# Power and Sample Size Analysis

Clay Ford

Fall 2017

# Topics

- Intro to power and sample size concepts
- Calculate power and sample size for various statistical tests using the `pwr` package in `R` and a few built-in `R` functions

## Hello, my name is...

- I suspect people place sticky-back name tags on the left side of their chest about 75% of the time (probably because most people are right handed). I create an experiment to verify this.
- I randomly sample $n$ people and determine the proportion $p$ of people who place a name tag on the left.
- I conduct a one-sample proportion test to see if the sample proportion is significantly greater than random chance (0.50).
- I will reject the null hypothesis of random chance if the p-value is below 0.05.
- *How many people should I sample?*
- Or, *I can only sample 30 people. Do I have sufficient power?*

# Determining sample size and power

A sufficient sample size for a statistical test is determined from:

1. Power
2. Effect size
3. Significance level
4. Alternative direction (*only for certain tests*)

Determining the power of a statistical test is determined from:

1. Sample Size
2. Effect size
3. Significance level
4. Alternative direction (*only for certain tests*)

# What is power?

- Power is the probability a statistical test will correctly detect a hypothesized effect (if it really exists).
- In a hypothesis test, we assume two possible realities:

1. Null Hypothesis: No effect (eg, random chance, 0.50)
2. Alternative Hypothesis: *some* effect (eg, 0.75)

- At the conclusion, we decide whether to reject or fail to reject #1, usually based on a *p-value* falling below a threshold such as 0.05.
- We would like to have a high probability (or high power) of rejecting #1 if #2 is true. The usual desired power is at least 0.80.

# What is effect size?

- One definition is "the degree to which the null hypothesis is false."
- Estimating 90% of people place name tags on the left is much larger than estimating 55% put name tags on the left.
- In the former scenario (90%), I don't have to sample that many people to confirm my suspicion. In the latter scenario (55%), I probably need to sample quite a few people to get a proportion that is significantly different from random chance (50%)
- In sample size and power analyses, we have to pick an effect size. We usually pick the *smallest effect we don't want to miss*.

# What is significance level?

- This is the cut-off for determining whether or not our p-value is significant.
- Typical values are 0.05, 0.01 and 0.001.
- Recall, a p-value is the probability under the null hypothesis that a statistical summary of data would be equal to or more extreme than its observed value.

# What is Alternative Direction?

- This refers to how we think our alternative hypothesis differs from the null.
- If I think the left preference is greater than (or less than) 50%, then I have a *one-sided* alternative.
- If I think the left preference is simply different from 50%, then I have a *two-sided* alternative.
- The one-sided alternative is a stronger assumption. Most power and sample size analyses will play it safe by assuming a two-sided alternative.
- Some hypothesis tests only have a two-sided alternative, such as ANOVA and chi-square.

# Type I and Type II errors

- If I conclude people prefer left when they actually don't I have made a *Type I error*. (Rejecting the null hypothesis in error)
- If I conclude people have no preference when they really do prefer left I have made a *Type II error*. (Failing to reject null hypothesis in error)
- We usually never know if we have made these errors.
- Our tolerance for a Type I error is the significance level. Usually 0.05, 0.01 or lower.
- Our tolerance for a Type II error is 1 - Power. Usually 0.20 or lower.

# Visualizing a one-sample proportion test

The following web app allows us to see how power is affected by sample size, effect size and significance level:

`https://clayford.shinyapps.io/power_nhst/`

Let's take a look.

# Calculating power and sample size

Power and sample size formulas have been derived for many statistical tests that allows us to...

- calculate **sample size** given power, effect size and significance level
- calculate **power** given sample size, effect size and significance level

The parameters in the formulas are related such that one is determined given the others.

# The `pwr` package

Today we'll use three base `R` functions and the `pwr` package.

```
install.packages("pwr")
library(pwr)
```

The `pwr` package implements power and sample size analyses as described in *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*, Cohen (1988).

One of the tricks to using the `pwr` package is understanding how it defines *effect size*.

# Effect size in the `pwr` package

- Cohen defines "effect size" as "the degree to which the null hypothesis is false."
- Example: If our null is 50%, and the alternative 75%, the effect size is 25%.
- But the functions in the `pwr` package require the effect size to be metric-free (unitless).
- **This means you need to calculate effect size before using `pwr` functions. Entering the wrong effect size leads to incorrect power and sample size estimates!**
- Fortunately the `pwr` package provides a few functions for this.

# The `pwr` functions and associated statistical tests (1)

- `pwr.p.test`: one-sample test for proportions (ES=h)
- `pwr.2p.test`: two-sample test for proportions (ES=h)
- `pwr.2p2n.test`: two-sample test for proportions, unequal sample sizes (ES=h)
- `pwr.t.test`: one-sample and two-sample t-tests for means (ES=d)
- `pwr.t2n.test`: two-sample t-test for means, unequal sample sizes (ES=d)

Notice the effect sizes: h and d. We'll define these shortly.

# The pwr functions and associated statistical tests (2)

- `pwr.chisq.test`: chi-squared tests; goodness of fit and association (ES=w)
- `pwr.r.test`: correlation test (ES=r)
- `pwr.anova.test`: test for one-way balanced anova (ES=f)
- `pwr.f2.test`: test for the general linear model (multiple regression) (ES=f2)

Notice the effect sizes: w, r, f and f2. We'll define these shortly.

# The `ES` functions

Functions to compute effect size:

- `ES.h`: compute effect size h for proportion tests
- `ES.w1`: compute effect size w1 for chi-squared test for goodness of fit
- `ES.w2`: compute effect size w2 for chi-squared test for association
- `cohen.ES`: return conventional effect size (small, medium, large) for all tests available in `pwr`

We will use these functions as needed in the examples that follow.

Other effect sizes (d, r, f, and f2) must be calculated by hand.

# Conventional effect size

- Sometimes we don't know the precise effect size we expect or hope to find. In this case we can resort to conventional effect sizes of "small", "medium", or "large".
- The `cohen.ES` function returns these for us according to the statistical test of interest.
- For example, a "medium" effect size for a proportion test: `cohen.ES(test="p", size="medium")`
- This returns 0.5.

# Base `R` power and sample size functions

Base `R` includes three functions for calculating power and sample size:

- `power.prop.test`: two-sample test for proportions
- `power.t.test`: one-sample and two-sample `t` tests for means
- `power.anova.test`: one-way analysis of variance tests

These functions **do not** require calculating a unitless effect size and assume equal sample sizes across groups.

# Leave one out

- The `pwr` functions and base `R` functions have `n` and `power` arguments.
- To calculate `power`, you **leave it out** of the function.
- To calculate sample size (`n`), you **leave it out** of the function.
- For example, to calculate the sample size needed for a one-sample proportion test to have 80% power assuming a "small" effect size of $h = 0.2$, significance level of 0.05 and a one-sided "greater" alternative:

```
pwr.p.test(h=0.2, power = 0.8, sig.level=0.05,
           alternative = "greater")
```

- This returns a sample size of 155. If there really is a "small" effect in the population, a sample size of $n = 155$ gives us a 80% chance of rejecting the null of no effect.

# Let's get started!

- We'll go through each function available to us in the `pwr` package and base `R`.
- We'll go to `R` and demonstrate how to use it.
- I'll give you a quick opportunity to practice.
- As we'll see, understanding power and sample size analyses requires understanding the statistical test we're using.

# one-sample test for proportions

- Test if a proportion is equal to some hypothesized value versus a null value, such as random chance, or 0.5.
- `pwr.p.test`
- Requires effect size `h`, which is the arcsine transformation. Use `ES.h` function.
- Why h? Observe 0.65 - 0.50 and 0.16 - 0.01 both equal 0.15. But 0.16 is 16 times larger than 0.01, while 0.65 is only 1.3 times larger than 0.50. The arcsine transformation basically captures these differences.
- Conventional effect sizes: 0.2 (small), 0.5 (medium) and 0.8 (large)
- Remember, effect size $h$ is not a proportion. It ranges in practical value from about 0.02 to 3.

# one-sample test for proportions - example

We think people place name tags on the left side of their chest 75% percent of the time versus random chance (50%). What sample size do we need to show this assuming a significance level (Type I error) of 0.05 and a desired power of 0.80?

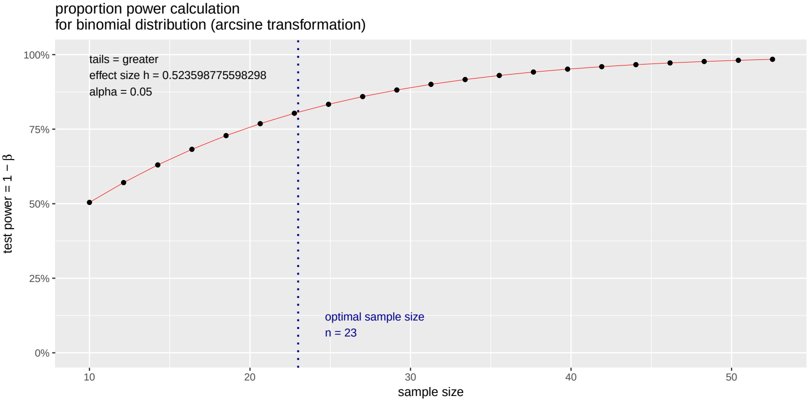
# one-sample test for proportions - code

```r
library(pwr) # do this once per session
h <- ES.h(p1 = 0.75, p2 = 0.50)
pwr.p.test(h = h, sig.level = 0.05, power = 0.80,
           alternative = "greater")
```

```
##
##      proportion power calculation for binomial distribut
##
##               h = 0.5235988
##               n = 22.55126
##       sig.level = 0.05
##           power = 0.8
##     alternative = greater
```

# one-sample test for proportions - plot

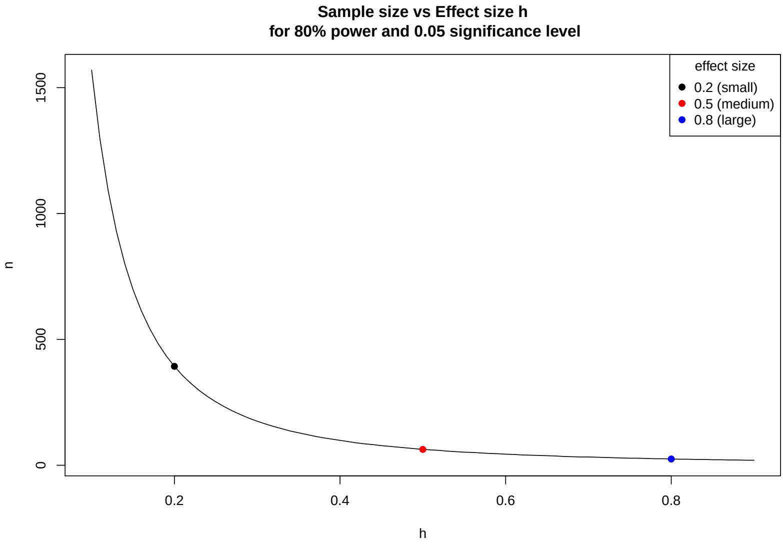
```
plot(pwr.p.test(h = h, sig.level = 0.05, power = 0.80,
           alternative = "greater"))
```



# one-sample test for proportions - results

- Always round up `n`. In our example, that gives us 23.
- Notice the argument `alternative = "greater"`. That was because we hypothesized greater than random chance (75% $>$ 50%)
- A safer and more common approach is to accept the default alternative: `alternative = "two.sided"`
- The `two.sided` alternative says we're not sure which direction the effect is in. It results in a larger sample size.
- For the remainder of the workshop we'll almost always use the default `alternative = "two.sided"`

# How effect size affects sample size




Let's go to R!

# two-sample test for proportions

- Test if two proportions are equal. The Null is no difference.
- `pwr.2p.test` or `power.prop.test`
- `pwr.2p.test` requires effect size `h`. Use `ES.h` function. (Effect size depends on the two proportions we compare.)
- `power.prop.test` allows you to use the raw proportions in the function.
- Both return sample size *per group*.

# two-sample test for proportions - example

We want to randomly sample male and female UVa undergrad students and ask them if they consume alcohol at least once a week. Our null hypothesis is no difference in the proportion that answer yes. Our alternative hypothesis is that there is a difference. (two-sided; one gender has higher proportion, I don't know which.) I'd like to detect a difference as small as 5%. How many students do I need to sample in each group if we want 80% power and a significance level of 0.05?

# two-sample test for proportions - code

These return different sample sizes!

```
# 55% vs. 50%
pwr.2p.test(h = ES.h(p1 = 0.55, p2 = 0.50),
            sig.level = 0.05, power = .80)
# 35% vs. 30%
pwr.2p.test(h = ES.h(p1 = 0.35, p2 = 0.30),
            sig.level = 0.05, power = .80)
# 15% vs. 10%
pwr.2p.test(h = ES.h(p1 = 0.15, p2 = 0.10),
            sig.level = 0.05, power = .80)
```

# two-sample test for proportions - code

The base R function is perhaps a little easier to use:

```
power.prop.test(p1 = 0.55, p2 = 0.50,
                sig.level = 0.05, power = .80)
power.prop.test(p1 = 0.35, p2 = 0.30,
                sig.level = 0.05, power = .80)
power.prop.test(p1 = 0.15, p2 = 0.10,
                sig.level = 0.05, power = .80)
```

# two-sample test for proportions - conventional effect size

- We may just want to use a conventional effect size if we're not comfortable specifying proportions
- Again, those are 0.2, 0.5, and 0.8
- Example

```
pwr.2p.test(h = 0.2, sig.level = 0.05, power = 0.8)
```

- We can only use conventional effect sizes with `pwr` functions

Let's go to R!

# two-sample test for proportions, unequal sample sizes

- Test if two proportions are equal with unequal sample sizes. The Null is no difference.
- `pwr.2p2n.test`
- Requires effect size `h`. Use `ES.h` function.
- It has two `n` arguments: `n1` and `n2`. Can be used to find a sample size for one group when we already know the size of the other.

# two-sample test for proportions, unequal sample sizes - example

Let's return to our undergraduate survey of alcohol consumption. It turns out we were able to survey 543 males and 675 females. What's the power of our test with a significance level of 0.05? Let's say we're interested in being able to detect a "small" effect size (0.2).

# two-sample test for proportions, unequal sample sizes - code

```
pwr.2p2n.test(h = 0.2,
              n1 = 543, n2 = 675,
              sig.level = 0.05)
```

Let's go to R!

# one-sample, two-sample and paired t tests for means

- Test if a mean is equal to specific value (one-sample), test if means of two different groups are equal (two-sample), or test if "paired" means are equal
- `pwr.t.test` requires effect size `d`. `d` is the difference in population means divided by the standard deviation of either population (since they are assumed equal). Now we have to make a guess at the standard deviation.
- There is no function for effect size `d`. We have to calculate this ourselves if we wish to use `pwr.t.test`.
- Conventional effect sizes: 0.2, 0.5 and 0.8
- Specify type of test with `type` argument: `"two.sample"`, `"one.sample"`, `"paired"`

# one-sample, two-sample and paired `t` tests for means

- The base R function `power.t.test` calculates effect size automatically given `delta` and `sd` arguments.
- `delta` is difference in means; `sd` is standard deviation
- Specify type of test with `type` argument: `"two.sample"`, `"one.sample"`, `"paired"`

# two-sample t test - example 1

I'm interested to know if there is a difference in the mean price of what male and female students pay at the library coffee shop. Let's say I randomly observe 30 male and 30 female students check out from the coffee shop and note their total purchase price. How powerful is this experiment if I want to detect a "medium" effect in either direction with a 0.05 significance level?

# two-sample t test - code

```
pwr.t.test(n = 30, d = 0.5, sig.level = 0.05)
```

```
##
##      Two-sample t test power calculation
##
##               n = 30
##               d = 0.5
##       sig.level = 0.05
##           power = 0.4778965
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

# two-sample t test - example 2

- Let's say we want to be able to detect a difference of at least 75 cents in the mean purchase price. How can we convert that to an effect size?
- We need to make a guess at the population standard deviation. If we have absolutely no idea, one rule of thumb: take the difference between the maximum and minimum values and divide by 4 (or 6).
- Let's say max is $10 and min is $1. So our guess at a standard deviation is $(10 - 1)/4 = 2.25$.
- $d = 0.75/2.25 \approx 0.333$

# two-sample t test - code

```r
# requires d
pwr.t.test(d = 0.333, power = 0.80, sig.level = 0.05)

# does not require d
power.t.test(delta = 0.75, sd = 2.25,
             power = 0.80, sig.level = 0.05, )
```

# one-sample and paired t test

- To calculate power and sample size for one-sample t test, set the `type` argument to `"one.sample"`
- A paired t-test is basically the same as a one-sample t test. Instead of one sample of individual observations, you have one sample of `pairs` of observations, where you take the difference between each pair to get a single sample of differences. These are commonly before and after measures on the same person.
- To calculate power and sample size for paired t test, set the `type` argument to `"paired"`

# one-sample t test - example

I think the average purchase price at the Library coffee shop is over \$3 per student. My null is \$3 or less; my alternative is greater than \$3. If the true average purchase price is \$3.50, I would like to have 90% power to declare my estimated average purchase price is greater than \$3. How many transactions do I need to observe assuming a significance level of 0.05?

Let's say max purchase price is \$10 and min is \$1. So our guess at a standard deviation is $9/4 = 2.25$.

# one-sample t test - code

```r
d <- 0.50/2.25
pwr.t.test(d = d, sig.level = 0.05, power = 0.90,
           alternative = "greater",
           type = "one.sample")

# or with power.t.test:
power.t.test(delta = 0.50, sd = 2.25, power = 0.90,
             sig.level = 0.05,
             alternative = "one.sided",
             type = "one.sample")
```

Let's go to R!

## two-sample t test for means, unequal sample sizes

- Test if means from different groups are equal with unequal sample sizes. The Null is no difference.
- `pwr.t2n.test`
- Requires effect size `d`.
- It has two `n` arguments: `n1` and `n2`. Can be used to find a sample size for one group when we already know the size of the other.

# two-sample t test for means, unequal sample sizes - example

Let's say we have data on 35 male customers and estimated a mean purchase price. How many females do we need to sample to detect a medium gender effect of 0.5 with a desired power of 0.80 and a significance level is 0.05?

# two-sample t test for means, unequal sample sizes - code

```r
pwr.t2n.test(n2 = 35, d = 0.5, power = 0.8)
```

```
##
##      t test power calculation
##
##              n1 = 321.7572
##              n2 = 35
##               d = 0.5
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
```

Let's go to R!

# chi-squared tests

Two kinds of chi-squared tests:

1. goodness of fit test
2. test for association

- `pwr.chisq.test`
- Uses effect size `w`, which differs depending on the test.
- Use `ES.w1` for goodness of fit and `ES.w2` for test for association
- Also requires degrees of freedom: `df`
- conventional effect sizes: 0.1, 0.3, 0.5

# chi-squared tests - goodness of fit

- A single dimension of proportions is tested against a prespecified set of proportions which constitutes the null hypothesis.
- Example: $H_0 : \frac{1}{3}, \frac{1}{3}, \frac{1}{3}$ vs $H_a : \frac{1}{2}, \frac{1}{4}, \frac{1}{4}$
- Rejecting the null means we have sufficient evidence to conclude the data don't appear to "fit" the prespecified set of proportions.
- If we were hoping to show our data "fit" the prespecified set of proportions, then failure to reject the Null is a good thing.
- `df` = number of categories - 1

# chi-squared tests - test of association

- A table of counts classified by two variables is tested against the expected table of counts given the two variables are independent.
- Rejecting the null means the data appear to be associated in some way.
- df $=$ (Var1 number of categories - 1) $\times$ (Var2 number of categories - 1)
- This test doesn't tell you anything about the strength or direction of association.

# chi-square goodness of fit test - example

A market researcher is seeking to determine preference among 4 package designs. He arranges to have a panel of 100 consumers rate their favorite package design. He wants to perform a chi-square goodness of fit test against the null of equal preference (25% for each design) with a significance level of 0.05. What's the power of the test if 3/8 of the population actually prefers one of the designs and the remaining 5/8 are split over the other 3 designs? *(From Cohen, example 7.1)*

# chi-square goodness of fit test - code

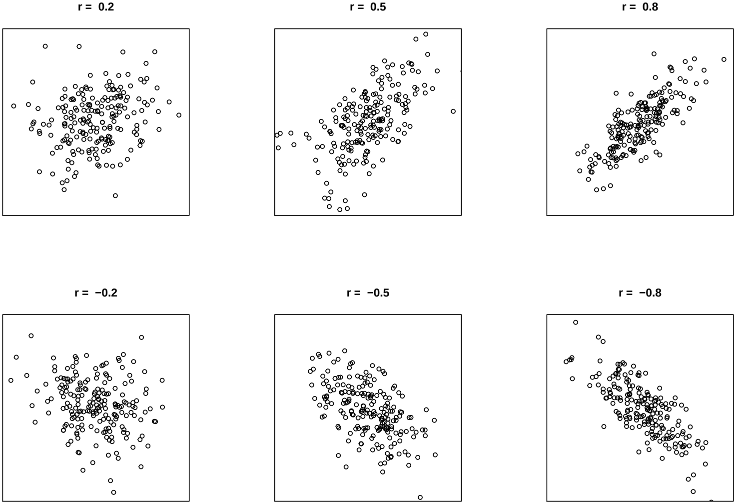
```r
# To calculate effect size, we need to create vectors
# of null and alternative proportions:
null <- rep(0.25, 4)
alt <- c(3/8, rep((5/8)/3, 3))
pwr.chisq.test(w=ES.w1(P0 = null,P1 = alt),
               N=100, df=(4-1), sig.level=0.05)
```

Let's go to R!

# Correlation test

- Test whether there is any linear relationship between two continuous variables. Null is correlation coefficient $r = 0$.
- `pwr.r.test`
- Testing if correlation is 0 is the same as testing if the slope in simple linear regression is 0.
- Correlation is already unitless, so we don't require a formula to calculate effect size.
- Conventional effect sizes: 0.1, 0.3, 0.5

# Correlation review



# Correlation test - example

I'm a web developer and I want to conduct an experiment with one of my sites. I want to randomly select a group of people, ranging in age from 18 - 65, and time them how long it takes them to complete a task, say locate some piece of information. I suspect there may be a "small" positive linear relationship between time it takes to complete the task and age. How many subjects do I need to detect this positive (ie, $r > 0$) relationship with 80% power and the usual 0.05 significance level?

# Correlation test - code

```
pwr.r.test(r = 0.1, sig.level = 0.05, power = 0.8,
           alternative = "greater")
```

```
##
##      approximate correlation power calculation (arctangh
##
##               n = 616.1032
##               r = 0.1
##       sig.level = 0.05
##           power = 0.8
##     alternative = greater
```

Let's go to R!

# balanced one-way analysis of variance test

- ANOVA, or Analysis of Variance, tests whether or not means differ between more than 2 groups.
- "One-way" means one explanatory variable.
- "Balanced" means we have equal sample size in each group.
- The null hypothesis is that the means are all equal.
- `pwr.anova.test` or `power.anova.test`
- The `power.anova.test` function that comes with base R is easier to use than `pwr.anova.test` and does not require calculating an effect size.

# balanced one-way analysis of variance test

- The `power.anova.test` function requires you to specify the number of `groups`, the between group variance (`between.var`), and the within group variance (`within.var`), which we assume is the same for all groups.
- The `pwr.anova.test` function requires you to provide an effect size, `f`.
- The effect size, `f`, for k groups is calculated as $SD_{means}$ / $SD_{populations}$ (Translation: standard deviation of the k means divided by the common standard deviation of the populations involved.)
- conventional effect sizes: 0.1, 0.25, 0.4

# balanced one-way analysis of variance test - example

I'm a web developer and I'm interested in 3 web site designs for a client. I'd like to know which design(s) help users find information fastest, or which design requires the most time. I design an experiment where I have 3 groups of randomly selected people use one of the designs to find some piece of information and I record how long it takes. (All groups look for the same information.) How many people do I need in each group if I believe two of the designs will take 30 seconds and one will take 25 seconds? Assume population standard deviation is 5 and that I desire power and significance levels of 0.8 and 0.05.

# balanced one-way analysis of variance test - code

```r
# The between group variance: var(c(30, 30, 25)) = 8.3
# The within group variance: 5^2
power.anova.test(groups = 3, between.var = 8.3,
                 within.var = 5^2, power = 0.8)
```

Let's go to R!

# test for the general linear model

- By "general linear model" we mean multiple regression.
- Test that the proportion of variance explained by the model predictors is 0. Equivalently, test whether all the model coefficients (except the intercept) are 0.
- `pwr.f2.test`
- This is a little tricky to use because not only do we have to supply an "effect size" (`f2`), we also have to supply numerator (`u`) and denominator (`v`) degrees of freedom instead of sample size.
- numerator (`u`) and denominator (`v`) degrees of freedom refer to the F test that tests whether all the model coefficients (except the intercept) are 0.
- conventional effect sizes: 0.02, 0.15, 0.35
- There is currently no built-in plot method.

# test for the general linear model - effect size

- The `f2` effect size is $R^2/(1 - R^2)$, where $R^2$ is the coefficient of determination, aka the "proportion of variance explained".
- To determine effect size you hypothesize the proportion of variance your model explains, or the $R^2$. For example, 0.45. This leads to an effect size of $0.45/(1 - 0.45) \approx 0.81$
- We can reverse this. Given an effect size, we can determine $R^2$: $ES/(1 + ES)$. For example, $0.81/(1 + 0.81) \approx 0.45$
- There is no function for this.

# test for the general linear model - degrees of freedom

- The numerator degrees of freedom, `u` is the number of coefficients you'll have in your model (minus the intercept).
- The denominator degrees of freedom `v` is the number of error degrees of freedom. `v` = `n` - `u` - 1.
- if we want to determine sample size for a given power and effect size, we have to find `v`, which we then use to solve `n` = `v` + `u` + 1. (!)
- There is no `n` argument!

# test for the general linear model - example

I'm hired to survey a company's workforce about job satisfaction. I ask employees to rate their satisfaction on a scale from 1 (hating life) to 10 (loving life). I know there will be variability in the answers, but I think two variables that will explain this variability are salary and age. In fact I think it will explain at least 30% ($R^2 = .30$) of the variance. How powerful is my "experiment" if I randomly recruit 40 employees and accept a 0.05 significance level?

## test for the general linear model - code

```r
# Two predictors, so u = 2
# 40 subjects, so v = 40 - 2 - 1 = 37
# R^2 = .30, so effect size f2 = 0.3/(1 - 0.3)

pwr.f2.test(u = 2, v = 37, f2 = 0.3/(1 - 0.3),
            sig.level = 0.05)
```

Let's go to R!

# Other Software

## Software

- PASS. `http://www.ncss.com/software/pass/` ($395/year or $795 perpetual)
- nQuery. `http://www.statsols.com/products/nquery-advisor-nterim/` ($440/year)
- PROC POWER in SAS. Power and sample size analyses for a variety of statistical analyses
- G*Power. `http://www.gpower.hhu.de/en.html` (Free)

## R packages

- `TrialSize`. Functions and examples from the book *Sample Size Calculation in Clinical Research*
- `samplesize`. Computes sample size for Student's t-test and for the Wilcoxon-Mann-Whitney test for categorical data
- `clinfun`. Functions for both design and analysis of clinical trials.

# References


Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. LEA.

Dalgaard, P. (2002). *Introductory Statistics with R*. Springer. (Ch. 2)

Hogg, R and Tanis, E. (2006). *Probability and Statistical Inference (7th ed.)*. Pearson. (Ch. 9)

Kabacoff, R. (2011). *R in Action*. Manning. (Ch. 10)

Ryan, T. (2013). *Sample Size Determination and Power*. Wiley.

# Thanks for coming today!

For help and advice with your statistical analysis:
`statlab@virginia.edu`

Sign up for more workshops or see past workshops:
`http://data.library.virginia.edu/statlab/`

Register for the Research Data Services newsletter to stay up-to-date on RDS events and resources:
`http://data.library.virginia.edu/newsletters/`