

Testing differences between two groups when your outcome is continuous

Example: We have a dataset called data1, which consists of 146 observations (patients) and 5 variables (id, treat, age, sex and wt). Participants have been randomized to receive an active treatment or placebo. We will only consider three baseline characteristics: age, sex and weight.

```
head(data1)           #prints the first few rows of the data set
```

id	treat	age	sex	wt
1	Treated	18	Female	62.6
2	Treated	50	Male	57.4
3	Treated	37	Male	104.6
4	Treated	25	Female	55.5
5	Placebo	60	Female	58.4
6	Treated	44	Female	41.9

Suppose we would like to examine whether weights are different on average between the two treatment groups. Weight is continuous, and it is observed to be normally distributed with no significant outliers or skewness. Furthermore, we expect the spread of weights to be roughly similar in the two treatment groups.

As each patient contributed one observation to the analysis and was randomized to one of two treatments, this allows us to use the two independent sample t-test.

To perform an independent two-group t-test

First, install the t.test package in R:

```
install.packages("t.test")  
library("t.test")
```

Then it is a simple line of code to generate the t-test to examine mean differences in weight according to treatment group. By typing var.equal=TRUE, we are assuming that the variance of weight is roughly similar in both treatment groups. If we cannot assume this, we would type var.equal=FALSE to get a p-value based on Welch's t-test instead.

```
t.test(wt ~ treat , data = data1 , var.equal=TRUE)
```

The R output is below:

```
Two Sample t-test
data: wt by sex
t = -1.1541, df = 139, p-value = 0.2504
alternative hypothesis: true difference in means between group Female and group Male is not equal to 0
95 percent confidence interval: -8.82488, 2.31977
sample estimates:
mean in group Female      mean in group Male
      66.88193             70.13448
```

The p-value is 0.25, which means the mean differences in weight are not statistically significant at the 0.05 level of significance. This corresponds with the 95% confidence interval for the difference in means, which is our plausible range for the true difference given our data and includes 0 (no difference).

To perform an independent two-group Wilcoxon Mann Whitney test (WMW)

Suppose we now want to run the nonparametric analog of the two-sample t-test called the **Wilcoxon Mann-Whitney test**.

```
wilcox.test(wt ~ treat , data = data1 , alternative = "two.sided")
```

The output:

```
Wilcoxon rank sum test with continuity correction
data: wt by treat
W = 2294, p-value = 0.9365
alternative hypothesis: true location shift is not equal to 0
```

Once again, as the p-value is > 0.05 , we do not reject the null hypothesis that the distribution of weights are equal between treatment groups.

Paired data

Paired or correlated data arise mostly when your outcome is measured multiple times on each individual. Other correlated data include measurements from clusters such as littermates, families, and hospitals.

Suppose we have recorded the weights of 10 mice pre and post treatment. Our hypothesis is that treatment will be associated with higher weights.

The *diet* data set:

ID	pre	post
1	200.1	392.9
2	190.9	393.2
3	192.7	345.1
4	213.0	393.0
5	241.4	434.0
6	196.9	427.9
7	172.2	422.0
8	185.5	383.9
9	205.2	392.3
10	193.7	352.2

```
t.test(diet$pre,diet$post , data = data2 ,paired=TRUE )
```

The R output is below:

```
Paired t-test

data: diet$pre and diet$post

t = -20.883, df = 9, p-value = 6.2e-09
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
-215.5581 -173.4219

sample estimates:
mean difference
-194.49
```

The p-value for the paired t-test is <0.0001 , so there is a statistically significant difference pre to post treatment with respect to weight. Note that the degrees of freedom for the t-test is 9 (10 mice – 1), which reflects that we have 10 independent observations (each mouse).