

# Tutorial 03

ENVX2001 – Applied Statistical Methods

Semester 1

## Welcome

In this tutorial we will learn the diagnostic tools for checking assumptions behind  $t$ -tests and ANOVA, then take apart an ANOVA table to understand what every number means.

You will learn how to:

1. Identify the three assumptions of  $t$ -tests and ANOVA.
2. Use R to produce diagnostic plots and numerical checks.
3. Read diagnostics and decide whether assumptions are met.
4. Read an ANOVA table and interpret the F ratio and degrees of freedom.

## Checking assumptions

What assumptions do  $t$ -tests and ANOVA make? There are three:

1. **Normality**: the residuals follow a normal distribution.
2. **Equal variance**: the spread of observations is roughly the same across groups.
3. **Independence**: the observations do not influence each other.

We can check the first two from the data. Independence comes from the study design (random assignment, no repeated measures on the same subject), so we cannot diagnose it from a plot.

### Exercise 1: Checking assumptions in R

The diagnostics are the same whether we fit a  $t$ -test or an ANOVA. Run each block and the tutor will walk through the output.

The `sleep` dataset records the change in hours of sleep for 10 patients under two drug treatments (20 observations total). It is built into R, so no file download is needed.

#### t-test

The `t.test()` object does not store residuals, so we check normality on the raw data in each group. This is equivalent to checking the residuals, since within each group the residuals are just the values minus the group mean (same shape, shifted to zero). With only two groups this is straightforward:

```
CODE
```

```
str(sleep)
```

CODE

```
# Subset the data into two groups
group1 <- subset(sleep, group == "1")
group2 <- subset(sleep, group == "2")

# QQ plot for Group 1
qqnorm(group1$extra, main = "Group 1")
qqline(group1$extra, col = "red")

# QQ plot for Group 2
qqnorm(group2$extra, main = "Group 2")
qqline(group2$extra, col = "red")

# Shapiro-Wilk test for normality in Group 1
shapiro.test(group1$extra)

# Shapiro-Wilk test for normality in Group 2
shapiro.test(group2$extra)
```

## ANOVA

ANOVA and the  $t$ -test are mathematically equivalent when there are only two groups, so we can fit an ANOVA to the same data:

CODE

```
model <- aov(extra ~ group, data = sleep)
```

With `aov()`, residuals are built in:

CODE

```
qqnorm(residuals(model), main = "")
qqline(residuals(model), col = "red")
shapiro.test(residuals(model))
```

**Equal variance** (this check is the same whether you used the  $t$ -test or ANOVA approach above):

CODE

```
boxplot(extra ~ group, data = sleep)

# Using group1 and group2 from the t-test tab (or re-create them here)
group1 <- subset(sleep, group == "1")
group2 <- subset(sleep, group == "2")
max(sd(group1$extra), sd(group2$extra)) /
  min(sd(group1$extra), sd(group2$extra))
```

The SD ratio rule of thumb: below 2 is generally acceptable.

**Independence** cannot be checked from data. It comes from the study design: random assignment, no repeated measures on the same subject.

R can also produce four diagnostic plots at once with `plot(aov(...))`, including the QQ plot and two others we have not covered yet. We will explore these in the lab next week.

### Reading diagnostics

Now that we know the code, can we read the output? The widget below generates random data with 2 to 4 groups, shows one diagnostic, and asks whether the assumption holds.

```
OUTPUT
{ojjs}
//| echo: false
import { assumptionCheckerWidget } from "../assets/js/assumption-checker.js"
assumptionCheckerWidget()
```

#### **i** Checkpoint

We should be able to run the diagnostic checks in R, read the output, and decide whether the assumption is met.

## The ANOVA table

When we run `summary(aov(...))` in R, we get an ANOVA table. Most people skip straight to the p-value, but every number in the table tells us something about the data and the experiment. The two most important things to understand are the **F ratio** (is there a real effect, or just noise?) and the **degrees of freedom** (how big was the experiment?).

ANOVA table interpretation and basic calculations (the F ratio and degrees of freedom) are examinable.

### Exercise 2: The F ratio and degrees of freedom

A researcher tested three planting densities of a maize fodder crop (low, medium, and high) across 15 plots, with 5 plots per density. The sample means and standard deviations (kg of dry matter per plot) are:

Table 1: Sample means and standard deviations of dry matter yield (kg per plot) for three planting densities of maize.

Density	Mean	Std.Dev
Low	17.58	2.70
Medium	27.18	1.89
High	27.14	2.02

The low-density group produced noticeably less dry matter than the other two. But is that difference real, or could it be noise? Here is the ANOVA table R produces:

Table 2: One-way ANOVA for the effect of planting density on dry matter yield of maize.

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treatment	2	305.92	152.96	30.77	< 0.001
Residual	12	59.65	4.97		

A second researcher tested three fertiliser blends on wheat yield using the same design (3 treatments, 5 plots each). Their table:

Table 3: One-way ANOVA for the effect of fertiliser blend on wheat yield.

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treatment	2	11.4	5.7	1.14	0.35
Residual	12	60.0	5.0		

### The F ratio

The F value is a signal-to-noise ratio:

$$F = \frac{MS_{trt}}{MS_{res}}$$

$MS_{trt}$  (treatment mean square) measures how much the group means differ from the overall mean.  $MS_{res}$  (residual mean square) measures the scatter within each group. When F is close to 1, the groups look no more different than random noise would produce. When F is large, the data suggest a real effect.

In the maize table,  $F = 152.96 \div 4.97 = 30.77$ . The between-group signal is about 30 times larger than the within-group noise, and the p-value is tiny. Planting density clearly affects dry matter yield.

In the wheat table,  $F = 5.70 \div 5.00 = 1.14$ . The treatment variation is about the same size as the residual noise. The fertiliser blends made no detectable difference.

The sum of squares (Sum Sq) in the treatment and residual rows add up to the total variation in the data. The ANOVA table partitions all variation into two sources: differences between groups and scatter within groups.

### Degrees of freedom

The degrees of freedom tell us about the size and shape of the experiment:

- **Treatment df** =  $t - 1$ , where  $t$  is the number of groups. Both tables show  $df = 2$ , so both experiments compared 3 groups.
- **Residual df** =  $N - t$ , where  $N$  is the total number of observations. Both tables show  $df = 12$ , so  $N = 12 + 3 = 15$  observations in each experiment.
- Each group therefore had  $15 \div 3 = 5$  replicates.

### Try it yourself

A colleague compared 4 watering regimes on tomato plants, with 6 plants per regime. Their ANOVA table shows  $MS_{trt} = 45.2$  and  $MS_{res} = 8.1$ , but they forgot to include the F value and degrees of freedom.

1. What are the treatment and residual degrees of freedom?
2. Calculate the F ratio.
3. The critical value at  $\alpha = 0.05$  for these degrees of freedom is 3.10. Is the result significant?

If the observed F exceeds the critical value, we reject the null hypothesis at that significance level. You can find critical values in R with `qf(0.05, df1, df2, lower.tail = FALSE)`.

#### **i** Checkpoint

We should be able to read an ANOVA table, explain what F measures (between-group variation divided by within-group variation), compute F from the mean squares, and work out the number of groups and observations from the degrees of freedom.

### Wrap-up

We covered the diagnostic tools for checking assumptions and learned to read an ANOVA table. In the lab, we will fit ANOVA models in R and use post-hoc tests to identify which groups differ.