Levene's Test

To test

$$\begin{aligned} H_0 &= \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 \\ H_1 &= \text{At least one pair of } \sigma^2 \text{ different.} \end{aligned}$$

Test statistic

$$W = \frac{(n-k)}{(k-1)} \frac{SST_Z}{SSE_Z} = \frac{MST_Z}{MSE_Z}$$

where SST_Z and SSE_Z are the usual sums of squares evaluated for the new data z_{ij} where

$$z_{ij} = |x_{ij} - \overline{x}_i|.$$

If H_0 is true

$$W \sim \text{Fisher-F}(k-1, n-k).$$

Example: PTSD Example (see handout).

$n = 45, k = 4.$

$$\begin{array}{rl} \text{F-statistic} & F = 3.046 \\ \text{Critical Value} & F_{0.05}(3,41) \doteq 2.84 \\ & F_{0.025}(3,41) \doteq 3.46 \\ & F_{0.01}(3,41) \doteq 4.31 \end{array}$$

Tables in McClave and Sincich give $F_\alpha(3, 40)$.

$$\implies \text{Reject } H_0 \text{ at } \alpha = 0.05 \quad (p = 0.039).$$

BUT Levene's Test suggests that the assumption of equal variances is **NOT** valid.

Why do we need the three assumptions ?

- independence
- Normality
- equal variances

- so that we can predict (under H_0) that

$$F \sim \text{Fisher-F}(k-1, n-k)$$

and complete the test (compute p-values and the rejection region).

But our hypothesis of interest is

$$H_0 \; : \; \text{No difference between treatments}$$

Under this hypothesis, the treatment labels

SHOULD NOT MATTER !

i.e. we should be able to exchange the labels, and not notice any major difference in the test statistic.

This leads us to consider **permutation** or **randomization** tests.

i.e. we compute the test statistic for all possible relabellings consistent with H_0, retaining the group sample sizes, and use these values to compute the rejection region.

Randomization/Permutation Tests

Suppose that there are N possible relabellings that give rise to test statistics

$$F_1, F_2, \ldots, F_N$$

Then the rejection region for significance level α is the interval to the right of

$$N(1-\alpha)\text{th largest of the values } F_1, F_2, \ldots, F_N$$

and the p-value is

$$\frac{\text{Number of } \mathrm{F}_1, F_2, \ldots, F_N \geq F}{N}$$

where

$$F = \frac{MST}{MSE}$$

is the true test statistic.

If the group sample sizes are $n_1, n_2, \ldots, n_k$ then

$$N = \frac{n!}{n_1!n_2!\ldots n_k!}$$

where

$$n! = n(n-1)(n-2)\ldots 3.2.1$$

("n factorial") - potentially very large.

Example: PTSD Example.

$$k = 4, n = 45 \qquad (n_1 = 14, n_2 = 10, n_3 = 11, n_4 = 10)$$

There are

$$\frac{45!}{14!10!11!10!} = 2.610 \times 10^{24}$$

possible relabellings: a very big number.

We compute $F = \frac{MST}{MSE}$ for each relabelling. For the real data, $F = 3.046$.

Example: PTSD Example (continued).

Using this approach, we compute for $\alpha = 0.05$

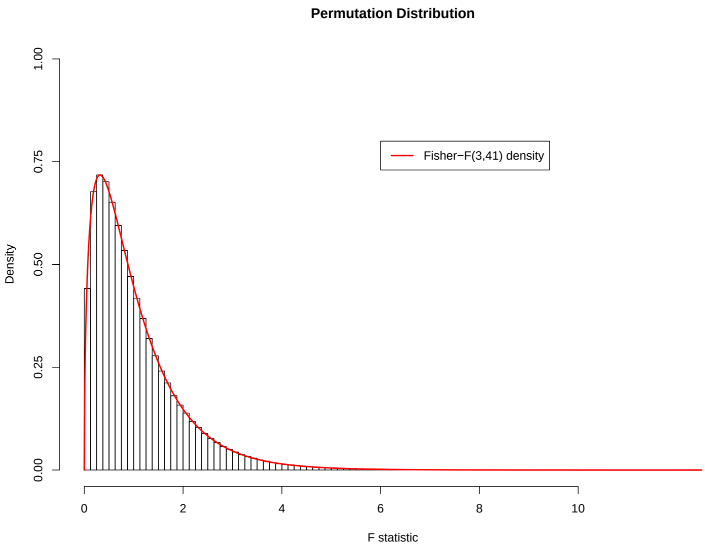
$$\begin{aligned} \text{CRITICAL VALUE} &: \quad C_R = 2.844 \\ p\text{-VALUE} &: \quad p = 0.040 \end{aligned}$$

Compare this with the ANOVA F-test values

$$\begin{aligned} \text{CRITICAL VALUE} &: \quad C_R = 2.833 \\ p\text{-VALUE} &: \quad p = 0.039 \end{aligned}$$

(using the Fisher-F(3,41) distribution.

Thus we obtain virtually identical results; **but the randomization test does not need the assumptions of normality or equal variances.**


Permutation Distribution
Fisher−F(3,41) density
Density
1.00
0.75
0.50
0.25
0.00
0
2
4
6
8
10
F statistic

Example: PTSD Example (continued).

Thus the null hypothesis (of equal means) is

REJECTED

under both procedures at the $\alpha = 0.05$ significance level.

In this case, the computations give similar conclusions. Here the truth or otherwise of the normality/equal variance assumptions **does not matter**.

Final Note on ANOVA F-test for a CRD

If $k = 2$, consider $F = MST/MSE$;

$$\begin{aligned} MST &= \frac{1}{k-1}\sum_{i=1}^{k} n_i(\overline{x}_i - \overline{x})^2 = n_1(\overline{x}_1 - \overline{x})^2 + n_2(\overline{x}_2 - \overline{x})^2 \\ &= \frac{n_1 n_2}{n_1 + n_2}(\overline{x}_1 - \overline{x}_2)^2 \end{aligned}$$

$$\begin{aligned} MSE &= \frac{1}{n-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \overline{x}_i)^2 = s_P^2 \\ &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \end{aligned}$$

Therefore

$$F = \frac{\left(\frac{n_1 n_2}{n_1 + n_2}\right)(\bar{x}_1 - \bar{x}_2)^2}{s_P^2} = \left(\frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right)^2$$

Thus $F = t^2$, where t is the two-sample t-test statistic.

Thus if $k = 2$, the ANOVA F-test and the two sample t-test are **EQUIVALENT**

$$\begin{aligned} t &\sim \text{Student-t}(n-2) \\ F &\sim \text{Fisher-F}(1, n-2) \end{aligned}$$

and we must get the same conclusion (to reject H_0 or otherwise) using either statistic.