Statistical Power

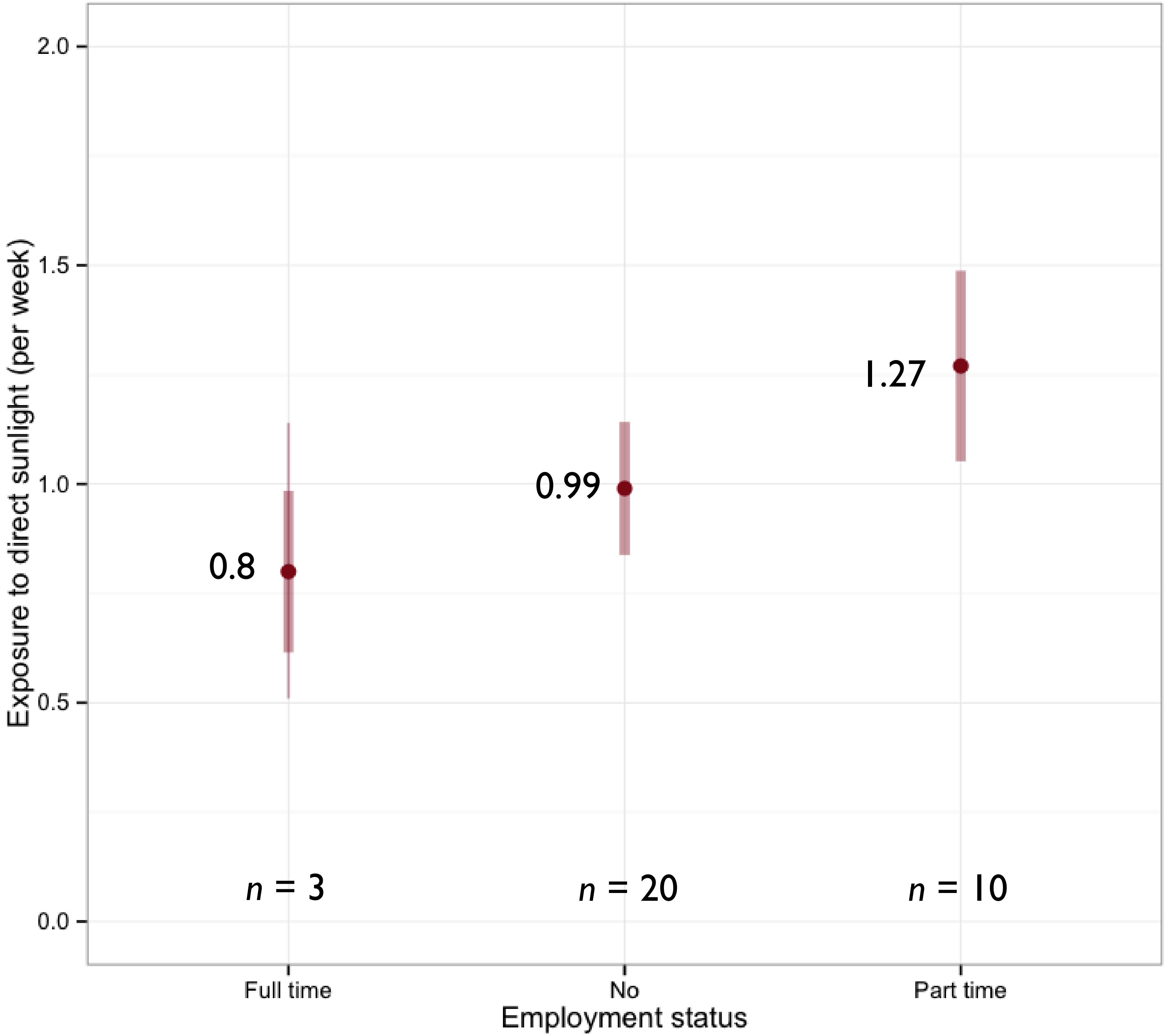
Andrew Zieffler

Educational Psychology

University of Minnesota

Driven to Discover℠


Does employment status have an effect on the average amount of sun exposure per week for elderly women?



2.0
1.5
1.0
0.5
0.0
Exposure to direct sunlight (per week)
1.27
0.99
0.8
n = 3
n = 20
n = 10
Full time
No
Part time
Employment status

```r
# Fit the ANOVA model
> lm.1 = lm(sun ~ employed, data = osteo)
> anova(lm.1)
```

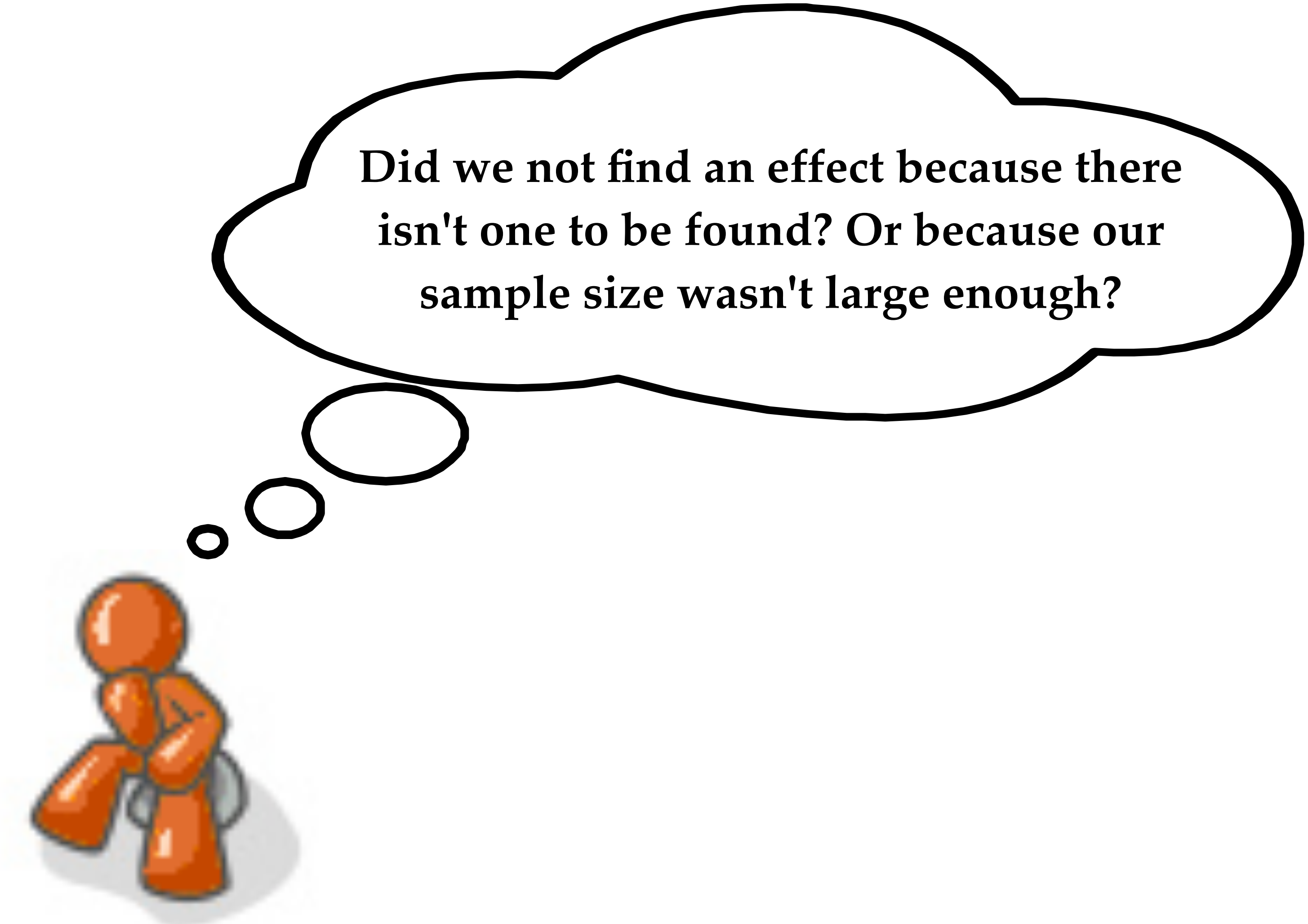
```
Analysis of Variance Table

Response: sun
          Df  Sum Sq Mean Sq F value Pr(>F)
employed   2  0.7465 0.37327  0.8514 0.4369
Residuals 30 13.1527 0.43842
```

```r
# Find the effect size
> summary(lm.1)
```

```
...
Residual standard error: 0.6621 on 30 degrees of freedom
Multiple R-squared:  0.05371,	Adjusted R-squared:  -0.009375
F-statistic: 0.8514 on 2 and 30 DF,  p-value: 0.4369
```

The direct exposure to sunlight of 33 elderly women were analyzed in a fixed-effects analysis of variance (ANOVA) with employment status (full time vs. part time vs. none) as a between-subjects factor. The effect of region was **not** found to be statistically significant, $F(2, 30) = 0.85$, $p = 0.437$. ***These data do not provide any evidence against the null hypothesis, $\eta^2 = 0.05$, 95% CI = [0, 0.24].***


Did we not find an effect because there isn't one to be found? Or because our sample size wasn't large enough?

Failing to Reject the Null Hypothesis

- No probability of making a type I error (you need to reject in order to make a type I error)

$$\alpha = 0$$

- Some probability of making a type II error

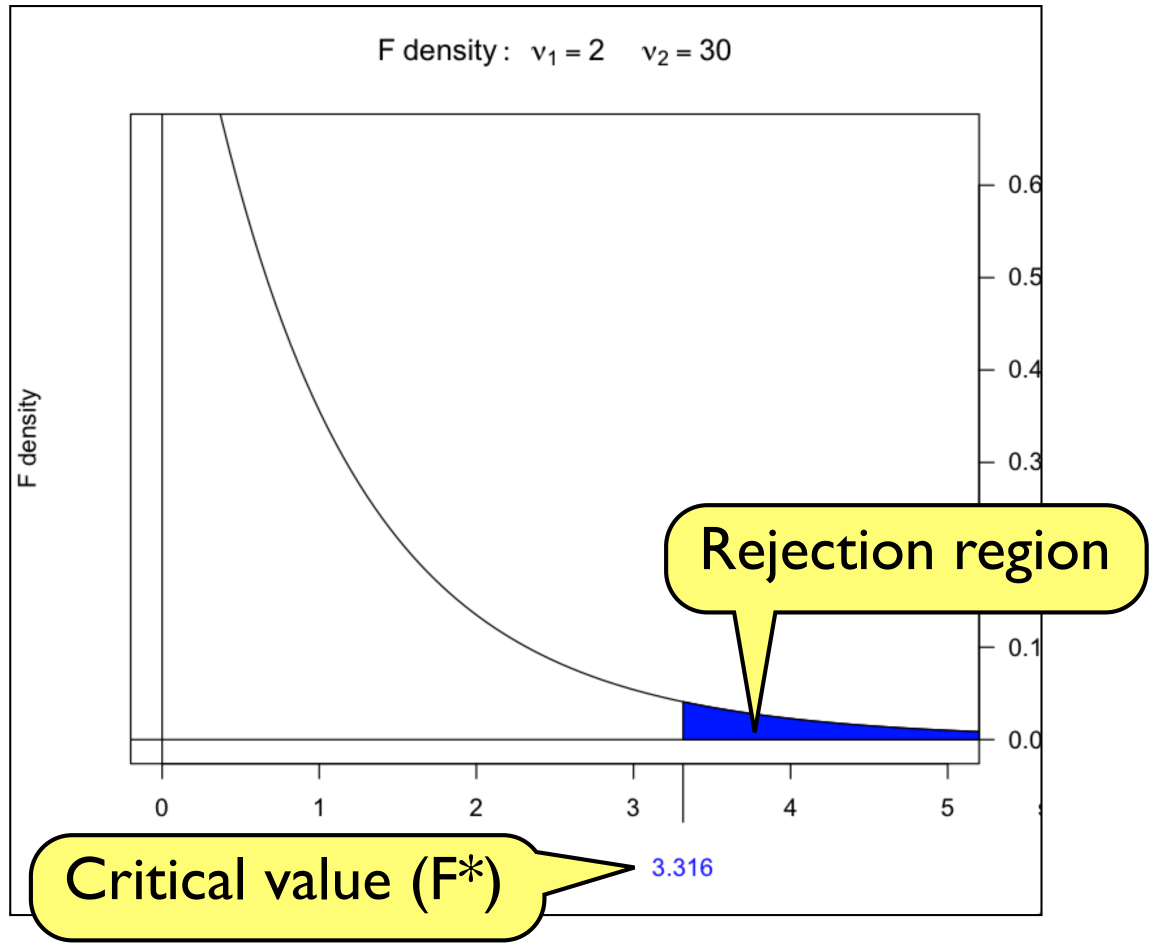
$$\beta \neq 0$$

To compute the probability of making a type II error more specifically we need to specify the direction and magnitude of the true population effect

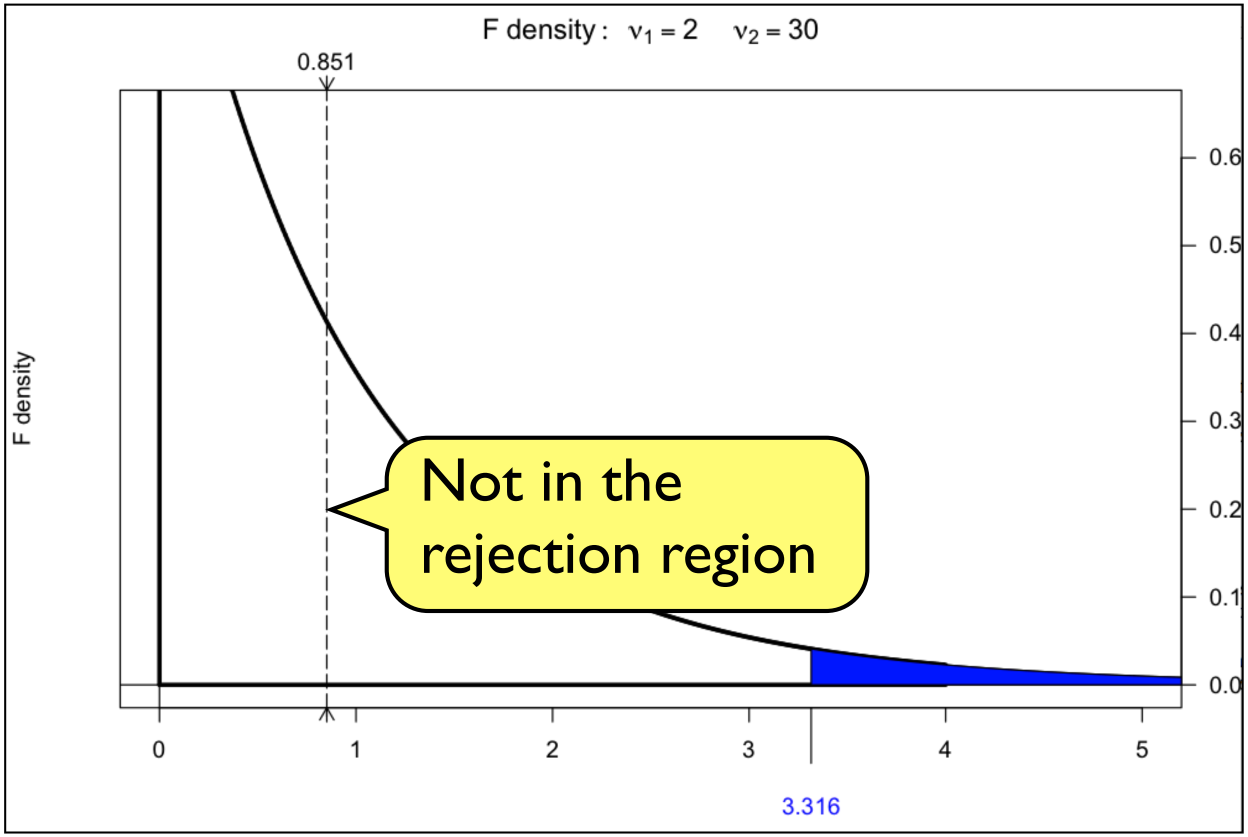
If the null hypothesis is TRUE (no effect), the *F*-distribution is identified completely based on the degrees of freedom (in this case namely 2 and 30)

```
# To find the critical value
> qf(p=0.95, df1=2, df2=30)
```

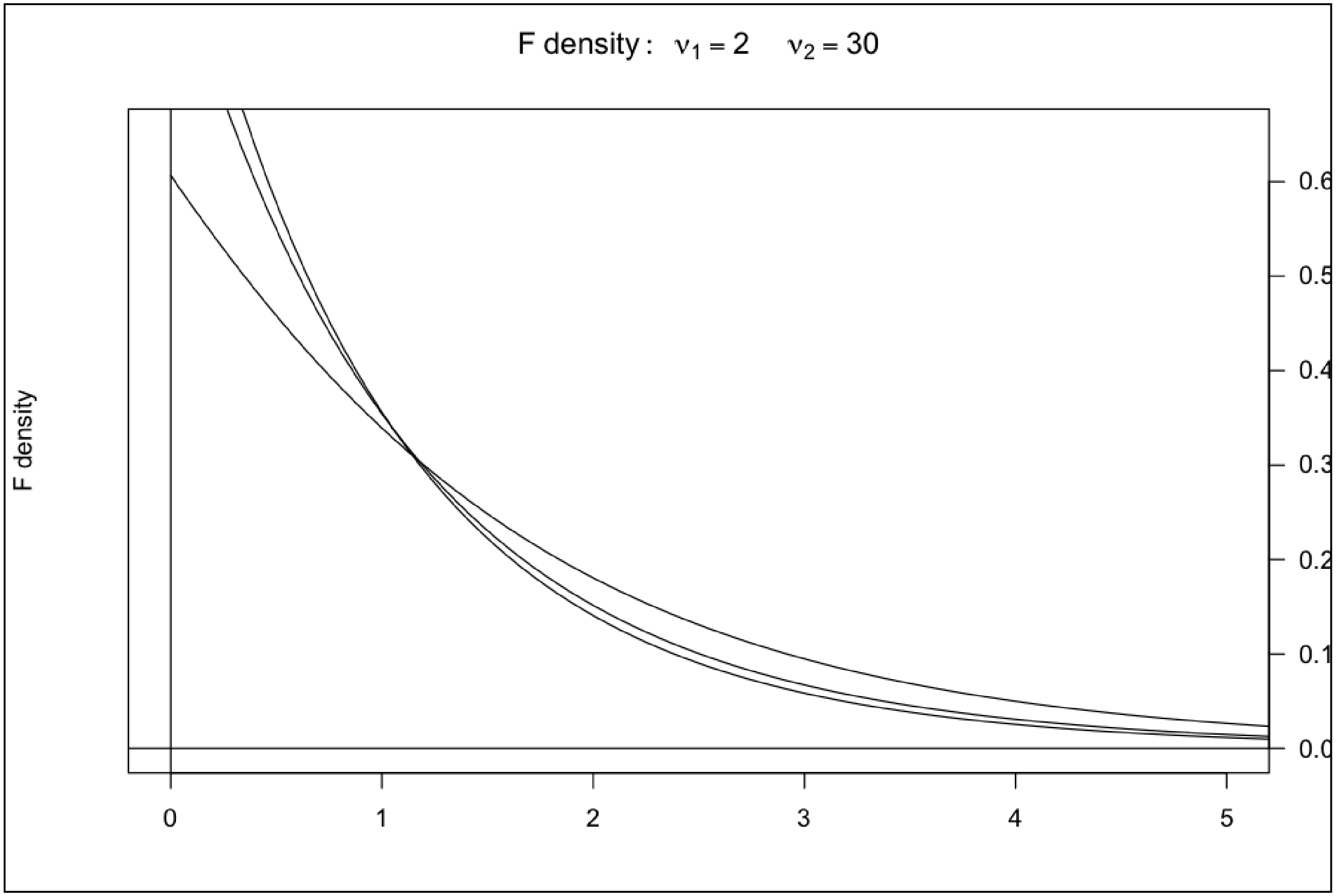
```
[1] 3.31583
```



The observed *F*-statistic is not in the rejection region (we fail to reject the null hypothesis)



What does the *F*-distribution look like if the null hypothesis is false?



...it depends.

It depends on the actual effect in the population. The degree of effect is specified by another parameter, λ, which is known as the **non-centrality parameter**.

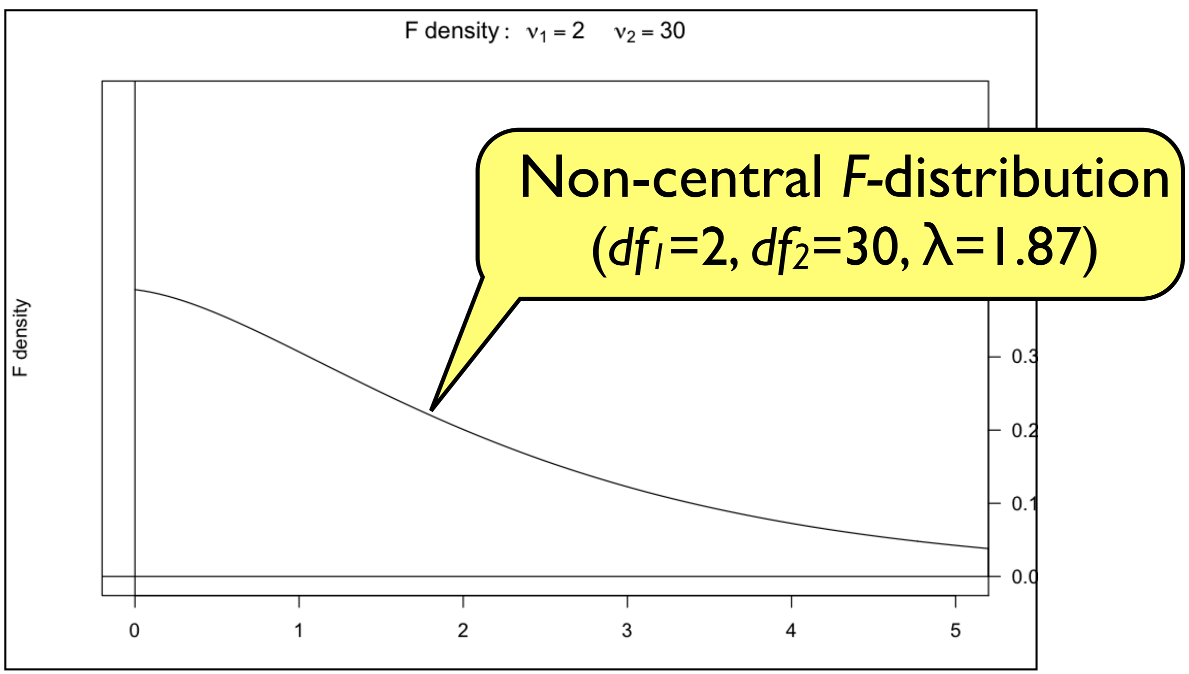
We can estimate λ from the data using

$$\hat{\lambda} = N \times \frac{R^2}{1 - R^2}$$

For our example,

```
...
Residual standard error: 0.6621 on 30 degrees of freedom
Multiple R-squared:  0.05371,   Adjusted R-squared:  -0.009375
F-statistic: 0.8514 on 2 and 30 DF,  p-value: 0.4369
```

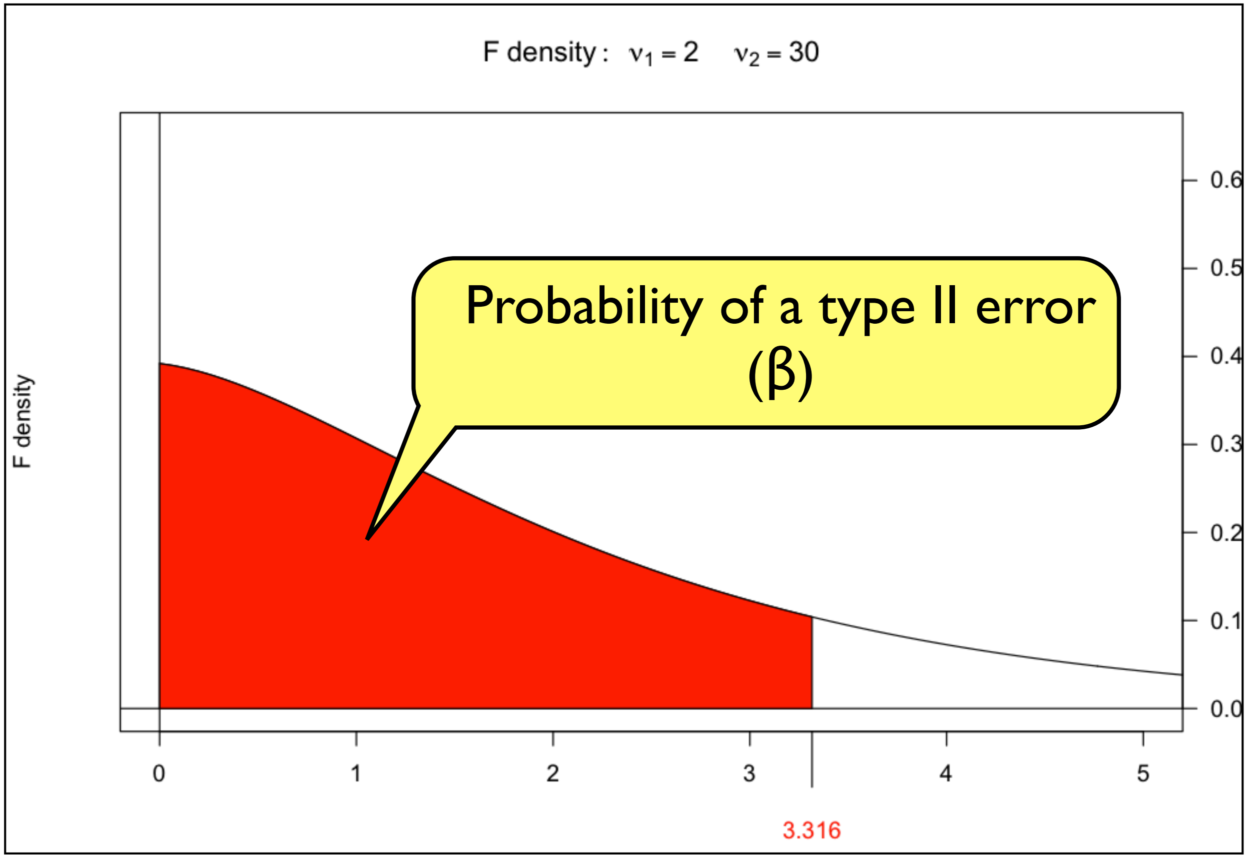
$$\hat{\lambda} = 33 \times \frac{0.05371}{1 - 0.05371} = 1.87303$$



Probability of a type II error

- A type II error is fail to reject the null hypothesis (not in the rejection region) but,

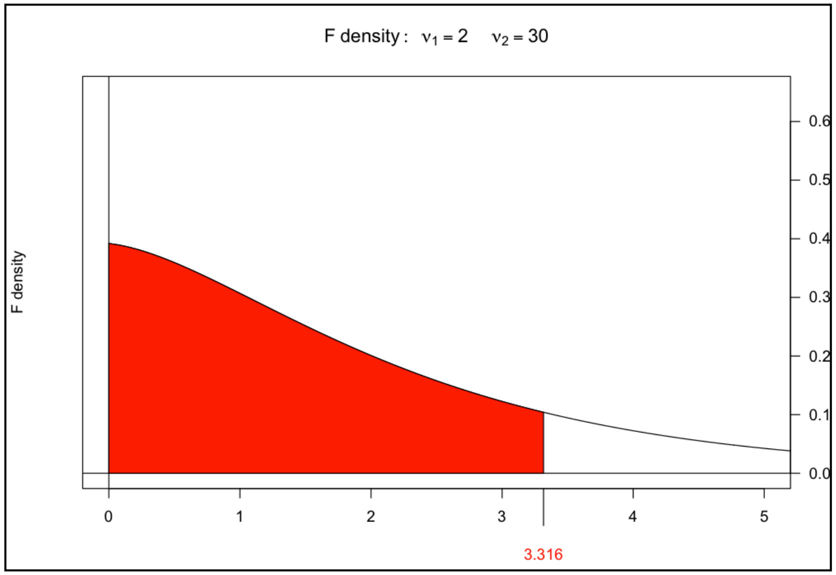
- there is an effect (you are in the non-central *F*-distribution)



To compute this we need to find the cumulative density of the Non-central *F*-distribution (df_1=2, df_2=30, λ=1.87) from $-\infty$ to the critical value of 3.316.

```
# To find the cumulative density
> pf(q=3.316, df1=2, df2=30, ncp=1.87303)
```

```
[1] 0.8033059
```

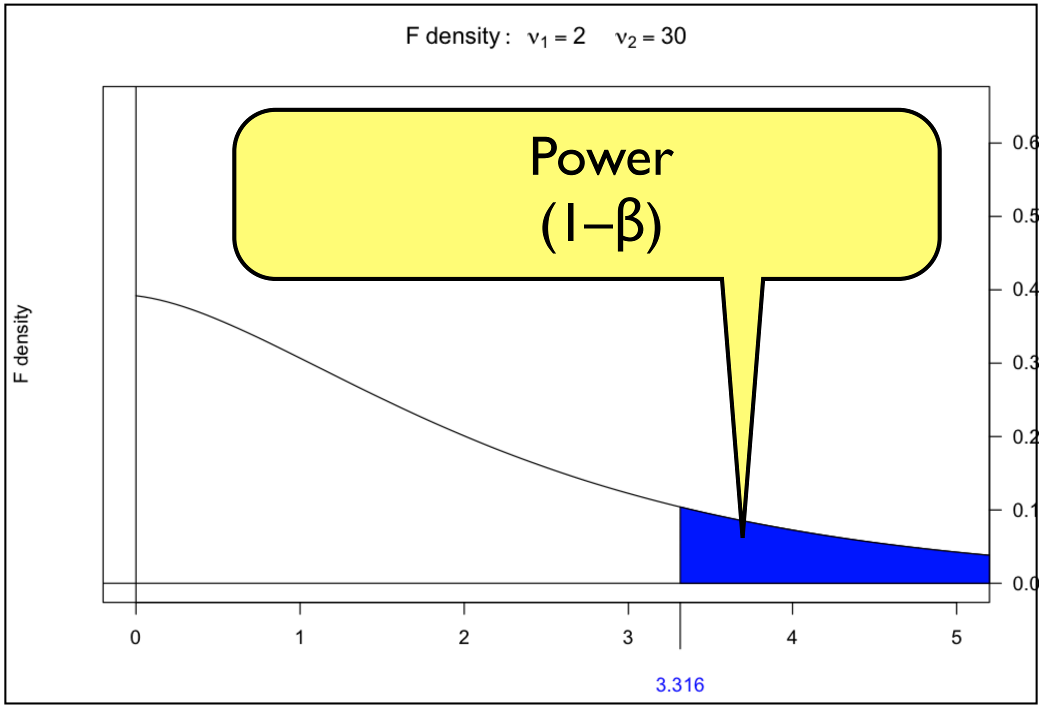


The probability of making a type II error, if the true effect in the population is R^2=0.054 and we use a fixed-effects ANOVA with a type I error rate of 0.05 to analyze the data is 0.80!

To compute power, we subtract the probability of making a type II error from 1.

```
# To find the cumulative density
> 1 - 0.8033059
```

```
[1] 0.1966941
```

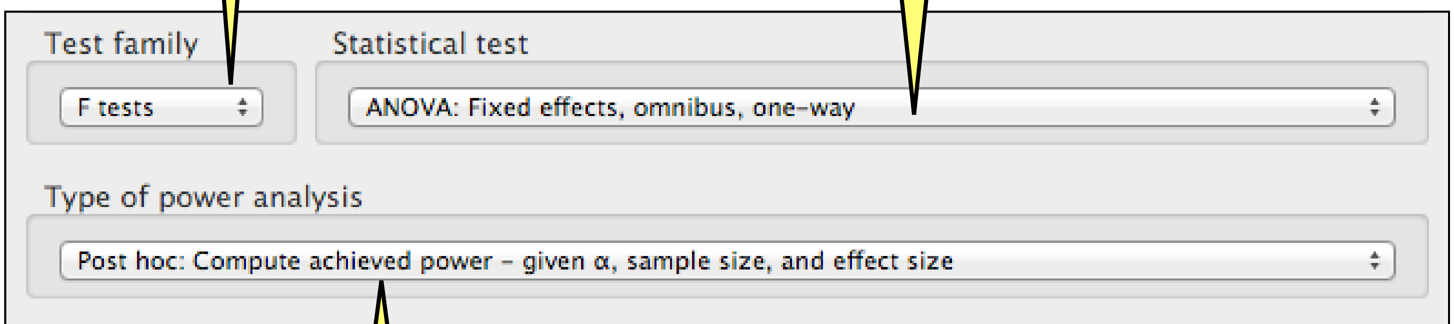


The power (probability) of detecting an effect of R^2=0.054 using a fixed-effects ANOVA with a type I error rate of 0.05 is 0.20!

Using *G*Power 3* to compute statistical power

1. Set the appropriate *Test family* (for ANOVA this is the F test)

2. Set the appropriate *Statistical test*



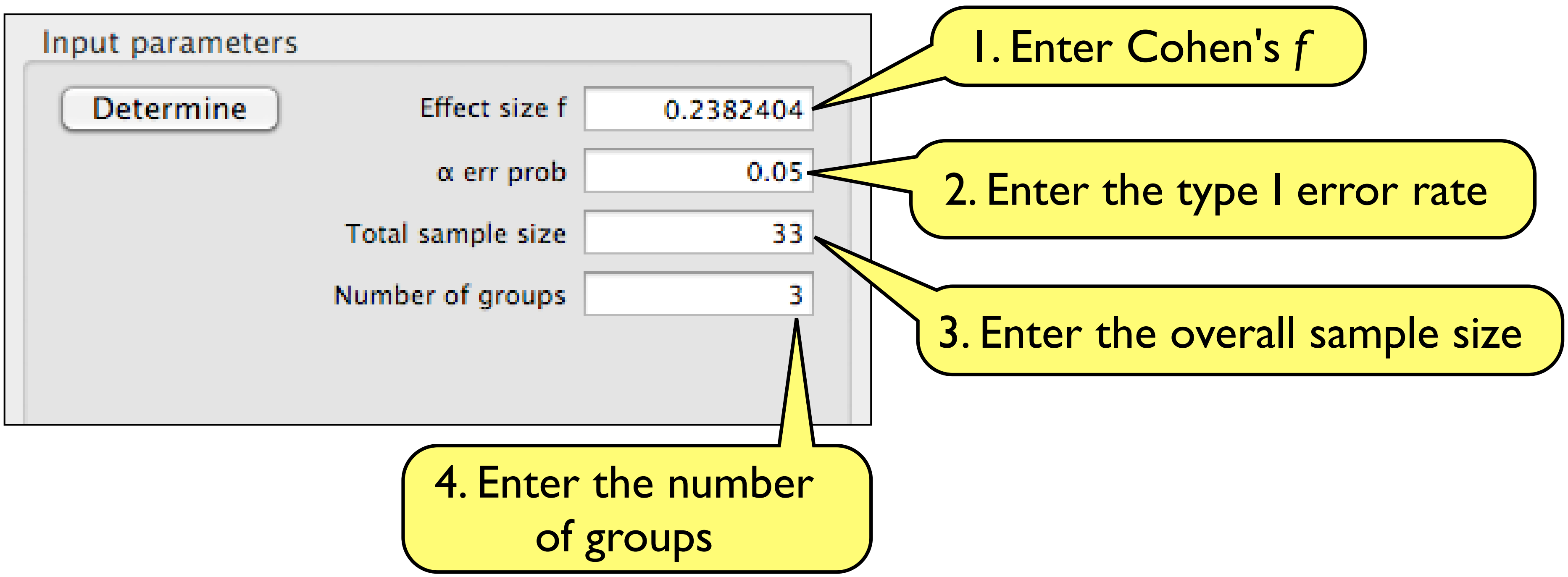
3. Set the type of *power analysis*

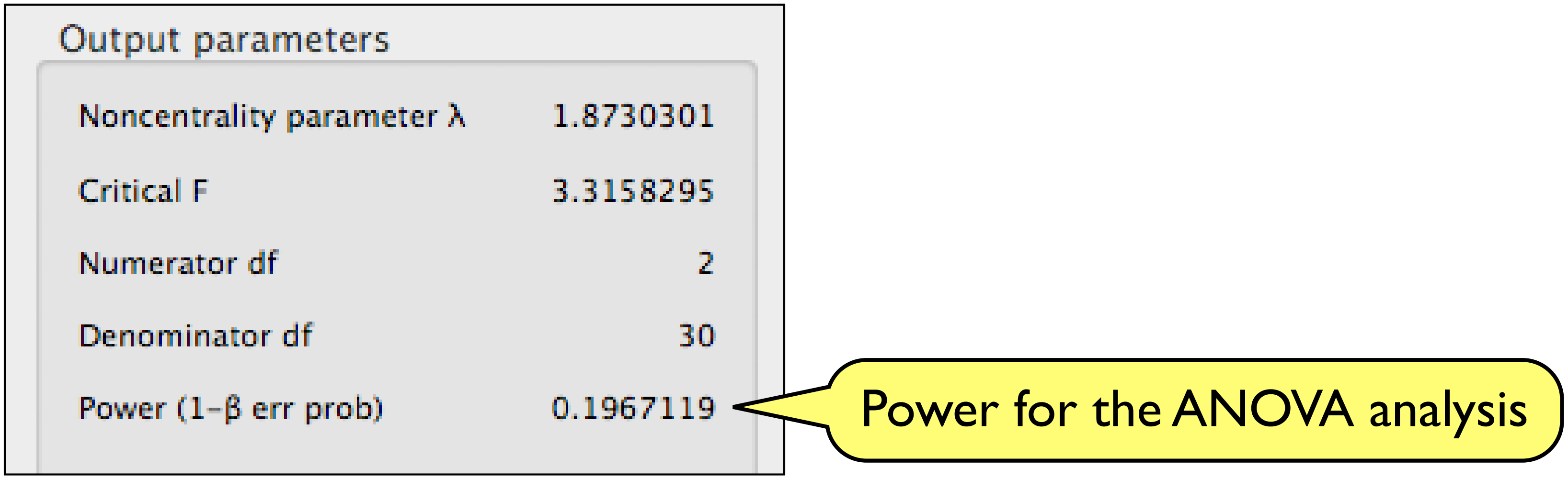
- **Post hoc:** compute power for a particular analysis
- **A priori:** compute sample size

To compute **post hoc** power...

$$f = \sqrt{\frac{R^2}{1 - R^2}}$$

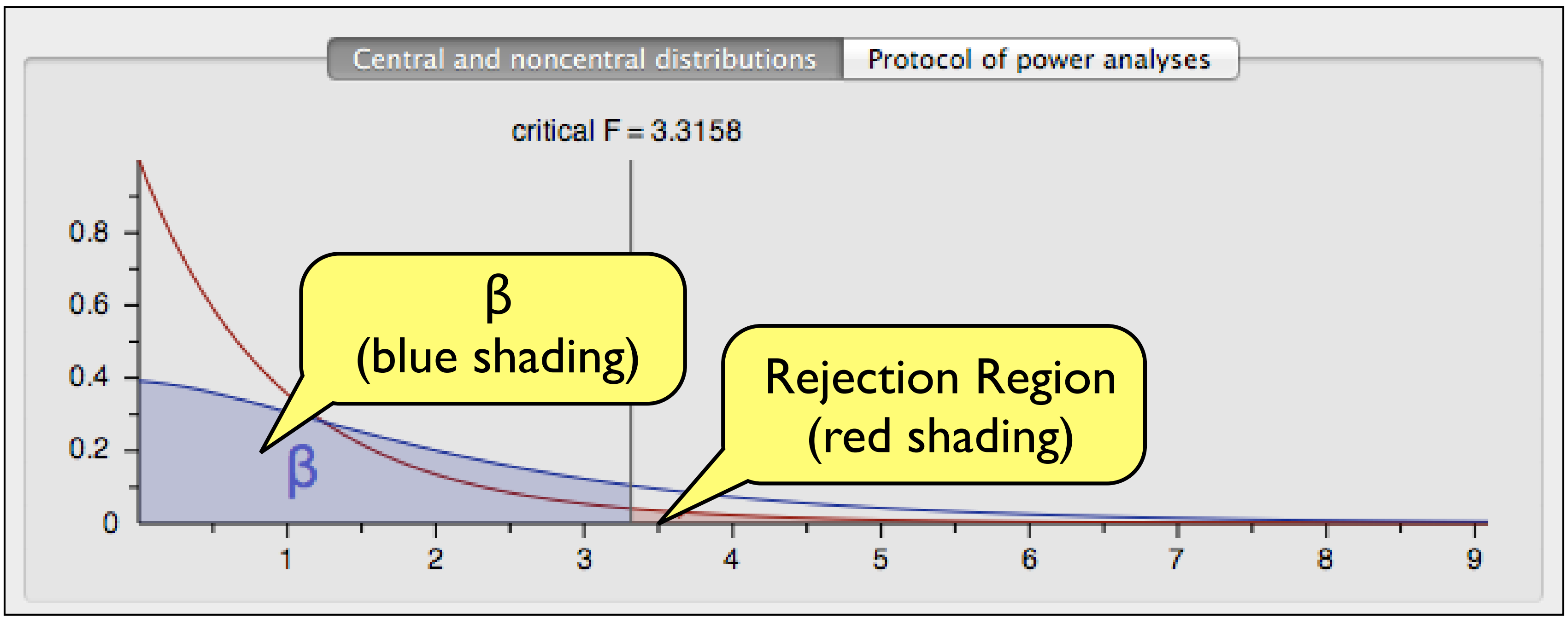
$$f = \sqrt{\frac{0.05371}{1 - 0.05371}} = 0.2382404$$

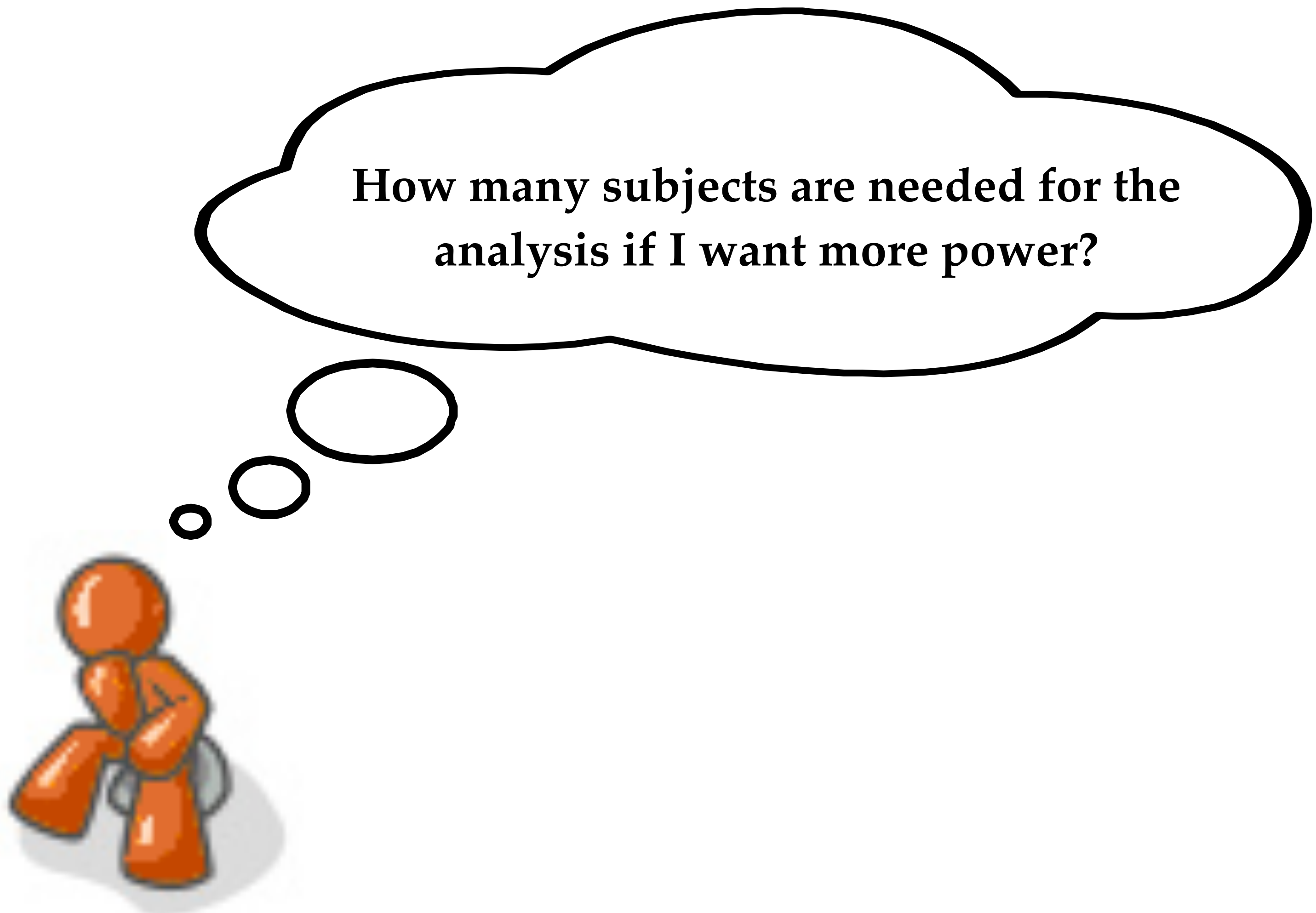




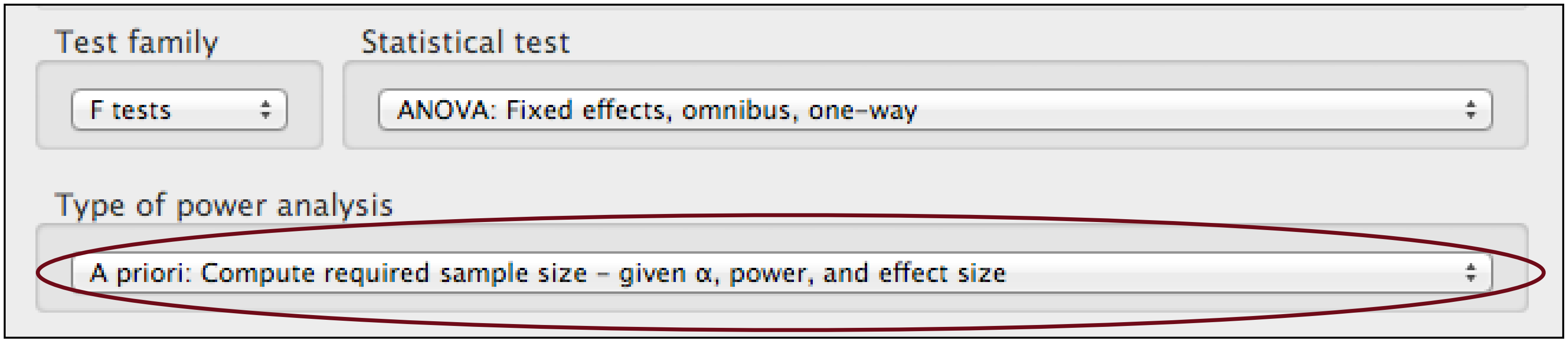
Output parameters

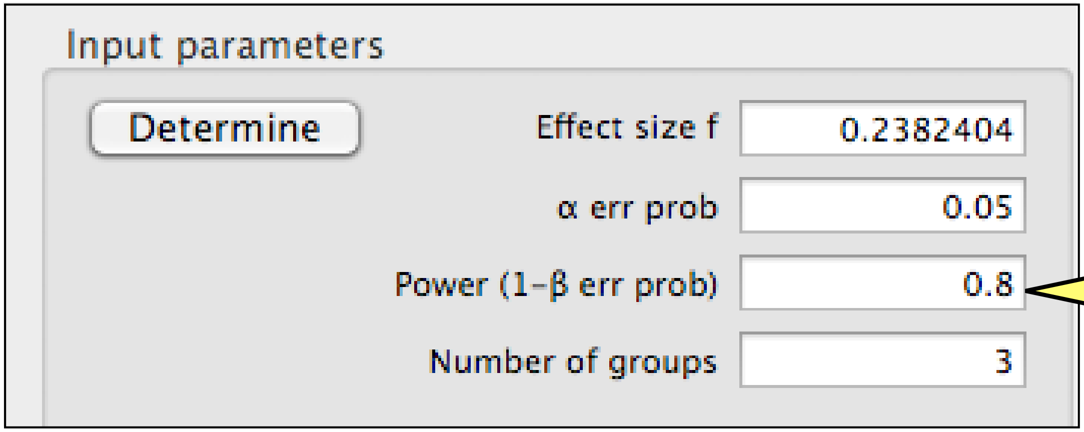
Noncentrality parameter λ	1.8730301
Critical F	3.3158295
Numerator df	2
Denominator df	30
Power (1–β err prob)	0.1967119




How many subjects are needed for the analysis if I want more power?

Compute **a priori** power (sample size)





Set power to 0.8

Output parameters

Noncentrality parameter λ	9.8759769
Critical F	3.0488326
Numerator df	2
Denominator df	171
Total sample size	174
Actual power	0.8030028

Sample size needed to carry out a fixed-effects, one-factor ANOVA to detect an effect of 0.2382

N = 174

Assumes a balanced design (equal group sizes) … n=58 subjects in each of the three groups

An *a priori* power analysis was carried out to determine the sample size needed to detect a main effect, $\eta^2 = 0.054$ (Cohen's $f = 0.238$), using a fixed-effects, one-factor ANOVA with power of 0.8. The required sample size for a factor with three levels assuming a balanced design, as computed by *G*Power 3* (Faul, Lang, & Buchner, 2007), was $N = 174$ ($n = 58$ per group).

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

```r
# To compute sample size using R
> library(pwr)
> pwr.anova.test(
     k=3,
     f=0.2382404,
     sig.level=0.5,
     power=0.8
     )
```

```
     Balanced one-way analysis of variance power calculation

              k = 3
              n = 57.59336
              f = 0.2382404
      sig.level = 0.05
          power = 0.8

NOTE: n is number in each group
```