

Sample Size Calculation

Athena Xiourouppa

The following document details the calculation for the sample size of Karl's experiment. We define the variables given:

```
R2 <- 0.1 # R^2 from previous experiment
power <- 0.9 # desired power
k <- 5 # number of regressors: concentration, cell age, treatment (two levels), cell type
alpha <- 0.05 # significance level
```

We also install the `pwr` package:

```
pacman::p_load(pwr)
```

To calculate sample size, n , using these variables, we use the following formula:

$$R^2 = \frac{kF}{n - k - 1 + kF} \quad (1)$$

where F is the inverse F-statistic with k and $n - k - 1$ degrees of freedom at significance level α . We define R^2 in terms of the effect size f_2 :

$$R^2 = \frac{f^2}{f^2 + 1} \implies f^2 = \frac{R^2}{1 - R^2}. \quad (2)$$

```
f2 = R2/(1-R2)
```

To solve for the parameter $n - k - 1$, and hence n , we use the `f2 test` in the `pwr` package.

```
pftest <- pwr.f2.test(u = k, v = NULL, f2 = f2, sig.level = 0.05, power = power)
pftest
```

Multiple regression power calculation

```
u = 5
v = 147.8645
f2 = 0.1111111
sig.level = 0.05
power = 0.9
```

```
n <- ceiling(pfctest$v + k + 1)
n
```

```
[1] 154
```

Therefore, we need $n = 154$ samples.