

# GE2101 METHODS & PRACTICES IN GEOGRAPHY 2025 Godfrey Yeung

## Practical Exercise 3: t-test and Analysis of Variance (ANOVA)

### Comparing sample means: stone roundness

Roundness is an expression of the corner curvature of a stone. Roundness ranges from 0 to 1, where 0 signifies stones that are wholly angular, and 1 those that are completely rounded. Stone roundness commonly varies in different types of glacial sediments. For example, stones from glaciofluvial sediments (e.g. in fans or sandur, where stones are deposited by glacial meltwater, causing abrasion of rough edges) often appear to be more rounded than stones from till, that is, sediment laid down directly by a glacier. This variation raises the possibility that depositional environments can be inferred from measurement of stone roundness. Hence, the simple research question arises: does stone roundness from different types of glacial sediments vary significantly? If it does, then the stones can be considered to come from different populations and roundness may therefore be useful for distinguishing between different glacial depositional environments. But if no significant differences occur, and the samples come from the same population, then roundness cannot be used to infer depositional environment.

The file '**Rgroup.omv**' contains the results of a study of variations in stone roundness in glacial deposits from Area W. Four categories of deposit were identified in the field:

- (1) subglacial till (glacial till is the unsorted, coarsely graded and heterogeneous sediments that was deposited directly by the glacier)
- (2) ice-front alluvial fan deposits (alluvial fans are fan-shaped deposits of water-transported material (alluvium))
- (3) proglacial sandur deposits (a sandur is a glacial outwash plain formed of sediments deposited by melt water at the terminus of a glacier), and
- (4) esker deposits (an esker is a sinuously curving, narrow deposit of coarse gravel that forms along a melt water stream channel, developing in a tunnel within or beneath the glacier).

A number of samples were taken from each category of glacial deposit and the mean roundness for each was calculated. The four variables in the file contain the mean roundness values for each category of deposit. Note that the number of rows, or samples within each category, varies between 33 and 60.

There are 2 variables in the datafile:

- Round: stone roundness
- Group: group 1 = samples collected in subglacial till, group 2 = samples collected in ice-front alluvial fan deposits, group 3 = samples collected in proglacial sandur deposits, and group 4 = samples collected in esker deposits

Let us firstly examine the samples using descriptive statistics and histograms, and determine quickly and qualitatively if there appear to be significant differences between them.

Load the data file and produce:

- 1 Informative descriptive statistics for all four variables
- 2 Histograms for all four sediment types, i.e., four variables
- 3 Boxplots for the four sediment types
- 4 Can you identify differences between the four sediments?

### ***Independent samples t-test***

The *t*-test is a parametric test that can be used to test hypotheses about single means (that the mean is equal to a particular value) or to compare pairs of means. In this case we are considering, the *t*-test can be used to compare each pair of categories of glacial deposit. To illustrate this, we shall consider whether the stone roundness differs significantly between till (Group 1) and fan deposits (Group 2).

Learning objective: Why we want to conduct the *t*-test?




The stages involved in carrying out a *t*-test are as follows:

1. Set up your hypothesis e.g.,:  
H<sub>0</sub>: there is no significant difference between the stone roundness of the samples of till and fan deposits.  
H<sub>1</sub>: there is a significant difference....
2. Select an appropriate test (in this case an independent *t*-test).
3. Select a 0.05 significance level for the test (this is a commonly used level).

### Procedures:


- As jamovi do not have a specific function that allow us to specify the specific groups for the *t*-test, we have to use the **<Filters>** function to select our preferred groups.
- Click **<Data><Filters>** to activate the filtering function.
- We use the **IF** function (listed under Logical when you click the  $f_x$  icon) to select cases that are listed under groups 1 & 2 in the 'Group' variable.
- As the code is written as 'IF(expression, value, else)', so you enter the code as shown below, e.g., **IF(Group<3)** means to filter out all cases in groups 3 & 4.
- Click the upward arrow icon after finished (to hide the command).
- You can disable the filtering function by clicking the 'switch' on the top right-hand corner (the 'switch' will change to 'inactive' & in brown colour).

ROW FILTERS

+  Filter 1 active  | 

$f_x = IF(Group < 3)$  +

Description

*Selected* 

	Filter 1	Round	Group
1	✓	0.402	1
2	✓	0.314	1
3	✓	0.319	1
4	✓	0.496	1
5	✓	0.452	1

2 0.2 0.4 Rou

- Select **<Analyses><T-test><Independent Samples T-test>**
- Put the 'Round' variable into 'Dependent Variables' and 'Group' variable into 'Grouping Variables'.
- Select **<Mean difference>** and **<Descriptives>** under 'Additional Statistics' and **<Homogeneity test>** under 'Assumption Checks'.

Independent Samples T-Test

		statistic	df	p
Round	Student's t	<u>-14.21<sup>a</sup></u>	98.00	<u>&lt; .001</u>

<sup>a</sup> Levene's test is significant ( $p < .05$ ), suggesting a violation of the assumption of equal variances

## Assumptions

Test of Equality of Variances (Levene's)

	F	df	df2	p
Round	4.24	1	98	<u>0.042</u>

*Note.* A low p-value suggests a violation of the assumption of equal variances

[3]

Group Descriptives

	Group	N	Mean	Median	SD	SE
Round	1	50	0.36	0.35	0.06	0.01
	2	50	0.57	0.55	0.08	0.01

Before examine the  $t$ -test result, we have to decide whether to reject or accept the hypothesis concerning equality of variance of the two groups. In

this example, the hypothesis about the equality of the variances of the two groups must be rejected because the significance (probability) value attached to the calculated value of  $F$  (in Levene's test) is 0.042, which is less than the selected level of significance at 0.05, i.e., heterogeneity of variances. (If the variances are not significantly different (homogeneity of variances), then the significance value associated with the  $F$  statistic is greater than your selected significance level.) You can also see a warning note about the heterogeneity of variances from the  $t$ -test table.

To test the null hypothesis that there is no difference between the two variables when the assumption of homogeneity of variances is violated, we have to ask jamovi to give us an alternative calculated value of  $t$ , e.g., Welch's  $t$ .

- Select the **<Welch's>** under 'Tests'.

Independent Samples T-Test

		Statistic	df	p	Mean difference	SE difference		Effect Size
Round	Student's t	-14.21 *	98.00	< .001	-0.21	0.01	Cohen's d	-2.84
	Welch's t	<u>-14.21</u>	91.04	<u>&lt; .001</u>	-0.21	0.01	Cohen's d	<u>-2.84</u>

\* Levene's test is significant ( $p < .05$ ), suggesting a violation of the assumption of equal variances

You can see the calculated value of Welch's  $t$  is the same as of the conventional calculated  $t$ -statistic in this case: -14.21. This is also the case of the calculated probability of committing type-I error but the degree of freedom is different. As the probability value is lower than the significance level (at  $p \leq 0.05$ ), we can reject null hypothesis.

You can also ask jamovi to estimate the effect size for you. Cohen's  $d$  is estimated through the difference between two means divided by a standard deviation for the data.

- Select the **<Effect size>** under 'Additional Statistics'.

A value of 2.84 indicates a (very) large effect between sample groups 1 & 2 & the chance of getting a mean difference of 0.21 on stone roundness is lower than 0.1% (ignore the negative sign as of the calculated value of  $t$ ).

#### Notes:

- Ignore the negative sign of calculated  $t$  (as the sign is determined by the mean of group 1).
- Remember to disable the filtering function before continuing (otherwise, you will only include cases under groups 1 & 2).

#### Questions for Practice (using the above example)

1. What is the probability associated with the computed value of the  $t$ -test statistic?

2. Is this probability greater or less than your chosen significance level?
3. Do your answers to 1 and 2 lead you to reject or accept the null hypothesis?
4. What conclusions can you draw about stone roundness in the glacial deposits examined?

### **Analysis of variance (ANOVA)**

ANOVA is a method of comparing several group means by testing whether the variance between the *groups (categories of sample)* is greater than the variance within them. If it is, then it is likely that the samples have been drawn from different populations. If, on the other hand, the variance between the groups (samples) is smaller than it is within them then it is likely that the samples have been drawn from the same population.

Learning objective: What are the major differences between *t*-test and ANOVA?

#### Procedures:

- Select **<Analyses><ANOVA><One-way ANOVA>**
- Put the 'Round' variable into the '**Dependent Variables**' & 'Group' variable into '**Grouping Variables**'
- Select **<Descriptive table>** under 'Additional Statistics'
- Select **<Assume equal (Fisher's)>** under 'Variances'
- Select **<Homogeneity test>** under 'Assumption Checks'

#### **One-Way ANOVA**

One-Way ANOVA

		F	df1	df2	p
Round	Welch's	130.76	3	95.34	< .001
	Fisher's	113.34	3	189	< .001

Group Descriptives

	Group	N	Mean	SD	SE
Round	1	50	0.36	0.06	0.01
	2	50	0.57	0.08	0.01
	3	60	0.64	0.10	0.01
	4	33	0.44	0.09	0.02

#### **Assumption Checks**

Homogeneity of Variances Tests

		Statistic	df	df2	p
Round	Levene's	3.52	3	189	0.016
	Bartlett's	11.04	3		0.012

*Note.* Additional results provided by *moretests*

You can see the mean value of stone roundness vary from 0.36 to 0.64.

Before we refer to the ANOVA table, we have to check whether the assumption of homogeneity of variances is violated. The Levene tests show that whether the variances of the four groups of sample are significantly different. As the Levene's test is clearly significant at 0.05 level, we cannot use the calculated Fisher's  $F$  ratio in the ANOVA table.

Note: Your copy of jamovi may not have the Bartlett's test (it is another test on the assumption of homogeneity of variances & to be interpreted the same way as Levene's test).

The question we are considering is whether the differences between the means arise from chance alone, or whether they represent significant differences between the sample categories such that they may be regarded as having been drawn from different populations.

To answer this question, we refer to the alternative  $F$  ratio (jamovi provides only one option and it is selected by default): the calculated Welch's  $F$  ratio is 130.76 and it is significant. Thus we reject the null hypothesis and conclude that there are significant differences in the roundness of the stones sampled from the four types of glacial sediments in Area W. In other words, we conclude that the four sample groups come from different populations.

The degrees of freedom for the between sample variance estimate is the number of categories minus 1, or  $4 - 1 = 3$  in this case. The degrees of freedom for the within sample variance estimate is the number of individual data items in all the categories minus the number of categories (or  $193 - 4 = 189$  in the Fisher's, & adjusted to 95.34 in the Welch's).

Learning objective: Why we may want to conduct the post-hoc test?

As you know that there are significant differences of stone roundness between four groups of sample, you want to identify the source(s) of such differences by conducting a post-hoc test.

- Click **<Post Hoc Tests>**, select **<Games-Howell (unequal variances)>** under 'Post-Hoc Test' as the assumption of homogeneity of variances is violated.
- Optional: you can select **<Test results (t and df)>** and **<Flag significant comparisons>** under 'Statistics'.

Games-Howell Post-Hoc Test – Round

		1	2	3	4
1	Mean difference	—	−0.21 ***	−0.28 ***	−0.08 ***
	p-value	—	< .001	< .001	< .001
2	Mean difference		—	−0.07 ***	0.13 ***
	p-value		—	< .001	< .001
3	Mean difference			—	0.20 ***
	p-value			—	< .001
4	Mean difference				—
	p-value				—

Note. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Now let us move on to examine the post hoc tests. The Games-Howell's tests indicate that the differences of stone roundness is due to highly statistical significant on all combinations of sample groups. For instance, the probability of getting the average differences in stone roundness of 0.211 between sample group 1 (subglacial till) and 2 (alluvial fan deposits) is smaller than 0.001, hence, they are likely to come from two different populations. They should be, obviously (well, one come from subglacial till and another one from alluvial fan deposits, ... You are probably complaining loudly that Godfrey is 'mental' ☹☹☹ as he asks you to spend the time to verify something so obvious, i.e., calculation for its own sake!!!)

### **ASSESSED QUESTIONS (15%):**

#### **Notes:**

- This assessed exercise account for 15% of your course mark.
- You are expected to find out the answers of informally assessed parts of this exercise independently before the next practical.

#### **Deadline:** 11 / 14 March, 5pm (based on your tutorial session)

- The answers of all assessed exercises must be submitted through the ExamSoft online before the deadline (your submission will be *time-stamped* so don't wait till the last moment for the submission as poor internet connection is not a legitimate reason for the late submission).
- Click on the <Assessments> folder after you logging into ExamSoft, then click **Exercise 2 ANOVA**.
- I suggest you work out the answers for ALL assessed questions and save it in a Word document before transferring/copying them in a single attempt – you can submit your answers only **ONCE**.

#### **Attention:**

- Most of the questions are in forms of MCQs or fill in the blank. You are expected to select the best answer for MCQs. For fill in the blank

questions, do pay attention to the question-specific requirements before key in your specific numerical answers).

- I suggest you can keep a record of your submission.
- You can find your mark via ExamSoft after the submission deadline (I may delay the release of marks if there are a number of late/non-submission). We will discuss the questions and answers during the next practical.

### **Earnings of Geographers**

The President of NUS Geographical Society (Geog Soc) asks for your help to examine evidence that could be used for the promotion of Geography in outreach activities. As a devoted geographer, you are keen to find the empirical evidence to debunk the myth that Geography Majors have lower earning power than other disciplines. You want to show to JC students that the transferable skills that you are learning will make you competitive in the post-NUS life 😊.

You have identified a survey that could be useful for your task. Parts of this survey is about the impact of undergraduate degrees on early-career earnings of fresh graduates at higher educational institutions in country X (it is parts of a real survey, not fabricated data).

We will use **Earnings\_Grouped.omv** for this exercise. There are four variables in the dataset:

Earnings: Average income of graduates (in \$)

Group: Sex (M/F)

Female\_earnings: Average income of female graduates (in \$)

Subject: Subjects/Majors at different higher educational institutions

Partly due to your interest on gender-related issue, your first research objective is as follow: Whether there is a significant difference of average income between male and female fresh graduates at higher educational institutions in country X.

\*\*\* We are **NOT** going to deal with the abnormal distribution of the dataset for this exercise, partly for the sake of ease of interpretation in Qs 5-6 and partly due to the limited effects from simple data transformation methods.

Learning objectives: the selection of statistical test, its operational procedures and the interpretation of results

1. Is the Levene's test significant? (1%)
2. What is the calculated value of  $t$ ? (enter the numerical answer & rounded it to two decimal places) (1.5%)
3. What is the mean difference of income between male and female graduates? (enter the numerical answer & rounded it to two decimal places & NO measurement unit) (1%)

4. Is the statistical test significant, and at what level? (enter 'significant' or 'insignificant' for the first part & the numerical answer for the second part of your answers; do include the zero before the decimal place, e.g., enter 0.05 or 0.01 if it is significance) (1%)

You then tackle the second research objective: Whether there is a significant difference of average income between female graduates of different Subjects/Majors in higher educational institutions in country X.

5. Calculate the *F* ratio (enter the numerical answer & rounded it to two decimal places) (2%)
6. Is the statistical test significant and at what level? (enter 'significant' or 'insignificant' for the first part & the numerical answer for the second part of your answers; do include the zero before the decimal place, e.g., enter 0.05 or 0.01 if it is significance) (2%)

You are keen to explore the results further by examining the average income of female graduates in Geography vis-à-vis female graduates in other subjects. The codes for all 30 subjects are as follows:

Subject	Code	Subject	Code	Subject	Code
Agriculture	1	Engineering	11	Philosophy	21
Allied to med	2	English	12	Physics	22
Architecture	3	Geography	13	Physsci	23
Biosciences	4	History	14	Politics	24
Business	5	Languages	15	Psychology	25
Comms	6	Law	16	Social care	26
Computing	7	Maths	17	Sociology	27
Creative arts	8	Medicine	18	Sportsci	28
Economics	9	Nursing	19	Technology	29
Education	10	Pharmacology	20	Vetsci	30

7. What is the average income of female Geography graduates? (enter the numerical answer & NO decimal place & measurement unit) (1.5%)
8. You can see that female graduates of different subjects have different level of income. Graduates of which two subjects have a significantly (at 0.01 level) **lower** average income than Geography graduates? (2%)
9. Graduates of which three subjects have a significantly (at 0.01 level) **higher** average income than Geography graduates? (3%)

Note: It may be easier for you to screen capture the jamovi's big table before identifying the results, if you don't have a large computer screen.

10. What conclusion can you draw from these findings, i.e., how does female Geography Majors perform in terms of average income vis-à-vis other disciplines? (this question is NOT assessed)