A Sample Size: More Than A Number

Karen Copeland, Ph.D.
Long Time JMP User
Boulder Statistics, LLC
JMP Statistical Writing Team

Copyright © JMP Statistical Discovery LLC. All rights reserved.

Agenda

- Introduction to sample size thinking
- Example 1: Sizing a study for comparing means of two populations
- Example 2: Sizing a study with a proportion endpoint
- Additional thoughts

Copyright © JMP Statistical Discovery LLC. All rights reserved.

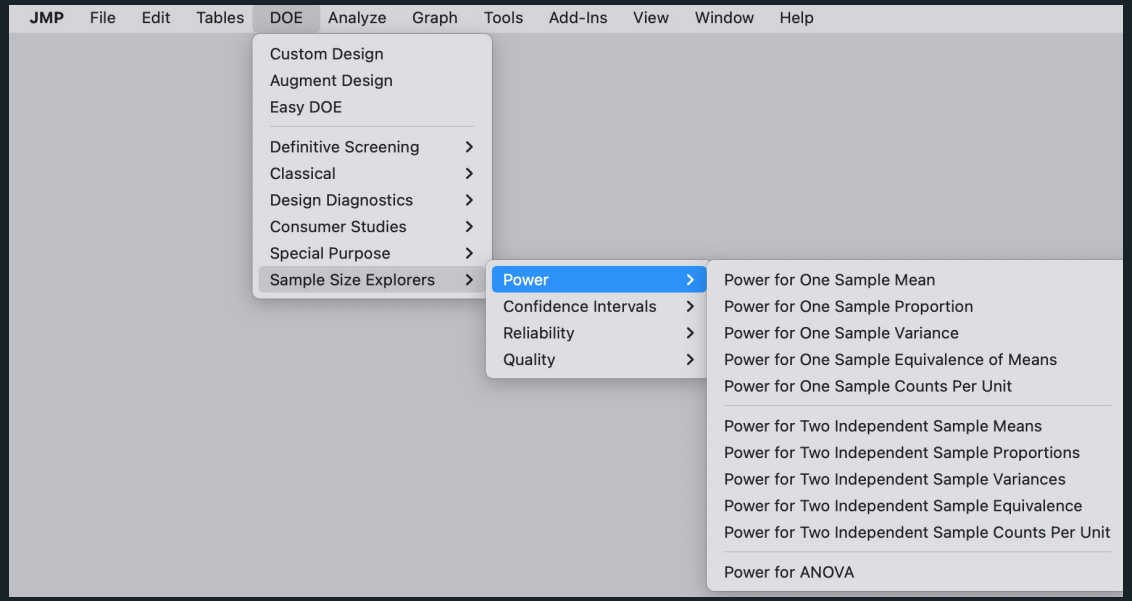
What sample size do I need?

Easy, use a sample size calculator.

Wait, why does JMP call sample size calculators explorers?

Why are they in the DOE menu?

Which one do I use?



Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Sample Size Basics

- Sample size is calculated prior to running a study
  - A study is an experiment, designed ahead of time, so DOE menu for sample size
- Sample size depends on:
  - Goal of the study (e.g., regulatory clearance, publication, R&D question)
  - Primary endpoint (what you are trying to show)
  - Study design
  - Outcome assumptions (can be based on prior knowledge; however, often we are guessing)
- Sample size is a risk/benefit exploration, so Sample Size Explorers
  - More is generally better
  - More costs more
  - More might not be possible

DOE > Sample Size Explorers

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 1: Sizing a study for comparing 2 means

- Fit Y by X Platform
- Power Explorer for Two Independent Sample Means

Copyright © JMP Statistical Discovery LLC. All rights reserved.

Example 1: Sizing a study for comparing 2 means

company is in R&D phase, study is a sample collection study, there is no primary endpoint…but…we need a power analysis

- How can we provide a power analysis without a primary endpoint?
- Generate a research endpoint
  - Can I distinguish the difference in means between my sick and healthy subjects for some primary biological markers?
- Use the sample size from the power analysis and expected prevalence of illness to justify the number of subjects we are requesting to enroll in the study

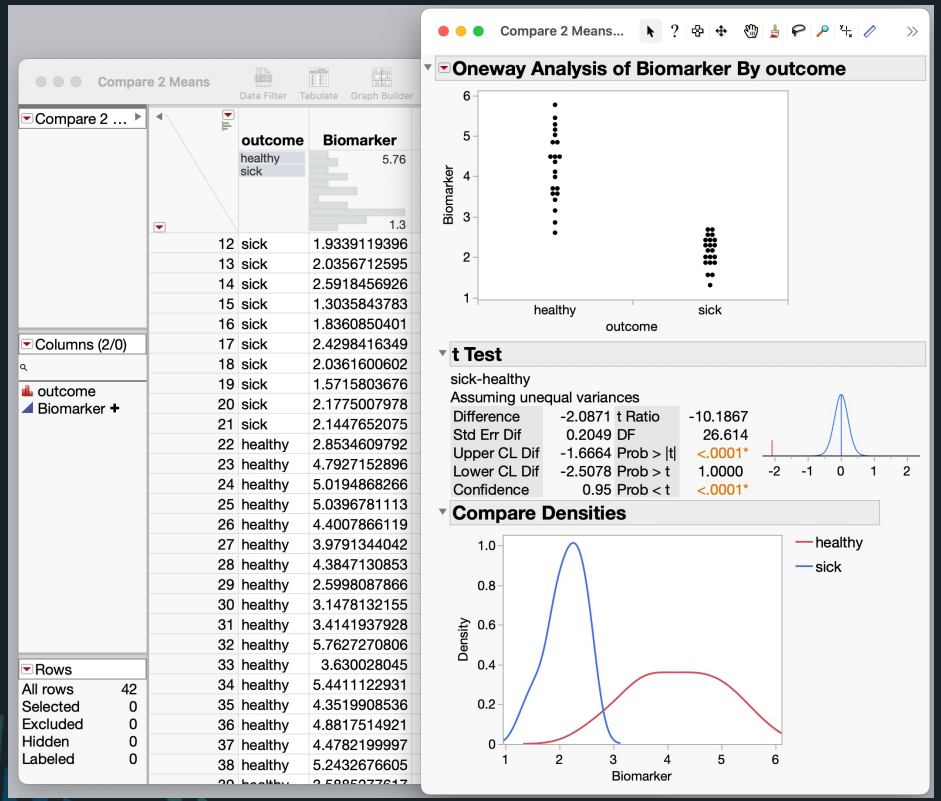
- I need to understand tests for comparing two independent means
- I need a calculator for the power of a test to compare two independent means

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 1: Sizing a study for comparing 2 means

If I had data….what would I do?



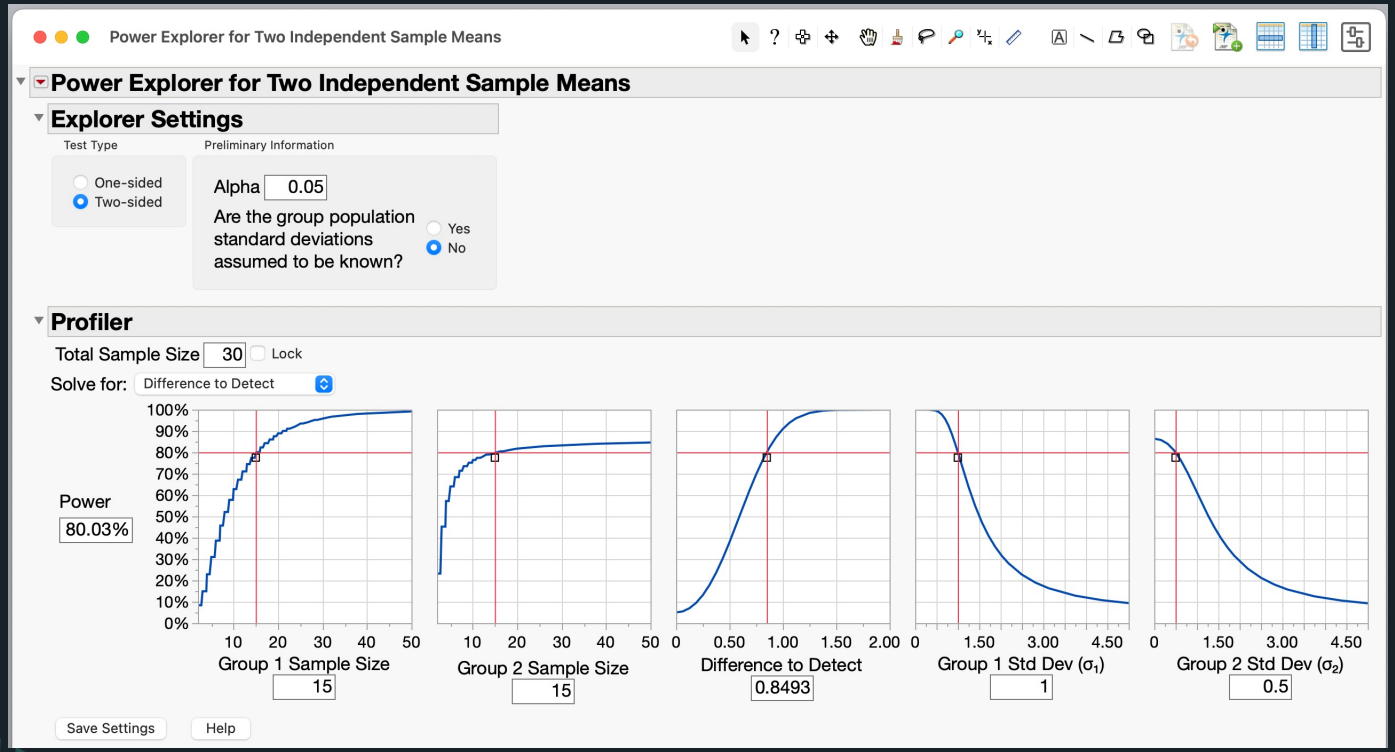
1. Fit Y by X (Oneway Analysis)
2. Red Triangle Menu
    1. t Test
    2. Compare Densities
    3. Display Options > Points Jittered

Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Example 1: Sizing a study for comparing 2 means

How much data should I collect?



Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 1: Sizing a study for comparing 2 means

How much data should I collect?

- To obtain 15 positive samples (from unique subjects), assuming a 10% prevalence rate of sickness over the study period would require 15/0.10 = 150 subjects. If the prevalence is lower, say 5% then we would need 15/0.05 = 300 subjects.

- Note: Sample size is a risk/benefit so we want to consider various sample sizes.

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 2: Sizing a study with a proportion endpoint

- Distribution Platform
- Interval Explorer for One Sample Proportion

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 2: Sizing a study with a proportion endpoint

How many samples do I need to demonstrate sensitivity and specificity for a regulatory filing?

- Sensitivity: the proportion of positive cases that test positive
- Specificity: the proportion of negative cases that test negative
- Sample size is calculated for each metric individually
  - Add for total sample size for retrospective study
  - For a prospective study calculate the total number of subjects to enroll based on the prevalence of the positive outcome.

Ex: If I need 100 positive subjects and 150 negative subjects and prevalence is 25%, then I need to enroll about 100/0.25 = 400 subjects to have roughly 100 positives subjects (and then 300 negatives).

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 2: Sizing a study with a proportion endpoint

Preliminary Information

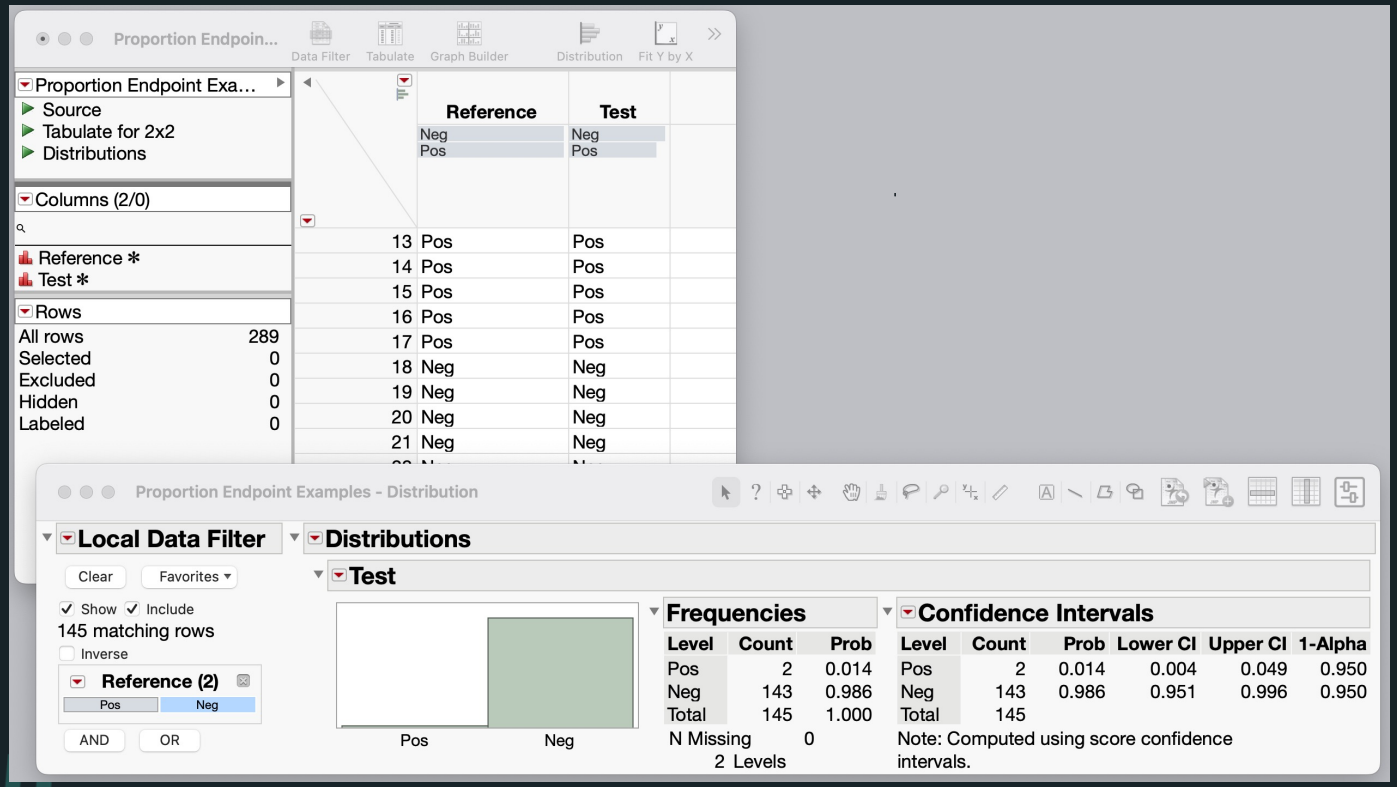
- Goal of the study: Regulatory Filing (high level of evidence needed)
- Primary endpoint: Demonstrate the 95%LCL for Se > 0.80 and Sp > 0.80
- Study design: Retrospective review of CT scans
- Assumptions:
  - Se > 0.9
  - Sp > 0.85
- I need to understand the confidence interval as an outcome
- I need a calculator for a confidence interval for a proportion

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 2: Sizing a study with a proportion endpoint

If I had data….what would I do?

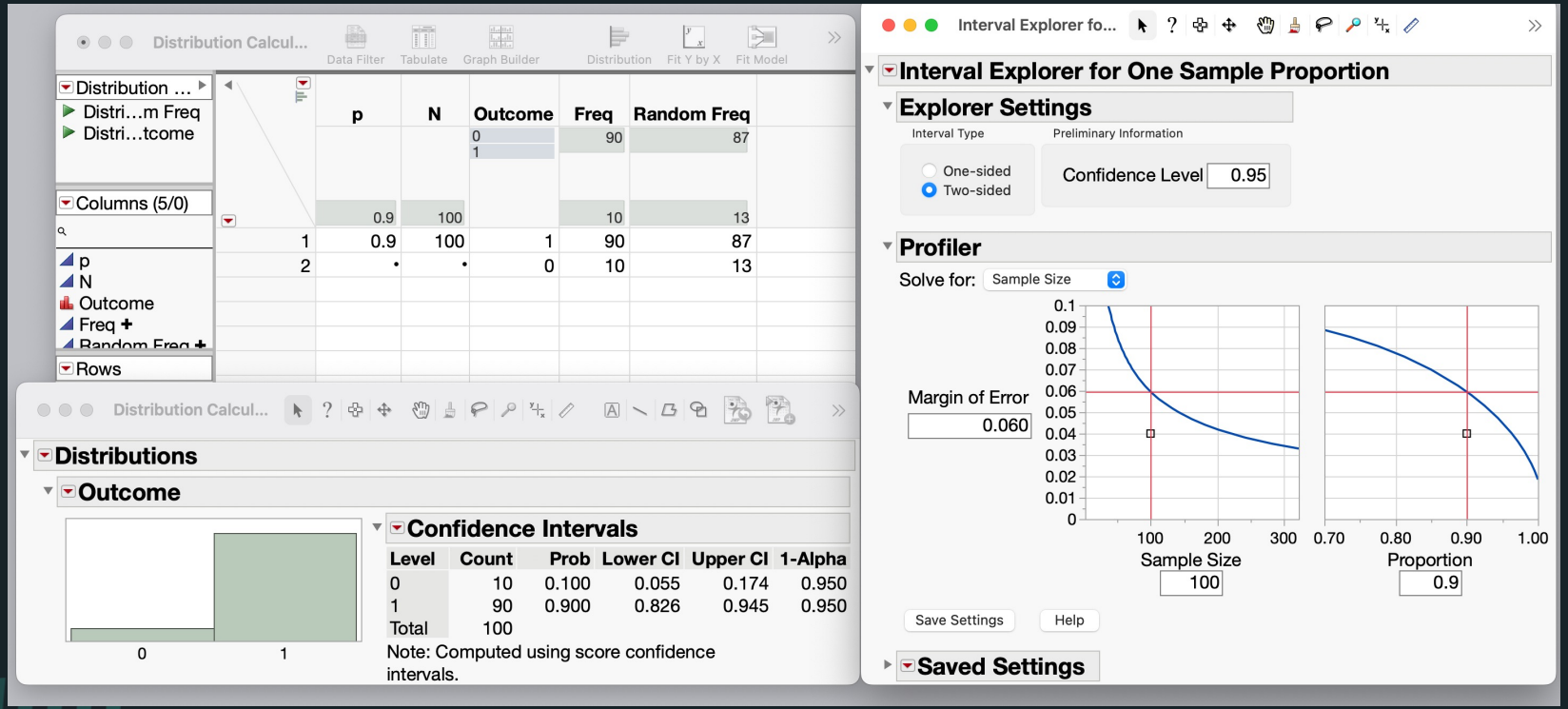


Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Example 2: Sizing a study with a proportion endpoint

How much data should I collect?



Copyright © JMP Statistical Discovery LLC. All rights reserved.

JMP STATISTICAL DISCOVERY

Additional Topics

Other Sample Size Methodologies

- Simulation:
  - Use pilot data to define distributions
  - Use random number generators to generate a study run (trial)
  - Analyze the data to see if your endpoint is met (yes/no)
  - Repeat some large number of times and calculate the proportion of times the endpoint is met (i.e., power)
- Best allocation of what you have:
  - Consider what samples/resources you have available
  - Use sample explorers to evaluate what you might be able to conclude
  - Use findings to decide if what you have is sufficient to proceed

Copyright © JMP Statistical Discovery LLC. All rights reserved.

jmp STATISTICAL DISCOVERY

Thank You

karen@boulderstats.com

jmp.com

Copyright © JMP Statistical Discovery LLC. All rights reserved.