statstutor

community project

encouraging academics to share statistics support resources

stcp-gilchristsamuels-2

The following resources are associated:

Checking normality for parametric tests worksheet

Independent Samples t-test

Research question type: Examining the difference between two unrelated (independent or unpaired) or sets of measurements

What kind of variables: Scale/continuous outcome (dependent) variable and nominal group (independent) variable

Common applications: Comparing the means of data from two groups, for example intervention and control (no intervention) or those receiving treatment A and treatment B

Example: Calorie content of hot dogs

[Data taken from Moore, D. and McCabe, G. (2002) *Introduction to the Practice of Statistics*, New York: WH Freeman.]

A US magazine, Consumer Reports, carried-out a survey of the calorie content of a number of different brands of hotdog. The calorie content of 20 beef and 17 poultry hotdogs was recorded as below:

Beef	186, 181, 176, 149, 184, 190, 158, 139, 175, 148, 152, 111, 141, 153, 190, 157, 131, 149, 135, 132
Poultry	129, 132, 102, 106, 94, 102, 87, 99, 170, 113, 135, 142, 86, 143, 152, 146, 144

Research question: Is there a **difference** in calorie content between beef and poultry hotdogs?

Calories are **scale data** and meat type is **nominal**. Continuous data are often summarised by giving a mean and standard deviation, and the **independent t-test** is used to compare means of two groups. The independent t-test depends on the means, standard deviations and sample sizes of each group. Various assumptions also need to be made – see validity section below.

Hypotheses

The null hypothesis is:

   H_0: There is no difference in mean calorie content between the beef and poultry hotdogs

And the alternative hypothesis is:

   H_1: There is a difference in mean calorie content between the beef and poultry hotdogs

Steps in SPSS

1. Enter the data into SPSS in 2 columns: the first column indicates the group (*Meat*, coded 1 for beef and 2 for poultry using the values field) and the other has the outcome variable (*CalorieContent*). The data view should then look something like the picture on the right.

2. Describe the data sets using Analyze – Descriptive Statistics – Explore. Select *CalorieContent* for the dependent variables list and *Meat* for the factor list. Under Plot… choose Normality plots with tests. This provides the following information:

	Meat	CalorieContent
1	1	186
2	1	181
3	1	176
4	1	149
5	1	184
6	1	190
7	1	158
8	1	139
9	1	175
10	1	148
11	1	152
12	1	111
13	1	141
14	1	153
15	1	190

Meat	Cases	Mean	Standard deviation	Confidence interval
Beef	20	156.9	22.6	146.3 – 167.5
Poultry	17	122.5	25.5	109.4 – 135.6

The lack of overlap between the confidence intervals for the two means indicates that it is likely that we will reject the null hypothesis. However, we must first test the dependent variable by group for normality – see the **Checking Normality for Parametric Tests worksheet**. We may conclude that the assumptions of the independent samples t-test are satisfied.

Running the test

Use Analyze – Compare Means – Independent Samples T-test:

- Select the outcome variable (*CalorieContent*) as the Test Variable, and the grouping variable (*Meat*) as the Grouping Variable

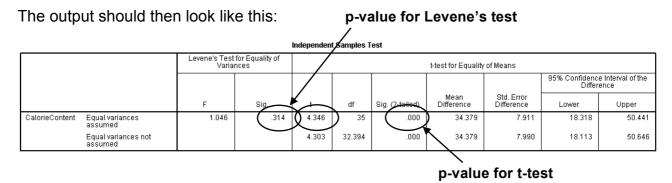- Define Groups as 1 and 2 (coded as above)

- Select Continue and OK

www.statstutor.ac.uk   © Mollie Gilchrist and Peter Samuels   Reviewer: Ellen Marshall
                                        Birmingham City University   University of Sheffield

Based on material provided by Loughborough University Mathematics Learning Support Centre and Coventry University Mathematics Support Centre

Results

The output should then look like this:

p-value for Levene's test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
CalorieContent	Equal variances assumed	1.046	.314	4.346	35	.000	34.379	7.911	18.318	50.441
	Equal variances not assumed			4.303	32.394	.000	34.379	7.990	18.113	50.646

p-value for t-test

The relevant results for the independent t-test are ringed. Levene's test checks the null hypothesis that the variances of the two groups are equal. In this example its p-value is 0.314, so the assumption of equal variances is not violated so we look at the top row of the output table. (If it is less than 0.05 then the bottom output row should be used instead.)

The value of the t statistic is 4.346, and the p-value is displayed 0.000 (meaning 0.000 rounded to 3 decimal places, i.e. less than 0.0005). This means that there is a very small probability of this result occurring by chance under the null hypothesis of no difference between the two groups. [Note: Never put $p = 0.000$ in your write-up].

Conclusion

The null hypothesis is formally rejected with 99.9% confidence since $p < 0.001$. Informally we state that there is very strong evidence that the calorie content of hot dogs for these two groups is different, with the mean calorie content of beef hot dogs approximately 34 calories higher than poultry hot dogs.

Of course, if we were to take other samples of hotdogs, we might get a different mean difference between the two groups. The lower and upper limits of the 95% confidence interval tell us that we can be 95% confident that the population mean content of beef hotdogs is between 11 and 58 calories higher than the mean calorie content of poultry hotdogs.

You would need to consider if this difference in calorie content is practically important, not just statistically significant.

Comments

- In general, multiple t-tests for comparing more than two groups should not be used – use Analysis of Variance (ANOVA) instead

- A statistically significant result from performing a t-test (i.e. $p \leq 0.05$) will imply that the 95% confidence intervals for the means will not overlap

- It is good practice to include a column with ID numbers to link participants with their data

- There does not have to be the same number of participants in each group, but the sample sizes should not be too different

- If the normality assumption is violated it is still possible to run a t-test provided that the group sizes are equal and greater than 25 – see Sawilowsky, S. S. and Blair, R. C. (1992) A more

realistic look at the robustness and Type II error properties of the t test to departures from population normality, *Psychological Bulletin*, 111(2), pp. 352–360.

- If neither the normality assumptions nor these robust assumptions are met, use the Mann-Whitney U test