L. Power and Sample Size

Recall how we computed power for a two-sample t-test:

letting \( \mu_0 \) represent the value of \( \mu_1 - \mu_2 \) under \( H_0 \) and \( \mu_D \) represent a value of \( \mu_1 - \mu_2 \) under \( H_a \).
We can write down the same thing for F-tests:

\[
\text{power} = P \left[ F^* = \frac{MST}{MSE} > F_{\alpha,t-1,N-t} \bigg| \lambda \right]
\]

What is \( \lambda \)? \( \lambda \) is the non-centrality parameter and it represents how different the \( \mu_i \) are from each other:

\[
\lambda = \frac{\sum_{i=1}^{t} r_i \tau_i^2}{\sigma^2} \quad \text{and} \quad \tau_i = \mu_i - \bar{\mu}.
\]

i.e., \( \lambda \) encompasses the \( \mu_i \) (or \( \tau_i \)) values under \( H_a \).
When $H_0 : \mu_1 = \mu_2 = \ldots = \mu_t$ is true, then

$$F^* \sim F_{t-1, N-t} (\lambda = 0, \text{central F distribution})$$

but when $H_0$ is false, then

$$F^* \sim F_{t-1, N-t, \lambda} (\text{non-central F distribution}).$$

We can use power tables that show curves of power, $\phi = \sqrt{\frac{\lambda}{t}}$ as the non-centrality parameter, numerator and denominator degrees of freedom and $\alpha$-level.
To use these, we must specify:

- \( \alpha \)-level

- numerator degrees of freedom, \( t - 1 \)

- effect sizes (\( \tau_i \)) and variance (\( \sigma^2 \))

and one of:

- denominator degrees of freedom (\( \sum_i r_i - t \))

- desired level of power.
EXAMPLE VOR in ataxia patients
(from Oehlert, 2000)

Spinocerebellar ataxias (SCAs) are inherited, degenerative, neurological diseases that may affect eye movements, among other things. Different types of SCAs may affect eye movements differently.

Researchers want to measure the amplitude of the vestibulo-ocular reflex for 20 deg/sec² velocity ramps (VOR). Patients will sit in a chair placed onto a turntable that rotates increasingly rapidly.
VOR measures how well patients are able to focus on a fixed target while the chair rotates. They will use patients who have SCA Types 1, 5, and 6, with equal numbers of patients of each type.

Previous research has shown $\sigma^2 \approx 0.15$ and the effect sizes are $\approx -0.43, 0.64,$ and $-0.21$. We want $\alpha = 0.05$. 
What is the power if we use \( r_i \equiv 3 \ \forall i \)?
What is the power if we use $r_i \equiv 4 \; \forall i$?
What if we ask the question the other way; how many patients per group do we need to achieve a power of at least 0.90?

By doing a bit of algebra and assuming equal group sizes, we see that \( r = \frac{\phi^2 \sigma^2 t}{\sum_{i=1}^{3} \tau_i} \). For this example data, then \( r = \frac{0.45\phi^2}{0.6386} \). Using the power curves and this equation, we can backsolve to get \( r \).
There are many, many web pages with power and sample size calculators. Stick with calculators provided by educational institutions if possible. Determine if the calculator does an easy calculation correctly before trusting it to do a more complex calculation.

For a list of links, see e.g.,
http://members.aol.com/johnp71/javastat.html#Power

There is also much commercial software (PASS, nQuery, ...).