

Testing for equality of means of multiple populations

September 23, 2026

Example: Paper Strength

- Tensile strength (TS) is an important characteristic for certain types of paper for industrial use;
- A reasonable conjecture is that this characteristic is influenced by the kind of wood fiber used in the manufacturing process.
- The process engineer wants to investigate whether four different wood fibers result in papers with relevant differences of TS, using a pilot plant as his experimental unit.
- Example adapted from Montgomery & Runger (2010), Ch. 13.

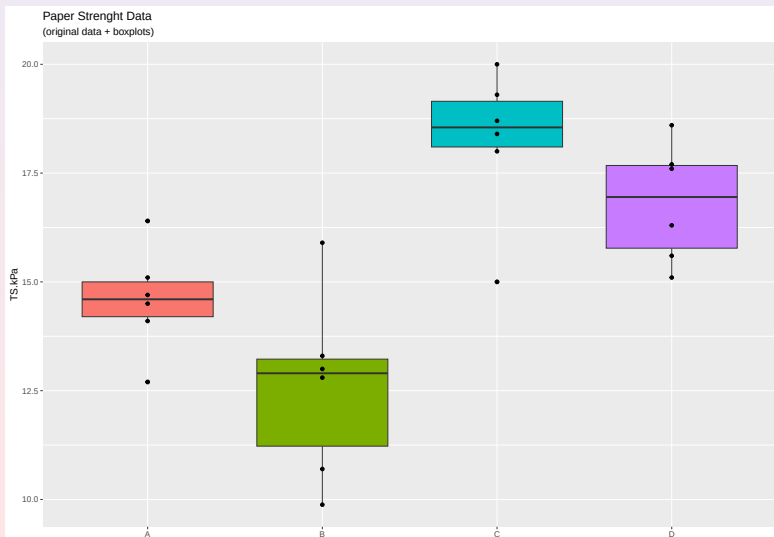
Example: Paper Strength

- Suppose that the total budget allocated for the experiment allows only six production runs for each kind of wood fiber.
- Under these specifications, the experiment has a single experimental *factor* (*wood fiber*) with $a = 4$ *levels* (fiber types *A*, *B*, *C* and *D*) and $n = 6$ *replicates* at each level.
- The response variable will be the tensile strength of paper (measured, e.g., in kPa). The engineering team is interested in finding out whether any fiber type leads to an increase in the mean TS value of the paper.

Example: Paper Strength

```
library(ggplot2)
ggplot(paper,
       aes(x    = Fiber.type,
           y    = TS.kPa,
           fill = Fiber.type)) +
  geom_boxplot() + geom_point() +
  ggtitle("Paper Strenght Data",
         "(original data + boxplots)") +
  theme(legend.position = "none")
```

Example: Paper Strength



Example: Paper Strength

```
> my.model <- aov(TS.kPa ~ Fiber.type,
+                 data = paper)
> summary.aov(my.model)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Fiber.type	3	110.77	36.92	13.62	4.56e-05 ***
Residuals	20	54.24	2.71		

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

Analysis of Variance

The name “analysis of variance” stems from a partitioning of the total variability in the response variable into components that are consistent with a model for the experiment.

$$SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \underbrace{n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2}_{SS_{Treatments}} + \underbrace{\sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2}_{SS_{Error}}$$

A large value of $SS_{Treatments}$ reflects large differences in treatment means.

Why ANOVA works?

Theorem:

- 1 $SS_{Error}/\sigma^2 \sim \chi^2_{N-a}$
- 2 $SS_{Treatments}/\sigma^2 \sim \chi^2_{a-1}$ if the null hypothesis $H_0 : \tau_i = 0$ for all i is true
- 3 By Cochran's theorem, these two are independent.
- 4 Therefore the ratio has an F distribution.

Why ANOVA works?

Theorem:

- 1 $SS_{Error}/\sigma^2 \sim \chi^2_{N-a}$
- 2 $SS_{Treatments}/\sigma^2 \sim \chi^2_{a-1}$ if the null hypothesis $H_0 : \tau_i = 0$ for all i is true
- 3 By Cochran's theorem, these two are independent.
- 4 Therefore the ratio has an F distribution.

$$E(MS_{Error}) = \sigma^2$$

$$E(MS_{Treatments}) = \sigma^2 \text{ under } H_0$$

If H_0 does not hold, then

$$E(MS_{Treatments}) = \sigma^2 + \frac{n \sum_{i=1}^a \tau_i^2}{a - 1} > \sigma^2$$

Thus, the ratio is larger when H_0 is violated. So we reject for only large values of the F statistic (right tailed rejection region).