Power and Sample Size (CRP 241) — Tutorial

CRP 241 Tutorial

Duke University Clinical Research Training Program

2025-11-05

Table of contents

Introduction

Audience: physicians in an intro statistics course. We use simple functions, short sentences, and wrap text to 80 characters. We transform design assumptions (effect size, alpha, power) into sample size, or solve for the missing quantity.

> **Learning Objectives**
>
> - Know the inputs for power/sample size problems
> - Use `pwr` for two-sample t-tests and two-proportion tests
> - Inflate enrollment to handle expected attrition

Data overview

- No external dataset is used. We work from assumed means, SD, and proportions drawn from published contexts.

Key questions

- How many participants per arm are needed for a planned effect?
- With fixed n, what effect size is detectable at chosen alpha and power?
- How do we adjust total enrollment for attrition?

0.1 1) Setup

Explanation

- Automatically check whether the `pwr` package is installed; install it only if missing, then load it. This keeps the tutorial runnable without manual edits.

```
if (!requireNamespace("pwr", quietly = TRUE)) install.packages("pwr")
library(pwr)
```

Interpretation

- `pwr` provides helpers like `pwr.t.test()` and `pwr.2p.test()` that solve for one missing quantity given the other three.

Clinical Interpretation

- These tools help plan feasible studies: estimate how many patients to enroll to detect a meaningful effect with a given false-positive rate.

0.2 2) Load data

Explanation

- No external files are needed. We compute from design inputs only.

0.3 3) Explore: key quantities

Explanation

- For power problems, you provide 3 of the 4: effect size, n, alpha, power.
- For two-proportion tests, effect size uses Cohen's h: h = 2*asin(sqrt(p1))* − *2*asin(sqrt(p2)).

Stat interpretation

- These formulas standardize differences so we can compare scenarios.

Clinical takeaway

- Define a minimal clinically important difference (MCID) up front; design choices should reflect what matters for patients.

0.4 4) Continuous outcome example (ORBITA)

Explanation

- Goal: detect a 30-second improvement in exercise time (SD = 75 s).
- Two-sample t-test, two-sided alpha = 0.05, power = 0.80.
- Cohen's d = 30 / 75 = 0.4.

```
pwr.t.test(sig.level = 0.05, power = 0.8, d = 0.4,
           type = "two.sample", alternative = "two.sided")
```

```

     Two-sample t test power calculation

              n = 99.08032
              d = 0.4
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

> **Interpretation**
>
> - Output suggests about 99 per group (200 total).

Clinical Interpretation

- Plan for ~100 patients per arm to detect a 30-second difference with 80% power at alpha 0.05.

0.4.1 Account for attrition

Explanation

- If 33% drop out, only 67% remain. Inflate total enrollment so the final analyzable sample stays ~200.

```
200 / 0.67
```

```
[1] 298.5075
```

Interpretation

- 200 / 0.67 298.5 → round up to 300 total.

Clinical Interpretation

- Enroll ~150 per arm to end with ~100 per arm after attrition.

0.5 5) Binary outcome example (Gulf War)

Explanation

- Two-sided test of proportions, alpha = 0.05, power = 0.95.
- Compare p1 = 0.30 vs p2 = 0.15. Use Cohen's h for effect size.

```
h <- 2*asin(sqrt(0.30)) - 2*asin(sqrt(0.15))
pwr.2p.test(h = h, power = 0.95, sig.level = 0.05)
```

```
     Difference of proportion power calculation for binomial distribution (arcsine transforma

              h = 0.3638807
              n = 196.2812
      sig.level = 0.05
```

```
          power = 0.95
    alternative = two.sided

NOTE: same sample sizes
```

> **Interpretation**
>
> - Required n 196 per group → round to 200 per group (400 total).

> **Clinical Interpretation**
>
> - Expect about 400 analyzable participants in total for this effect size.

0.5.1 Account for 10% attrition

Explanation

- Enroll more so that after 10% loss, ~400 remain.

```r
totalsubjects <- 400 / 0.90

# show total and per-group target (example rounding to 450 total)
totalsubjects
```

```
[1] 444.4444
```

```r
450 / 2
```

```
[1] 225
```

Stat interpretation

- Total 444.4 → round to 450; 225 per group.

Clinical takeaway

- Start with 225 per arm to end near 200 per arm.

0.6 6) Detectable effect with fixed n (continuous)

Explanation

- Two-sample t-test with alpha = 0.05, power = 0.90, and n = 200 per group.
- Solve for the minimal detectable Cohen's d.

```
pwr.t.test(power = 0.90, sig.level = 0.05, n = 200,
           type = "two.sample", alternative = "two.sided")
```

```

     Two-sample t test power calculation 

              n = 200
              d = 0.3249364
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

Stat interpretation

- Detectable standardized effect is d 0.32.

Clinical takeaway

- With 200 per arm, you can detect a small-to-moderate effect. Translate to patient units with delta = d * sigma.

0.7 7) Interpret and conclude

Summary (stat)

- Continuous outcome: ~100 per arm (200 total) for d = 0.4 at 80% power; inflate to ~300 total for 33% attrition.
- Binary outcome: ~200 per arm (400 total) at 95% power; inflate to ~450 total for 10% attrition.
- With n = 200 per arm, minimal detectable d 0.32.

Summary (clinical)

- Plan enrollment with expected drop-out to hit analyzable targets.
- Discuss MCID with clinicians so targets match patient value.

- Express results in patient-centered units (seconds, absolute risk change), not only standardized metrics.

0.8 Glossary

- p-value: Probability, if the null is true, of seeing results this extreme or more extreme.
- Confidence interval (CI): A range of values that likely includes the true effect; a 95% CI would capture the true value in 95% of repeated samples.
- Power: Probability of detecting an effect of a given size if it truly exists.
- Alpha (significance level): Tolerated false-positive rate (commonly 0.05).
- Cohen's d: Standardized mean difference for continuous outcomes.
- Cohen's h: Standardized difference between two proportions.