Power Calculations with R

BIOS 6611

CU Anschutz

Week 4

Background

In many cases do not have a closed form of an equation to use for power, sample size, etc. For example, if we are comparing the means of two samples and do not assume the standard deviation is known, we do not have the nice closed form solutions based on the standard normal distribution. Instead, we have to use an iterative process based on the t-distribution.

Fortunately, many packages in R exist that facilitate power calculations!

In fact, most statistical software (e.g., SAS) will include the ability to do power calculations. There are also specialized software programs that almost exclusively conduct power calculations (e.g., PASS, PS).

Our Examples

In this lecture set we will use the base `stats` package and also introduce the `pwr` package. We will see examples for:

- two proportions (`stats` package function `power.prop.test()`)
- one mean with unknown SD (`stats` package function `power.t.test()`)
- two means with unknown SD (`stats` package function `power.t.test()`)
- two means with unequal sample sizes and unknown SD (`pwr` package function `pwr.t2n.test()`)

Comparing Two Proportions

The `power.prop.test` function

```
power.prop.test(n = NULL, p1 = NULL, p2 = NULL, sig.level = 0.05,
                power = NULL,
                alternative = c("two.sided", "one.sided"),
                strict = FALSE, tol = .Machine$double.eps^0.25)
```

From the "Details" of the documentation:

Exactly one of the parameters `n`, `p1`, `p2`, `power`, and `sig.level` must be passed as `NULL`, and that parameter is determined from the others.

`p1` and `p2` represent the probability in the two groups that we expect for a given scenario. If you wish to solve for `p2` it will search for a possible solution, if it exists, between `p1` and 1. Likewise, solving for `p1` searches from 0 to `p2`.

Proportion Example for Power

Prof. Vacca Seene wants to conduct a study that compares the proportion of participants in a vaccine trial who ultimately contract the disease. They ask what their power is if the expect $p_1 = 0.1$, $p_2 = 0.2$, $\alpha = 0.05$, with $n = 250$ per group?

```
power.prop.test(n=250, p1=0.1, p2=0.2, sig.level=0.05)
```

```
##
##      Two-sample comparison of proportions power calculation
##
##               n = 250
##              p1 = 0.1
##              p2 = 0.2
##       sig.level = 0.05
##           power = 0.881558
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

Proportion Example for Detectable Difference

Dr. Kiebl R. Elfe is conducting a study where individuals are randomized to either receive an oatmeal chocolate chip cookie or an oatmeal raisin cookie that looks like chocolate chips. They are measuring if a participant reports "surprise" at the cookie they ate. They expect only 5% of chocolate chip participants to be surprised and want to know the minimum level of surprise they would detect if they desire 80% power, a 10% type I error rate, and randomize 22 individuals to each group.

```
power.prop.test(n=22, p1=0.05, p2=NULL, sig.level=0.10, power=0.8)
```

```
##
##      Two-sample comparison of proportions power calculation
##
##               n = 22
##              p1 = 0.05
##              p2 = 0.3400777
##       sig.level = 0.1
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

Comparing One Mean with Unknown SD

The `power.t.test` function

```
power.t.test(n = NULL, delta = NULL, sd = 1, sig.level = 0.05,
             power = NULL,
             type = c("two.sample", "one.sample", "paired"),
             alternative = c("two.sided", "one.sided"),
             strict = FALSE, tol = .Machine$double.eps^0.25)
```

From the "Details" of the documentation: Exactly one of the parameters `n`, `delta`, `power`, `sd`, and `sig.level` must be passed as `NULL`, and that parameter is determined from the others.

The `type` can be specified as either a two-sample, one-sample, or paired t-test.

One-Sample t-test Example for Sample Size

Marathoner Erin Runferryfar has an average marathon time of 5 hours with an estimated standard deviation of 30 minutes. They are trying a new training plan and want to know how many marathons they need to run to see if it improves their average marathon time by 15 minutes with $\alpha = 0.1$ and 80% power.

```
power.t.test(n=NULL, delta=15, sd=30, sig.level=0.1, power=0.8,
             type='one.sample', alternative='one.sided')
```

```
##
##      One-sample t test power calculation
##
##               n = 18.90542
##           delta = 15
##              sd = 30
##       sig.level = 0.1
##           power = 0.8
##     alternative = one.sided
```

Comparing Two Means with Unknown SD

Two-Sample t-test Example for Power

Dr. Hart B. Urn, a gastroenterologist, wants to compare if patients quality of life differs for those who practice breathing exercises versus taking a PPI based on a 100 point validated scale. They expect a difference between the groups of 10 ($s = 20$) and can enroll up to $n = 20$ per group. What is their power if they desire an $\alpha = 0.01$?

```
power.t.test(n=20, delta=10, sd=20, sig.level=0.01, power=NULL,
             type='two.sample', alternative='two.sided')
```

```
##
##      Two-sample t test power calculation
##
##               n = 20
##           delta = 10
##              sd = 20
##       sig.level = 0.01
##           power = 0.1439336
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

Two-Sample t-test Example for SD

Dr. Hart B. Urn's heart sank when they heard the estimated power. However, they know of other validated instruments that may have less variability but still use a 100 point scale. If the difference is 10, $\alpha = 0.01$, $n = 20$ per group, and they desire 80% power, what is the largest SD?

```
power.t.test(n=20, delta=10, sd=NULL, sig.level=0.01, power=0.8)
```

```
##
##      Two-sample t test power calculation
##
##               n = 20
##           delta = 10
##              sd = 8.846129
##       sig.level = 0.01
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

Two-Sample t-test Example for Detectable Difference

Dr. Hap E. Gauss is an anesthesiologist who wishes to compare the length of time needed for the anesthesiologist versus the surgeon for a given procedure over the course of a week. If $n = 21$ procedures, $s = 16$ minutes, $\alpha = 0.05$, $1 - \beta = 0.8$, what is the difference they could detect if it exists?

```
power.t.test(n=21, delta=NULL, sd=16, sig.level=0.05, power=0.8)
```

```
##
##      Two-sample t test power calculation
##
##               n = 21
##           delta = 14.17681
##              sd = 16
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

Comparing Two Means with Unknown SD and Unequal n

The `pwr.t2n.test` function

```
library(pwr)
pwr.t2n.test(n1 = NULL, n2= NULL, d = NULL, sig.level = 0.05,
             power = NULL, alternative = c("two.sided","less","greater"))
```

From the "Details" of the documentation: Exactly one of the parameters `d`, `n1`, `n2`, `power`, and `sig.level` must be passed as `NULL`, and that parameter is determined from the others.

`d` here is **Cohen's d** and measures the **effect size**: $d = \frac{|\mu_1 - \mu_2|}{s}$. $d > 0$ with larger values indicating a greater difference.

Unequal n for Two-Sample t-test Example for Power

Dr. Kiebl R. Elfe noticed in his grocery store that raisins are on a different shelf than the chocolate chips. A previous study suggested 20% of the population prefers raisin cookies, but perhaps they have different heights!

He wishes to test if the heights of the two groups are different by measuring the heights of 50 men. His H_1 is a difference of three inches (s=3 inches) with $\alpha = 0.05$. What is his power to revolutionize cookie marketing?

```
library(pwr)
pwr.t2n.test(n1=10, n2=40, d=1, sig.level=0.05, power=NULL)
```

```
##
##      t test power calculation
##
##              n1 = 10
##              n2 = 40
##               d = 1
##       sig.level = 0.05
##           power = 0.7914513
##     alternative = two.sided
```

Other Considerations

The previous slides illustrated a set of examples for both proportions and means with unknown standard deviations.

Other packages and functions exist that address other statistical tests or assumptions (e.g., if we needed to assume the two groups had different SDs).

These calculations help us to identify the feasibility of answering a hypothesis. If the power is too low or sample size too high, we may be able to modify our assumptions and still identify meaningful scenarios.