

Testing differences between multiple groups with a continuous outcome

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

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

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

We can generate a descriptive table for the data using the [table1 package](#):

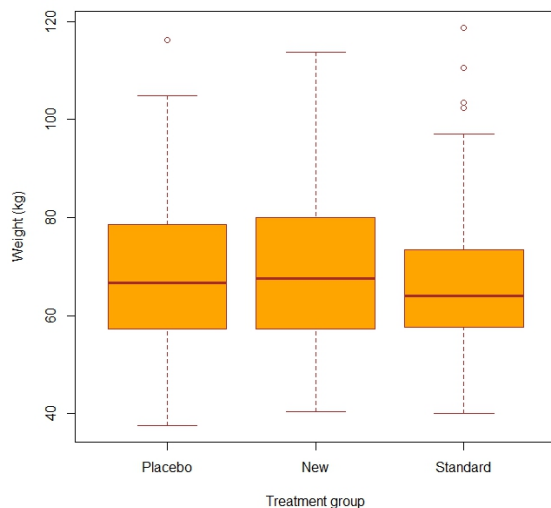
	Placebo (N=52)	New (N=48)	Standard (N=46)	Overall (N=146)
Age (years)				
Median [Min, Max]	37.5 [18.0, 65.0]	39.0 [18.0, 65.0]	40.5 [18.0, 64.0]	39.0 [18.0, 65.0]
Sex				
Female	34 (65.4%)	30 (62.5%)	23 (50.0%)	87 (59.6%)
Male	18 (34.6%)	18 (37.5%)	23 (50.0%)	59 (40.4%)
Weight (kg)				
Mean [SD]	68.4 [16.3]	69.0 [16.5]	67.8 [17.4]	68.4 [16.6]

Suppose we would like to formally test whether weight (`wt`) is different on average between the three 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 three treatment groups.

As each patient contributed one observation to the analysis and was randomized to one of three treatments, this allows us to use an **Analysis of Variance (ANOVA)**.

We can also get a sense of the data by creating a **boxplot**, which shows the spread of weights within each treatment group as well as average group differences.

```
boxplot(wt~treat,  
data=dat,  
xlab="Treatment group",  
ylab="Weight (kg)",  
col="orange",  
border="brown")
```



To perform an ANOVA

```
summary(aov(wt ~ treat, data = dat))
```

As wt (weight) is the dependent variable, it is specified on the left hand side of the ~. Your independent variable treat (treatment) is specified on the right hand side.

The R output is below:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treat	2	36	18.02	0.064	0.938
Residuals	143	39998	279.70		

This initial output is a bit confusing at first but let's break it down. An ANOVA seeks to partition the total variability in the dependent variable (wt) – variability due to treatment (the differences between the average weight across the three treatment groups) and variability due to residual

error (the differences between weights observed within a group and the average weight in that group).

We see in the output that the mean of squares due to treatment (MSTR) is 18 and the mean sum of squares due to error (MSE) is 279.7. The F-test, looks at the ratio of MSTR/MSE which is 0.064. This low ratio shows that much of the variability in patient weight is coming from residual variability, as opposed to treatment differences. We can see this in the above figure – the average weight is similar *between* treatment groups, but there is a lot of variability around average weight *within* each treatment group. This lack of between group differences relative to within group variability is also reflected in the p-value – which is 0.938 and clearly not statistically significant. We, of course, expected this lack of effect on weight due to treatment as patients are randomized to treatment – on average, weights are the same across treatment group.

Although it is not necessary here given the overall F-test was not statistically significant, we can compare mean differences in both treatment groups with the placebo group using Dunnett's test.

```
DunnettTest(x=dat$wt, g=dat$treat)
```

```
Dunnett's test for comparing several treatments with a control :
```

```
95% family-wise confidence level
```

```
$Placebo
```

	diff	lwr.ci	upr.ci	pval
New-Placebo	0.5951923	-6.893525	8.083910	0.9774
Standard-Placebo	-0.6433946	-8.216294	6.929505	0.9743

```
---
```

```
Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

To perform an independent Kruskal-Wallis test (KW)

Suppose we now want to run the nonparametric analog of ANOVA called the **Kruskal-Wallis test**.

```
kruskal.test(wt ~ treat , data = dat)
```

The output:

```
Kruskal-Wallis rank sum test
```

```
data: wt by treat
```

```
Kruskal-Wallis chi-squared = 0.30435, df = 2, p-value = 0.8588
```

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.