Summary

If the assumptions

- independence (holds by design in a CRD)
- Normal populations
- equal variances

hold, use

ANOVA F-test

If the assumptions **do not hold**

- use Randomization/Permutation test
- use Non-parametric test (see Section 3)

1.3 Multiple Comparison of Means

If the ANOVA F-test null hypothesis

$$H_0 : \mu_1 = \cdots = \mu_k$$

is **rejected**, then it is of interest to discover which of the means are different. For k groups, there are $c = k(k-1)/2$ pairs of group means that can be compared.

Consider a "family" of hypothesis tests - a collection of tests of different hypotheses carried out independently on different data sets. For each test in the family, we consider testing the hypothesis at significance level α.

Notation

Label the tests $i = 1, \ldots, c$, and for each i, label

- the null hypotheses H_{0i}
- the test statistics T_i
- the rejection regions $\mathcal{R}_i$

that are potentially different for each i.

We specify for each i,

$$\alpha = P[T_i \in \mathcal{R}_i | H_{0i} \text{ is } \textbf{TRUE}]$$

which implicitly defines $\mathcal{R}_i$. Note that α is the

"Test Type-I Error Rate" or "Comparisonwise Error Rate"

Now consider the results of all tests in the family; what is the "Familywise" Type-I error rate ?

Using the laws of probability

$$P[T_i \in \mathcal{R}_i | H_{0i} \text{ is } \textbf{TRUE}] = \alpha$$

means that

$$P[T_i \notin \mathcal{R}_i | H_{0i} \text{ is } \textbf{TRUE}] = 1 - \alpha$$

giving the probability that the test **does not reject** H_{0i}, if H_{0i} is in fact true, is $1 - \alpha$.

Now we consider all tests together;

$$P[\text{Each } T_i \notin \mathcal{R}_i | \text{Each } H_{0i} \text{ is } \textbf{TRUE}] = (1 - \alpha)^c$$

This is the probability that each test results in the null hypothesis **not** being rejected, that is, the probability that we **never** commit a Type-I error.

Therefore the probability of **at least one** Type-I error is

$$\alpha_F = 1 - (1 - \alpha)^c$$

α_F is the **Familywise Error Rate**.

	$\alpha = 0.05$	$\alpha = 0.01$
c	α_F	α_F
5	0.226	0.049
10	0.401	0.096
50	0.923	0.395
100	0.994	0.634

Therefore, whenever we carry out a "family" of tests, we should not use the traditional choices of $\alpha = 0.05$ or 0.01 on each test.

To fix $\alpha_F = 0.05$, say, we need to use α on each test where

$$\alpha_F = 1 - (1-\alpha)^c \quad \Longleftrightarrow \quad \alpha = 1 - (1-\alpha_F)^{1/c}$$

For example, if $\alpha_F = 0.05$ and $c = 10$, use

$$\alpha = 1 - (1 - 0.05)^{1/10} = 0.0051$$

Bonferroni Method

It can be shown that

$$1 - (1 - \alpha)^c \approx c\alpha$$

Therefore, if α_F is the familywise error rate, we can set

$$\alpha = \frac{\alpha_F}{c}$$

to get the comparisonwise error rate.

α_F / c is known as the Bonferroni Correction.