1: Consider a one-way ANOVA design with 5 factors and 10 replicates per factor. Suppose that factor variance \( \sigma^2_\tau \) is ten percent of the total variance \( \sigma^2_T \) (i.e., \( \frac{\sigma^2_\tau}{\sigma^2_T} = 0.10 \)).

(a) Given that the total variance equals the treatment plus error variance \( \sigma^2_T = \sigma^2_\tau + \sigma^2_e \), what is \( \frac{\sigma^2_\tau}{\sigma^2_e} \)?

Note that \( \sigma^2_T = \sigma^2_\tau + \sigma^2_e \) implies \( \frac{\sigma^2_\tau}{\sigma^2_e} = 1 + \frac{\sigma^2_\tau}{\sigma^2_T} \), or that \( \frac{\sigma^2_\tau}{\sigma^2_e} = \frac{\sigma^2_T}{\sigma^2_\tau} - 1 \), and hence

\[
\frac{\sigma^2_\tau}{\sigma^2_e} = \frac{1}{\frac{\sigma^2_T}{\sigma^2_\tau} - 1} = \frac{1}{1/0.1 - 1} = \frac{1}{9} = 0.1111
\]

(b) What is the 95% critical value for the F-test?

The 95% critical value \( f \) satisfies \( \Pr(F_{N-1,N(n-1)} \leq f) = 0.95 \). Here \( N = 5, n = 10 \). Using R, we find that

\[
> qf(0.95, 5-1,5*(10-1))
\]

[1] 2.578739

(c) What is the power of this design (assuming a test of \( \alpha = 0.05 \)) for a fixed-effects ANOVA?

From Equation A5.23 in the notes, the power is given by

\[
\Pr(F_{N-1,N(n-1),\lambda} \geq 2.578739)
\]

where the noncentrality parameter \( \lambda = n(N - 1)\frac{\sigma^2_\tau}{\sigma^2_e} = 10 * 4 * (1/9) \). In R, we can compute this by using

\[
> 1-pf(2.578739,5-1,5*(10-1),10*4/9)
\]

[1] 0.3207070

Note that we used \( 1-pf \), as \( pf \) returns the probability of being \( \leq x \), while we wish the probability of being \( \geq x \). Note we could also have used

\[
> 1-pf(qf(0.95,5-1,5*10-1),5-1,5*(10-1),10*4/9)
\]

directly having R recompute the critical value.

(d) What is the power of this design under a random-effects ANOVA?

Recalling Equation A5.30b in the notes, the power is given by

\[
\Pr\left[ F_{N-1,N(n-1)} > \frac{F_{N-1,N(n-1),(1-\alpha)}}{1 + n \left( \frac{\sigma^2_\tau}{\sigma^2_e} \right)} \right] = \Pr\left[ F_{N-1,N(n-1)} > \frac{2.578739}{1 + n \left( \frac{\sigma^2_\tau}{\sigma^2_e} \right)} \right]
\]

In R,

\[
> 1- pf(2.578739/(1+10/9),5-1,5*(10-1))
\]

[1] 0.3150735

Giving a power close too, but slight less than, a fixed-effects design.

(e) Given these sample sizes, what is the smallest value of \( \frac{\sigma^2_\tau}{\sigma^2_e} \) that gives a (fixed-effects) 95% test a power of 0.90? (You will need to do this, and some the remaining problems, by trial and error.)

We try various values of \( \frac{\sigma^2_\tau}{\sigma^2_e} \):

\[
> 1-pf(2.578739,5-1,5*(10-1),10*4*(1/2)) \quad \# \text{trying } 1/2
\]

[1] 0.943019

\[
> 1-pf(2.578739,5-1,5*(10-1),10*4*(0.4)) \quad \# \text{trying } 0.4
\]

[1] 0.8768408

\[
> 1-pf(2.578739,5-1,5*(10-1),10*4*(0.43)) \quad \# \text{trying } 0.43
\]

1
Hence, this design has the power to detect an effect accounting for 30% or more of the total variation.

Note that we could also solve this more quickly by using the graphics commands in R, plotting power as a function of varying values of $\sigma_T^2/\sigma_e^2$. In particular to example power for $0.1 \leq \sigma_T^2/\sigma_e^2 \leq 0.9$, the R code is

```r
> curve(1-pf(2.578739,5-1,5*(10-1),10*4*x), 0.1,0.9)
```

...which returns a nice graph. We can focus in on finer regions of the curve by restricting the interval to smaller regions, e.g.,

```r
> curve(1-pf(2.578739,5-1,5*(10-1),10*4*x), 0.4,0.5)
```

(f) Given these sample sizes, what is the smallest value of $\sigma_T^2/\sigma_e^2$ that gives a random-effects 95% test a power of 0.90?

Here, we need to solve for $\sigma_T^2/\sigma_e^2$ in

$$\Pr \left[ F_{N-1,N(n-1)} > \frac{2.578739}{1+n(\sigma_T^2/\sigma_e^2)} \right] = 0.9$$

Using R [ qf (0.1, 5-1, 5* (10-1) ) ], we find that $\Pr \left[ F_{N-1,N(n-1)} > 0.26323 \right] = 0.9$, hence we solve for

$$\frac{2.578739}{1+10(\sigma_T^2/\sigma_e^2)} = 0.26323$$

giving $\sigma_T^2/\sigma_e^2 = 0.8796524$, so that $\sigma_T^2$ must account for at least 46.8% of the total variance (using the expression in part (f) to covert $\sigma_T^2/\sigma_e^2$ into $\sigma_T^2/\sigma_e^2$.)

(g) How many replicates per factor are needed to give the fixed-effects ANOVA a power of 90% under a test of significant with $\alpha = 0.05$?

Once again, we can use trail and error, varying $n$ to solve

$$\Pr(F_{5-1,5(n-1),n(5-1)/9} \geq F_{5-1,5(n-1),[0.95]}) = 0.9$$

in R, we first try $n = 30$.

```r
> n <- 30; 1-pf(qf(0.95, 5-1, 5* (n-1) ), 5-1, 5* (n-1) , n*4/9)
[1] 0.8339192
```

```r
> n <- 35; 1-pf(qf(0.95, 5-1, 5* (n-1) ), 5-1, 5* (n-1) , n*4/9)
[1] 0.8940105
```

```r
> n <- 36; 1-pf(qf(0.95, 5-1, 5* (n-1) ), 5-1, 5* (n-1) , n*4/9)
[1] 0.9034603
```

```r
n = 36 it is.
```

(h) How many replicates per factor are needed to give the random-effects ANOVA a power of 90% under a test of significant with $\alpha = 0.05$? (Again, need to use trail and error)

Here we need to find $n$ such that

$$\Pr \left[ F_{5-1,5(n-1)} > \frac{F_{5-1,5(n-1),[0.95]}}{1+n/9} \right] = 0.9$$

in R, we first try $n = 80$,

```r
> n <- 80; 1-pf(qf(0.95, 5-1, 5* (n-1) ) / (1+n/9) , 5-1, 5* (n-1) )
```

[1] 0.9015374
Optimal design for a random-effects ANOVA. Suppose you have a total $T = 100$ measurements that you can make, and you have to decide how best to allocate them over $N$ and $n$ in a random-effects design. Should one choose more factors (increase $N$) at the expense of fewer replicates $n$ per factor? Obviously, there is some intermediate trade-off between the two. Suppose that the factor variance is $\sigma^2_\tau = 10$ and the error variance $\sigma^2_e = 20$.

(a) Compute the power of this design for the following combinations of $N$ and $n$:

\[
\begin{array}{cccccccc}
50 & 23 & 33 & 25 & 4 & 20 & 5 & 10 & 10 \\
5 & 10 & 20 & 4 & 25 & 3 & 33 & 2 & 50 \\
\end{array}
\]

Hint: It might make sense to first write an R function to do this for arbitrary $N$, $n$.

\[
f_{\text{crit}} \leftarrow \text{function}(N, n, \alpha) \ qf(1-\alpha, N-1, N*(n-1))
\]

is the R code for the $\alpha$-level critical value, $F_{N-1,N(n-1)[1-\alpha]}$ under the null hypothesis. Recall (Equation A5.30b in the power notes) that the power is

\[
\Pr \left( F_{N-1,N(n-1)} > \frac{F_{N-1,N(n-1)[1-\alpha]}}{1 + n\sigma^2_\tau / \sigma^2_e} \right)
\]

Using the above function, a new function, \texttt{fpower} can be written in R as

\[
\text{fpower} \leftarrow \text{function}(N, n, \alpha, \text{var}) \ 1-\text{pf}(f_{\text{crit}}(N, n, \alpha)/(1+n*\text{var}), N-1, N*(n-1))
\]

where $\text{var} = \sigma^2_\tau / \sigma^2_e$. For example, \texttt{fpower(33, 3, 0.05, 10/20)} returns [1] 0.9107213. Running through the above values of $N$, $n$,

\[
\begin{array}{cccccccc}
50 & 23 & 33 & 25 & 4 & 20 & 5 & 10 & 10 \\
5 & 10 & 20 & 4 & 25 & 3 & 33 & 2 & 50 \\
\end{array}
\]

(b) What is the optimal design (i.e., which combination of $N$ and $n$ gives the largest power)?

\[
N = n = 10
\]

(c) Repeat (a) and (b) assuming $\sigma^2_\tau = 20, \sigma^2_e = 10$

\[
\begin{array}{cccccccc}
50 & 23 & 33 & 25 & 4 & 20 & 5 & 10 & 10 \\
5 & 10 & 20 & 4 & 25 & 3 & 33 & 2 & 50 \\
0.99995 & 0.99999 & 0.99999 & 0.99998 & 0.99996 & 0.99932 & 0.9838 & 0.9549 & 0.8459 \\
\end{array}
\]