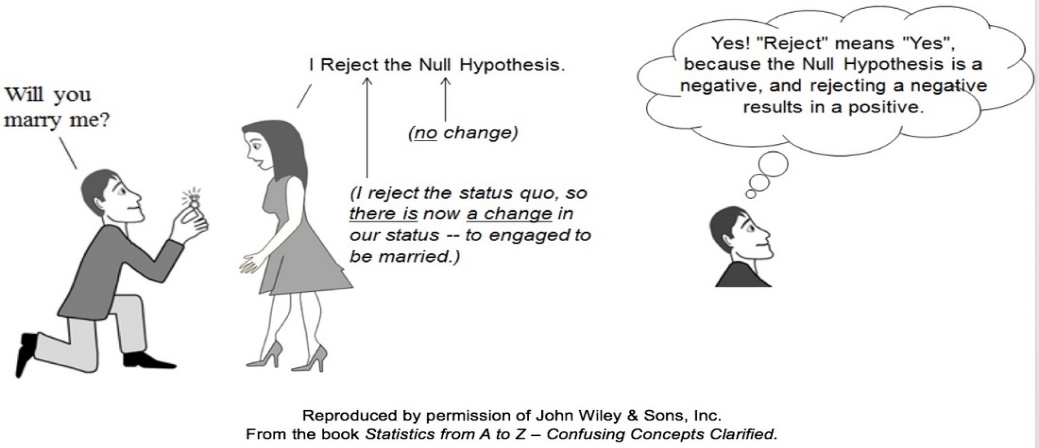
**MAT 3103: Computational Statistics and Probability**

**Chapter 10: Test of Hypothesis**



**Hypothesis test:**

A hypothesis test is a formal way to make a decision based on statistical analysis. A hypothesis test has the following general steps:

* Set up two contradictory hypotheses. One represents our assumption.
* Perform an experiment to collect data.
* Analyze the data using the appropriate distribution.
* Decide if the experimental data contradicts the assumption or not.
* Translate the decision into a clear, non-technical conclusion.

We will build up all the pieces we need, then put them together to test hypothesis.

**Parameter:** It is the unknown constant characteristic of the population observations. Function of population, population mean (*µ*), population variance () are parameters.

**Statistic:** It is the function of sample observations. Sample mean (), sample variance () are statistic.

**Hypothesis:** A hypothesis is a statement about one or more of parameter(s) of a population which we want to verify on the basis of information contained in a sample.

**Example:** Internet server claims that computer users in **AIUB** spend on the average 15 hours per week on browsing. We conduct a survey based on a sample of 250 users to arrive at a correct decision. Here, the server's claim is referred to as a hypothesis.

**Null hypothesis:** It is a statement which tells us that no difference exits between the parameter and the statistic, i.e.,given the test scores of two random [samples](http://en.wikipedia.org/wiki/Statistical_sample) , does one group differ from the other? A possible null hypothesis is,, [ = mean of population 1,  = mean of population 2].

**Alternative hypothesis:** The alternative hypothesis is the logical opposite of the null hypothesis. The rejection of a null hypothesis leads to the acceptance of the alternative hypothesis, i.e., Possible alternative hypothesizes are, , or, or, .

**Test statistic:** It is the function of sample observations which is used to verify the null hypothesis.

**Level of significance:** It is the probability with which we want to risk rejecting the null hypothesis even though it is true. We denote it by *α*; usually *α* = 0.05.

|  |  |
| --- | --- |
|  |  |

**Acceptance region:** If the value of the test statistic falls into the probability space of the distribution of the test statistic and lead us to accept the null hypothesis then the probability space is called the acceptance region.

**Critical Region:** The probability space in which the test statistic falls and leads us to reject the null hypothesis is called critical region or rejection region. In the given figure the critical and acceptance region are shown.

If *Z* ≥ Zα/2, *H*0 is rejected in favor of *H*1.

If Z ≤ Zα/2, *H*0 is rejected in favor of *H*1.

If l*Z*l ≥ Zα/2, *H*0 is rejected in favor of *H*1.

**Type I error:** It is the probability of rejecting the null hypothesis when the null hypothesis is true.

**Type II error:** It is the probability of accepting the null hypothesis when the null hypothesis is false.

**Example:** Consider a defendant in a trial. The null hypothesis is "defendant is not guilty;" the alternate is "defendant is guilty." A Type I error would correspond to convicting an innocent person; a Type II error would correspond to setting a guilty person free.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Reality** | |
| Not guilty | Guilty |
| **Verdict** | Guilty | **Type I Error:** Innocent goes jail | Correct Decision |
| Not guilty | Correct Decision | **Type II Error:** Guilty goes free |

**Test of hypothesis:**

It is the statistical process of verifying the null hypothesis using any test statistic. The steps are:

* State the null hypothesis, .
* State the alternative hypothesis, .
* Choose the level of significance, *α*.
* Select an appropriate test statistic.
* Calculate the value of the test statistic.
* Determine the critical region.
* Reject if the value of the test statistics falls in the critical region; otherwise accept .

**Test regarding one mean:**

Let be n sample observations drawn independently from a population with mean and variance . Let us assume that the sample observations follow normal distribution, i.e. *x ,* ). The problem is to test the null hypothesis : = against : .

**Assumptions:**

1. is unknown and n is small (), the test statistic is:
2. is unknown and n is large (), the test statistic is:
3. is known ( is small or large), the test statistic is:

**Example 10.1:** A random sample of 10 persons is selected and their level of education (in completed years of schooling) is recorded. The sample observations are, Do you think that the average schooling years of the persons in population is 5?

**Solution:** Let,, is unknown.

We need to test, vs .

.

Test statistic:. Since <. So is accepted.

Hence, we can conclude that the average schooling years can be considered as 5.

**Test of equality of two means:**

**Example 10.2:** The number of computer graduates coming out from two different universities A and B are employed in different organizations to do job related to computer.

|  |  |
| --- | --- |
| University | Number of graduates employed in computer related job |
|  |  |
|  | : |

Do you think that the employment facility for both the universities is similar?

**Solution:** Let, , . Also, let = =.

We need to test, vs.

Bothandare small and is not known.

Test statistic:.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ∑ | | | [ ∑–] (1898 – ) | |
| ∑ | | | [∑–] (2569 – ) | |
|  | | |  | |
| Since <, is accepted. Employment facility for students of both universities is same. | | | | |
| **Note** | If both and are large , then test statistic is: | |

**Test of equality of several means:**

**Example 10.3:** There are 25 computers in an office. These computers are supplied by 5 companies in lots of 5. During working hours each computer fails to work for some time (in hour). The failure time of different computers are shown below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | () | | | | | () | () |
| A | 1.5 | 1 | 0.5 | 2 | 3 | 8 | 1.6 |
| B | 1 | 0.5 | 0.2 | 0.5 | 0.5 | 2.7 | 0.54 |
| C | 0 | 2 | 2 | 0 | 0 | 4 | 0.8 |
| D | 2 | 2.5 | 1 | 2 | 3 | 10.5 | 2.1 |
| E | 3 | 3 | 2 | 2.5 | 2.5 | 13 | 2.6 |

Are the computers of different companies alike in failure time?

**Solution:** We need to test: = = = = vs At least one of them doesn’t hold.

|  |  |
| --- | --- |
|  |  |
|  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ANOVA TABLE** | | | | | |
| Sources of variation |  |  |  |  |  |
|  |  |  | = 3.73 |  |  |
|  |  |  | = 0.58 |
|  |  |  | | | |
| Hence, the computers of different companies are not similar in respect of failure time. | | | | | |

**Test regarding one proportion:**

Let be the values of the variable observed from population units, where

=

Let, number of units in the **population** who possess a particular characteristic

number of **sample** units possessing the characteristic under study.

So, proportion of **population** units who possess a particular characteristic.

proportion of **sample** units possessing the characteristic.

The problem is to test, : = : .

Test statistic,

**Example 10.4:** A sample of 15 students is selected from a group of 100 students and their grade in SSC examination is recorded as follows:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Students | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| Grade | B | C | A | D | B | C | D | A | B | C | D | B | C | C | D |

Do you think that 10% students get grade A?

**Solution:** We need to test, : vs : .

Test statistic:

Since < 1.96, is accepted. It can be considered that 10% students got grade A.

**Test regarding two proportions:**

**Example 10.5:** Is the severity of the drug problem in high school the same for boys and girls?  85 boys and 70 girls were questioned and 34 of the boys and 14 of the girls admitted to having tried some sort of drug.  What can be concluded at the 0.05 level?

**Solution:** We need to test vs .

Test statistic,

  0.4,     0.2,

Since > 1.96, is rejected. We conclude that gender does make a difference for drug use.

**Test regarding several proportions:**

**Example 10.6:** The following are the number of defective computers of different laboratories:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Laboratories |  |  |  |  | Total |
| No. of defective computers | 8 | 15 | 5 | 12 | 40 |

Are the proportions of defective computers of different laboratories similar?

**Solution:** We need to test = = = vs At least one of them doesn’t hold. Test Statistic,



Since <7.81, is accepted. Hence, the proportions of defective computers of different laboratories are similar.

**Test of independence:**

Let us consider that a researcher has n units in a sample. The sample observations are classified according to qualitative characters, say A and B as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| A B | B | Not B |  |
| A |  |  |  |
| Not A |  |  |  |
| Total () |  |  |  |

Here, are observed frequencies in different cells.

= observation of row and column recorded from the experiment.

= = expected frequency corresponding to row and column under

We need to test, : The characters A and B are independent vs : They are not independent.

Test statistic: . Here, *r* = no. of rows and *c* = no. of columns.

For *r* = 2 and *c* = 2, = ~ .

**Example 10.7:** 150 computer graduates are interviewed and are classified according to their result and job satisfaction. Do you think that the graduates with good result are satisfied with their job?

|  |  |  |  |
| --- | --- | --- | --- |
| Result | Job satisfaction | | Total |
| Yes | No |
| Good | 22 | 58 | 80 |
| Not good | 20 | 50 | 70 |
| Total | 42 | 108 | 150 |

**Solution:** : Job satisfaction does not depend on good result

: The good result and job satisfaction are associated

Test statistic: = =

Since < , is accepted. So, job satisfaction does not depend on good result.

**Example 10.8:** The following are number of emails of different organizations, where emails are classified according to local and foreign emails. The classified results are shown below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Origin of emails | Emails of organizations ( | | | | |
| O1 | O2 | O3 | O4 | Total |
| Local | 8 | 22 | 25 | 20 | 75 |
| Foreign | 17 | 73 | 25 | 10 | 125 |
| Total | 25 | 95 | 50 | 30 | 200 |

Is there any association between origin of emails and organizations?

**Solution:** : Origin of emails does not depend on organization

: Origin of emails depends on organization.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Origin of emails | Emails of organizations ( | | | | |
| O1 | O2 | O3 | O4 | Total |
| Local | 9.375 | 35.625 | 18.75 | 11.25 | 75 |
| Foreign | 15.625 | 59.375 | 31.25 | 18.75 | 125 |
| Total | 25 | 95 | 50 | 30 | 200 |

Test statistic:

Since > ,is rejected. Hence, origin of emails depends on organization.

|  |  |
| --- | --- |
| **Hints** | = = = 9.375 |

**Exercise 10**

**10.1** Among 157 African-American men, the mean systolic blood pressure was 146 mm Hg with a standard deviation of 27. We wish to know if on the basis of these data, we may conclude that the mean systolic blood pressure for a population of African-American is 140.

**10.2** Are the proportions of road accidents similar in various highways of Bangladesh?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Highways | 1 | 2 | 3 | 4 | Total |
| No. of road accidents (*Oi*) | 50 | 42 | 32 | 82 | 206 |

**10.3** Is there any association between subjects taught and job satisfaction?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Job satisfaction | Subjects taught (Oij) | | | | | Total |
| BBA | EEE | CS | CSE | SE |
| Yes | 12 | 22 | 18 | 15 | 20 | 87 |
| No | 18 | 32 | 30 | 15 | 25 | 120 |
| Total | 30 | 54 | 48 | 30 | 45 | 207 |

**10.4** Are the proportions of female students similar in various departments of AIUB?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Departments | 1 | 2 | 3 | 4 | Total |
| No. of female students (*Oi*) | 250 | 450 | 150 | 150 | 1000 |

**10.5** For a sample of size 36, ∑*x* = 761.6, ∑*x2* = 16125.5. Is the population mean 21?



**10.6** A company claims that its batteries have a mean life of 100 hours. You try to verify this for a sample of size 21 with mean 97 hours and variance 9 hours.



**10.7** Out of 25 students, 8 are female. Is the overall proportion of female students 0.40 in AIUB?

**10.8** A researcher claims that 10-year old children watch 6.6 hours of TV daily on average. You try to verify this for a sample of size 100 with mean 6.1 hours and standard deviation 2.5 hours.

**10.9** Is the probation problem the same for boys and girls at AIUB?  CGPA of 100 boys and 125 girls were randomly checked. 25 boys and 18 girls were under probation.

**10.10** The information about daily temperature (in 0 Celsius) of two months are as:

|  |  |  |  |
| --- | --- | --- | --- |
| Month 1 | = 31 | ∑ = 1032 | = 1.41 |
| Month 2 | = 30 | ∑ = 1035 | = 1.09 |

Do you think that the temperature of both the months are similar?



**10.11** Is high blood pressure associated with heart problem?

|  |  |  |
| --- | --- | --- |
| Blood Pressure | Heart Problem | |
| Yes | No |
| High | 150 | 120 |
| Not High | 122 | 158 |

**10.12** A group of students are classified by their residential origin and full attention to learning:

|  |  |  |  |
| --- | --- | --- | --- |
| Residential origin | Full attention | | Total |
| Yes | No |
| Rural | 138 | 64 | 202 |
| Urban | 64 | 84 | 148 |
| Total | 202 | 148 | 350 |

Is there any association between origin and full attention?

**10.13** Are the computers of different companies alike in failure time?

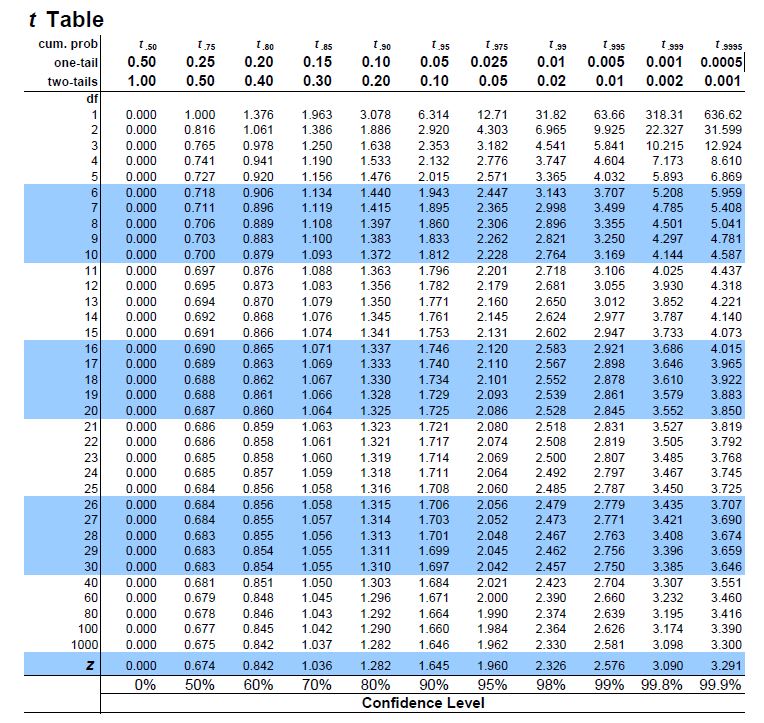
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Company | Failure time (in hour) of computers () | | | | |  |
| A | 120 | 122 | 110 | 125 | 118 | *∑∑*= 2167 |
| B | 100 | 105 | 108 | 110 | 100 |  |
| C | 95 | 90 | 80 | 90 | 85 | *∑∑*=238943 |
| D | 125 | 130 | 114 | 125 | 115 |  |

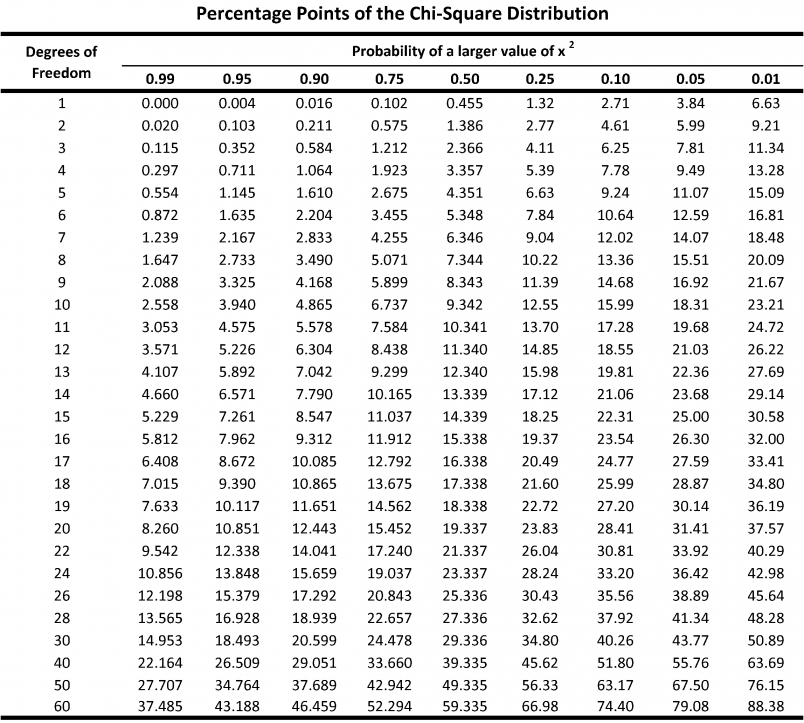
**10.14** The information about salary of two different institutions (in taka) are as:

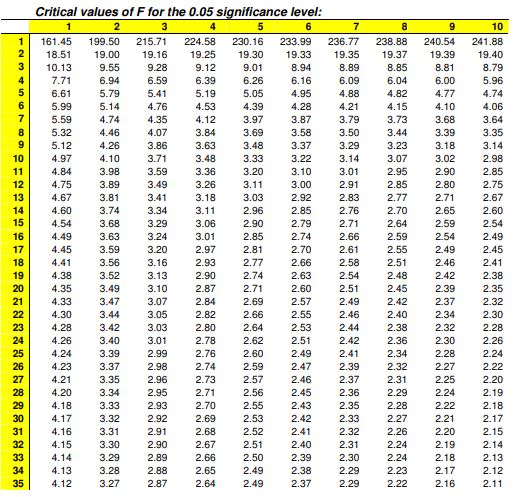
|  |  |  |  |
| --- | --- | --- | --- |
| Institution 1 | = 15 | ∑ = 16875 | = 5625 |
| Institution 2 | = 20 | ∑ = 26500 | = 50625 |

Do you think that the salary information of both institutions are similar?

**10.15** A toothpick manufacturer wants every box to contain exactly (on average) 500 toothpicks. In a random sample of 25 boxes, mean is 498 toothpicks and standard deviation is 9 toothpicks.







**Sample MCQs**

1. A quality control specialist took a random sample of n = 10 pieces of gum and measured their thickness and found the mean 7.6 and standard deviation 0.10. Do you think that the mean thickness of the spearmint gum it produces is 7.5?

a) Reject the null hypothesis

b) Accept the null hypothesis

c) Not concluded

d) None of the above

2. The information about daily temperature (in ° Celsius) of two months are as:

|  |  |  |  |
| --- | --- | --- | --- |
| Month 1 | 𝑛1 = 31 | Σ𝑥1 = 2042 | 𝑠12= 2.31 |
| Month 2 | 𝑛2 = 30 | Σ𝑥2 = 2045 | 𝑠22= 2.19 |

Do you think that the temperature of both the months are similar?

a) Null hypothesis is accepted

b) Null hypothesis is rejected

c) Both a and b

d) None of the above

3. The information about daily temperature (in ° Celsius) of two cities of different days are as:

|  |  |  |  |
| --- | --- | --- | --- |
| City 1 | 𝑛1 = 11 | Σ𝑥1 = 204 | 𝑠12= 4 |
| City 2 | 𝑛2 = 15 | Σ𝑥2 = 209 | 𝑠22= 5 |

Do you think that the temperature of both the cities are similar?

a) Null hypothesis is accepted

b) Null hypothesis is rejected

c) Both a and b

d) None of the above

4. Is gender independent of education level? A random sample of 395 people were surveyed, and each person was asked to report the highest education level they obtained. The data that resulted from the survey is summarized in the following table:

|  | **High School** | **Bachelors** | **Masters** | **Ph.d.** | **Total** |
| --- | --- | --- | --- | --- | --- |
| **Female** | 60(50.89) | 54(49.87) | 46(50.38) | 41(49.87) | 201 |
| **Male** | 40(49.11) | 44(48.13) | 53(48.62) | 57(48.13) | 194 |
| **Total** | 100 | 98 | 99 | 98 | 395 |

a) Null hypothesis is accepted

b) Null hypothesis is rejected

c) Both a and b

d) None of the above