

Parameters	Sample type	Section	Excel / Tools / Data Analysis
Means	Independent	I. II. V.	z-test Two Sample for Means t-test Two-Sample Assuming Equal Variances t-test Two-Sample Assuming Unequal Variances
Means	Paired	III.	t-test Paired Two Sample for Means
Proportions	Independent	IV.	
Proportions	Paired	not covered	
Variances	Independent	VI.	F-test Two Sample for Variances
Variances	Paired	not covered	

I. Section 7.3.1 Two phenomena with unknown means, μ_1 and μ_2 , but the standard deviations, σ_1 and σ_2 have known values. Two independent samples n_1 and n_2 have been selected from the respective phenomena for inference about the difference in the unknown means.

$$100(1-\alpha)\% \text{ Confidence Interval for } \mu_1 - \mu_2: \bar{x}_1 - \bar{x}_2 \pm z_{1-\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$H_0: \mu_1 - \mu_2 = D_0$ versus $H_a: \mu_1 - \mu_2 (< \text{ or } \neq \text{ or } >) D_0$

$$\text{Test Statistic for the test is } TS = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Use the standard normal distribution for p-values (NORMSDIST) or critical values (NORMSINV).
Excel Data Analysis procedure: **z-test Two Sample for Means**.

II. Sections 7.3.(2, 3 & 4) Two phenomena with unknown means, μ_1 and μ_2 , and unknown standard deviations, σ_1 and σ_2 , but it is assumed that $\sigma_1 = \sigma_2 = \sigma$. Two independent samples n_1 and n_2 have been selected from the respective phenomena for inference about the difference in the unknown means. From the sample data the estimator of the common variance σ^2 is

$$s_p^2 = \frac{SST_1 + SST_2}{(n_1 - 1) + (n_2 - 1)} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$100(1-\alpha)\% \text{ Confidence Interval for } \mu_1 - \mu_2: \bar{x}_1 - \bar{x}_2 \pm t_{1-\alpha/2, n_1+n_2-2} \sqrt{s_p^2 \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}$$

$H_0: \mu_1 - \mu_2 = D_0$ versus $H_a: \mu_1 - \mu_2 (< \text{ or } \neq \text{ or } >) D_0$

$$\text{Test Statistic for the test is } TS = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2 \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

Use the t distribution with n_1+n_2-2 degrees of freedom for p-values (TDIST) or critical values (TINV) (an acceptable practice is to use the standard normal for p-values if the degrees of freedom is greater than 30).

Note: this test is robust to the assumption of equal variances if $n_1 \approx n_2$.

Excel Data Analysis procedure: **t-test Two-Sample Assuming Equal Variances**.

III. Section 7.4 Two phenomena with unknown means, μ_1 and μ_2 , data are gathered by **paired sampling**. A difference is calculated for each of the n pairs of observations. The mean of the sample differences is denoted \bar{D} and the sample standard deviation of the differences is s_D .

100(1- α)% Confidence Interval for $\mu_1 - \mu_2$: $\bar{D} \pm t_{1-\alpha/2, n-1} * (s_D / \sqrt{n})$

$$TS = \frac{\bar{D} - D_0}{\frac{s_D}{\sqrt{n}}}$$

$H_0: \mu_1 - \mu_2 = D_0$ versus $H_a: \mu_1 - \mu_2 (< \text{ or } \neq \text{ or } >) D_0$ The Test Statistic is

Use the t distribution with $n - 1$ degrees of freedom for p-values (TDIST) or critical values (TINV). (An acceptable practice is to use the standard normal for p-values if the degrees of freedom is greater than 30.)

Excel Data Analysis procedure: **t-test Paired Two Sample for Means**.

IV. Section 7.5 Two phenomena with unknown proportions, π_1 and π_2 . Two **independent samples** n_1 and n_2 have been selected from the respective phenomena and the sample proportions p_1 and p_2 are to be used for inference about the difference in the unknown proportions. Both n_1 and n_2 should be 50 or more and preferably $n_1 * p_1 \geq 5$ & $n_1 * (1-p_1) \geq 5$ & $n_2 * p_2 \geq 5$ & $n_2 * (1-p_2) \geq 5$.

100(1- α)% Confidence Interval for $\pi_1 - \pi_2$: $p_1 - p_2 \pm z_{1-\alpha/2} * \sqrt{\frac{p_1 * (1-p_1)}{n_1} + \frac{p_2 * (1-p_2)}{n_2}}$

$H_0: \pi_1 - \pi_2 = 0$ versus $H_a: \pi_1 - \pi_2 (< \text{ or } \neq \text{ or } >) 0$.

$$p = \frac{n_1 * p_1 + n_2 * p_2}{n_1 + n_2}$$

First use all the sample data to find a pooled proportion, p , where

$$TS = \frac{p_1 - p_2}{\sqrt{p(1-p) \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

is the Test Statistic for the test.

Use the standard normal distribution for p-values (NORMSDIST) or critical values (NORMSINV).

There is no Excel procedure.

V. Section 7.3.3 B in the 1995 text (not covered in the 1999 text). Two phenomena with unknown means, μ_1 and μ_2 , and **unknown standard deviations**, σ_1 and σ_2 , 2 **independent samples** n_1 and n_2 , inference about the difference in the unknown means. For this section the degrees of freedom are defined by the complicated formula (7.11) on page 390 of the 1995 text.

Excel calculates the value for degrees of freedom.

The 100(1- α)% Confidence Interval for $\mu_1 - \mu_2$ is

$$TS = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$H_0: \mu_1 - \mu_2 = D_0$ versus $H_a: \mu_1 - \mu_2 (< \text{ or } \neq \text{ or } >) D_0$ The Test Statistic is

Use the t distribution with df defined by (7.11) on page 390 of 1995 text.

Excel procedure: **t-test Two-Sample Assuming Unequal Variances**.

VI. Section 7.6.2 in the 1995 text (not covered in the 1999 text). Two normally distributed phenomena with variances, σ_1^2 and σ_2^2 . Two **independent samples** n_1 and n_2 have been selected from the respective phenomena for inference about the ratio of the unknown variances. $100(1-\alpha)\%$

Confidence Interval for σ_1^2 / σ_2^2 :

Lower Limit = $f_{\alpha/2, n_1-1, n_2-1} * (s_1^2 / s_2^2)$ **Upper Limit** = $f_{1-\alpha/2, n_1-1, n_2-1} * (s_1^2 / s_2^2)$.

H₀: $\sigma_1^2 / \sigma_2^2 = 1$ versus **H_a: σ_1^2 / σ_2^2 (< or ≠ or >) 1**

Test Statistic for the test is **TS = s_1^2 / s_2^2** . Use the f distribution with n_1-1 numerator degrees of freedom and n_2-1 denominator degrees of freedom for p-values (FDIST) or critical values (FINV).

Excel Data Analysis procedure: **F-test Two Sample for Variances**.