A. Testing for Equal Population Means: One-Way ANOVA

1. Model: k (k≥2) populations, outcome quantitative and normally distributed, equal spread (σ²), label the means μ₁ to μₖ, name the populations with a nominal variable.

2. H₀: μ₁=…=μₖ. H₁: at least one population mean differs from the others

3. Experiment: Randomly select nᵢ subjects from population i. For simplicity we focus on the case with n from each of the k populations.

4. A useful statistic: F=MS_{between}/MS_{within}.

5. MS_{within}=SS_{within}/df_{within} is an estimate of σ² that is unaffected by whether or not the null hypothesis is true. df_{within}=k(n-1).

6. MS_{between}=SS_{between}/df_{between} is an estimate of σ² if the null hypothesis is true, but larger otherwise. df_{between}=k-1.

7. Under the null hypothesis, experimental repetitions will give F statistics that vary, but center around 1. The exact “null distribution” of this F statistic, if the assumptions are true, is the theoretical distribution called F_{a,b} where a is the numerator df and b is the denominator df.

8. From the position of the F statistic in its theoretical null sampling distribution, we can find the p-value (significance).

9. If the p-value is less than our pre-chosen alpha (α), e.g. 0.05, then our results are “surprisingly uncommon” for similar experiments in which the null hypothesis is true. We therefore reject the null hypothesis and conclude that there is good evidence that at least one treatment corresponds to a population with a mean different from the others. Because this is an experiment, we are justified in concluding that the treatment causes a change in the mean population outcome.

10. We always keep in mind that if we reject the null hypothesis, we might be making a type 1 error (rejected the null hypothesis when it is true), and if we do not reject it, we might be making a type 2 error (accepting the null when it is false).

11. For complicated mathematical reasons, the t-test for independent samples gives a statistic, t, which when squared exactly equals the F of ANOVA for k=2. Therefore the t-test is unnecessary.
B. Some ANOVA theory

1. Intermediate Theory of One-Way ANOVA: Whatever the group means are in the populations, the k different within-group variances are independent estimates of the true common population variance. Each variance is calculated as follows: Let $X_{ij}$ represent the outcome for the j'th subject in treatment group i, and let $\bar{X}_i$ be the mean of the n (or $n_i$) subjects in group i. For group i, calculate the n deviations, $(X_{ij} - \bar{X}_i)$. Square and sum them to get $SS_i$. The degrees of freedom are n-1. Each variance estimate is $SS_i/(n-1)$. The trick is to pool the k estimates as follows:

$$SS_{within} = \sum_{i=1}^{k} SS_i$$ and $df_{within} = \sum_{i=1}^{k} df_i$ which is k(n-1) if each group has size n.

The additional information about the $\mu$'s (which is independent of the information already used) is in the $X_i$ values. Let $\bar{X}$ be the grand mean. Then

$$SS_{between} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$. Because we can choose k-1 means freely while still maintaining the same grand mean there are (k-1) df in the numerator of the F statistic. Intuitively, can you see that even when the $\mu$'s are all equal, the value of $MS_{between}$ will not tend to be zero? Can you see the role of “n” in $SS_{between}$?

2. Advanced Theory of One-Way ANOVA: If $X$ follows a standard normal distribution, then $X^2$, which is always non-negative, can’t be normally distributed. It follows a chi squared distribution with 1 degree of freedom ($\chi^2(1)$), which happens to have mean 1 and variance 2. To avoid having lots of similar distributions, instead of looking at $X^2$, if X follows a $N(\mu, \sigma^2)$ distribution we look at $(X - \mu)^2/\sigma^2$ which will follow $\chi^2(1)$. This extends directly to the fact that any Mean Square divided by the true variance that it is estimating follows the $\chi^2(df)$ distribution, which has mean df and variance 2df. Finally, the ratio of two independent random variables that each follow a chi square distribution follows the F distribution (another non-negative distribution). So, in ANOVA, when $MS_1$ and $MS_2$ are independent, the ratio

$$F = \frac{MS_1/\sigma_1^2}{MS_2/\sigma_2^2}$$

follows the F distribution with numerator and denominator degrees of freedom equal to the numbers of degrees of freedom in the corresponding Mean Squares. In our standard usage, under the null distribution the two unknown variances are assumed to be equal, so they cancel out.