Statistics 1.

T-Test Review

2014. Prepared by Lauren Pincus

With input from Mark Bell

Agriculture Innovation Program

Table of contents

Questions for review 3
Material to Review 5
1. One-Sample T-Test 5
1.1 Establish Hypotheses 5
1.2 Calculate Test Statistic 5
1.3 Use This Value to Determine P-Value 6
1.4 Test your understanding. 7
2. Independent (Two Sample) T-Test 9
2.1 Establish Hypotheses 9
2.2 Calculate Test Statistic 9
2.3 Use This Value To Determine P-Value 10
2.4 Test your understanding. 10
3. Paired T-Test. 12
3.1 Establish Hypotheses 12
3.2 Calculate Test Statistic 12
3.3 Use This Value to Determine P-Value 13
3.4 Test your understanding. 13
Definition of terms 14

Questions for review

Situation 1

1. The researcher knows that, nationally, wheat yields are 2,600 kg per ha^{-1}. The researcher measures wheat yields after side-dressing and finds that yields are 3,085 kg per ha^{-1} with a standard deviation of 506. If the researcher is interested in comparing the yields of the side-dressed wheat with national average yields, then the researcher would want to use a

   a) CRD
   b) independent t-test
   c) paired t-test
   d) one-sample t-test
   e) RCBD
   f) Split plot

2. Which of the following is an appropriate null hypothesis for this study?

   a) National yields are greater than the side-dressed wheat yields.
   b) The difference between the national average yield and the yields of side-dressed wheat is zero.
   c) Side-dressed wheat yields are greater than the national yields.

3. Write an appropriate alternative hypothesis.

   ____
   ____
   ____

Answer 1

1. d) One-sample t-test The one-sample t-test compares the mean score of a sample to a known value, usually the population mean.
2. b) “The difference between the national average yield and the yields of side-dressed wheat is zero.”
3. The difference between the national average yield and the yields of side-dressed wheat is NOT zero.

Situation 2

1. A livestock researcher has 20 cows which he randomly divides into two groups. One group receives a feed supplement and the other does not. Body weight is measured for both groups after 6 months. Since the researcher is comparing the means of these two sample groups, the researcher wants to use a

a) CRD
b) independent t-test
c) paired t-test
d) one-sample t-test
e) RCBD
f) Split plot

2. If the researcher stated that the null hypothesis is "body weight of the supplemented cattle are GREATER than weight from cattle that do not receive a supplement," then the researcher would want to use a

one-tailed t-test	☐ Yes ☐ No
two-tailed t-test	☐ Yes ☐ No

Answer 2

1. b) Independent t-test **Use an independent t-test** when you want to compare the mean of one sample with the mean of another sample to see if there is a statistically significant difference between the two. As the name suggests, you **use an independent t-test** when your samples are **independent**
2. One-tailed A one-tailed test is used for hypotheses which state that the observed mean is either less than or greater than the true mean (H_1: $\mu > \mu_0$ ***OR*** H_1: $\mu < \mu_0$), but not both. If you want to test both greater than and/or less than (i.e., H_1: $\mu \neq \mu_0$), then you need to use a **two-tailed test**.

Situation 3

A researcher believes that tomato variety has an effect on yield response to side-dressed nitrogen. To account for this, the researcher grows two seedlings each of 10 tomato varieties. One seedling of each variety gets side-dressed nitrogen and the other receives no side-dressing. In this study, the observation for each variety in a sample is matched with the observation for the same variety in the other sample. To properly compare these two groups (with and without side-dressing), the researcher would want to use a

a) CRD
b) independent t-test
c) paired t-test
d) one-sample t-test
e) RCBD
f) Split plot

Answer 3

3. c) paired t-test A **paired t-test** is used to compare two population means where you have two samples in which observations in one sample can be **paired** with observations in the other sample.

Material to Review

Test	Summary
1. One-Sample T-Test	Compare the mean score of a sample to a known value
2. Independent (two-sample) T-Test	Compare the mean of one sample with the mean of another sample
3. Paired T-Test	Determine whether there is a difference between the average values of paired samples subjected to two different conditions

1. One-Sample T-Test

The one-sample t-test compares the mean score of a sample to a known value, usually the population mean (the average for the outcome of some population of interest). The test is a comparison of <u>the mean of the sample</u> (observed average) and the <u>population mean</u> (expected average), with an adjustment for the number of cases in the sample and the standard deviation of the average.

1.1 Establish Hypotheses

The null hypotheses can take one of two forms:

- difference between observed and expected is 0, or
- difference between observed and expected is <u>not</u> 0

1.2 Calculate Test Statistic

Calculation of the test statistic requires four components:

1. The average of the sample (observed average)
2. The number of observations (n).
3. The population average or other known value (expected average)
4. The standard deviation of the sample, or observed, average

With these four pieces of information, we calculate the following test statistic, t:

$$t = \frac{(\text{observed} - \text{expected})}{\text{SD}_{\text{observed}} \times \sqrt{(\text{number of observations in sample} \,/\, \text{number of observations} - 1)}}$$

Standard deviation can be calculated using the following equation.

$$S = \sqrt{\frac{\Sigma(X - \overline{X})^2}{N}}$$

where S = the standard deviation of a sample,
Σ means "sum of,"
X = each value in the data set,
$\overline{X}$ = mean of all values in the data set,
N = number of values in the data set.

To calculate the t statistic for a sample with 2, 4, 4, 4, 5, 5, 7, 9 as the observations and 4.5 as the known population average.

N (number of observations) = 8
Sample average = (2+4+4+4+5+5+7+9)/8 = 5
Sample standard deviation = (sum(obs-av)^2 / # obs)

✓(($(2-5)^2 + (4-5)^2 + (4-5)^2 + (4-5)^2 + (5-5)^2$ + etc./8)
✓((9+1+1+1+0+0+4+16)/8) = 4
✓ √4
= 2

$$t = \frac{(\text{observed} - \text{expected})}{\text{SD}_{\text{observed}} \times \sqrt{(\text{number of observations in sample} / \text{number of observations} - 1)}}$$

t = (5 - 4.5) / (2 x √ (8 / 7))
= 0.109

1.3 Use *t* Value to Determine P-Value

First chose a significance level, often referred to as alpha (α). A critical value of 5% is usually considered acceptable. This means that we are willing to accept a 5% chance, or probability, of **incorrectly rejecting** the null hypothesis. A critical value is a point on the test distribution that is compared to the test statistic to determine whether to reject the null hypothesis. Critical values correspond to α, so their values become fixed when you choose the test's α. If the absolute value of your test statistic is greater than the critical value, you can declare statistical significance and reject the null hypothesis.

The *p* value is the probability of the test statistic being at least as extreme as the one observed given that the null hypothesis is true. If the *p* value is equal or smaller than our chosen significance level, it suggests that the observed data (the values we collected in our particular sample, is inconsistent with the assumption that the null hypothesis is true. The null hypothesis must be rejected and the alternative hypothesis accepted as true. A p value of 0.001 tells us that there is a 0.001% probability that we mistakenly concluded that our alternative hypothesis is true. Rejecting the null hypothesis when it is actually true is known as a Type I error.

Use an online calculator such as the one linked here to find the *p* value corresponding to our calculated *t* statistic.

http://www.socscistatistics.com/pvalues/tdistribution.aspx

1.4 Test your understanding.

Question 1: You are interested in knowing how the maize yields in your region compare to the national average. The national average is 4 ton ha^{-1}. You sample 10 fields in your region and construct the data set below. Are yields in your region statistically different than the national average?

Sampled Yields
3.5
2.7
6.4
3.4
2.9
3.7
3.1
4.3
2.6
3.0

Answer 1:
Sample average = 32.6 / 10 = 3.56
Sample SD = $\sqrt{}$ ($(3.5 – 3.56)^2 + (2.7 – 3.56)^2 + (6.4 – 3.56)^2 + \ldots + (3.0 – 3.56)^2$ / 10) = 1.0622
N = 10

$$t = \frac{(\text{observed} - \text{expected})}{\text{SD}_{\text{observed}} \times \sqrt{(\text{number of observations in sample} \;/\; \text{number of observations} - 1)}}$$

t = (3.56 – 4.0) / (1.0622 x $\sqrt{(10 / 9)}$) = -0.3929

Using an online calculator tells us the *p* value is 0.704179. The result is *not* significant at $p < 0.05$.

Question 2: You are using SPSS to run a one-sample t-test and receive this output. How would you report the results?

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
maize_yield	40	3.7225	.73709	.11654

One-Sample Test						
	Test Value=4					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
maize_yield	-2.381	39	.022	-.27750	-.5132	-.0418

Answer 2:
Mean regional maize yields (3.73 +/- 0.74) were lower than the national average of 4.0 kg ha^{-1}, a statistically significant difference of 0.26 kg ha^{-1} (95% CI, 0.04 to 0.51), $t(39) = -2.831$, $p = .022$.

2. Independent (two-sample) T-Test

You use an independent t-test when you want to compare the mean of one sample with the mean of another sample to see if there is a statistically significant difference between the two. As the name suggests, you use an independent t-test when your samples are independent. The two-sample t-test is a hypothesis test for answering questions about the mean where the data are collected from two random samples of independent observations, each from an underlying normal distribution.

2.1 Establish Hypotheses

Null hypothesis is that the difference between the two groups is 0. Another way of stating the null hypothesis is that the difference between the mean of the treatment group and the mean of the control group is zero.

Alternative hypothesis - the difference between the two sample means is not zero.

2.2 Calculate Test Statistic

Calculation of the test statistic requires three components:

1. The average of both samples (observed averages). Statistically, we represent these as

$$\overline{x_1} \text{ and } \overline{x_2}$$

2. The standard deviation (SD) of both averages. Statistically, we represent these as

$$SD_1 \text{ and } SD_2$$

3. The number of observations in both populations, represented as

$$n_1 \text{ and } n_2$$

Let's say an analysis of data comparing side-dressed tomatoes and non side-dressed tomatoes showed the following:

	Side-dressed tomatoes	Non side-dressed tomatoes
Average weight	3100 g	2750 g
SD	420	425
N	75	75

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\left(\frac{SD_1^2}{\eta_1} + \frac{SD_2^2}{\eta_2}\right)}}$$

$$t = \frac{3100 - 2750}{\sqrt{\left(\frac{420^2}{75} + \frac{425^2}{75}\right)}}$$

$$t = \frac{350}{\sqrt{2352 + 2408.3}}$$

$$t = 5.07$$

2.3 Use This Value To Determine P-Value

Having calculated the t-statistic, follow the same procedure as above to determine whether the t-statistic reaches the threshold of statistical significance. For independent samples however, degrees of freedom are calculated by (N_1 – 1) + (N_2 – 1).

2.4 Test your understanding.

Independent random samples selected from two melon variety populations produced the results shown below:

	Melon Variety A	Melon Variety B
Average weight	5 kg	7 kg
SD	1.2 kg	0.4 kg
N	42	35

1. What is the null and alternative hypothesis?

2. What is the t statistic?

3. Can the researcher say there is a statistically significant difference between the weight of these two melon varieties?

Answers:

1. Null: There is no difference between the average melon weights.

   Alternative: There is a difference between the average melon weights.

2. *t* statistic = $(5 – 7) / \sqrt{(1.2 / 42) + (0.4 / 35)}$ = -10.15

3. The mean weights between Variety A and Variety B were significantly different ($t(75) = -10.15$, $p < 0.00001$)

3. Paired T-Test

A paired sample t-test is used to determine whether there is a significant difference between the average values of the same measurement made under two different conditions. Both measurements are made on each unit in a sample, and the test is based on the paired differences between these two values.

3.1 Establish Hypotheses

Null hypothesis for the paired sample t-test is H_0: d = μ1 - μ2 = 0

where d is the mean value of the difference.

This null hypothesis is tested against one of the following alternative hypotheses, depending on the question posed:

H_1: d > 0
H_1: d < 0

The paired sample t-test can only be used when we have matched samples.

3.2 Calculate Test Statistic

We get the following data in our study where a researcher takes 2 transplants from 10 varieties and randomly assigns them to either receive side-dressing or not.

Variety ID	Yield of Side-dressed tomatoes	Yield of non side-dressed tomatoes	Differences	Differences^2
1	120	94	26	676
2	156	87	69	4761
3	148	118	30	900
4	95	97	-2	4
5	107	110	-3	9
6	98	82	16	256
7	93	105	-12	144
8	114	95	19	361
9	127	134	-7	49
10	134	121	13	169
Means	119.2	104.3	Sum of differences: 149	Sum of differences^2: 7329

To work out the differences, the yield of side-dressed tomatoes is subtracted from the yield of the non side-dressed tomatoes. This value is then squared to produce the differences^2 values.

The test statistic for the paired sample t-test is given by:

$t = \Sigma$ differences / $\sqrt{}$((N*Σ differences^2 - (Σ differences)^2) / (N-1)).

From the table above it can be seen that the Σ differences = 149 and the Σ differences^2 = 7329.

N = 10

t can be computed as:
$t = 149 / \sqrt{(((10 * 7329) - 149^2) / 9)} = 0.0262$

3.3 Use This Value to Determine P-Value

Having calculated the t-statistic, follow the same procedure as above to determine whether the t-statistic reaches the threshold of statistical significance. For paired samples however, degrees of freedom are computed as N - 1.

3.4 Test your understanding.

Question 1:
Calculate the t statistic for the following paired samples and determine if there is a statistically significant difference between the two samples

Sample 1	Sample 2
7	24
5	22
12	23
9	26
5	29
12	22

Question 2:
Is a paired t test needed for the following scenarios? Answer Yes or No.

1. Comparing the average yields of two legume varieties?
2. Comparing weight gain in goats before and after being fed a special diet?
3. Testing whether shoot length and leaf width are related?

Answer 1:

The value of *t* is 7.745967. The value of *p* is 0.000573. The result is significant at $p \leq 0.05$.

Answer 2:

1. No
2. Yes
3. No

Definition of terms

Variability is a characteristic of biological material. Hence we need to decide whether differences between experimental units result from unaccounted variability or real treatment effects.

An **experimental unit** refers to the unit of material to which a treatment is applied. It can be a single leaf, a whole plant, an area of ground containing many plants, a pot, or in the greenhouse. The term plot is synonymous with experimental unit.

A **treatment** may be an amount of material or a method that is to be tested and compared with other treatments in the experiment; e.g., cultivar fertilizer doses, etc.

Variable: A measurable characteristic of an experimental unit is a variable; plant height, days to flowering, panicle length, or grain yield etc. Individual measurements of a variable are data; 150 cm plant height, 45 days to 50% flowering, 12 cm panicle length, or 3240 kg/ha of grain etc.

Data: A set of observations or measurements of a particular variable in an experiment; 105 cm, 95 cm, 120 cm, 75 cm, 100 cm, ..., represent data about the variable (plant height).

Population: In a statistical sense, a population is a set of measurements or counts of a single variable on all the units in the specified population.

Sample: A sample is a set of measurements (observations) obtained from part of the specified population. For example: All 150 plants in a plot form the population. Ten plants used for recording the plant height from this population form the sample. We obtain information from the sample.

Mean: Mean is the simple arithmetic average (of the sample or population).

Standard Deviation (SD): The measure of dispersion of the data around their mean is the Standard Deviation. SD tells us how scattered the sample observations are around the mean.

Variance: Square of the Standard Deviation.

Standard Error (SE): The SD of the population of means is its standard error. SE tells us how scattered the treatment means (for example) will be.

Degrees of freedom (df): Represents the freedom with which the variability in a data set could be accounted for. Usually, but not necessarily, df is one less than the number of observations (n-1).

Distribution of data

Data on a variable may commonly follow one of the following distributions.

Binomial distribution

This type of distribution is expected in data, which describes the proportion of occurrences in which each occurrence can only be one of two possible outcomes. For example in data representing percent survival of insects, one can expect either dead or alive insects.

Poisson Distribution

This is a distribution, which represents occurrences of rare events. For example, count data such as the number of infested plants, the number of lesions per leaf, or number of weeds per unit area.

Normal Distribution

These data represent a continuous distribution. Most biological data, when plotted in a frequency curve, represent a bell shaped and symmetrical curve. Such data are said to be normally distributed. For example, grain yield, plant height, etc.

Test of Significance:

Variability is a characteristic feature of nature. Two plants growing side by side are not alike even under similar conditions. We also know that data on plant growth characters reveal this variability. For example, from data on five samples plant height such as 115 cm, 95 cm, 82 cm, 108 cm, and 72 cm, ..., can we say the growth was good (115 cm plant height) or the growth was poor (72 cm plant height)? What is the truth?