Lean Six Sigma Training Ltd

Lean Six Sigma

Black Belt

Online Study Guide

Section 4.2 : Hypothesis Testing

Testing of Hypothesis

- A hypothesis is an assumption about the population parameter.
- Examples of parameters are population mean or proportion.

Null hypothesis

- Hypothesis about the population to be verified
- It is always about a population parameter and not about a sample statistic
- Denoted by H_o

Alternative hypothesis

- Hypothesis to be accepted when null hypothesis is not true. It is the opposite of null hypothesis
- Denoted by H_1 or H_a

Example: To check mean is equal to 15

- Null hypothesis: H_o: Mean = 15
- Alternative hypothesis: H_1: Mean ≠ 15 or Mean > 15 OR Mean < 15

Testing of Hypothesis

There are four steps in statistical inference namely:

a. Formulation of hypothesis regarding the population.

b. Collection of sample observations from the population.

c. Calculation of statistics based on the sample.

d. Acceptance or rejection of the hypothesis depending on the predetermined acceptance criterion.

Statistical inference generally renders itself to two types of errors. They are:

- Type I Error : Also known as an **error of the first kind** or an **α error**, the **probability** of rejecting a null hypothesis when it is actually true. (Also known as **producer's risk**)
- Type II Error : Also known as an **error of the second kind** or a **β error**, the probability of failing to reject a null hypothesis when it is false. (Also known as **consumer's risk**)

Testing of Hypothesis

Definition of terms

- Statistical test : A decision function that takes its values in the set of hypothesis.
- Region of acceptance : The set of values for which we fail to reject the null hypothesis.
- Critical region : The set of values of the test statistic for which the null hypothesis is rejected.
- Power of a test (1 – β) : the probability of not failing to reject a null hypothesis when it is false.
- Significance level of a test (α) : this the same as we defined in type 1 error.
- P – Value: In statistical hypothesis, testing the p-value is the probability of obtaining a result at least as extreme as the one that was actually observed given that the null hypothesis is true. The fact that p-values are based on this assumption is crucial to their correct interpretation.

****Note: One rejects the null hypothesis if the p-value ≤ the significance level.**

- Hypothesis test enables us to make an inference about the true population value at a desired level of confidence.

Testing of Hypothesis

Statistical Vs Practical significance

- In some situations, it may be possible to detect a statistically significant difference between two populations when there is no practical difference.

Example:

- Suppose that a test is devised to determine whether there is a significance difference in the surface finish when a lathe is operated at 400 rpm and at 700 rpm.
- If large sample sizes are used, it may be possible to determine that 400 rpm population has a tiny but statistically significant improved surface.
- However, if both speeds produce surface finishes that are capable of meeting the specifications, the best decision might be to go with the faster speed because of its associated increase in throughput.
- Thus, the difference between two populations, although statistically significant, must be weighted against other economic and engineering considerations.

Ref: The Certified Six Sigma Black Belt Handbook by Donald W. Benbow & T M Kubaik

Testing of Hypothesis

Point and Interval Estimation

- Suppose an estimate is required for the average diameter for a population of 1200 bolts received form a supplier.
- It is realistic that in spite of measuring all the 1200 bolts, one might randomly select a sample of 50 for measurements.
- The average and standard deviation turns out to be 1.24 and 0.006 respectively.
- Hence, the estimate for the average diameter on the entire lot of 1200 is around 1.24. This value is called the **point estimate**.
- In this case the sample mean (statistic) is an **estimator** of the population mean (parameter). Recall that the **statistic** is a value obtained from a sample while a **parameter** is a value from the population.
- The standard deviation of the distribution of means indicates the **amount of error** that will occur when a sample mean is used for estimating a population mean, known as the **standard error** (SE).

Testing of Hypothesis

Confidence Intervals for the mean μ of a population

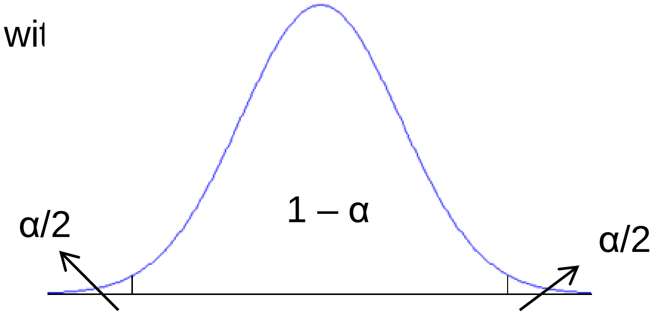
- Consider the previous example, is the population mean exactly 1.24?
- Probably not because of sampling error i.e. if some other sample is taken the average may not be the same as 1.24
- Hence a technique is required to determine how good this point estimate is. That technique is called the **confidence interval (CI)**.
- Example: after some calculation, it can be stated that “We are **95% confident** that the population mean is between 1.238 and 1.242” or equivalently 95% confidence interval for the population mean is (1.238, 1.242).

Testing of Hypothesis

Symbols used

Six Sigma Score	Sample	Population
Mean	$\bar{x}$	μ
Number of values (size)	n	N
Standard deviation	s	σ

- α = probability that the population mean is not falling in the interval (α-risk)
- $1 - \alpha$ = probability that the population mean is in the interval (confidence level)
- $Z_{\alpha/2}$ = the value from the Z-table (standard normal table) wit



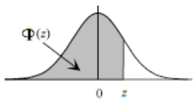
Testing of Hypothesis

Solution of the problem:

- In this problem, $\bar{x} = 1.24$, $n = 50$, $N = 1200$, $1-\alpha = 0.95$, $\alpha = 0.05$, $s = 0.006$
- Formulas for the endpoints of the CI are: $\mathbf{\bar{x} + Z_{\alpha/2}\,(s/\sqrt{n})}$
- $\bar{x} + Z_{\alpha/2}\,(s/\sqrt{n}) = 1.24 + Z_{\alpha/2}\,(0.006/\sqrt{50}) = 1.24 + 1.96 \times (0.006/\sqrt{50}) = (1.2383, 1.2417)$

[From the Z-table, the z-value with 0.025 to its right is 1.96]

- Therefore we are 95% confident that the population mean μ is between 1.2383 and 1.2417



z	Φ(z)	z	Φ(z)	z	Φ(z)	z	Φ(z)	z	Φ(z)	z	Φ(z)	z	
1.20	0.88493	1.60	0.94520	2.00	0.97725	2.40	0.99180	2.80	0.99744	3.20	0.99931	3.60	0
1.21	0.88686	1.61	0.94630	2.01	0.97778	2.41	0.99202	2.81	0.99752	3.21	0.99934	3.61	0
1.22	0.88877	1.62	0.94738	2.02	0.97831	2.42	0.99224	2.82	0.99760	3.22	0.99936	3.62	0
1.23	0.89065	1.63	0.94845	2.03	0.97882	2.43	0.99245	2.83	0.99767	3.23	0.99938	3.63	0
1.24	0.89251	1.64	0.94950	2.04	0.97932	2.44	0.99266	2.84	0.99774	3.24	0.99940	3.64	0
1.25	0.89435	1.65	0.95053	2.05	0.97982	2.45	0.99286	2.85	0.99781	3.25	0.99942	3.65	0
1.26	0.89617	1.66	0.95154	2.06	0.98030	2.46	0.99305	2.86	0.99788	3.26	0.99944	3.66	0
1.27	0.89796	1.67	0.95254	2.07	0.98077	2.47	0.99324	2.87	0.99795	3.27	0.99946	3.67	0
1.28	0.89973	1.68	0.95352	2.08	0.98124	2.48	0.99343	2.88	0.99801	3.28	0.99948	3.68	0
1.29	0.90147	1.69	0.95449	2.09	0.98169	2.49	0.99361	2.89	0.99807	3.29	0.99950	3.69	0
1.30	0.90320	1.70	0.95543	2.10	0.98214	2.50	0.99379	2.90	0.99813	3.30	0.99952	3.70	0
1.31	0.90490	1.71	0.95637	2.11	0.98257	2.51	0.99396	2.91	0.99819	3.31	0.99953	3.71	0
1.32	0.90658	1.72	0.95728	2.12	0.98300	2.52	0.99413	2.92	0.99825	3.32	0.99955	3.72	0
1.33	0.90824	1.73	0.95818	2.13	0.98341	2.53	0.99430	2.93	0.99831	3.33	0.99957	3.73	0
1.34	0.90988	1.74	0.95907	2.14	0.98382	2.54	0.99446	2.94	0.99836	3.34	0.99958	3.74	0
1.35	0.91149	1.75	0.95994	2.15	0.98422	2.55	0.99461	2.95	0.99841	3.35	0.99960	3.75	0
1.36	0.91308	1.76	0.96080	2.16	0.98461	2.56	0.99477	2.96	0.99846	3.36	0.99961	3.76	0
1.37	0.91466	1.77	0.96164	2.17	0.98500	2.57	0.99492	2.97	0.99851	3.37	0.99962	3.77	0
1.38	0.91621	1.78	0.96246	2.18	0.98537	2.58	0.99506	2.98	0.99856	3.38	0.99964	3.78	0
1.39	0.91774	1.79	0.96327	2.19	0.98574	2.59	0.99520	2.99	0.99861	3.39	0.99965	3.79	0
1.40	0.91924	1.80	0.96407	2.20	0.98610	2.60	0.99534	3.00	0.99865	3.40	0.99966	3.80	0
1.41	0.92073	1.81	0.96485	2.21	0.98645	2.61	0.99547	3.01	0.99869	3.41	0.99968	3.81	0
1.42	0.92220	1.82	0.96562	2.22	0.98679	2.62	0.99560	3.02	0.99874	3.42	0.99969	3.82	0
1.43	0.92364	1.83	0.96638	2.23	0.98713	2.63	0.99573	3.03	0.99878	3.43	0.99970	3.83	0
1.44	0.92507	1.84	0.96712	2.24	0.98745	2.64	0.99585	3.04	0.99882	3.44	0.99971	3.84	0
1.45	0.92647	1.85	0.96784	2.25	0.98778	2.65	0.99598	3.05	0.99886	3.45	0.99972	3.85	0
1.46	0.92785	1.86	0.96856	2.26	0.98809	2.66	0.99609	3.06	0.99889	3.46	0.99973	3.86	0
1.47	0.92922	1.87	0.96926	2.27	0.98840	2.67	0.99621	3.07	0.99893	3.47	0.99974	3.87	0
1.48	0.93056	1.88	0.96995	2.28	0.98870	2.68	0.99632	3.08	0.99896	3.48	0.99975	3.88	0
1.49	0.93189	1.89	0.97062	2.29	0.98899	2.69	0.99643	3.09	0.99900	3.49	0.99976	3.89	0
1.50	0.93319	1.90	0.97128	2.30	0.98928	2.70	0.99653	3.10	0.99903	3.50	0.99977	3.90	0
1.51	0.93448	1.91	0.97193	2.31	0.98956	2.71	0.99664	3.11	0.99906	3.51	0.99978	3.91	0
1.52	0.93574	1.92	0.97257	2.32	0.98983	2.72	0.99674	3.12	0.99910	3.52	0.99978	3.92	0
1.53	0.93699	1.93	0.97320	2.33	0.99010	2.73	0.99683	3.13	0.99913	3.53	0.99979	3.93	0
1.54	0.93822	1.94	0.97381	2.34	0.99036	2.74	0.99693	3.14	0.99916	3.54	0.99980	3.94	0
1.55	0.93943	1.95	0.97441	2.35	0.99061	2.75	0.99702	3.15	0.99918	3.55	0.99981	3.95	0
1.56	0.94062	1.96	0.97500	2.36	0.99086	2.76	0.99711	3.16	0.99921	3.56	0.99981	3.96	0
1.57	0.94179	1.97	0.97558	2.37	0.99111	2.77	0.99720	3.17	0.99924	3.57	0.99982	3.97	0
1.58	0.94295	1.98	0.97615	2.38	0.99134	2.78	0.99728	3.18	0.99926	3.58	0.99983	3.98	0
1.59	0.94408	1.99	0.97670	2.39	0.99158	2.79	0.99736	3.19	0.99929	3.59	0.99983	3.99	0
1.60	0.94520	2.00	0.97725	2.40	0.99180	2.80	0.99744	3.20	0.99931	3.60	0.99984	4.00	0

Testing of Hypothesis

Margin of Error and Sample size

- The margin of error (E) is defined as $E = Z_{\alpha/2} (\sigma/\sqrt{n})$
- It can be noted that the margin of error (E) decreases as the sample size (n) increases which in turn improves the precision of the estimate.
- If the margin of error (E) and confidence level (1-α) is known in advance, then the required sample size (n) can be determined by the formula for n which is given below:

$$n = (\sigma Z_{\alpha/2} / E)^2$$

- rounded up to a whole number since sample size can not be a fraction

Testing of Hypothesis

Confidence Interval conditions

- What we should do if population standard deviation is not known (which is very realistic)?
- In those cases we use the sample standard deviation **s** as an estimate of the population standard deviation **σ**. (Note that in the previous problem of CI, this concept is explained)
- Equations of CI are given below for various conditions:

CI for	Conditions	Equations
Population Mean (μ)	Large sample (n ≥ 30) when σ known	$\bar{x} - z_{\frac{\alpha}{2}} \times \left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \left(\frac{\sigma}{\sqrt{n}}\right)$
Population Mean (μ)	Large sample (n ≥ 30) when σ unknown	$\bar{x} - z_{\frac{\alpha}{2}} \times \left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \left(\frac{s}{\sqrt{n}}\right)$
Population Mean (μ)	Small sample (n < 30) when σ unknown and the population is normal	$\bar{x} - t_{\frac{\alpha}{2}, n-1} \times \left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{\frac{\alpha}{2}, n-1} \times \left(\frac{s}{\sqrt{n}}\right)$
Population standard deviation (σ)	Population is normal	$\sqrt{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}, n-1}}} \le \sigma \le \sqrt{\frac{(n-1)s^2}{\chi^2_{\left(1-\frac{\alpha}{2}\right), n-1}}}$

Testing of Hypothesis

Types of tests

Continuous data

- Population mean equal to a specified value
- Two population means are equal or not
- Population standard deviation equal to a specified value
- Two population standard deviations (or variances) are equal or not

Discrete data

- Proportion equal to a specified value
- Two proportions are equal or not

Testing of Hypothesis

Name	Formula	Assumptions	When to use
One sample z -test	$z = \frac{\overline{x} - \mu_0}{(\sigma/\sqrt{n})}$	Normal distribution or n ≥ 30 and population variance σ is known	To test Mean = μ_0 when population variance σ is known
One sample t-test	$t = \frac{\overline{x} - \mu_0}{(s/\sqrt{n})}$, With $df = n - 1$	Normal distribution and population standard deviation σ is unknown Sample size (n) < 30	To test Mean = μ_0 when population variance σ is unknown
Two sample z-test	$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	Normal populations , independent observations, and both σ_1 & σ_2 are known When population variances (σ_1 & σ_2) are not known use sample standard deviation as an estimate of the population standard deviation (i.e. use s_1 & s_2 instead of σ_1 & σ_2) Large samples (i.e. n ≥ 30)	To test $\mu_1 = \mu_2$ both population variances are known

Testing of Hypothesis

Name	Formula	Assumptions	When to use
Two sample pooled t-test	$t = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, With $df = n_1 + n_2 - 2$ Where $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$,	Normal populations, independent observations & $\sigma_1 = \sigma_2$ (= σ say, which is unknown, estimated by s_p)	To test $\mu_1 = \mu_2$ both population variance is unknown
Paired t-test	$t = \frac{\overline{d}}{\frac{s_d}{\sqrt{n}}}$ with df = n - 1	Paired sample Large samples (i.e. n ≥ 30) or differences are normally distributed	To test the equality of two population mean

Lean Six Sigma Training Ltd

Testing of Hypothesis

Name	Formula	Assumptions	When to use
One proportion test	$z = \frac{p' - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$np_0 \geq 5$ and $n(1 - p_0) \geq 5$	To test proportion is equal to a specified value
Two proportion test	$z = \frac{p_1' - p_2'}{\sqrt{p_p'(1-p_p')\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ Where $p_p' = \frac{x_1 + x_2}{n_1 + n_2}$	Independent observations $x_1 \geq 5$, $n_1 - x_1 \geq 5$, $x_2 \geq 5$, $n_2 - x_2 \geq 5$	To test two proportions are equal or not
Single variance chi-square test	$\chi^2 = \frac{s^2(n-1)}{\sigma_0^2}$	Normal population	To test a population standard deviation (or variance) is equal to a specific value
F Test	$F = s^2_1/s^2_2$	Normal populations, independent random samples	To test two variances are equal or not

Testing of Hypothesis

Steps for testing of hypothesis:

- Check for the conditions required for the test to be met or not
- State the null and the alternative hypothesis and identify whether it is a one-tailed or two-tailed test
- Determine the α-value (significance level) which is similar to the use of α in confidence intervals.
- Identify the test statistic applicable for the hypothesis test and calculate its value. Some of the inputs to the formulas come from the sample data.
- Determine the critical values, typically found in tables such as Z, t, F or χ^2 . Use these values to define the critical region (in turn reject region).
- Determine whether the null hypothesis should be rejected at the pre-determined level of significance (α).
    - If the value of the test statistic falls in the reject region then the null hypothesis is rejected and the alternative hypothesis is accepted.
    - If the value of the test statistic does not fall in the reject region then the null hypothesis is not rejected.

1. State the inference in terms of the original problem

Testing of Hypothesis

Examples of hypothesis testing that can be used in Six Sigma applications

1) To test Mean = Specified value

- A Six Sigma GB project was carried out to reduce the average cycle time to 14 minutes. The following data on cycle time is collected after the project is completed. Test whether the project team is successful in reducing the average cycle time to 14 minutes with 95% confidence. Assume that the cycle times are normally distributed.

14.1	14.4	13.8	14.2	14.3
13.7	14.2	13.6	14.5	14.1

Steps

1. Population is normal, sample size (n) = 10 < 30 and population standard deviation (σ) is unknown, all the conditions for one-sample t test are satisfied
2. Ho: Mean = 14 (Null hypothesis), Ha: Mean > 14 (Alternative hypothesis). This is a one tailed test.
3. From the problem, 1 - α = 0.95 which gives α = 0.05
4. The test statistic for this test is t with (n-1) degrees of freedom. Test statistic t: t = 0.949269

Testing of Hypothesis

Determining the Critical Region

Table value of t corresponding to the t statistic with DF = 9 & α = 0.05 is 1.833

Critical Region: t > 1.833

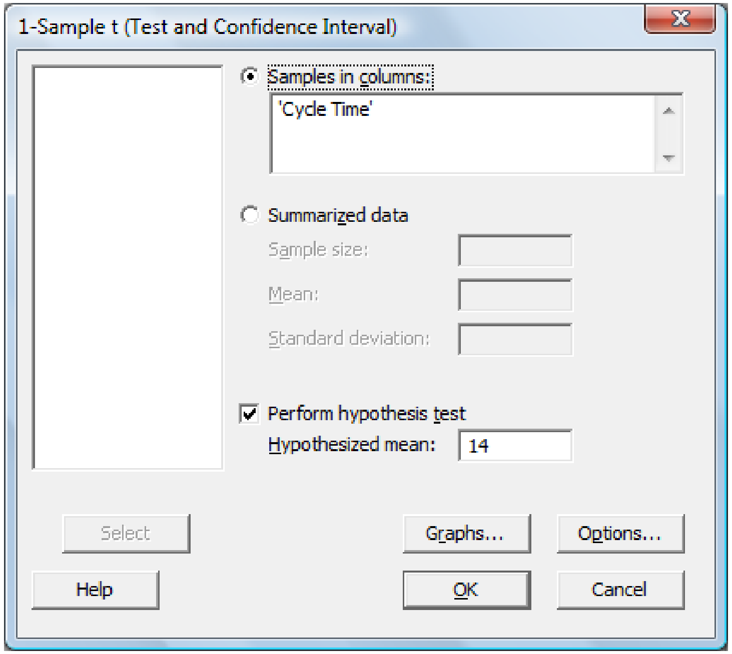
Inference: Since the computed value of t is not falling in the critical region, we accept the null hypothesis at the 5% level of significance. Hence, the GB project team is able to succeed in reducing the average cycle time.

With the help of MINITAB

Enter the data in the MINITAB worksheet

Select Stat → Basic Statistics → 1 sample t

Click OK. Minitab will give the output as :



One-Sample T: Cycle Time

Test of mu = 14 vs not = 14

Variable	N	Mean	StDev	SE Mean	95% CI	T	P
Cycle Time	10	14.0900	0.2998	0.0948	(13.8755, 14.3045)	0.95	0.367

Testing of Hypothesis

Note that p-value is 0.367 > 0.05, we accept the null hypothesis.

2) To test if two means are equal

A new methodology is introduced in a process to reduce errors. Though the new methodology considerably reduces errors, there is a feeling among the agents that the time in new methodology is more than that of the old one. 10 random samples on time is taken from both the methodologies. Check whether the average time is the same across methodologies.

Old	89.5	90	91	91.5	92.5	91	89	89.5	91	92
New	89.5	91.5	91	89	91.5	92	92	90.5	90	91

Here, $n_1 = n_2 = 10$

x_1-bar = 90.7, x2-bar = 90.8 (1 – Old, 2 – New)

$S_1 = 1.159502$, $s_2 = 1.032796$

To test H0: μ1 = μ2

Against H1: μ1 ≠ μ2

Testing of Hypothesis

α = 0.05 (level of significance)

Test statistic t:

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \qquad s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2},$$

$$df = n_1 + n_2 - 2$$

Computed value of t = -0.20365

Table value of t = 2.101

Critical Region: I t I > 2.101

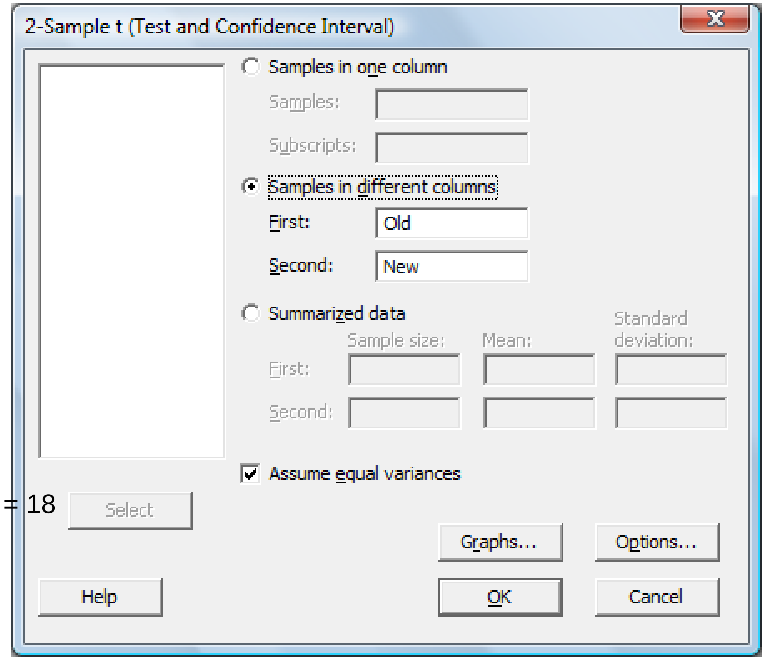
- Inference: Since the computed value of t is not falling in the critical region, we accept the null hypothesis. Hence, we can conclude that average time is the same across methodologies.

With the help of MINITAB

- Enter the data in the Minitab worksheet
- Select Stat → Basic Statistics → 2-Sample t
- A dialog box appears where we need to select the appropriate options as shown in the next slide

Testing of Hypothesis

Click OK. Minitab will give the following output:



Two-Sample T-Test and CI: Old and New

Two-sample T for Old vs New

	N	Mean	StDev	SE Mean
Old	10	90.70	1.16	0.37
New	10	90.80	1.03	0.33

Difference = mu (Old) - mu (New)
Estimate for difference: -0.100
95% CI for difference: (-1.132, 0.932)
T-Test of difference = 0 (vs not =): T-Value = -0.20 P-Value = 0.841 DF = 18

Inference: Since the p-value = 0.841 > 0.05, we do not have enough evidence to reject H0.

Testing of Hypothesis

3) To test whether two variances are equal or not

- To study the variation among two auditors, 7 samples are recorded. The client rating for all these samples is 1.5. Two auditors are requested to rate the samples on a 1 to 2 scale. Test whether the variation between the two auditors is equal.

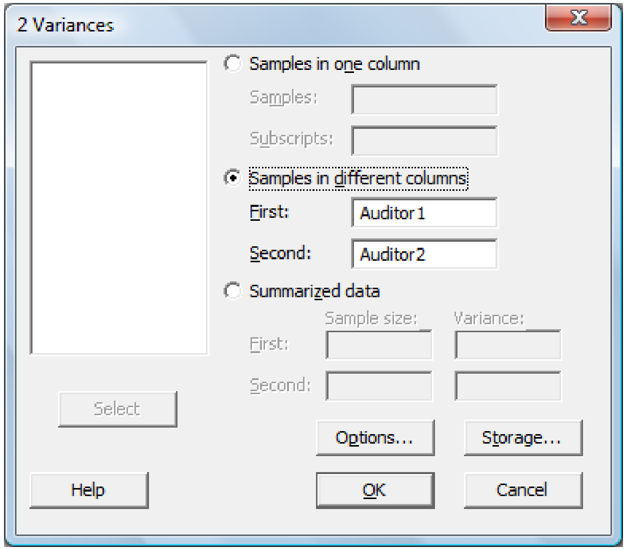
To test Ho: $\sigma_1 = \sigma_2$

Against H1: $\sigma_1 \neq \sigma_2$

Test statistic $F = s^2_1/s^2_2$

Minitab Steps

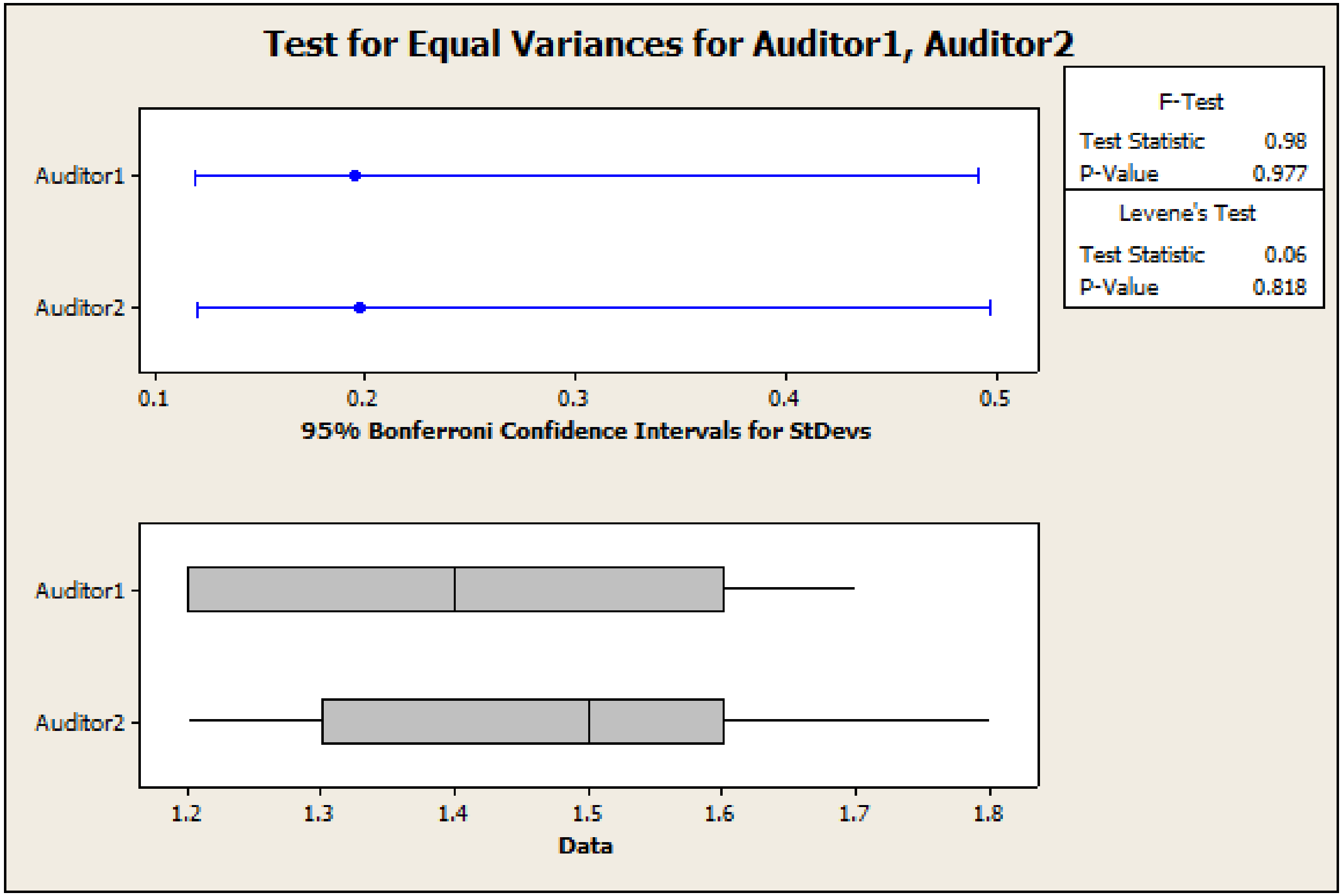
- Enter the samples in the Minitab in different columns
- Select Stat → Basic Statistics → 2-Variances



Auditor 1	Auditor 2
1.4	1.5
1.3	1.4
1.2	1.5
1.5	1.3
1.6	1.8
1.2	1.2
1.7	1.6

Testing of Hypothesis

- Click OK. Minitab will give the following output: (The following showing the variation in the two data set)



Testing of Hypothesis

Test for Equal Variances: Auditor1, Auditor2

95% Bonferroni confidence intervals for standard deviations

	N	Lower	StDev	Upper
Auditor1	7	0.118620	0.195180	0.490977
Auditor2	7	0.120093	0.197605	0.497077

F-Test (Normal Distribution)
Test statistic = 0.98, p-value = 0.977

- **Inference:** Since the p-value is greater than 0.05, we do not reject the null hypothesis. Hence, it is concluded that the variation between the two auditors is equal.

4) To test Proportion = specified value

- In a random sample of 250 purchase orders, 41 are found to be defective. Test whether the % of defectives is 10.
- To test $p = 0.1$

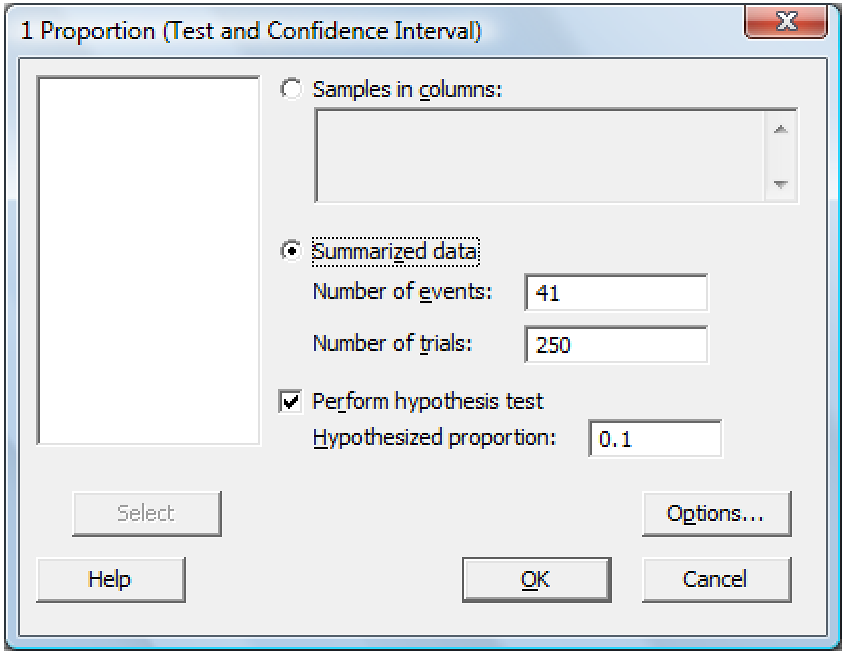
  Against $p \neq 0.1$

Testing of Hypothesis

Test statistic z: $z = \dfrac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

Minitab Steps

- Select Stat → Basic Statistics → 1 Proportion
- Enter the necessary information as shown
- Minitab will give the following output:



Test and CI for One Proportion

Test of p = 0.1 vs p not = 0.1

Sample	X	N	Sample p	95% CI	Exact P-Value
1	41	250	0.164000	(0.120333, 0.215836)	0.002

- **Interpretation:** Since the p-value is less than 0.05, we reject the null hypothesis at the 5% significance level. Hence the defective rate has changed statistically & it is not 10%.

Testing of Hypothesis

5) To test whether two proportions are equal or not

- Some measures are taken to improve the accuracy (reduce the number of defectives) of a process. Based on the data given below, test whether the measures really resulted in improving the accuracy.

	Number Audited	Defectives Found
Before improvement	75	12
After Improvement	68	8

To test Ho: p1 = p2

Against H1: p1 ≠ p2

Test statistic z:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}} \qquad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Minitab Steps

- Select Stat → Basic Statistics → 2 Proportions
- Enter the necessary information as shown in the next slide

Testing of Hypothesis

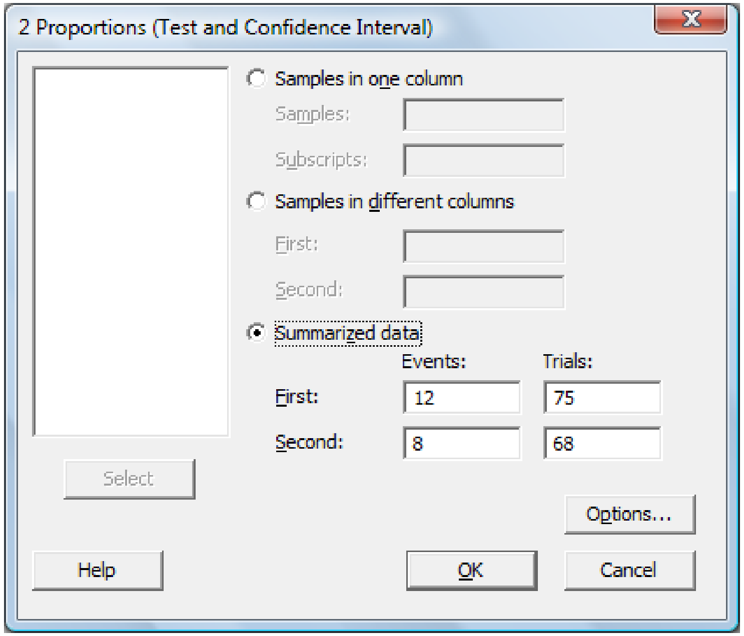
Minitab will give the following output:

Test and CI for Two Proportions

Sample	X	N	Sample p
1	12	75	0.160000
2	8	68	0.117647

Difference = p (1) - p (2)
Estimate for difference: 0.0423529
95% CI for difference: (-0.0705546, 0.155261)
Test for difference = 0 (vs not = 0): Z = 0.74 P-Value = 0.462

Fisher's exact test: P-Value = 0.630



Inference: Since p-value is greater than 0.05, we do not reject the null hypothesis at the 5% level. Hence, the measures do not really result in improving the accuracy.

Testing of Hypothesis

One way ANOVA – single factor ANOVA

- Example: A manufacturing company has three plants namely X, Y, and Z. A random sample of scrap generated in pounds was collected for 6 days in the following manner. Do they differ in producing waste?

	Waste (in lbs)		
Observation No.	Plant X	Plant Y	Plant Z
1	85	71	59
2	75	75	64
3	82	73	62
4	76	74	69
5	71	69	75
6	85	82	67

- Assuming equal variances across plants, test whether the waste produced are same at 5% level of significance.
- Hence, we are to test the null hypothesis H_o: All means are equal against H_1: At least two of the means are different.

Testing of Hypothesis

- Step 1: Calculate the sum, number, and averages of different levels

Level	N	Sum	Average
Plant X	6	474	79
Plant Y	6	444	74
Plant Z	6	396	66

- Step 2: Calculate N = 6+6+6 = 18
- Step 3: Calculate Grand Total (T) = 1314
- Step 4: Calculate Correction Factor (CF) = T^2 / N = 95922
- Step 5: Total sum of squares (TSS) = Sum of the squares of all individual observations – CF = 946
- Step 6: Sum of squares of Factors (SS_{Factor}) = (474^2 + 444^2 + 396^2) / 6 – CF = 516
- Step 7: Sum of squares of Error (SS_{Error}) = TSS – SS_{Factor} = 430
- Step 8: Total DF (degrees of freedom) = N-1 = 17, DF of Factors = 3 -1 = 2, DF of Errors = 17 – 2 = 15

Testing of Hypothesis

- Step 9: ANOVA table

Sources	SS	DF	MS	F	F Table Value
Factor	516	2	258	9	3.68
Error	430	15	28.66667		

- MS = SS / DF
- F = MS of Factor / MS of Error
- $F_{\text{Table value}} = F_{(2,15)}$ with $\alpha = 5\%$
- Interpretation: Since F > $F_{\text{Table value}}$, we reject the null hypothesis (i.e., the average test scores are the same for all the three plants). At 5% level the plants are not differ in producing waste.

One-way ANOVA: Single factor ANOVA

- Using Minitab:
- Copy the data in the Minitab worksheet the way it is given
- Select Stat > ANOVA > One-Way (Unstacked)
- Select the responses as plant X, Plant Y, and Plant Z in the Minitab dialog box

Testing of Hypothesis

- Minitab output

One-way ANOVA: Plant X, Plant Y, Plant Z

```
Source  DF     SS     MS     F      P
Factor   2  516.0  258.0  9.00  0.003
Error   15  430.0   28.7
Total   17  946.0

S = 5.354   R-Sq = 54.55%   R-Sq(adj) = 48.48%


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level    N    Mean  StDev  --------+---------+---------+---------+-
Plant X  6  79.000  5.831                        (-------*------)
Plant Y  6  74.000  4.472                (------*-------)
Plant Z  6  66.000  5.657  (-------*-------)
                           --------+---------+---------+---------+-
                                66.0      72.0      78.0      84.0

Pooled StDev = 5.354
```

- Note that p-vale = 0.003 < 0.05. Hence, we reject the null hypothesis. Hence the waste produced form 3 plants differ at 5% significance level.

Testing of Hypothesis

Two-way ANOVA

- Example: On a feeding experiment, a farmer has four types of hogs denoted by A, B, C, and D. Each of these types are divided into three groups which are fed on different kinds of rations 1, 2, and 3. The following results are obtained, the numbers in the table being the weight gain (in lbs) by the various groups and there being two hogs in each group.

Hogs	1	2	3
A	3	6.1	4.5
	3.6	4.9	5.6
B	7	7	8.5
	8.2	6.1	7.1
C	4.5	7	4.2
	6	5	4
D	9	14	9
	11	13	10

Testing of Hypothesis

- Minitab steps: Enter the data in the same format as discussed in the previous problem in Minitab worksheet
- Select Stat > ANOVA > Two-way
- Enter row factor as breed and column factor as rations
- Minitab output

Two-way ANOVA: Response versus Breed and Rations

Source	DF	SS	MS	F	P
Breed	3	151.961	50.6538	62.31	0.000
Rations	2	9.210	4.6050	5.66	0.019
Interaction	6	20.660	3.4433	4.24	0.016
Error	12	9.755	0.8129		
Total	23	191.586			

S = 0.9016   R-Sq = 94.91%   R-Sq(adj) = 90.24%

- Interpretation: From the above table, we conclude that observed F for interaction is significant (note that p-value = 0.016 < 0.05). The null hypothesis is not accepted for breed or rations because they affect the gains in weights significantly (note the p-values in both the cases < 0.05).

Testing of Hypothesis

Chi-square test

- In Six Sigma, there are numerous instances when the analyst requires to compare the percentage of items distributed among several categories such as operators, materials, methods, or any other group of interest. A sample is selected, estimated, and positioned into one of several categories from each of the groups. The results can be illustrated as a table with m rows which signify the group of interest and k columns which signify the categories. This table can be evaluated to answer the question “Do the groups differ with regard to the proportion of items in the categories”?

Example of chi-square test

- The manager of a nationalized bank is examining the mortgage payments made by the customers of the bank. A payment is classified as ‘good’ if it arrives on or before time, ‘delinquent’ if it arrives late or is not paid. In addition, the customers’ incomes are classified as low, medium, or high. The distribution of a group of randomly selected 200 customers is as follows:

Testing of Hypothesis

- Test the hypothesis that payment being good or delinquent is independent of income level at significance level 0.05.

	Income Level	Okay Percentage	
Payment	Low	Medium	High
Good	45	93.32%	65
Delinquent	5	99.38%	15

Solution

- Chi-square is computed by first finding the expected frequencies in each cell. This can be done by using the equation:
- Frequency expected = (Row sum * Column sum) / Overall sum

	Low	Medium	High	Tota
Good	45	50	65	160
Delinquent	5	20	15	40
Total	50	70	80	200

Testing of Hypothesis

- The table below shows the expected frequency for all cells.

	Low	Medium	High
Good	40	56	64
Delinquent	10	14	16

- The next step is to compute X^2 as follows:

$$X^2 = \sum (O - E)^2 / E$$

- Where O = frequency observed and E = frequency expected. The summation is taken over all the cells.
- Computed X^2 = 6.417411
- Given that $\alpha = 0.05$
- The degrees of freedom for the chi-square test is (3-1) * (2-1) = 2
- Critical value of X^2 = 5.991
- Since the computed value of chi-square exceeds the critical value, we reject the null hypothesis that payment being good or delinquent is independent of a person’s income.

Testing of Hypothesis

Non parametric tests

- Parametric method uses the assumption that the underlying distribution is **normal**, whereas non parametric procedure makes **no assumption** regarding the underlying distribution except that the distribution is **continuous**.
- In this case you will be testing other measures of population parameters like **mode**, **median** etc.

Comparison of parametric & non-parametric test

Parametric Test	Non-parametric Test
If assumptions are met, parametric tests provide greater power than non-parametric tests with equal sample sizes	Non-parametric test results are more robust against violation of assumptions

Testing of Hypothesis

When to use non-parametric tests

According to *"Nonparametric Statistics: An Introduction"* by Jean D. Gibbons:

Use **non-parametric tests** if *any* of the following conditions are true:

1. The data are **counts** or frequencies of different types of outcomes.
2. The data are measured on a **nominal scale**.
3. The data are measured on an **ordinal scale**.
4. The assumptions required for the validity of the corresponding parametric procedure are **not met** or cannot be verified.
5. The **shape of the distribution** from which the sample is drawn is **unknown**.
6. The **sample size** is **small**.
7. The measurements are **imprecise**.
8. There are **outliers** in the data making the **median** more representative than the **mean**.

Testing of Hypothesis

Non-parametric test	What it does	Parametric Analogs
Mann-Whitney U Test	Performs a hypothesis test of the equality of two population medians and calculates the corresponding point estimates and confidence interval	Two-sample t-test
Kruskal-Wallis (H Test)	• Performs a hypothesis test of the equality population medians for two or more populations. • It is the generalized version of Mann-Whitney test. • Statistically stronger for data many distributions than Mood's Median Test	One-way ANOVA

Testing of Hypothesis

Non-parametric test	What it does	Parametric Analogs
Mood's Median test	• Performs a hypothesis test of the equality of population medians. • This test is robust against outliers and errors in data and is particularly appropriate in the preliminary stages of analysis. • Mood's Median test is more robust against outliers than the Kruskal-Wallis test. • But it is less powerful (the confidence interval is wider on the average) for analyzing data from many distributions including data from normal distribution.	One-way ANOVA
Leven's Test	• Performs a statistical test of the equality of two population standard deviations.	

Testing of Hypothesis

Summary of topic covered in this section

- Null / alternative hypothesis, Type 1 & Type 2 Error, and power of a test
- Point and interval estimate, confidence interval
- Margin of error, sample size
- Test for means, variances, and proportions (one sample and two sample)
- Goodness of Fit test
- ANOVA (One-way ANOVA and Two-way ANOVA)
- Non-parametric tests