho:  >= 20 one-tailed  Ha:  < 20 ‘lower’ |
| greater than 20 ( > 20 )  no more than 20 ( <= 20 )  at most 20 ( <= 20 ) | Ho:  <= 20 one-tailed  Ha:  > 20 ‘upper’ |
| different from 20 ( / 20 )  equal to 20 ( = 20 ) | Ho:  = 20 two-tailed  Ha:  / 20 |

|  |  |  |
| --- | --- | --- |
| Example 1 | Example 2 | Example 3 |
| Lower One-Tailed Test  Ho:  =   Ha:  < 20    CV: –Z(0.10) = –1.282  RR: < –1.282  TS: Zt = +1.813  Do Not Reject Ho:  =    |  |  |  |  | | --- | --- | --- | --- | |  |  |  |  | |  |  |  |  |   CV  –1.282  TS: Zt = +1.813 | Upper One-Tailed Test  Ho:  =   Ha: 20    CV: +Z(0.10) = +1.282  RR: > +1.282  TS: Zt = +1.813  Reject Ho:  =    |  |  |  |  | | --- | --- | --- | --- | |  |  |  |  | |  |  |  |  |   CV  +1.282  TS: Zt = +1.813 | Two-Tailed Test  Ho:  =   Ha:  / 20    CV: +/- Z(0.05) = +/- 1.645  RR: < –1.645 and > +1.645  TS: Zt = +1.813  Reject Ho:  =    |  |  |  |  | | --- | --- | --- | --- | |  |  |  |  | |  |  |  |  |   CV CV  –1.645 +1.645  TS: Zt = +1.813 |

**Equivalence**: The strongest inference would select the value of the parameter in the null hypothesis closest to the alternative hypothesis for comparison. Thus, an equivalent set of hypotheses are frequently used with no loss of accuracy.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | Ho:  >= 20  Ha:  < 20 | Is equivalent to | Ho:  = 20  Ha:  < 20 |  |
|  |  |  |  |  |
|  | Ho:  <= 20  Ha:  > 20 | Is equivalent to | Ho:  = 20  Ha:  > 20 |  |
|  |  |  |  |  |

|  |  |  |
| --- | --- | --- |
|  | **Formulation Examples** |  |

|  |  |
| --- | --- |
| 1. Weight Reducers International advertises that those who join will lose, on the average, at least 10 pounds the first two weeks. A random sample of people was drawn to test this claim. | Ho:  = 10  Ha:  > 10 |
| 2. Dole is concerned that its 16-ounce cans of sliced peaches are being overfilled. The quality-control department took a random sample to test this claim. | Ho:  = 16  Ha:  > 16 |
| 3. The new director of the local office of the Utah state unemployment service thought the mean wait time in line of 28 minutes to file a claim was much too long. Therefore, she instituted a number of changes to speed up the filing process. Three weeks later a random sample was selected to test the changes. | Ho:  = 28  Ha:  < 28 |
| 4. Peterson Automotive, a Honda dealership, recently stated in an advertisement that Honda owners average more than 85,000 miles before trading in or selling their Hondas. A random sample of Honda owners was selected to test this claim. | Ho:  = 85000  Ha:  > 85000 |
| 5. The director of a state agency claims that the average starting annual salary for clerical employees in the state is less than $30,000. She collected a random sample to test her claim. | Ho:  = 30000  Ha:  < 30000 |
| 6. Bowman Electronics sells electronic components for car stereos. They claim that the average life of a component exceeds 4,000 hours. They selected a random sample to test this claim. | Ho:  = 4000  Ha:  > 4000 |
| 7. The makers of a new home furnace system claim that if the furnace is installed, homeowners will observe an average fuel bill of less than $80.00 per month during January if their house has between 2,200 and 2,400 square feet of heated living space. A consumer agency plans to test this claim by taking a random sample of homes of this size where the new furnace has just been installed. | Ho:  = 80  Ha:  < 80 |

|  |  |  |
| --- | --- | --- |
|  | **p-Value or Observed Significance Level** |  |

|  |  |  |
| --- | --- | --- |
| Example 1 | Example 2 | Example 3 |
| Lower One-Tailed Test  Ho:  =   Ha:  < 20    CV: –Z(0.10) = –1.282  RR: < –1.282  TS: Zt = +1.813  Do Not Reject Ho:  =    |  |  |  |  | | --- | --- | --- | --- | |  |  |  |  | |  |  |  |  |   CV  –1.282  TS: Zt = +1.813 | Upper One-Tailed Test  Ho:  =   Ha: 20    CV: +Z(0.05) = +1.282  RR: > +1.282  TS: Zt = +1.813  Reject Ho:  =    |  |  |  |  | | --- | --- | --- | --- | |  |  |  |  | |  |  |  |  |   CV  +1.282  TS: Zt = +1.813 | Two-Tailed Test  Ho:  =   Ha:  / 20    CV: +/- Z(0.05) = +/- 1.645  RR: < –1.645 and > +1.645  TS: Zt = +1.813  Reject Ho:  =    |  |  |  |  | | --- | --- | --- | --- | |  |  |  |  | |  |  |  |  |   CV CV  –1.645 +1.645    TS: Zt = +1.813 |
| P-value:  P[Z<1.813]=0.9651  P-value is 0.9651 | P-value:  P[Z>1.813]=0.0349  P-value is 0.0349 | P-value:  2\*P[Z>1.813]=2\*0.0349  P-value is 0.0698 |

**Discussion.**

|  |
| --- |
| The P-value, or “observed significance level”, represents a boundary of the rejection region and can be compared with the level of significance, , for interpretation:  If P-value is less than , then reject Ho.  If P-value is greater than , then do not reject Ho. |
| A conclusion using a p-value usually states the inference and provides the p-value as support. To illustrate, the statement for each example above could be:  Example 1. “The null hypothesis Ho:= is not rejected (p=0.9651).”  Example 2. “The null hypothesis Ho:= is rejected (p=0.0349).”  Example 3. “The null hypothesis Ho:= is rejected (p=0.0698).” |
| Therefore,  Small p-values support “Reject the Null Hypothesis”  Large p-values support “Do Not Reject the Null Hypothesis” |

|  |  |  |
| --- | --- | --- |
|  | **Type I Error and Type II Error** |  |

|  |  |
| --- | --- |
| Null Hypothesis 🡪  Alternative Hypothesis 🡪  Level of Significance 🡪  Critical Value(s) 🡪  Rejection Region 🡪  Test Statistic 🡪  Inference 🡪 | Ho:  =   Ha:  / 20    CV: +/- Z(0.05) = +/- 1.645  RR: < –1.645 and > +1.645  TS: Zt = (23-20)/(3.7/sqrt(5))=+1.813  Reject Ho:  =  |
| Decision Rule: | If TS is in RR defined by CV, then reject Ho. |

|  |
| --- |
| If the Null Hypothesis, Ho, is TRUE and you Reject Ho, then that is a Type I Error.  And, P[Type I Error] = P[ TS is in RR | Ho is TRUE ] = a |
| If the Null Hypothesis, Ho, is NOT TRUE and you Do Not Reject Ho, then that is a Type II Error.  And, P[Type II Error] = P[ TS is NOT in RR | Ho is NOT TRUE ] = b |
| For the Inference of REJECT Ho :  (1) probability of being incorrect is the “Level of Significance” of the test or a  (2) probability of being correct is the “Power” of the test or 1–b |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | States of Nature | |  |
|  |  | Ho is True   =  | Ho is False   / 20 |  |
| Decision | Reject Ho | Type I Error  =P[Type I Error] | Correct Decision  (1–) | Level of Significance of a test is .  Power of a test is 1– |
| Do Not Reject Ho | Correct Decision  (1–) | Type II Error  =P[Type II Error] |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| For example, suppose Ho is false and m = 21. What is b=P[Type II Error]?  Thus, b=P[Type II Error] is the probability that the Test Statistic is in the Rejection Region when m = 21.  Find b = P[ 17.28 ≤ `X ≤ 22.72 | m = 21 ]  = P[ (17.28 – 21)/(3.7/sqrt(5)) ≤ (`X–m )/( s /sqrt(n)) ≤ (22.72–21)/ (3.7/sqrt(5)) ]  ≈ P[ –2.248 ≤ Z ≤ +1.039 ] ≈ 0.8383 | | | | |
| Ho: m = 20 a/2=0.05 |  |  | | a/2=0.05 |
|  |  |  | |  |
| LCV m = 20 UCV `X  17.28 22.72 | | | | |
|  | | | | |
| Ha:  / 20 ( m = 21 ) |  |  |  |  |
|  |  |  |  |  |
| LCV m = 21 UCV `X  17.28 22.72 | | | | |