

Comparing for differences between two samples

Dependent variable: continuous (NB. discrete data can be used - such as counts, time measured in days, etc. – however binary data cannot)

Independent variable: one categorical variable with two levels/groups

Check data:

- Look at plot (e.g. individual values plot) for any general trends or 'bunching' of data that may violate assumptions
- Boxplot to check for outliers – deal with these as required.

Check assumptions (see checking data sheet for further detail):

- Data are quantitative
- Samples are independent
- Data are normally distributed
- Variances of groups are homogeneous (equal)

*Note: If data is normally distributed but you have unequal variances, follow the parametric test instructions but do not tick the box for 'Assume equal variances'.

If no assumption is violated:

PARAMETRIC DATA = Two-sample t-test (Independent or unpaired samples t-test)

- Stat
- Basic Statistics
- 2-Sample t
 - Check 'Each sample is in its own column' is selected if data are in separate columns
 - Put each set of data into the Sample 1 and Sample 2 boxes
 - Under 'Options' check the box next to 'Assume equal variances'
 - OK
- P-value: <0.05 indicates a significant difference between the two groups
 ≥0.05 indicates no significant difference between the two groups

"There was [no / a] significant difference between the [*dependent variable*] at [*independent variable group 1*] (mean = ..., SD = ...) and [*group 2*] (mean = ..., SD = ...) (t(df) = ...; P[</=] ...)."

e.g. There was no significant difference between the height of ghost orchids at sites A (mean = 35.7, SD = 13.1) and B (mean = 40.7, SD = 10.6) (t(28) = -1.14; P=0.265).

If normality or equal variances assumptions are violated:

NON-PARAMETRIC DATA = Mann Whitney U-test

- Stat
- Nonparametrics
- Mann-Whitney
- Put one set of data in 'First sample' box and the other in 'Second sample' box
- Make sure 'Confidence level' is 95.0, and 'Alternative' is 'not equal'
- P-value:
 - <0.05 indicates a significant difference between the two groups
 - ≥ 0.05 indicates no significant difference between the two groups

"There was [no / a] significant difference between the [*dependent variable*] at [*independent variable group 1*] (median = ...) and [*group 2*] (median = ...) ($W = \dots$; $P[< / =] \dots$)."

e.g. There was no significant difference between the median lengths of sandsmelt surveyed in 2017 (median = 71) and 2019 (median = 70) ($W = 962$; $P < 0.05$).

Note: When there is a significant difference between two groups, it is good practice to state which group is larger/smaller than the other (i.e. the direction of the difference).

e.g. Median sandsmelt length was significantly greater in 2016 (median = 77) than in 2019 (median = 70) ($W = 6943$; $P < 0.001$).