

Assigned Problem

A researcher wants to determine if biofeedback therapy reduces stress in hospital workers exposed to extremely stressful situations. Sixteen participants were involved in the study. Eight participants participated in ten biofeedback therapy sessions versus the other eight participants who did not participate in any biofeedback therapy sessions.

The researcher hypothesized that the participants participating in biofeedback therapy sessions should have lower stress scores than the participants who did not participate in any biofeedback therapy sessions. Assume that alpha is 0.005.

The results of the study are presented below:

Stress Scores of Participants Given Biofeedback Therapy - Experimental Group (X_1)	Stress Scores of Participants Not Given Biofeedback Therapy - Control Group (X_2)
53	64
44	68
61	84
50	65
40	50
39	60
58	81
42	76

Example Three:

A researcher wants to determine if an experimental drug will help morbidly obese patients (more than 100 lbs over ideal weight) lose weight. The patient's weights were measured before treatment and then again six months after administration of treatment. The researcher hypothesizes that the patients' weights will be less after treatment. The results of the study are presented below. Assume alpha is 0.005.

Patient	Weight before treatment (lbs)	Weight after treatment (lbs)
1	245	202
2	340	284
3	298	244
4	355	285
5	321	325
6	414	340
7	309	249

Using Microsoft Excel to calculate the mean, variance, and standard deviation

Using Microsoft Excel to calculate the mean, variance, and standard deviation:

Students may want to use Microsoft Excel to perform some of their calculations rather than complete each calculation by hand. The steps necessary to use Excel to calculate the mean, variance, and standard deviation for a sample population are outlined below. If you are familiar with using Excel to complete a formula, the following formulas are used to calculate the mean, variance, and standard deviation for a sample population:

To calculate mean: “=AVERAGE(data set)”

To calculate variance for sample population: “=VAR(data set)”

To calculate standard deviation for sample population: “=STDEV(data set)”

Step One: Copy and paste data into Microsoft Excel.

	A	B
	Females Resting Heart Rate (in bpm)	Males Resting Heart Rate (in bpm)
1		
2	67	90
3	95	74
4	93	83
5	85	94
6	100	72
7	99	87
8	104	86
9	101	84
10	108	70
11	93	89
12	94	92
13	102	84
14	84	84
15	85	97
16	100	100
17	93	96
18	90	90
19	105	74
20	99	83
21	91	94
22	101	72
23	104	90
24	107	74
25	107	83
26	98	94
27	107	72
28	109	87
29	83	86
30	100	84

Step Two: Type in =AVERAGE(into the box below Column A.

	A	B
	Females Resting Heart Rate (in bpm)	Males Resting Heart Rate (in bpm)
1		
2	67	90
3	95	74
4	93	83
5	85	94
6	100	72
7	99	87
8	104	86
9	101	84
10	108	70
11	93	69
12	94	92
13	102	84
14	84	84
15	85	97
16	100	100
17	93	96
18	90	90
19	105	74
20	99	83
21	91	94
22	101	72
23	104	90
24	107	74
25	107	83
26	98	94
27	107	72
28	109	87
29	83	86
30	100	84
31	=AVERAGE()	

Step Three: Highlight the data for which you want to calculate the mean.

	A	B	C
	Females Resting Heart Rate (in bpm)	Males Resting Heart Rate (in bpm)	
1			
2	67	90	
3	95	74	
4	93	83	
5	85	94	
6	100	72	
7	99	87	
8	104	86	
9	101	84	
10	108	70	
11	93	69	
12	94	92	
13	102	84	
14	84	84	
15	85	97	
16	100	100	
17	93	96	
18	90	90	
19	105	74	
20	99	83	
21	91	94	
22	101	72	
23	104	90	
24	107	74	
25	107	83	
26	98	94	
27	107	72	
28	109	87	
29	83	86	
30	100	84	
31	=AVERAGE(A2:A30)		
32	=AVERAGE(number1, [number2], ...)		

Step Four: Type) to complete the average formula after the data covered and press return. Excel will complete the calculation.

=AVERAGE(A2:A30)

	A	B
	Females Resting Heart Rate (in bpm)	Males Resting Heart Rate (in bpm)
1		
2	67	90
3	95	74
4	93	83
5	85	94
6	100	72
7	99	87
8	104	86
9	101	84
10	108	70
11	93	69
12	94	92
13	102	84
14	84	84
15	85	97
16	100	100
17	93	96
18	90	90
19	105	74
20	99	83
21	91	94
22	101	72
23	104	90
24	107	74
25	107	83
26	98	94
27	107	72
28	109	87
29	83	86
30	100	84
31	96.68965517	

Step Five: Repeat all steps to calculate the mean for the data in column B.

Step Six: Follow the same steps as above, but substitute the average formula with the formula for calculating the variance for a sample population. The formula used to calculate the variance for a sample population is **=VAR(data set)**

=VAR(A2:A30) **=VAR(B2:B30)**

A	B
Females Resting Heart Rate (in bpm)	Males Resting Heart Rate (in bpm)
67	90
95	74
93	83
85	94
100	72
99	87
104	86
101	84
108	70
93	69
94	92
102	84
84	84
85	97
100	100
93	96
90	90
105	74
99	83
91	94
101	72
104	90
107	74
107	83
98	94
107	72
109	87
83	86
100	84
96.68965517	84.31034483
=VAR(A2:A30)	=VAR(B2:B30)

A	B
Females Resting Heart Rate (in bpm)	Males Resting Heart Rate (in bpm)
1	
2	90
3	74
4	83
5	94
6	72
7	87
8	86
9	84
10	70
11	69
12	92
13	84
14	84
15	97
16	100
17	96
18	90
19	74
20	83
21	94
22	72
23	90
24	74
25	83
26	94
27	72
28	87
29	86
30	84
31	84.31034483
32	79.29310345

Mean For Data

Variance For Data

Step Seven: Follow the same steps as above, but substitute the variance formula with the formula for calculating the standard deviation for a sample population. The formula used to calculate the standard deviation for a sample population is **=STDEV(data set)**

=STDEV(A2:A30)

=STDEV(B2:B30)

A	B
Females Resting Heart Rate (in bpm)	Males Resting Heart Rate (in bpm)
67	90
95	74
93	83
85	94
100	72
99	87
104	86
101	84
108	70
93	69
94	92
102	84
84	84
85	97
100	100
93	96
90	90
105	74
99	83
91	94
101	72
104	90
107	74
107	83
98	94
107	72
109	87
83	86
100	84
96.68965517	84.31034483
89.86453202	79.29310345
=STDEV(A2:A30)	=STDEV(B2:B30)

Resource Sheet

Conducting a Two Sample T-test

This document presents a step-by-step approach to conducting a two sample t-test. It includes instructions for using a calculator and Microsoft Excel (optional). While statistical analysis can be conducted by hand, scientists often process large data sets that require the use of software for accuracy and speed. The Excel commands are shown as alternative method for the calculations.

Step One: Identify the null hypothesis and alternative hypothesis (Researcher will have setup both hypotheses before beginning experiment).

- **Null Hypothesis (H_0):** This hypothesis is the starting point for the investigation. The researcher may believe that there is no effect of the treatment or no relationship between the two means (μ).
 - For our analysis, we will use an equality for the null hypothesis ($H_0: \mu_1 = \mu_2$). This means that the mean of sample 1 is equal to the mean of sample 2.
- **Alternative Hypothesis (H_a):** This hypothesis is what the researcher initially assumes to be true. If we use the example from the null hypothesis above, the researcher may believe there is an effect of the treatment or a relationship between two means.
 - The alternative hypothesis is usually in the form of $\mu_1 > \mu_2$, $\mu_1 < \mu_2$, or $\mu_1 \neq \mu_2$. The means that the mean of sample 1 is greater than, less than, or not equal to the mean of sample 2.
- **Example:** A pharmaceutical company has developed a new drug that enhances growth in children with human growth hormone deficiency. They will conduct a study that involves two groups of children to evaluate the effectiveness of a new drug. One group will receive the experimental drug and the other a placebo drug to study their height.
 - Study variables: μ_1 = Mean height of experimental drug group, μ_2 = Mean height of placebo drug group
 - Null hypothesis ($H_0: \mu_1 = \mu_2$) There is no difference in the heights of the two groups.
 - Alternative hypothesis ($H_A: \mu_1 > \mu_2$) Group 1 will be taller than Group 2.

Step Two: Establish the level of significance, which is known as the alpha value.

- The alpha value (α) is the level of significance and is set by the experimenter at the beginning of the investigation. Typical alpha values are 0.10, 0.05 and 0.01. An α value equal to 0.01 means that the probability that the difference is due to chance is less than or equal to 1%. The α value is also used to determine if the null hypothesis should be rejected.

Step Three: Calculate the mean, variance, and standard deviation for both samples.

a. Calculate the mean for both samples.

$$\bar{x} = \frac{\sum x_i}{n}$$

$\sum x_i$ is the sum of the value of the samples
n is the number of samples

Excel command (Optional)

- Note that all excel commands start with the equal sign
- In an empty cell, type the command **=average(**
- Once the program recognizes the command, the cursor will blink next to the bracket.
- Select your data with your mouse, then click Enter.

b. Calculate the variance (s^2) for both samples and round your answers to the nearest thousandth.

- First: Calculate the deviation (or difference of the data point from the mean) for each data point.
- Second: Square these deviations to ensure that all values are positive.
- Third: Calculate the sum of all of these squared deviations.
- Fourth: Divide by the number of data points in the data set minus one.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Excel command (Optional)

- In an empty cell, type the command **=var(**
- Once the program recognizes the command, the cursor will blink next to the bracket.
- Select your data with your mouse, then click Enter.

c. Calculate the standard deviation (s) by taking the square root of the variance for both samples.

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Excel command (Optional)

- In an empty cell, type the command **=stdev(**
- Once the program recognizes the command, the cursor will blink next to the bracket.
- Select your data with your mouse, then click Enter.

Step Four: Determine which type of two sample t-test is most appropriate (independent t-test or dependent t-test).

- Independent t-test: Used to determine whether the difference between the means of two independent groups (both of which are being tested for the same dependent variable) is statistically significant.
 - If the independent t-test is most appropriate, determine which formula variation is most appropriate.
- Dependent t-test (Paired t-test): Used to determine whether the difference between the means of two groups (each containing the same participants and being tested at two different points) is statistically significant.

Type of t-test	Formula	When to Use	Degrees of Freedom (v)
Independent t-test	$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \right] \left(\frac{n_1+n_2}{n_1n_2} \right)}}$	Sample sizes are unequal AND either one or both samples are small ($n < 30$).	n_1+n_2-2
Independent t-test	$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{1}{n}(s_1^2 + s_2^2)}}$	Sample sizes are equal (regardless of size)	$2(n-1)$ (This is the same as n_1+n_2-2 because $n_1=n_2$)
Independent t-test	$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)}}$	Sample sizes are unequal AND both sample sizes are large ($n > 30$)	n_1+n_2-2
Dependent t-test (Paired t-test)	$t = \frac{ \Sigma D }{\sqrt{\left[\frac{n(\Sigma D^2) - (\Sigma D)^2}{(n-1)} \right]}}$	Use with paired samples (i.e. subjects pretest and posttest)	$n-1$

Step Five: Use the t-test formula to calculate the t-statistic and determine the degrees of freedom.

a. Calculate the t-statistic using the appropriate formula.

- If completing a dependent t-test, you will also need to calculate the Difference Score (D) for each participant in the study group. See Example 3 on the Example T-Test Calculations Resource Sheet.
 $D = \text{Patient 1's mean before treatment} - \text{Patient 1's mean after treatment}$

b. Calculate the degrees of freedom (v).

$$v = n - 1$$

Step Six: Use the t-table to compare the t-statistic with the critical values to determine whether to reject or fail to reject the null hypothesis. If you have a two tailed test ($H_a: \mu_1 \neq \mu_2$), remember to use $\alpha = \alpha/2$ for the alpha value.

To find the critical value for the data

- Locate the column that corresponds to the alpha value.
- Locate the row that corresponds to the degrees of freedom (v).
- The intersection of the two is the critical value for the data.

v	α						
	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Source: Devore, Jay, and Kenneth Berk. Modern Mathematical Statistical with Applications. Belmont: Brooks/Cole, 2007.

- If the t-statistic is greater than the critical value, then you will reject the null hypothesis.
- If the t-statistic is smaller than the critical value, then you will fail to reject the null hypothesis.

Step Seven: Based on your results in step six, state whether your results are or are not statistically significant.

- If the t-statistic is greater than the critical value, then you will reject the null hypothesis. This means that the results **ARE STATISTICALLY SIGNIFICANT**.
- If the t-statistic is smaller than the critical value, then you will fail to reject the null hypothesis. This suggests that there is no difference between the two groups and the results are **NOT STATISTICALLY SIGNIFICANT**.

Statistics Terminology Resource Sheet

Alternative Hypothesis: What the experimenter thinks may be true or wishes to be true before he or she begins an experiment. Also called the research hypothesis.

Causal Relationship: A relationship in which one factor can be said to be the cause of another. This relationship can be tested with controlled experiments.

Confounding Variable: A variable that is not accounted for in the experimental design, varies systematically with the dependent variable, and prevents a clear interpretation of the effect of the independent variable on the dependent variable.

Controlled, Randomized Experiment: The most powerful experimental design because it allows for the inference of causation. The participants are randomly chosen from a population and randomly assigned to one of two (or more) groups, usually an experimental group that receives some treatment and a control group that receives a placebo.

Correlation: The degree of relationship between two variables.

Cross Sectional Study: A research study that involves a single observation (such as an interview or questionnaire), which may be useful to determine

Statistics Terminology Resource Sheet

Alternative Hypothesis: What the experimenter thinks may be true or wishes to be true before he or she begins an experiment. Also called the research hypothesis.

Causal Relationship: A relationship in which one factor can be said to be the cause of another. This relationship can be tested with controlled experiments.

Confounding Variable: A variable that is not accounted for in the experimental design, varies systematically with the dependent variable, and prevents a clear interpretation of the effect of the independent variable on the dependent variable.

Controlled, Randomized Experiment: The most powerful experimental design because it allows for the inference of causation. The participants are randomly chosen from a population and randomly assigned to one of two (or more) groups, usually an experimental group that receives some treatment and a control group that receives a placebo.

Correlation: The degree of relationship between two variables.

Cross Sectional Study: A research study that involves a single observation (such as an interview or questionnaire), which may be useful to determine how variables affect each other at the same time and period.

Degrees of Freedom: A parameter that is equal to the number of observations or groups in a study minus some value(s) that limit the observations' or groups' freedom to vary.

Dependent Variable: In an experiment, a measure expected to vary across different levels of the independent variable. It is also called the response variable.

Directional Alternative Hypothesis: Where the alternative hypothesis is specifically stated beforehand; for example, Group 1's mean is greater than Group 2's mean. This is also called a one-tailed test of significance.

Double-Blind Experiment: Neither the participant nor the treatment allocator knows whether the participant has been allocated to a treatment or control group.

Experimental Study: A type of study in which all variable measurements and manipulations are under research control. The goal of an experiment is to control the environment in such a way that manipulation of the independent variable yields a direct, corresponding change in the dependent variable.

Independent t-test: A statistical test to determine whether there are significant differences between two independent groups' means being tested on the same dependent variable.

ine how variables affect each other at the same time and period.

Degrees of Freedom: A parameter that is equal to the number of observations or groups in a study minus some value(s) that limit the observations' or groups' freedom to vary.

Dependent Variable: In an experiment, a measure expected to vary across different levels of the independent variable. It is also called the response variable.

Directional Alternative Hypothesis: Where the alternative hypothesis is specifically stated beforehand; for example, Group 1's mean is greater than Group 2's mean. This is also called a one-tailed test of significance.

Double-Blind Experiment: Neither the participant nor the treatment allocator knows whether the participant has been allocated to a treatment or control group.

Experimental Study: A type of study in which all variable measurements and manipulations are under research control. The goal of an experiment is to control the environment in such a way that manipulation of the independent variable yields a direct, corresponding change in the dependent variable.

Independent t-test: A statistical test to determine whether there are significant differences between two independent groups' means being tested on the same dependent variable.

Independent Variable: In an experiment, the variable that the experimenter manipulated. It may also be called the treatment variable.

Mean: The arithmetical average of a group of scores.

Median: The center of a distribution of score, such that half of the scores are above that number and half of the score in the distribution are below that number.

Mode: The most frequently occurring score.

Nondirectional Alternative Hypothesis: Where the null hypothesis will be rejected if either Group 1's mean exceeds Group 2's mean, or vice versa, or where the null hypothesis will be rejected if a relationship exists, regardless of its nature. This is also called a two-tailed test of significance.

Null Hypothesis: The starting point in scientific research where the experimenter assumes there is no effect of the treatment or no relationship between the two variables.

Observational Study: A type of study in which individuals are observed or certain outcomes are measured. No attempt is made to affect the outcome (for example, no treatment is given).

P Value:The probability of results of the experiment being attributed to chance.

Dependent t-test:A test designed to determine the statistical difference between two groups' means where the participants in each group are either the same or matched pairs. Dependent t-tests are also referred to as paired t-tests.

Prospective Cohort Study:A research study that follows over time groups of individuals who are alike in many ways but differ by a certain characteristic (for example, female nurses who smoke and those who do not smoke) and compares them for a particular outcome (such as lung cancer).

Reliability:The likelihood that the same experiment will yield identical results if performed many times over, by the same team or by other teams.

Replication:The repeated ability to duplicate the results of a scientific experiment by different experimenters, which helps establish a hypothesis's usefulness (or non-usefulness in cases where findings cannot be replicated).

Retrospective Cohort Study:A research study in which the medical records of groups of individuals who are alike in many ways but differ by a certain characteristic (for example, female nurses who smoke and those who do not smoke) are compared for a particular outcome (such as lung cancer). This is also called a historic cohort study.

Single-Blind Experiment:The participant does not know whether he or she has been allocated to a treatment or control group.

Standard Deviation: A parameter of variability of data about the mean score.

Statistically Nonsignificant: Describes a mathematical measure of difference between groups. The difference is said to be statistically nonsignificant if it is not greater than what might be expected to happen by chance alone. This means that there is not enough evidence to reject the null hypothesis, and the results of the experiment are attributed to chance.

Statistically Significant:Describes a mathematical measure of difference between groups. The difference is said to be statistically significant if it is greater than what might be expected to happen by chance alone. This means that there is enough evidence to reject the null hypothesis and the results of the experiment are not attributed to chance.

Triple-Blind Experiment:Neither the participant, the treatment allocator, nor the response gatherer knows whether the participant has been allocated to a treatment or control group.

Variance: A parameter of variability of data about the mean score, which is the square of the standard deviation.

Validity:The extent to which the result cannot be attributed to error in sampling or measurement.

Example One:

A researcher wants to determine if the claim that males are less likely than females to develop a cold is true. To test this claim, the researcher followed 30 subjects for the course of three years (17 males and 13 females) to determine how many colds each subject developed.

The researcher hypothesized that males will develop fewer colds during a period of three years than females. The results of the study are presented below. The researcher will use an alpha value equal to 0.05.

Number of Colds (in Three Years)	
Males (X_1)	Females (X_2)
3	8
8	8
6	7
7	10
4	11
9	4
2	3
5	7
7	6
11	8
10	4
8	10
7	6
6	
7	
8	
5	

Step One: Identify null hypothesis and alternative hypothesis.

- Null hypothesis (H_0): Females develop the same number of colds as males.
 - $H_0: \mu_F = \mu_M$
- Alternative hypothesis (H_a): Males develop fewer colds than females.
 - $H_a: \mu_F > \mu_M$

Step Two: Set the level of significance

For this experiment assume that the alpha value is 0.05.

Step Three: Calculate the mean, variance, and standard deviation for both samples. Check your answers using Excel (optional).

a. Calculate the mean.

$$\bar{x} = \frac{\sum x_i}{n}$$

Note: Σ is the sum of the sample and n is the sample size

x_1 :
 $n=17$

x_2 :
 $n=13$

$$\Sigma x_1 = 3+8+6+7+4+9+2+6+7+11+10+8+7+6+7+8+5=113$$

$$\Sigma x_2 = 8+8+7+10+11+4+3+7+6+8+4+10+6=92$$

$$\bar{x}_1 = \frac{113}{17} = 6.647$$

$$\bar{x}_2 = \frac{92}{13} = 7.077$$

b. Calculate the variance (s^2) for both samples and round the answer to the nearest thousandth.

- Calculate the deviation from the mean for each data point.
- Square these deviations.
- Calculate the sum of all of these squared deviations.
- Divide by the number of data points in the data set minus one.

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}$$

n=17

$$\bar{x}_1 = 6.647$$

$$3-6.647=(-3.647)^2=13.301$$

$$8-6.647=(1.353)^2=1.831$$

$$6-6.647=(-0.647)^2=0.419$$

$$7-6.647=(0.353)^2=0.125$$

$$4-6.647=(-2.647)^2=7.001$$

$$9-6.647=(2.353)^2=5.537$$

$$2-6.647=(-4.647)^2=21.595$$

$$5-6.647=(-1.647)^2=2.713$$

$$7-6.647=(0.353)^2=0.125$$

$$11-6.647=(4.353)^2=18.949$$

$$10-6.647=(3.353)^2=11.243$$

$$8-6.647=(1.353)^2=1.831$$

$$7-6.647=(0.353)^2=0.125$$

$$6-6.647=(-0.647)^2=0.419$$

$$7-6.647=(0.353)^2=0.125$$

$$8-6.647=(1.353)^2=1.831$$

$$5-6.647=(-1.647)^2=2.713$$

n=13

$$\bar{x}_1 = 7.077$$

$$8-7.077=(0.923)^2=0.852$$

$$8-7.077=(0.923)^2=0.852$$

$$7-7.077=(-0.077)^2=0.006$$

$$10-7.077=(2.923)^2=8.544$$

$$11-7.077=(3.923)^2=15.39$$

$$4-7.077=(-3.077)^2=9.468$$

$$3-7.077=(-4.077)^2=16.622$$

$$7-7.077=(-0.077)^2=0.006$$

$$6-7.077=(-1.077)^2=1.16$$

$$8-7.077=(0.923)^2=0.852$$

$$4-7.077=(-3.077)^2=9.468$$

$$10-7.077=(2.923)^2=8.544$$

$$6-7.077=(-1.077)^2=1.16$$

$$\Sigma(x_1 - \bar{x}_1)^2 = 89.883$$

$$\Sigma(x_2 - \bar{x}_2)^2 = 72.924$$

$$\text{Sample 1 Variance } (s_1^2) = \frac{89.883}{17-1} = 5.618 \quad \text{Sample 2 Variance } (s_2^2) = \frac{72.924}{13-1} = 6.077$$

- c. Calculate the standard deviation (s) by taking the square root of the variance for both samples.

$$s = \sqrt{\frac{(\mathbf{x} - \bar{\mathbf{x}})^2}{\mathbf{n} - 1}}$$

$$\begin{array}{ll} \text{sample 1 variance } (s_1^2)=5.618 & \text{sample 2 variance } (s_2^2)=6.077 \\ \text{standard deviation } (s_1) = \sqrt{5.618} = 2.370 & \text{standard deviation } (s_2) = \sqrt{6.077} = 2.467 \end{array}$$

Step Four: Determine which type of t-test is most appropriate (independent t-test versus dependent t-test).

The samples are independent of each other (i.e., different subjects were used in the two samples). Therefore, an two-sample independent t-test is the most appropriate t-test to use. The two samples are unequal AND both sample sizes are small (<30), therefore the following formula should be used:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \right] \left(\frac{n_1+n_2}{n_1n_2} \right)}}$$

\bar{x} = mean
 s^2 = variance
 n = sample size

Step Five: Use the t-test formula to calculate the t-statistic and the degrees of freedom.

- a. Calculate the t-statistic.

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \right] \left(\frac{n_1+n_2}{n_1n_2} \right)}}$$

$$t = \frac{(6.647 - 7.077)}{\sqrt{\left[\frac{(17-1)5.618 + (13-1)6.077}{17+13-2} \right] \left(\frac{17+13}{17 \times 13} \right)}}$$

$$t = \frac{0.43}{\sqrt{\left[\frac{89.888 + 72.924}{28} \right] \left(\frac{30}{221} \right)}}$$

$$t = \frac{0.43}{\sqrt{\left[\frac{162.812}{28} \right] (0.136)}}$$

$$t = \frac{0.43}{\sqrt{(5.815)(0.136)}}$$

$$t = \frac{0.43}{\sqrt{(0.791)}}$$

$$t = \frac{0.43}{0.889} = 0.484$$

b. Calculate the degrees of freedom (df).

Degrees of Freedom (df) = $n_1 + n_2 - 2$

$$v = 17 + 13 - 2 = 28$$

Step Six: Use the t-table to compare the t-statistic with the critical values to determine whether to reject or fail to reject the null hypothesis. If you have a two tailed test ($H_a: \mu_1 \neq \mu_2$), remember to use $\alpha = \alpha/2$ for the alpha value.

ν	α			
	.10	.05	.025	.01
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602
16	1.337	1.746	2.120	2.583
17	1.333	1.740	2.110	2.567
18	1.330	1.734	2.101	2.552
19	1.328	1.729	2.093	2.539
20	1.325	1.725	2.086	2.528
21	1.323	1.721	2.080	2.518
22	1.321	1.717	2.074	2.508
23	1.319	1.714	2.069	2.500
24	1.318	1.711	2.064	2.492
25	1.316	1.708	2.060	2.485
26	1.315	1.706	2.056	2.479
27	1.314	1.703	2.052	2.473
28	1.313	1.701	2.048	2.467
29	1.311	1.699	2.045	2.462

Because the t-statistic(0.484) is less than the critical value in the t-table (1.701), we **fail to reject** the null hypothesis.

Step Seven:Based on your results in Step Six, state whether your results are or are not statistically significant.

Based on Step 6, we can conclude that there is no difference between the numbers of colds developed in males versus females. The difference between the means is **not statistically significant**.

Example Two:

A researcher wants to determine if biofeedback therapy reduces stress in hospital workers exposed to extremely stressful situations. Sixteen participants were involved in the study. Eight participants participated in ten biofeedback therapy sessions versus the other eight participants who did not participate in any biofeedback therapy sessions.

The researcher hypothesized that the participants participating in biofeedback therapy sessions should have lower stress scores than the participants who did not participate in any biofeedback therapy sessions. Assume that alpha is 0.005.

The results of the study are presented below:

Stress Scores of Participants Given Biofeedback Therapy - Experimental Group (X_1)	Stress Scores of Participants Not Given Biofeedback Therapy - Control Group (X_2)
53	64
44	68
61	84
50	65
40	50
39	60
58	81
42	76

Step One: Identify null hypothesis and alternative hypothesis.

- Null hypothesis: The stress score of those participating in biofeedback therapy sessions (μ_1) will be the same as those not participating in biofeedback therapy sessions (μ_2).
 - $H_0: \mu_1 = \mu_2$
- Alternative hypothesis: The stress score of those participating in biofeedback therapy sessions will be less than those not participating in biofeedback therapy sessions.
 - $H_a: \mu_1 < \mu_2$

Step Two: Set the level of significance

For this experiment assume that alpha value is 0.005.

Step Three: Calculate the mean, variance, and standard deviation for both samples. Check your answers using Excel (optional).

a. Calculate the mean.

$$\bar{x} = \frac{\sum x_i}{n}$$

Note: Σ is the sum of the sample and n is the sample size

X_1 :
 $n=8$

X_2 :
 $n=8$

$$\Sigma X_1 = 53 + 44 + 61 + 50 + 40 + 39 + 58 + 42 = 387$$

$$\Sigma X_2 = 64 + 68 + 84 + 65 + 50 + 60 + 81 + 76 = 548$$

$$\bar{x}_1 = \frac{387}{8} = 48.375$$

$$\bar{x}_2 = \frac{548}{8} = 68.5$$

b. Calculate the variance (s^2) for both samples and round the answer to the nearest thousandth.

- Calculate the deviation for each data point.
- Square these deviations.
- Calculate the sum of all of these squared deviations.
- Divide by the number of data points in the data set minus one.

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}$$

$$n=8$$

$$\bar{x}_1 = 48.375$$

$$53 - 48.375 = (4.625)^2 = 21.391$$

$$44 - 48.375 = (-4.375)^2 = 19.141$$

$$61 - 48.375 = (12.625)^2 = 159.188$$

$$50 - 48.375 = (1.625)^2 = 2.641$$

$$40 - 48.375 = (-8.375)^2 = 70.141$$

$$39 - 48.375 = (-9.375)^2 = 87.891$$

$$58 - 48.375 = (9.625)^2 = 92.641$$

$$42 - 48.375 = (-6.375)^2 = 40.641$$

$$\Sigma(x_1 - \bar{x}_1)^2 = 493.878$$

$$s^2 = \frac{493.878}{8 - 1} = 70.554$$

$$n=8$$

$$\bar{x}_1 = 68.5$$

$$64 - 68.5 = (-4.5)^2 = 20.25$$

$$68 - 68.5 = (-0.5)^2 = 0.25$$

$$84 - 68.5 = (15.5)^2 = 240.25$$

$$65 - 68.5 = (-3.5)^2 = 12.25$$

$$50 - 68.5 = (-18.5)^2 = 342.25$$

$$60 - 68.5 = (-8.5)^2 = 72.25$$

$$81 - 68.5 = (12.5)^2 = 156.25$$

$$76 - 68.5 = (7.5)^2 = 56.25$$

$$\Sigma(x_2 - \bar{x}_2)^2 = 900$$

$$s^2 = \frac{900}{8 - 1} = 128.571$$

b. Calculate the standard deviation (s) by taking the square root of the variance.

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$

$$\text{Variance } X_1 = 70.554$$

$$s = \sqrt{70.554} = 8.4$$

$$\text{Variance } X_2 = 128.571$$

$$s = \sqrt{128.571} = 11.339$$

Step Four: Determine which type of t-test is most appropriate (independent t-test versus dependent t-test).

The samples are independent (different subjects were used in sample 1 than in sample 2). Therefore, the independent t-test is the most appropriate t-test to use. The two samples are equal; therefore, the following formula should be used:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{1}{n}(s_1^2 + s_2^2)}}$$

\bar{x} = mean
 s^2 = variance
 n = sample size

Step Five: Use the t-test formula to calculate the t-statistic and determine the degrees of freedom.

a. Calculate the t-statistic.

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{1}{n}(s_1^2 + s_2^2)}}$$

$$t = \frac{(48.375 - 68.5)}{\sqrt{\frac{1}{8}(70.554 + 128.571)}}$$

$$t = \frac{20.125}{\sqrt{\frac{1}{8}(199.125)}}$$

$$t = \frac{20.125}{\sqrt{24.891}}$$

$$t = \frac{20.125}{4.989} = 4.034$$

b. Calculate the degrees of freedom.

$$\text{Degrees of Freedom} = 2(n-1)$$

$$dF = 2(8-1) = 2(7) = 14$$

Step Six: Use the t-table to compare the computed t-statistic with the critical values to determine whether the results are statistically significant.

ν	α				
	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977

The t-statistic of 4.034 is larger than the critical value (2.977), so we can **reject** the null hypothesis.

Step Seven:Based on your results in Step Six, state whether your results are or are not statistically significant.

Based on the results from Step 6, we can conclude that the results are **statistically significant.**

Example Three:

A researcher wants to determine if an experimental drug will help morbidly obese patients (more than 100 lbs over ideal weight) lose weight. The patient's weights were measured before treatment and then again six months after administration of treatment. The researcher hypothesizes that the patients' weights will be less after treatment. The results of the study are presented below. Assume alpha is 0.005.

Patient	Weight before treatment (lbs)	Weight after treatment (lbs)
1	245	202
2	340	284
3	298	244
4	355	285
5	321	325
6	414	340
7	309	249

Step One: Identify null hypothesis and alternative hypothesis.

- Null hypothesis: There is no difference in the weights of the patients before and after treatment.
- Alternative hypothesis: The weights of the patients after treatment will be less than before treatment.

Step Two: Set the level of significance.

For this experiment assume that alpha value is 0.005.

Step Three: Calculate the mean, variance, and standard deviation for both samples. Check your answer using Excel (optional).

a. Calculate the mean.

$$\bar{x} = \frac{\sum x_i}{n}$$

Note: Σ is the sum of the sample and n is the sample size

X_1 :

X_2 :

$n=7$

$n=7$

$$\Sigma X_1 = 245 + 340 + 298 + 355 + 321 + 414 + 309 = 2282$$

$$\Sigma X_2 = 202 + 284 + 244 + 285 + 325 + 340 + 249 = 1926$$

$$\bar{x}_1 = \frac{2282}{7} = 326$$

$$\bar{x}_2 = \frac{1926}{7} = 275.571$$

b. Calculate the variance (s^2) for both samples and round the answer to the nearest thousandth.

- Calculate the deviation for each data point.
- Square these deviations.
- Calculate the sum of all of these squared deviations.
- Divide by the number of data points in the data set minus one.

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}$$

n=7

$$\bar{x}_1 = 326$$

$$245-326=(-81)^2=6561$$

$$340-326=(14)^2=196$$

$$298-326=(-28)^2=784$$

$$355-326=(29)^2=841$$

$$321-326=(-5)^2=25$$

$$414-326=(88)^2=7744$$

$$309-326=(-17)^2=289$$

$$\Sigma(x_1 - \bar{x}_1)^2 = 16,440$$

$$s_1^2 = \frac{16,440}{7 - 1} = 2,740$$

n=7

$$\bar{x}_1 = 275.571$$

$$202-275.571=(-73.571)^2=5412.692$$

$$284-275.571=(8.429)^2=71.048$$

$$244-275.571=(-31.571)^2=996.728$$

$$285-275.571=(9.429)^2=88.906$$

$$325-275.571=(49.429)^2=2443.226$$

$$340-275.571=(64.429)^2=4151.096$$

$$249-275.571=(-26.571)^2=706.018$$

$$\Sigma(x_2 - \bar{x}_2)^2 = 13,869.714$$

$$s_2^2 = \frac{13,869.714}{7 - 1} = 2,311.619$$

c. Calculate the standard deviation for both samples by taking the square root of the variance.

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$

$$s_1^2 = 2,740$$

$$s_2^2 = 2,311.619$$

$$s_1 = \sqrt{2,740} = 52.345$$

$$s_2 = \sqrt{2,311.619} = 48.079$$

Step Four: Determine which type of t-test is most appropriate (independent t-test versus dependent t-test).

The samples are paired (the same subjects were used in sample 1 as in sample 2). Therefore, a dependent t-test is the most appropriate t-test to use.

Step Five: Use the t-test formula to calculate the t-statistic and degrees of freedom.

a. Calculate the t-statistic.

$$t = \frac{|\Sigma D|}{\sqrt{\left[\frac{n(\Sigma D^2) - (\Sigma D)^2}{(n-1)} \right]}}$$

n=sample size

D=Patient before treatment – patient after treatment

Step One: Calculate the difference scores (D) for each patient

Patient	Weight before treatment (lbs)	Weight after treatment (lbs)	D
1	245	202	245-202=43
2	340	284	340-284=56
3	298	244	298-244=54
4	355	285	355-285=70
5	321	325	321-325=-4
6	414	340	414-340=74
7	309	249	309-249=60

Step Two: Calculate D² for each patient

Patient	Weight before treatment (lbs)	Weight after treatment (lbs)	D	D ²
1	245	202	245-202=43	43 ² =1849
2	340	284	340-284=56	56 ² =3136
3	298	244	298-244=54	54 ² =2916
4	355	285	355-285=70	70 ² =4900
5	321	325	321-325=-4	(-4) ² =16
6	414	340	414-340=74	74 ² =5476
7	309	249	309-249=60	60 ² =3600

$$t = \frac{|\Sigma D|}{\sqrt{\left[\frac{n(\Sigma D^2) - (\Sigma D)^2}{(n-1)} \right]}}$$

$$t = \frac{|353|}{\sqrt{\left[\frac{7(21,893) - (124,609)^2}{(7-1)} \right]}}$$

$$t = \frac{|353|}{\sqrt{\left[\frac{153,251 - 124,609}{6} \right]}}$$

$$t = \frac{|353|}{\sqrt{\left[\frac{28642}{6} \right]}}$$

$$t = \frac{|353|}{\sqrt{4773.667}} = 5.109$$

b. Calculate the degrees of freedom (v).

$$v = n - 1$$

$$v = 7 - 1 = 6$$

Step Six: Use the t-table to compare the computed t-statistic with the critical values to determine whether the results are statistically significant.

v \ α	α				
	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977

Based on the comparison of the t-statistic of 5.109 to the critical value (3.707), we can conclude that the null hypothesis should be rejected.

Step Seven: Based on your results in Step Six, state whether your results are or are not statistically significant.

Based on the results from Step 6, we can conclude that the results are **statistically significant**.

Part I: Example One

A researcher wants to determine if the claim that males are less likely than females to develop a cold is true. To test this claim, the researcher followed 30 subjects for the course of three years (17 males and 13 females) to determine how many colds each subject developed.

The researcher hypothesized that males will develop fewer colds during a period of three years than females. The results of the study are presented below. Assume alpha is 0.05.

Number of Colds (in Three Years)	
Males (X_1)	Females (X_2)
3	8
8	8
6	7
7	10
4	11
9	4
2	3
5	7
7	6
11	8
10	4
8	10
7	6
6	
7	
8	
5	

Step One: Identify null hypothesis and alternative hypothesis.

- Null hypothesis (H_0):

Females develop the same number of colds as males

- Alternative hypothesis (H_a):

Males develop fewer colds than females

Step Two: Establish the level of significance, which is known as the alpha value. For the examples, the alpha value will be given.

Step Three: Calculate the mean, variance, and standard deviation for both samples. Check your answer using Excel (optional).

a. Calculate the mean.

$$\bar{x} = \frac{\sum x_i}{n}$$

Note: Σ is the sum of the sample and n is the sample size.

$$\bar{x}_1 = n=17,$$

$$6.64$$

$$\bar{x}_2 =$$

b. Calculate the variance (s^2) for both samples and round your answer to the nearest thousandth.

- Calculate the deviation for each data point.
- Square these deviations.
- Calculate the sum of all of these squared deviations.
- Divide by the number of data points in the data set minus one.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

$$\text{Sample 1 Variance } (s_1^2) =$$

$$\text{Sample 2 Variance } (s_2^2) =$$

c. Calculate the standard deviation (s) by taking the square root of the variance for both samples.

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Sample 1 standard deviation (s_1)=

Sample 2 standard deviation (s_2)=

Step Four: Determine which type of two sample t-test is most appropriate (independent t-test or dependent t-test).

The samples are independent of each other (i.e., different subjects were used in the two samples). Therefore, an two-sample independent t-test is the most appropriate t-test to use. The two samples are unequal AND both sample sizes are small (<30), therefore the following formula should be used:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \right] \left(\frac{n_1+n_2}{n_1n_2} \right)}}$$

\bar{x} = mean
 s^2 = variance
 n = sample size

Step Five: Use the t-test formula to calculate the t-statistic and determine the degrees of freedom.

a. Calculate the t-statistic.

b. Calculate the degrees of freedom (v).

$$df = 17 + 13 - 2 = 28$$

Step Six: Use the t-table to compare the t-statistic with the critical values to determine whether to reject or fail to reject the null hypothesis. If you have a two tailed test ($H_a: \mu_1 \neq \mu_2$), remember to use $\alpha = \alpha/2$ for the alpha value.

ν	α						
	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Source: Devore, Jay, and Kenneth Berk. Modern Mathematics Statistical with Applications. Belmont: Brooks/Cole, 2007.

t=_____

Step Seven: The t-value of is less than the critical value of at df=28 and level of significance of .05. So, We fail to reject null hypothesis. There is no difference between the number of colds developed in males versus females. The results are NOT statistically significant.

Part II: Assigned Study

Step One: Identify null hypothesis and alternative hypothesis.

- Null hypothesis (H_0):
- Alternative hypothesis (H_a):

Step Two: Establish the level of significance, which is known as the alpha value. For the examples, the alpha value will be given.

Step Three: Calculate the mean, variance, and standard deviation for both samples. Check your answer using Excel (optional).

a. Calculate the mean.

$$\bar{x} = \frac{\sum x_i}{n}$$

Note: Σ is the sum of the sample and n is the sample size.

$$\bar{x}_1 =$$

$$\bar{x}_2 =$$

b. Calculate the variance (s^2) for both samples and round your answer to the nearest thousandth.

- Calculate the deviation for each data point.
- Square these deviations.
- Calculate the sum of all of these squared deviations.
- Divide by the number of data points in the data set minus one.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

$$\text{Sample 1 Variance } (s_1^2) =$$

Sample 2 Variance (s_2^2)=

c. Calculate the standard deviation (s) by taking the square root of the variance for both samples.

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Sample 1 standard deviation (s_1)=

Sample 2 standard deviation (s_2)=

Step Four: Determine which type of two sample t-test is most appropriate (independent t-test or dependent t-test).

Step Five: Use the t-test formula to calculate the t-statistic and determine the degrees of freedom.

c. Calculate the t-statistic.

d. Calculate the degrees of freedom (v).

Step Six: Use the t-table to compare the t-statistic with the critical values to determine whether to reject or fail to reject the null hypothesis. If you have a two tailed test ($H_a: \mu_1 \neq \mu_2$), remember to use $\alpha = \alpha/2$ for the alpha value.

ν	α						
	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Source: Devore, Jay, and Kenneth Berk. Modern Mathematics Statistical with Applications. Belmont: Brooks/Cole, 2007.

t= _____

Step Seven:Based on your results in Step Six, state whether your results are or are not statistically significant.