

EXPLORING THE DIFFERENCES BETWEEN RESPONDENT GROUPS¹

CONDUCTING T-TESTS AND ANALYSIS OF VARIANCE



An overview of the relevant statistical concepts used in these instructions is provided in the [Getting Started with Statistics](#) guide.

The differences between Adult Social Care Survey (ASCS) or Carers Survey (PSS SACE) respondent groups² can be explored using **independent t-tests** and **analysis of variance (ANOVA)**. Step-by-step guides and pre-recorded presentations are included in the MAX toolkit to help you conduct these calculations in Excel. The purpose of this summary is to provide users with an overview of independent t-tests and ANOVA, and the available tools. Examples will be provided throughout and will be based on the following two questions:

Do overall social care-related quality of life (SCRQoL) scores (ASCOF 1A) differ between men and women?

Do overall social care-related quality of life (SCRQoL) scores (ASCOF 1A) differ by rating of satisfaction with services?

While this example is based on ASCS data, the principles and steps outlined in this guide are also relevant to the analysis of PSS SACE data.

¹ Please note that independent t-tests and ANOVAs can also be used to explore the relationship between survey variables and, conversely, that cross-tabulations and chi-square tests can be used to explore the differences between respondent groups. The format of these guides and presentations were used to introduce increasingly complex concepts to MAX toolkit users and attendees of the 2016 webinars.

² Respondent groups can be characterised by a range of features, including gender, age, ethnicity and primary support reason.

INTRODUCTION

An **independent t-test** can be used to determine whether there is a statistically significant difference between the means of two unrelated groups (e.g. men and women) on the same continuous³ dependent⁴ variable (e.g. SCRQoL). While multiple t-tests can be used to explore more than two groups (e.g. SCRQoL and satisfaction with services), such a strategy is resource-intensive and may generate a **type I error** [see hypothesis testing box below]. In these cases, **Analysis of Variance (ANOVA)** can be used to determine whether there is a statistically significant difference.

Independent t-tests and ANOVA are both types of **inferential statistic** and can be used to **infer (or deduce) characteristics about a population from a sample of data drawn from that population**. Each test generates a **p-value** which enables you to determine whether or not to reject the null hypothesis. This p value is the probability that the null hypothesis is true. If the p value is less than a set significance level (usually 5% [0.05] or 1% [0.001]), it is said that the pattern is 'statistically significant'. This means that there is a low probability that the observed pattern arose by chance.

Hypothesis testing

Hypothesis testing involves using statistical tests such as chi square to determine whether the patterns observed in a sample of data (in this instance, the sample of respondents who completed the ASCS or PSS SACE) are likely to reflect those that would be found in the entire population (in this instance, the entire population of adult social care service user or carers within a given LA). A hypothesis explores two hypotheses about a population: the **null hypothesis** (there is no difference between variables) and the **alternative hypothesis** (there is a difference between variables).

	Relationship	Example
Null	No difference	There is no difference in the proportion of men and women who say they have as much control over their daily life (ASCOF 1B) as

³ A **continuous variable** is a variable that can take on any value within a given range. For example, social care-related quality of life (SCRQoL) and carer social care-related quality of life (Carer SCRQoL) scores [max range 0-24] and age of service user or carer. The rest of the variables explored in the ASCS and PSS SACE are categorical.

⁴ A **dependent variable**, also known as an outcomes variable, is a variable that varies in value in response to the effect of another variable (e.g. SCRQoL). An **independent variable**, in contrast, is a variable that may affect another variable (e.g. gender, age).

		they would like. In other words, the frequency counts observed in the data have arisen 'by chance'
Alternative	Difference	There is a difference in the proportion of men and women who say they have as much control over their daily life (ASCOF 1B) as they would like. In other words, the frequency counts observed in the data have arisen 'by something other than chance' (e.g. an underlying factor)

A word of caution: Hypothesis testing is not 100% certain so there is a risk of drawing an incorrect conclusion.

Type I errors: occur when the null hypothesis is rejected when, in fact, it is true (i.e. you conclude your data shows there is an effect or difference when there is none). The risk of generating a type I error can be reduced by using a lower significance value (e.g. 0.001 rather than 0.05).

Type II errors: occur when the null hypothesis is accepted when, in fact, it is false (i.e. you conclude your data shows no effect or difference when it actually does). The risk of generating a type II error can be reduced by ensuring your sample size is sufficient. Guidance notes are provided alongside the statistical tests explored in the MAX toolkit.

INDEPENDENT T-TEST

Use: an **independent t-test** can be used to determine whether there is a statistically significant difference between the means of two unrelated groups (e.g. men and women) on the same continuous dependent variable (e.g. SCRQoL). Independent t-tests can be used to explore a range of questions including:

Do overall social care-related quality of life (SCRQoL) scores (ASCOF 1A) differ between men and women?

Assumptions: the suitability and accuracy of an independent t-test are dependent on the following assumptions being fulfilled.

1. **The dependent variable can be measured on a continuous scale:** in this instance, SCRQoL scores are measured on a scale of 0 to 24.
2. **The independent variable consists of two categorical, independent groups:** in this instance, gender includes two mutually exclusive groups – men and women.
3. **Independence of observations:** each respondent must provide data for one group only and cannot be in both groups. If this assumption is violated, a **paired-samples t-test** may be more suitable. Please note, however, that this test is not covered in the MAX toolkit.
4. **Normal distribution:** the dependent variable should be normally distributed for each group of the independent variable. In this instance, this means that the frequencies of SCRQoL responses would be symmetrical and bell-shaped for both men and women. The symmetry of a distribution is measured by skewness [see the **Getting Started with Statistics summary**] and is automatically calculated using the Descriptives Statistics function in Excel. **Please note:** SCRQoL is usually negatively skewed⁵ so this assumption is normally violated. In this instance, it may be more appropriate to conduct a **Mann-Whitney U-Test**, the non-parametric equivalent of an independent t-test. Details on how to do this are included in the step-by-step instructions.
5. **Homogeneity of variances:** the variance of the dependent variable (e.g. SCRQoL) is approximately equal in each group (e.g. men, women). This is assessed using a **Levene's Test** and is automatically calculated using the Descriptives Statistics function in Excel. If the generated p-value is greater than 0.05, the assumption is not violated and you can run a standard independent t-test with **equal variances assumed**. If, on the other hand, the generated p-value is less than 0.05, the assumption is violated and you will need to run a standard independent t-test with **equal variances not assumed**. Details on how to do this are included in the step-by-step instructions.

Interpretation: the independent t-test generates a p-value. This shows whether the null hypothesis, which assumes that there is no significant difference between the means of the two

⁵ SCRQoL scores are usually negatively skewed with most scores falling at the higher end or right side of the distribution.

groups, can be rejected. If the p-value less is less than the **critical value**⁶ (usually 0.05 or 0.001), the null hypothesis can be rejected and the alternative hypothesis, that there is a significant difference between the means of the two groups, can be accepted. **Please note:** unless you are specifying the direction of this difference (e.g. men report a higher rating of SCRQoL than women), you will need to use the two-tailed value generated by Excel.

ANALYSIS OF VARIANCE (ANOVA)

Use: **analysis of variance (ANOVA)** can be used to determine whether there is a statistically significant difference between the means of more than two unrelated groups (e.g. characterised by rating of satisfaction with services) on the same continuous dependent variable (e.g. SCRQoL). ANOVA overcomes many of the issues associated with conducting multiple t-tests (e.g. resource-intensive, prone to generating Type I errors) and can be used to explore a range of questions including:

Do overall social care-related quality of life (SCRQoL) scores (ASCOF 1A) differ by rating of satisfaction with services?

Assumptions: the suitability and accuracy of an ANOVA are dependent on the following assumptions being fulfilled.

1. **The dependent variable can be measured on a continuous scale:** in this instance, SCRQoL scores are measured on a scale of 0 to 24.
2. **Independence of observations:** each respondent must provide data for one group only and cannot be in both groups. If this assumption is violated, a different type of ANOVA will need to be conducted. Please note, however, that this test is not covered in the MAX toolkit.
3. **Normal distribution:** the dependent variable should be normally distributed for each group of the independent variable. The symmetry of a distribution is measured by skewness [see

⁶ To determine whether the critical value is significant, you need to know the **Degrees of Freedom (df)**. The df is the number of values in a dataset that are free to vary without influencing the result and are calculated as follows: (rows – 1) x (columns – 1)

the [Getting Started with Statistics summary](#)] and is automatically calculated using the Descriptives Statistics function in Excel. If this assumption is violated, it may be more appropriate to conduct a [Kruskal-Wallis test](#), the non-parametric equivalent of an ANOVA. Details on how to do this are included in the step-by-step instructions.

4. **Homogeneity of variances:** the variance of the dependent variable (e.g. SCRQoL) is approximately equal in each group. This is assessed using a [Levene's Test](#) and is automatically calculated using the Descriptives Statistics function in Excel. If the generated p-value is greater than 0.05, the assumption is not violated and you can run the ANOVA with **equal variances assumed**. If, on the other hand, the generated p-value is less than 0.05, the assumption is violated and you will need to run the ANOVA with **equal variances not assumed**. Details on how to do this are included in the step-by-step instructions.

Interpretation: the ANOVA test generates a p-value. This shows whether null hypothesis, which assumes that there are no significant differences between the means of the groups, can be rejected. If the p-value less is less than the [critical value](#)⁷ (usually 0.05 or 0.001), the null hypothesis can be rejected and the alternative hypothesis, that there are significant differences between the means of the groups, can be accepted.

TOOLS IN THE MAX TOOLKIT

A number of tools are provided in the MAX toolkit to help you conduct independent t-tests and analysis of variance (ANOVA). These include:

- Step-by step instructions (based on ASCS data); and a
- Series of pre-recorded training presentation which provide an overview of the analysis tool and guidance on how to input data, conduct analysis and interpret the results.

These can be accessed from the [Statistical tests and tools page](#) of the MAX toolkit.

⁷ To determine whether the critical value is significant, you need to know the [Degrees of Freedom \(df\)](#). The df is the number of values in a dataset that are free to vary without influencing the result and are calculated as follows: (rows – 1) x (columns – 1)

DISCLAIMER

The MAX toolkit and website are based on independent research commissioned and funded by the NIHR Policy Research Programme (Maximising the value of survey data in adult social care (MAX) project and the MAX toolkit implementation and impact project). The views expressed on the website and in publications are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care or its arm's length bodies or other government departments.