

20

Two Sample t-Tests

t-Test for Two Independent Samples
t-Test for Two Matched Samples
Confidence Interval for Two-Sample Test

The next logical step in the analysis of differences between parametric variables (that is, interval/ratio data) is **testing differences between two sample means**. In one-sample tests, we compared a single sample mean (\bar{X}) to a known or hypothesized population mean (μ). In two-sample tests, we compare two sample means (\bar{X} , \bar{Y}). The goal is to infer whether the populations from which the samples were drawn are significantly different.

Dr. Mark Cook studied the impact of active learner participation on adult learning in Sunday School in 1994.¹ Using twelve adult classes at First Baptist Church, Palestine, he randomly grouped six classes into a six-session Bible study using active participation methods and six that did not.² Because he used intact Sunday School classes, he applied an independent samples t-test to insure that the two experimental groups did not differ significantly in Bible knowledge prior to the course. Treatment and control groups averaged 6.889 and 7.436 respectively. They did not ($t = -0.292$, $t_{cv} = \pm 2.009$, $df = 10$, $\alpha = .05$).³ At the conclusion of the six-week course, both groups again took a test of knowledge. Treatment and control groups averaged 10.111 and 8.188 respectively. Applying the independent samples t-test, Dr. Cook discovered that the **active participation group scored significantly higher than the control** ($t = 2.048$, $t_{cv} = \pm 1.819$, $df = 10$, $\alpha = .05$).⁴ The average class achievement of those taught with active learner participation was significantly higher than those taught without it.

Descriptive or Experimental?

Two-sample tests can be used in either a **descriptive** or **experimental design**. In a **descriptive** design, two samples are randomly drawn from **two different** populations. If the analysis yields a significant difference, we conclude that **the populations from which the samples were drawn are different**. For example, a sample of pastors and a sample of ministers of education are tested on their leadership skills. The results of this study will describe the difference in leadership skill between pastors and ministers of education.

In an **experimental** design (like Dr. Cook's above), two samples are drawn from the **same population**. One sample receives a treatment and the other (control) sample

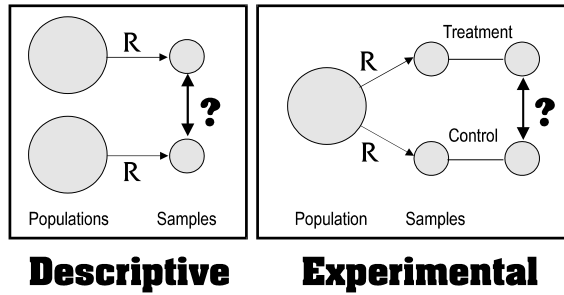
¹Marcus Weldon Cook, "A Study of the Relationship between Active Participation as a Teaching Strategy and Student Learning in a Southern Baptist Church," (Fort Worth, Texas: Southwestern Baptist Theological Seminary, 1994)

²Ibid., pp. 1-2

³Ibid., p. 45

⁴Ibid., p. 50

either receives no treatment or a different treatment. If the analysis of the post-test scores yields a significant difference, we conclude that the **experimental treatment will produce the same results in the population from which the sample was drawn.**



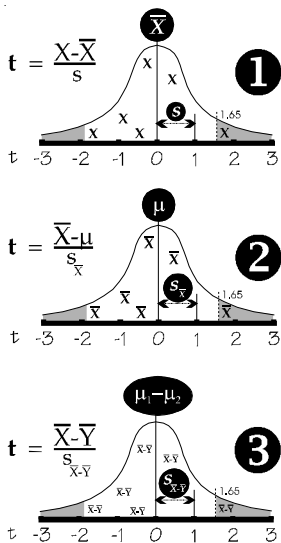
The results from Dr. Cook's study can be **directly** generalized to all adult Sunday School classes at First Baptist Church, Palestine -- since this was his population -- and further, by implication, to all similar adult Sunday School classes: active learner participation in Sunday School improves learning achievement.

t-Test for Independent Samples

Samples are “independent” when subjects are randomly selected and assigned individually to groups.⁵ The formula for the t-test for independent samples is

$$t = \frac{\bar{X} - \bar{Y}}{s_{\bar{X}-\bar{Y}}}$$

where the numerator equals the difference between two sample means (\bar{X}, \bar{Y}), and the denominator, $s_{\bar{X}-\bar{Y}}$ -- called the **standard error of difference** -- equals the combined standard deviation of both samples.



Review: We have studied a normal curve of **scores** (the frequency distribution) where the standard score (z, t) equals the ratio of the **difference between** score and mean ($X - \bar{X}$) to the distribution's **difference within** -- standard deviation (s). See diagram 1 at left.

We have studied a normal curve of **means** (the sampling distribution) where the standard score (z, t) equals the ratio of the **difference between** sample mean (\bar{X}) and population mean (μ) to the distribution's **difference within** -- the standard error of the mean ($s_{\bar{X}}$). See diagram 2.

In the independent-samples t-test, we discover a third normal curve -- a normal curve of **differences between samples** drawn from two populations. Given two populations with means μ_x and μ_y , we can draw samples at random from each population and compute the difference between them (\bar{X}, \bar{Y}). The mean of this distribution of differences equals $\mu_x - \mu_y$ (usually assumed to be zero) and the standard deviation of this distribution is the standard error of difference: $s_{\bar{X}-\bar{Y}}$. See diagram 3.

Any **score** falling into the shaded area of diagram 1 is declared significantly different from \bar{X} . Any **mean** falling into the shaded area of diagram 2 is significantly

⁵When subjects are drawn from a population of pairs, as in married couples, and assigned to husband and wife groups, the groups are not independent, but “matched” or “correlated.” We will discuss the correlated-samples t-test later in the chapter.

different from μ . Any **difference between means** falling into the shaded area of diagram 3 is significantly different from zero, meaning that \bar{X} , \bar{Y} are significantly different from each other. You will find the importance of this discussion in the logic which flows through all of these tests.

The Standard Error of Difference

The t-Test for Independent Samples tests a difference between two sample means against a hypothesized difference between populations (usually assumed to be zero). The most difficult part of the test is **computing the standard error of difference**, the combined standard deviation value for both samples together. The formula for the standard error of difference is:

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{\Sigma x^2 + \Sigma y^2}{n_x + n_y - 2} \cdot \left[\frac{1}{n_x} + \frac{1}{n_y} \right]}$$

The purpose of this formula is to **combine the standard deviations** of the two samples (s_x , s_y) into a single "standard deviation" value. We can rewrite the first part of the equation into a **more recognizable** (*but mathematically incorrect*) form like this:

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{\Sigma x^2}{n_x - 1} + \frac{\Sigma y^2}{n_y - 1} \cdot \left[\frac{1}{n_x} + \frac{1}{n_y} \right]}$$

Note:
The terms are separated to show link to standard deviation. It is mathematically incorrect.

The shaded area above highlights the formula we established for the estimated variance (s^2) in chapter 16. Here it computes s_x^2 . A similar element computes s_y^2 . These are added together. The gray [brackets] combines the n's of both samples into a harmonic mean and has the same mathematical function in the formula as "dividing by n," producing a combined variance. Mathematically, we add the sums of squares, divide by degrees of freedom, and multiply by the [harmonic mean]. The square root radical produces the standard-deviation-like **standard error of difference**.

The equation above spells out what the standard error of difference does, but it is tedious to calculate. If you have a **statistical calculator** that computes mean (\bar{X}) and variance (s^2) from sets of data, you can use an equivalent equation much more easily. The equivalent equation substitutes the term $(n_x - 1)s_x^2$ for Σx^2 , and $(n_y - 1)s_y^2$ for Σy^2 . A quick review of the equation for s^2 above shows these two terms to be equivalent. The altered equation looks like this:

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2} \cdot \left[\frac{1}{n_x} + \frac{1}{n_y} \right]}$$

Substituting $(n-1)$, s^2 (from the calculator), and n terms from both samples quickly simplifies the above equation. Let's illustrate the use of these equations with an example.

Example Problem

The following problem will illustrate the computational process for the independent-samples t-test. Suppose you are interested in developing a counseling technique to reduce stress within marriages. You randomly select two samples of married indi-

viduals out of ten churches in the association. You provide Group 1 with *group counseling* and study materials. You provide Group 2 with *individual counseling* and study materials. At the conclusion of the treatment period, you measure the level of marital stress in the group members. Here are the scores:

Group 1	Group 2
25 17 29 29 26 24 27 33	21 26 28 31 14 27 29 23
23 14 21 26 20 27 26 32	18 25 32 23 16 21 17 20
20 32 17 23 20 30 26 12	26 23 7 18 29 32 24 19

Using a statistical calculator, we enter the 24 scores for each group and compute means and standard deviations. Here are the results:

	Group 1	Group 2
\bar{X} :	24.13	22.88
s:	5.64	6.14

Are these groups **significantly different** in marital stress?

Or, asking it a little differently:

Are the two means, **24.13** and **22.88**, **significantly different**? Use the independent-samples t-test to find out. Here is the procedure.

Step 1: State the research hypothesis. $H_a: \mu_1 \neq \mu_2$ or $H_a: \mu_1 - \mu_2 \neq 0$

Step 2: State the null hypothesis. $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$

Step 3: Compute $\bar{X} - \bar{Y}$. (Numerator of t-ratio)

$$\bar{X} - \bar{Y} = 24.13 - 22.88 = 1.25$$

Step 4: Compute $s_{\bar{X}-\bar{Y}}$. (Denominator of t ratio)

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{(n_x-1)s_x^2 + (n_y-1)s_y^2}{n_x + n_y - 2} \cdot \left[\frac{1}{n_x} + \frac{1}{n_y} \right]}$$

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{(23) 5.64^2 + (23) 6.14^2}{24 + 24 - 2} \cdot \left[\frac{1}{24} + \frac{1}{24} \right]}$$

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{731.63 + 867.10}{46} \cdot [0.083]}$$

$$s_{\bar{X}-\bar{Y}} = \sqrt{2.885}$$

$$s_{\bar{X}-\bar{Y}} = 1.698$$

Step 5: Compute t

$$t = \frac{\bar{X} - \bar{Y}}{s_{\bar{X}-\bar{Y}}} = \frac{1.25}{1.698} = 0.736$$

Step 6: Determine t_{cv} (.05, df=46, 2 tail)

Degrees of freedom (df) for the independent samples t-test is given by $n_x + n_y - 2$, or $24 + 24 - 2 = 46$ in this case. Since the t-distribution table does not have a row for $df=46$, we must choose either $df=40$ or $df=50$. The value for $df=40$ is higher, and convention says it is better to choose the higher value (to guard against a Type I error). So,

$$t_{cv} (df=40, 2 \text{ tail}, 0.05) = \pm 2.021$$

Step 7: Compare: Is the computed t greater than the critical value?

$$t = 0.736 \qquad t_{cv} = 2.021$$

Step 8: Decide

The computed value for t is *less than* the critical value of t. It is not significant. We retain the null hypothesis.

Step 9: Translate into English

There was no difference between the two approaches to reducing marital stress.

The t-Test for Independent Samples allows researchers to determine whether two randomly selected groups differ significantly on a ratio (test) or interval (scale) variable. At the end of the chapter you will find several more examples of the use of Independent-samples t in dissertations written by our doctoral students.

You will often see journal articles that read like this:

The experimental group scored twelve points higher on post-test measurements of subject mastery ($t=2.751^$, $\alpha=0.05$, $df=26$, $t_{cv}=1.706$) and eight points higher in attitude toward the learning experience ($t=1.801^*$) than the control group. (The asterisk [*] indicates "significant difference").*

If the two samples are formed from *pairs of subjects* rather than two independently selected groups, then the *t-test for correlated samples* should be used.

t-Test for Correlated Samples

The Correlated Samples (or "Matched Samples") t-test is used when scores come from (1) *naturally occurring pairs*, such as husband-wife pairs, parent-child pairs, sibling pairs, roommates, pastor-staff minister pairs, or (2) *two measures taken on a single subject*, such as pre- and post-tests, or (3) *two subjects that have been paired* or "matched" on some variable, such as intelligence, age, or educational level.

Effect of Correlated Samples

The t-test for correlated samples uses a different formula to take into account the correlation which exists between the two groups. *Any positive correlation between the pairs reduces the standard error of difference between the groups.*⁶ The higher the correlation, the greater the reduction in the standard error of difference. This makes the *Correlated Samples t-Test a more powerful test than the Independent Samples t-Test*. Let's take a moment to analyze this.

The correlation between scores in matched samples controls some of the extraneous variability. In *husband-wife pairs*, for example, the differences in attitudes, socioeconomic status, and experiences are less for the pair than for husbands and wives in general.

Or, when *two scores are obtained from a single subject*, the variability due to individual differences in personality, physical, social, or intellectual variables is less in the correlated samples design than with independent samples.

Or, when a study *pairs subjects* on a given variable, it generates a relationship through that extraneous variable.

In each case, more of the extraneous "error" (uncontrolled variability) in the study is accounted for. This reduction in overall error reduces the standard error of difference. This, in turn, increases the size of *t* making the test more likely to reject the null hypothesis when it is false (i.e., *more powerful test*).

On the other hand, the degrees of freedom are reduced when using the correlated pairs formula. For independent samples, *df* is equal to $n_x + n_y - 2$. In the correlated samples, *n is the number of pairs of scores and df is equal to n (pairs) - 1*. Two groups of 30 subjects in an independent-samples design would yield $df=30+30-2=58$, but in a correlated samples design, $df=30 \text{ pairs}-1=29$. The effect of reducing *df* is to *increase* the critical value of *t*, and therefore we have a *less powerful test*.

The decision whether to use independent or correlated samples comes down to this: Will the loss of power resulting from a smaller *df* and larger critical value for *t* be offset by the gain in power resulting from the larger computed *t*? The answer is found in the degree of relationship between the paired samples. *The larger the correlation between the paired scores, the greater the benefit from using the correlated samples test.*

The Standard Error of Difference

The correlated samples t-Test is based on the *differences between paired scores* ($d=X-Y$) rather than the scores themselves. The formula looks like this:

$$t = \frac{\bar{d}}{s_d}$$

where \bar{d} is the mean of the differences and s_d is the standard error of difference. The value of s_d is given by

$$s_d = \sqrt{\frac{s_d^2}{n}} \quad \leftarrow s_d^2 = \frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n - 1}$$

I provide here a two-step process for computing the standard error of difference to reduce the complexity of the formula. First, compute the s_d^2 using the formula **above right**. Note the distinction between $\sum d^2$ (*differences between paired scores which are squared and then summed*) and $(\sum d)^2$ (*differences between paired scores which are summed and then squared*).

Secondly, use the s_d^2 term in the formula **above left** to compute the standard error of difference (s_d). Substitute \bar{d} and s_d in the matched-t formula and compute *t*. Compare the computed *t*-value with the critical value (t_{cv}) to determine whether the difference between the correlated-samples means is significant. Let's look at the proce-

⁶"Correlation" is another category of statistical analysis which measures the strength of association between variables (rather than the difference between groups). We'll study the concept in Chapter 22.

ture in an example problem.

Example Problem

A professor wants to empirically measure the impact of stated course objectives and learning outcomes in one of her classes. The course is divided into four major units, each with an exam. For units I and III, she instructs the class to study lecture notes and reading assignments in preparation for the exams. For units II and IV, she provides clearly written instructional objectives. These objectives form the basis for the exams in units II and IV. Test scores from I,III are combined, as are scores from II,IV, giving a total possible score of 200 for the “with” and “without” objectives conditions. Do students achieve significantly better when learning and testing are tied together with objectives? ($\alpha=0.01$)

Here is a random sample of 10 scores from her class:

Subject	Instructional Objectives		d	d ²
	Without (I,III)	With (II,IV)		
1	165	182	-17	289
2	178	189	-11	121
3	143	179	-36	1296
4	187	196	-9	81
5	186	188	-2	4
6	127	153	-26	676
7	138	154	-16	256
8	155	178	-23	529
9	157	169	-12	144
10	171	191	-20	400
	$\Sigma X=1607$	$\Sigma Y=1779$	$\Sigma d = -172$	$\Sigma d^2= 3786$

means: 160.7 - 177.9 = -17.2

*Insight: Compare the difference between the means with the mean of differences. What do you find?**

Subtract the “with” scores from the “without” (e.g., 165-182), producing the difference score d (-17).

Adding all the d -scores together yields a Σd of -172.

Squaring each d -score yields the d^2 score (-17 \times -17=289) in the right-most column.

Adding all the d^2 -scores together yields a Σd^2 of 3786. These values will be used in the equations to compute t . Here are the steps for analysis:

Step 1: State the directional research hypothesis. $H_a: \mu_{\text{without}} - \mu_{\text{with}} < 0$

Step 2: State the directional null hypothesis. $H_0: \mu_{\text{without}} - \mu_{\text{with}} \geq 0$

Step 3: Compute the mean difference (\bar{d})

$$\bar{d} = \frac{\Sigma d}{n} = \frac{-172}{10} = -17.2$$

*Answer: The difference between the means equals the mean of score differences.

Step 4: Compute s_d^2

$$s_d^2 = \frac{\Sigma d^2 - \frac{(\Sigma d)^2}{n}}{n-1} = \frac{3786 - \frac{(-172)^2}{10}}{10-1} = \frac{3786 - 2968.4}{9} = 91.96$$

Step 5: Compute $s_{\bar{d}}$, the standard error of difference

$$s_{\bar{d}} = \sqrt{\frac{s_d^2}{n}} = \sqrt{\frac{91.96}{10}} = \sqrt{9.196} = 3.032$$

Step 6: Compute t

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{-17.2}{3.032} = -5.673$$

Step 7: Determine t_{cv} from the t-distribution table. In this study, $\alpha = 0.01$.

$$t_{cv} \text{ (df=9, 1-tail test, } \alpha=.01) = \mathbf{2.821}$$

Step 8: Compare: Is the computed t greater than the critical value?

$$t = \mathbf{-5.673} \qquad t_{cv} = \mathbf{\pm 2.821}$$

Step 9: Decide:

The calculated t -value is greater than the critical value. It is significant. Therefore, we reject the null hypothesis and accept the research hypothesis. The “with objectives” test scores ($\bar{X}=177.9$) were significantly higher than the “without objectives” test scores ($\bar{Y}=160.7$).

Step 9: Translate into English

Providing students learning objectives which are tied to examinations enhances student achievement of course material.

The Two Sample Confidence Interval

Just as we developed a confidence interval about a sample mean (chapter 19), we can develop a confidence interval about the difference in means. Do you recall the formula for a *single-sample confidence interval*. The equation for a 95% confidence interval is

$$CI_{95} = \mu \pm t_{cv}s$$

where t_{cv} is the t critical value at df degrees of freedom and $\alpha=0.05$. We can also build a *confidence interval about a sample mean* (using it as a benchmark against several other means). This slightly different form looks like this:

$$CI_{95} = \bar{X} \pm t_{cv}s$$

The equation for the same 95% confidence interval about *differences between two independent means* follows the same form. The two-sample confidence interval is given by

$$CI_{95} = \bar{X} - \bar{Y} \pm t_{cv}s_{\bar{X}-\bar{Y}}$$

We can also build a confidence interval around a matched samples design:

$$CI_{95} = \bar{d} \pm t_{cv} s_{\bar{d}}$$

Confidence intervals for two-sample tests are rarely found. This section is included more for completeness of the subject than for practical considerations.

Summary

Two-sample t-tests are among the most popular of statistical tests. The t-test for independent samples assumes that the group scores are randomly selected individually (no correlation between groups) while the t-test for correlated samples assumes that the group scores are randomly selected as pairs.

Examples

Dr. Roberta Damon, cited in Chapter 19 for her use of the one-sample z-Test, also used the independent-samples t-Test to determine differences between missionary husbands and wives on three variables: communication, conflict resolution, and marital satisfaction. Her findings were summarized in the following table:⁷

TABLE 4

t SCORES AND CRITICAL VALUES FOR THE SCALES OF
COMMUNICATION, CONFLICT RESOLUTION AND MARITAL
SATISFACTION ACCORDING TO GENDER

Scale	t	Critical Value	Alpha Level
Communication	2.757 ⁸	2.617	.005
Conflict Resolution	0.470 ⁹	1.645	.05
Marital Satisfaction	-0.714 ¹⁰	1.645	.05

N = 165

df = 163

Dr. Don Clark studied the statistical power generated in dissertation research by doctoral students in the School of Educational Ministries, Southwestern Seminary, between 1978 and 1996.¹¹ Clark's interest in this subject grew out of a statement I made in research class: "Many of the dissertations written by our doctoral students fail to show significance *because the designs they use lack sufficient statistical power* to declare real differences significant." He wanted to test this assertion statistically.

One hundred hypotheses were randomly selected from two populations: hypotheses proven significant (Group X) and those not (Group Y). Seven of these were eliminated because they did not fit Clark's study parameters, leaving ninety-three.¹²

⁷Damon, 41.

⁸Husbands were significantly more satisfied than wives with the way they communicate. "This indicates a felt dissatisfaction among the wives which their husbands do not share" (p. 41).

⁹No difference in satisfaction with how couples resolve conflicts.

¹⁰No difference in marital satisfaction.

¹¹Clark, 30

¹²*Ibid.*, 39

¹³*Ibid.*, p. 44

The scores for this study were the **computed power values** for all 93 hypotheses of difference. The mean power score (n=47) for "significant hypotheses" was **0.856**. The mean power score (n=46) for "non-significant hypotheses" was **0.3397**. The standard error of difference was **0.052**, resulting in a **t-value of 9.904**,¹³ which was tested against t_{cv} of **1.660**. The "power of the statistical test is significantly higher in those dissertations' hypotheses proven statistically significant than those. . .not proven statistically significant."¹⁴

At first glance, the findings seem trivial. Wouldn't one expect to find higher statistical power in dissertations producing significant findings? The simple answer is no. A research design and statistic produces a level of power unrelated to the findings. A fine-mesh net *will catch fish if they are present*, but a broad-mesh net will not. We seek fine-mesh nets in our research. **The average power level of 0.34 for non-significant hypotheses shows that these studies were fishing with broad-mesh nets.** As I had said, "These studies were doomed from the beginning, before the data was collected." Dr. Clark found this to be true, at least for the 93 hypotheses he analyzed.

We have been growing in our research design sophistication over the years. Computing power levels of research designs as part of the dissertation proposal has not been required in the past. Based on Dr. Clark's findings, we need to seriously consider doing this in the future, if for no other reason than to spare our students from dissertation research which is futile in its very design from the beginning.

Dr. John Babler conducted his dissertation research on spiritual care provided to hospice patients and families by hospice social workers, nurses, and spiritual care professionals.¹⁵ In two of his minor hypotheses, he used the Independent-Samples t-Test to determine **whether there were differences in "provision of spiritual care"**¹⁶ **between male and female providers, and between hospice agencies related to a religious organization or not.** Males and females (n= 21, 174) showed **no significant difference in the provision of spiritual care** (Mean_m=49.6, Mean_f = 51.2, t= -0.83, p=0.409). Agencies related to religious organizations (n=26) and those which were not (n=162) showed **no difference in provision of spiritual care** (Mean_r =51.5, Mean_{nr}= 51.1, t=0.22, p=0.828).

Vocabulary

\bar{d}	Average difference between two matched groups
s_{x-y}	Standard error of difference (independent samples)
$s_{\bar{d}}$	Standard error of difference (matched samples)
Correlated samples	Sample paired subjects — test differences between pairs
Independent samples	Samples randomly drawn independently of each other
Standard error of difference	combined variability (standard deviation) of two groups of scores
t-Test for Correlated Samples	tests difference between means of two matched samples
t-Test for Independent Samples	tests difference between means of two independent samples

Study Questions

1. Compare and contrast "standard deviation," "standard error of the mean," and "standard error of difference."

¹⁴Ibid., p. 45

¹⁵Babler, 7

¹⁶Scores were generated by an instrument which Babler adapted by permission from the "Use of Spirituality by Hospice Professionals Questionnaire" developed by Millison and Dudley (1992), p. 34

2. Differentiate between “descriptive” and “experimental” use of the t-test.
3. List three ways we can design a “correlated-samples” study. Give an example of each type of pairing.
4. Discuss the two factors involved in choosing to use one test over the other. What is the major factor to consider?
5. A professor asked his research students to rate (anonymously) how well they liked statistics. The results were:

	Males	Females
\bar{X}	5.25	4.37
s^2	6.57	7.55
n	12	31

- a. State the research hypothesis
 - b. State the null hypothesis
 - c. Compute the standard error of difference
 - d. Compute t
 - e. Determine critical value ($\alpha=0.01$)
 - f. Compare
 - g. Decide
 - h. Translate into English
 - i. Develop CI_{99} for the differences.
7. Test the hypothesis that there is no difference between the average scores of husbands and their wives on an attitude-toward-smoking scale. Use 0.05 and the process listed in #6 (NOTE: Develop a CI_{95} confidence interval)

Husband	16	10	8	14	20	15	19	15
Wife	15	13	12	11	18	10	16	12

Sample Test Questions

1. The term which reflects the “spreadedness of differences between sample means” is called
 - A. standard deviation
 - B. standard error of the mean
 - C. standard error of difference
 - D. standard range
2. The purpose of $s_{\bar{x}-\bar{y}}$ or $s_{\bar{d}}$ is to combine the _____ of both samples into one number.
 - A. variances
 - B. df
 - C. standard deviations
 - D. sampling error
3. Which of the following is the same for the 1-sample and 2-sample tests we’ve studied so far?
 - A. All are ratios of “variability between” to “variability within.”
 - B. All calculate degrees of freedom the same way
 - C. All use the same statistical tables.
 - D. All yield the same level of power.
4. T F Jane is studying the attitudes of pastors and ministers of education toward Christian counseling. You would be right to advise her to randomly sample a group of churches and use the independent samples t-test to analyze the difference between pastors and educators on the same staffs.
5. T F The design which draws random samples from two populations and tests the two sample means for significant difference in the populations is called an experimental design.

