



# **IQRA NATIONAL UNIVERSITY**

**Name: Muhammad Azam Khan**

**ID: 13130**

**Subject: Statistical Inference**

**Semester: 8th**

**Submitted to: Ma'am Wajiha Amin**

## Q1#

A sample of 25 boxes is taken with a standard deviation of \$237.52, construct 95% and 50% confidence interval for the variance of the whole population of boxes.

**Answer:**

$$n = 25, \quad S = 237.52 \quad S^2 = (237.52)^2 = 56415.7504$$

$$\text{For 95\% C.I} = 1 - 0.05 = 0.95$$

$$\alpha = 0.05$$

$$\alpha/2 = 0.025$$

$$1 - 0.025 = 0.975$$

$$n - 1 = 25 - 1 = 24$$

Critical Values

$$\alpha/2 = 0.025 = 39.364$$

$$1 - \alpha = 0.975 = 12.40$$

Now we use Formula:

$$\frac{[(n-1)\sqrt{S}]^2}{\sqrt{x}^{\alpha/2}} \leq \text{?}^2 \leq \frac{[(n-1)\sqrt{S}]^2}{\sqrt{x}^{1-\alpha/2}}$$

$$\frac{24 \times 56415.7504}{39.364} \leq \text{?}^2 \leq \frac{24 \times 56415.7504}{12.401}$$

$$\frac{1353978.01}{39.364} \leq \text{?}^2 \leq \frac{1353978.01}{12.401}$$

$$34396.35225 \leq \text{?}^2 \leq 109182.9699$$

---

**For 50% C.I:**


---

$$\alpha = 0.5$$

$$\alpha/2 = 0.25$$

$$1 - \alpha = 0.75$$

$$\text{Critical Values} \quad \longrightarrow \quad \begin{array}{l} 0.25 = 28.241 \\ 0.75 = 19.037 \end{array}$$

Using Formula:

$$\frac{[(n-1)\sqrt{S}]^-}{\sqrt{\bar{x}}^{\frac{\alpha}{2}}} \leq \diamond^2 \leq \frac{[(n-1)\sqrt{S}]^-}{\sqrt{\bar{x}}^{1-\frac{\alpha}{2}}}$$

$$\frac{24 \times 56415.7504}{28.241} \leq \diamond^2 \leq \frac{24 \times 56415.7504}{19.037}$$

$$\frac{1353978.01}{28.241} \leq \diamond^2 \leq \frac{1353978.01}{19.037}$$

$$47943.69923 \leq \diamond^2 \leq 71123.496$$

**Q2#**

Variability in computer usage of two departments in a large company is tested. A random sample of 5 with the standard deviation of 3.4 is taken from accounting department while a sample of 6 with standard deviation of 2.6 is taken from research department. Construct 99% and 80% confidence intervals for two population variances.

**Answer:**

$$n_1 = 5, \quad S_1 = 3.4$$

$$n_2 = 6, \quad S_2 = 2.6$$

$$(S_1)^2 = (3.4)^2 = 11.56 \quad V_1 = n_1 - 1 = 5 - 1 = 4$$

$$(S_2)^2 = (2.6)^2 = 6.76 \quad V_2 = n_2 - 1 = 6 - 1 = 5$$

---

**For 99% C.I**


---

$$\alpha = 0.01$$

$$\alpha/2 = 0.005$$

$$V_1, V_2 = 15.56$$

$$V_2, V_1 = 22.46$$

## Q3#

A retailer found that the mean credit card sales for the last year was \$ 60 with the standard deviation of \$12. Manager wants to analyze average per transaction cost for the sales. To test for any change in the mean sales, retailer took a sample of 225 credit cards and their mean was \$65. If level of significance was at 2%, test the hypothesis for change in mean sales. Also compute the p value.

**Answer:**

$$\mu = 60$$

$$\sigma = 12$$

$$n = 225$$

$$\bar{x} = 65$$

$$\alpha = 2\% = 0.02$$

$$1) H_0: \mu = 60$$

$$H_1: \mu \neq 60$$

$$2) \alpha = 0.02$$

$$\alpha/2 = 0.01$$

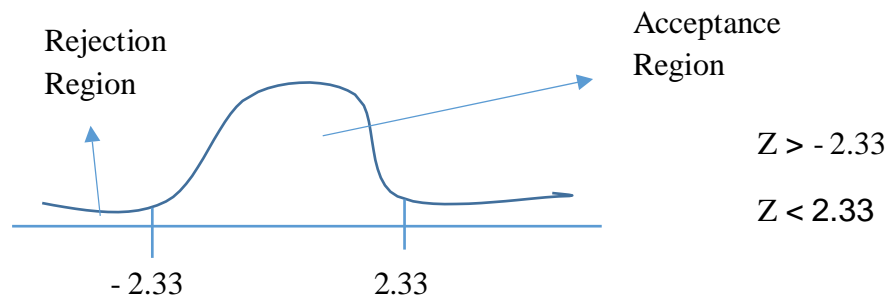
$$1 - 0.01 = 0.99$$

Finding "Z" From Table  $Z = 2.33$

$$3) Z - \text{Test}$$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$4)$$



$$5) Z = \frac{65 - 60}{12/\sqrt{225}} = \frac{5}{0.8} = 6.25$$

this value 6.25 as its greater than 2.33 so it is in the rejection region, so we reject  $H_0$  accept  $H_1$  for example ( $\mu \neq 60$ ) at 2% significant level.

Now P Value:

$$Z = 6.00$$

$$P(Z > 6.00) = 0.9999$$

$$P(Z < 6.00) \text{ 6.25 is not in table}$$

$$1 - 0.9999 = 0.0001$$

$$P \text{ value} < \alpha$$

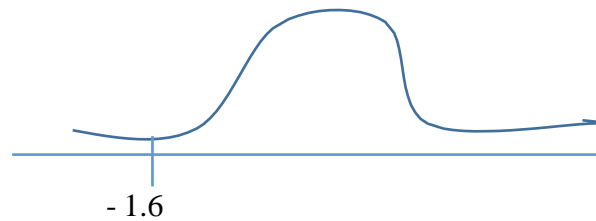
1.  $P < 0.02$  so we reject  $H_0$  and accept  $H_1$ .

## Q4#

If following are the results for hypothesis testing, analyze whether you will accept or reject null hypothesis.

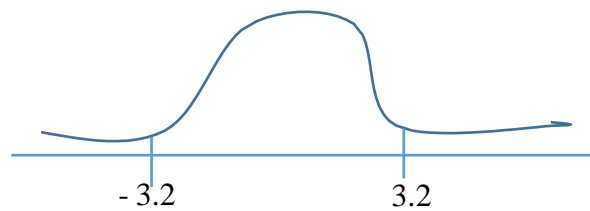
Value of test statistic	Value for critical region	Type of test
-1.5	-1.6	Left tail
2.9	3.2	Two tail

**Answer:**



$$Z = -1.5$$

we will accept  $H_0$  because value is in acceptance region.



$$Z = 2.9$$

As value is in acceptance region so we accept the  $H_0$

**Q5#**

The manager of a paint supply store claims that the mean paint volume in their containers is 1 gallon. If 20 containers are tested to check that mean content in reality is less than 1 gallon with sample standard deviation of 0.02 gallon. State

- ✚ The null and alternate hypothesis
- ✚ Type of test statistic used and why?
- ✚ Nature of test (1 tail or 2 tail) and why? draw diagram.

**Answer:**

$\mu = 1, \quad n = 20, \quad S = 0.02$

Now  $H_0: \mu = 1$

$H_1: \mu < 1$

t-test will be used because “ $\sigma$ ” is unknown and sample standard deviation is known.

Now we use:

$$\frac{11.56}{6.76} \times \frac{1}{15.56} < \frac{?}{2} < \frac{11.56}{6.76} \times 22.46$$

$$0.1099 < \frac{?}{2} < 38.408$$

**For 90% C.I**

$\alpha = 0.10$   
 $\alpha/2 = 0.05$

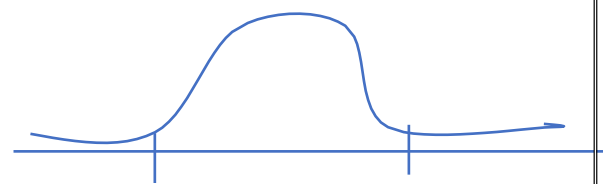
$V_1, V_2 = 7.39$   
 $V_2, V_1 = 9.36$

$$\frac{11.56}{6.76} \times \frac{1}{7.39} < \frac{?}{2} < \frac{11.56}{6.76} \times 9.36$$

$$0.231 < \frac{?}{2} < \frac{11.56}{6.76} < 16.006$$

Nature of test is one tail because  $H_1$  value is on the negative side of the diagram or distribution it is left tail.

t or z value





**Q6#**

A fast food chain has developed a new process to improve the service of drive through orders for their customers. Previous process filled the orders 88% of the time. A sample of 100 orders using new process shows that around 93 were filled correctly. At 0.01 level of significance, can you conclude that the new process has increased the proportion of orders filled correctly? Also compute the p value and suggest your decision.

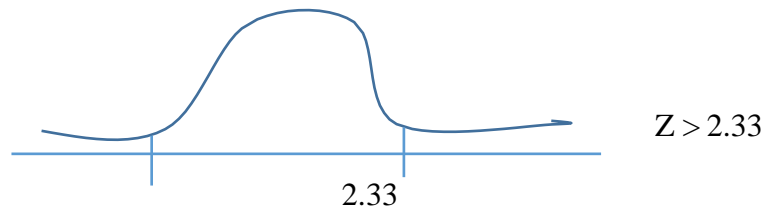
**Answer:**

$\pi = 88\% = 0.88$        $n = 100$        $P = \frac{93}{100} = 0.93$

1.  $H_0 \pi = 0.88$   
 $H_1 \pi > 0.88$

2.  $\alpha = 0.01$        $1 - 0.01 = 0.99$   
 $Z = 2.33$

3. Z - test       $Z = \frac{P - \pi}{\sqrt{\pi(1-\pi)}}$



4. 
$$Z = \frac{0.93 - 0.88}{\frac{\sqrt{0.88(1-0.88)}}{100}}$$

$$\frac{0.05}{\frac{\sqrt{0.88 \times 0.12}}{100}} = \frac{0.05}{\sqrt{0.001056}} = 1.5386$$

we will accept  $H_0$  and reject  $H_1$  because value 1.538 is less than 2.33 and it is in acceptance region.  $\pi > 0.88$

**5. P - value**

$P(Z > 1.5386)$   
 $P(Z > 1.54) = 0.938$   
 $1 - 0.938 = 0.062$

P value  $> \alpha$   
 So we accept  $H_0$  and reject  $H_1$

## Q7# Theory

Q1# What is “paired t-test” and when it is used?

### Answer:

Two-sample t-tests are statistical tests used to compare the means of two populations. This procedure provides several reports for making inference about the difference between two population means based on a paired sample. These reports include confidence intervals of the mean difference, the paired sample t-test, and non-parametric tests including the randomization test, the quantile (sign) test, and the Wilcoxon Signed-Rank test. Tests of assumptions and distribution plots are also available in this procedure.

A paired t-test (also known as a dependent or correlated t-test) is a statistical test that compares the averages means and standard deviations of two related groups to determine if there is a significant difference between the two groups.

A significant difference occurs when the differences between groups are unlikely to be due to sampling error or chance.

The groups can be related by being the same group of people, the same item, or being subjected to the same conditions.

### When to use a paired t-test?

Paired t-tests are used when the same item or group is tested twice, which is known as a repeated measures t-test. Some examples of instances for which a paired t-test is appropriate include:

- The before and after effect of a pharmaceutical treatment on the same group of people.
- Body temperature using two different thermometers on the same group of participants.
- Standardized test results of a group of students before and after a study prep course.

Q2# Difference between a normal t-test and a paired t-test in statistics?

### Normal t – test:

A Normal t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. It is mostly used when the data sets like the data set recorded as the outcome from flipping a coin 100 times would follow a normal distribution and may have unknown variances. A Normal t-test is used as a hypothesis testing tool which allows testing of an assumption applicable to a population.

A Normal t-test looks at the t-statistic the t-distribution values and the degrees of freedom to determine the statistical significance. To conduct a test with three or more means one must use an analysis of variance.

### Paired t – test:

A Paired t – test is also known as a correlated t-test which performed when the samples typically consist of matched pairs of similar units or when there are cases of repeated measures.

**For example:** there may be instances of the same patients being tested repeatedly before and after receiving a particular treatment. In such cases each patient is being used as a control sample against themselves.

This method also applies to cases where the samples are related in some manner or have matching characteristics like a comparative analysis involving children, parents or siblings. Correlated or paired t-tests are of a dependent type as these involve cases where the two sets of samples are related.

## Differentiate with the help of Example

### Normal t –test

Suppose we want to test the hypothesis that two samples have the same mean values for example:  $H_0: \mu_0 = \mu_1$ . In the following discussion we assume the data follows bivariate normal distribution. The normal t-test is of the form sample mean difference divided by sample standard deviation of the sample mean difference

$$\frac{\text{sample mean difference}}{\text{sample standard deviation of the sample mean difference}}$$

### Paired t – test:

The paired t-test is of the form

$$T_2 = \frac{\bar{X}_d}{\sqrt{\frac{1}{n(n-1)} \sum_{j=1}^n (X_{j\#} - \bar{X}_d)^2}}$$

It's obvious that the paired t-test is exactly the one-sample t-test based on the difference within each pair. Under the null hypothesis  $T_2$  always follows t-distribution with  $df = n-1$ .