Power Analysis

- The purpose of Power Analysis and Sample Size Estimation is to provide you with the statistical methods to assess validity of your statistical design.

α

Potential Outcomes from Analyses

		State of the World	
		HO	H1
Decision	H0	Correct Acceptance	Type II Error
	H1	Type I Error	Correct Rejection

Relationship between Potential Outcomes

	DECISION	
STATE OF NATURE	DO NOT REJECT Ho:	REJECT Ho:
Ho: TRUE	1. CORRECT (1 - ALPHA)	2. TYPE I ERROR (ALPHA) SUM=1
Ho: FALSE	3. TYPE II ERROR (BETA)	4. CORRECT (1-BETA)(POWER) SUM=1

Type I Error

- a Type I error represents, in a sense, a "false positive" for the researcher's theory. From society's standpoint, such false positives are particularly undesirable. They result in much wasted effort, especially when the false positive is interesting from a theoretical or political standpoint (or both), and as a result stimulates a substantial amount of research. Such follow-up research will usually not replicate the (incorrect) original work, and much confusion and frustration will result.

Type I Error

- the Type I error rate, must be kept at or below 0.05, and that, if at all possible, , the Type II error rate, must be kept low as well

Type II Error

- Type II error is a tragedy from the researcher's standpoint, because a theory that is true is, by mistake, not confirmed. So, for example, if a drug designed to improve a medical condition is found (incorrectly) not to produce an improvement relative to a control group, a worthwhile therapy will be lost, at least temporarily, and an experimenter's worthwhile idea will be discounted.

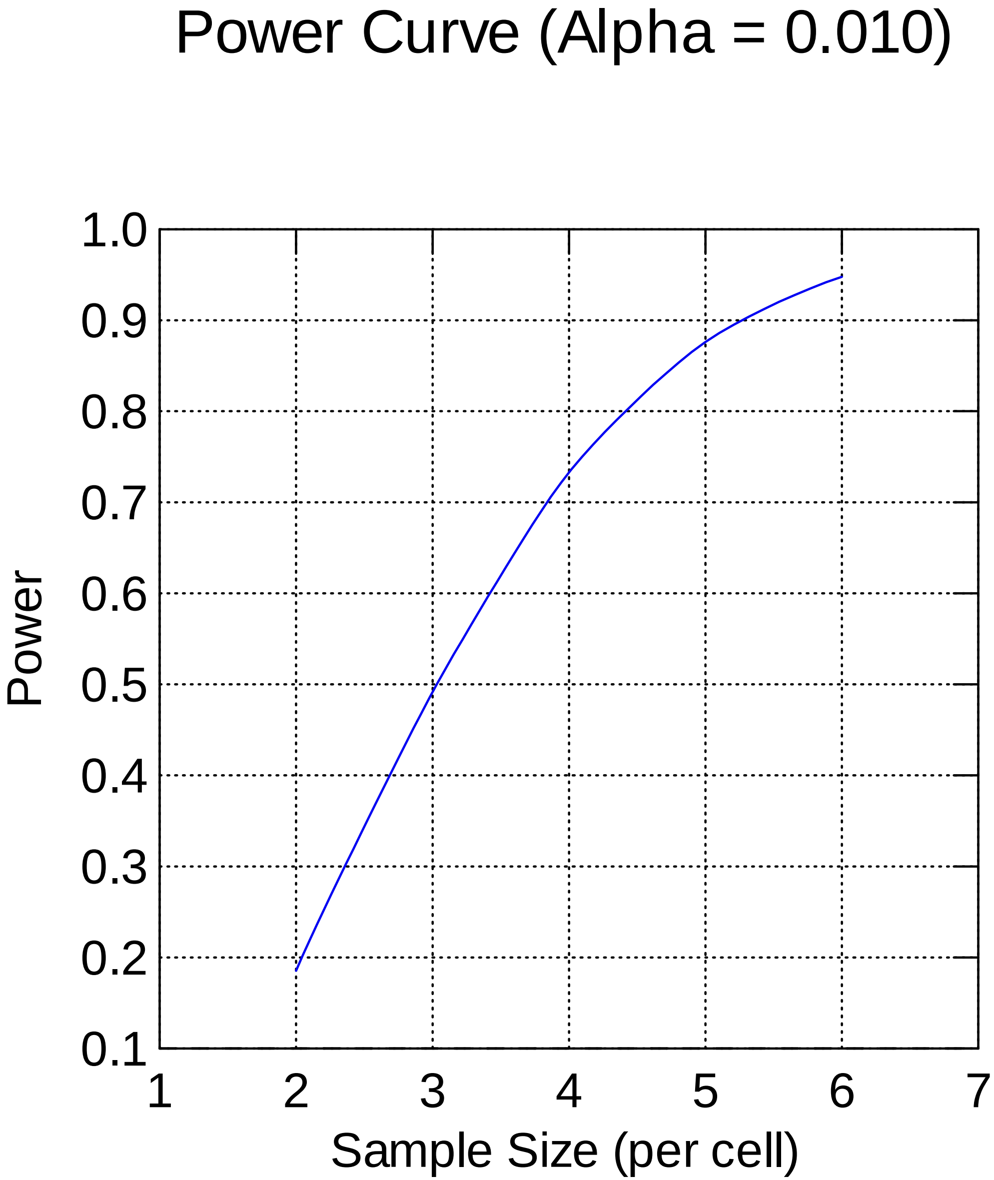
- “Statistical Power”, which is equal to 1 - Beta, must be kept correspondingly high. Power should be at least 0.80 to detect a reasonable departure from the null hypothesis.

Where does this come from?

- HAVING SPECIFIED ALPHA, THE INVESTIGATOR NEEDS NEXT TO SPECIFY THE CHANCES 1-BETA OF DETECTING THE PROPORTIONS AS DIFFERENT IF, IN THE UNDERLIEING POPULATIONS, THE PROPORTIONS ARE p1 AND p2. THE CRITERION SUGGESTED BY COHEN (1977) SEEMS REASONABLE. HE SUPPOSES IT TO BE THE TYPICAL CASE THAT A TYPE I ERROR IS SOME 4 TIMES AS SERIOUS AS A TYPE II ERROR. THIS IMPLIES THAT ONE SHOULD SET BETA, THE PROBABILITY OF A TYPE II ERROR, ABOUT EQUAL TO 4 ALPHA, SO THAT THE POWER BECOMES, ABOUT 1 - BETA = 1 - 4 ALPHA. THUS WHEN ALPHA = 0.01, 1 - BETA MAY BE SET AT 0.95; FOR ALPHA = 0.02, SET 1-BETA = 0.90; AND FOR ALPHA = .05, SET 1 - BETA = 0.80. WHEN ALPHA IS LARGER THAN 0.05 IT SEEMS SAFE TO TAKE 1 -BETA = 0.75 OR LESS.

Example: Two-Way ANOVA

- Alpha = 0.010
  Power = 0.800
  Model = Twoway
  Number of rows = 2
  Number of columns = 3
  Avg. std. sq. effect = 0.667
  Estimate to be based on column main effects.
  Noncentrality parameter = 4.000 * sample size
  SAMPLE
  SIZE POWER
  (per cell)
  2 0.185
  3 0.492
  4 0.733
  5 0.876
  Total Sample Size = 30



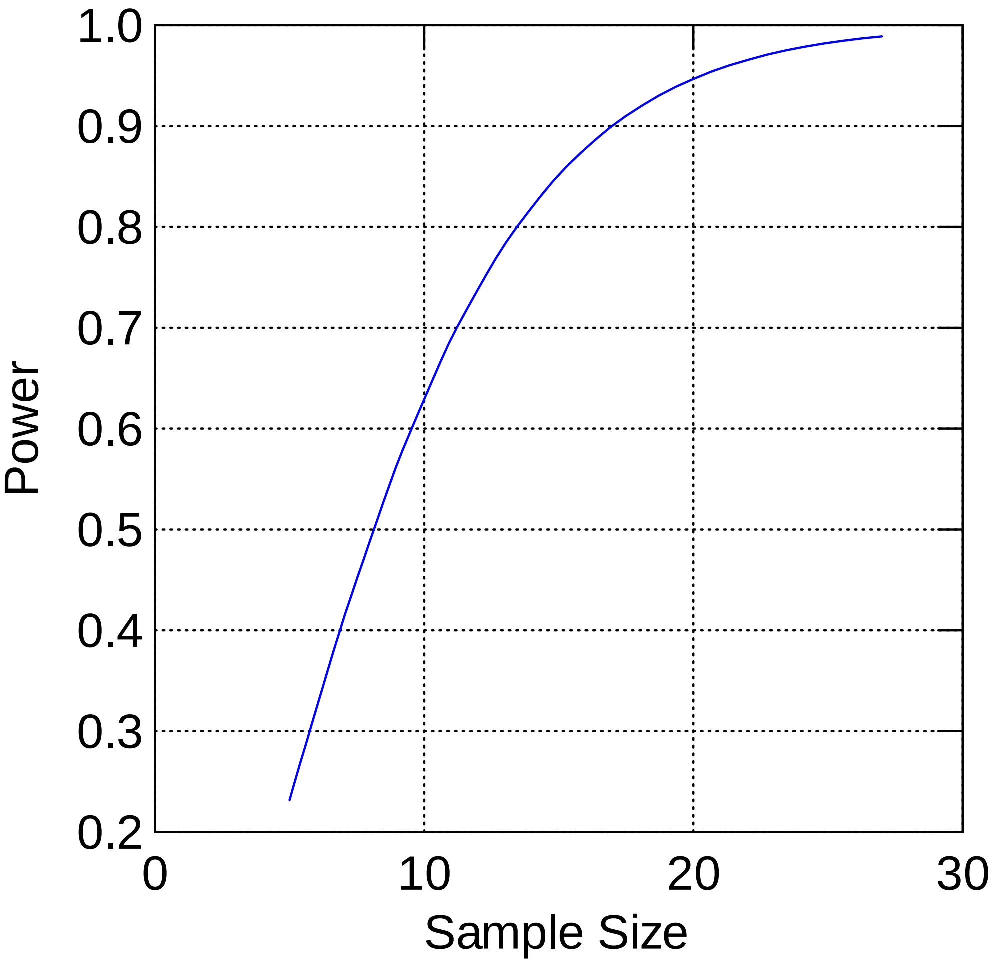
Power Curve (Alpha = 0.010)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Power
1
2
3
4
5
6
7
Sample Size (per cell)

Example: Correlation

Alpha = 0.050
Power = 0.800
Model = Single Correlation Alternative 'not equal'
Correlation coefficient = 0.700
Null value = 0.000
Effect Size = 0.700

SAMPLE SIZE (per cell)	POWER
11	0.689
12	0.740
13	0.783
14	0.820
15	0.852

Power Curve (Alpha = 0.050)



Example: Two-Sample t-test

Alpha = 0.050

Sample size: Low = 4

High = 30

Increment = 1

Model = Two Sample t-test with alternative 'not equal'

Effect Size = 0.500

Pooled S.D. = 4.000

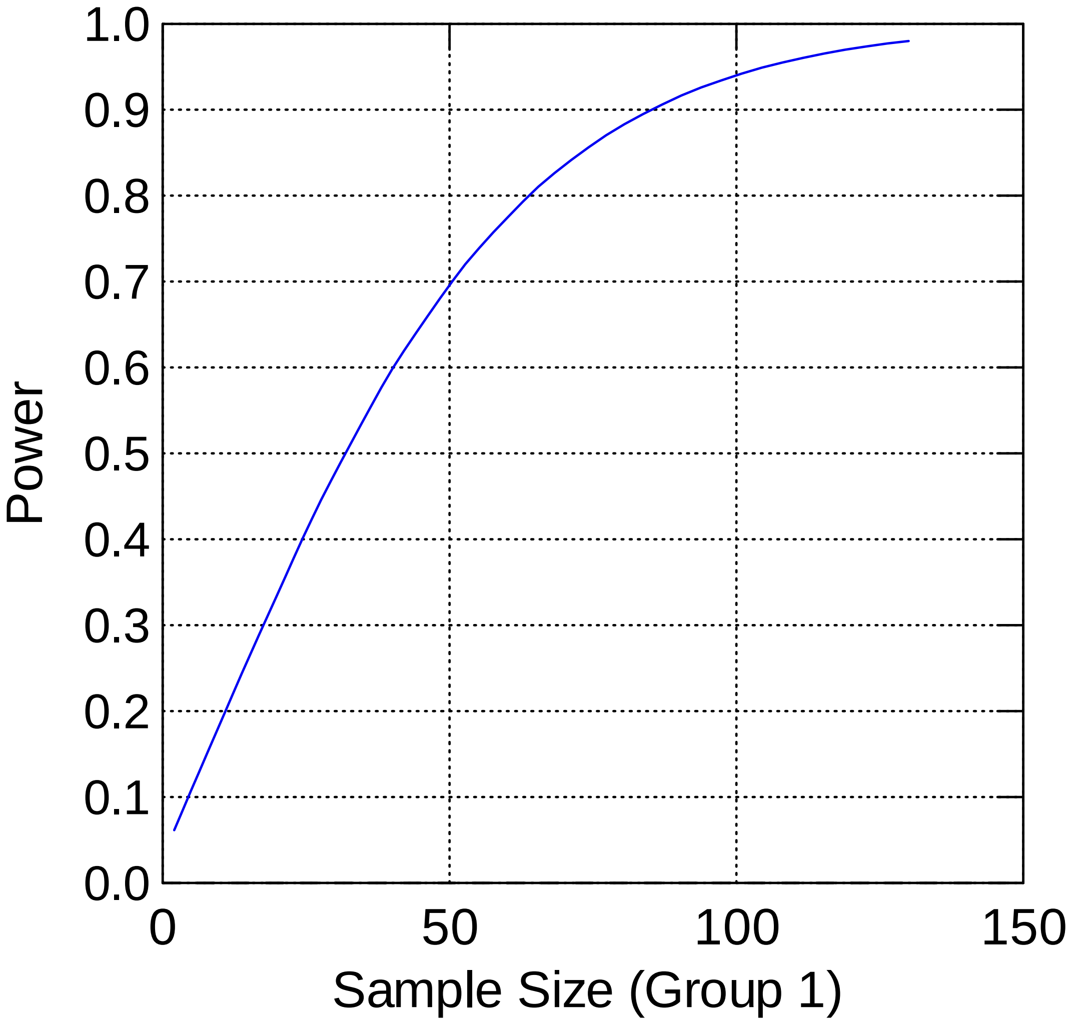
Mean(01) = 4.000

Mean(02) = 6.000

Noncentrality parameter = -0.500 * sqrt(sample size / 2)

SAMPLE SIZE (per cell)	POWER
4	0.092
5	0.108
6	0.123
7	0.139
8	0.154
9	0.170
10	0.185
11	0.201
12	0.216
13	0.232
14	0.247
15	0.262
16	0.278
17	0.293
18	0.308
19	0.323
20	0.338
21	0.353
22	0.367
23	0.382
24	0.396
25	0.410
26	0.424
27	0.438
28	0.451

Power Curve (Alpha = 0.050)

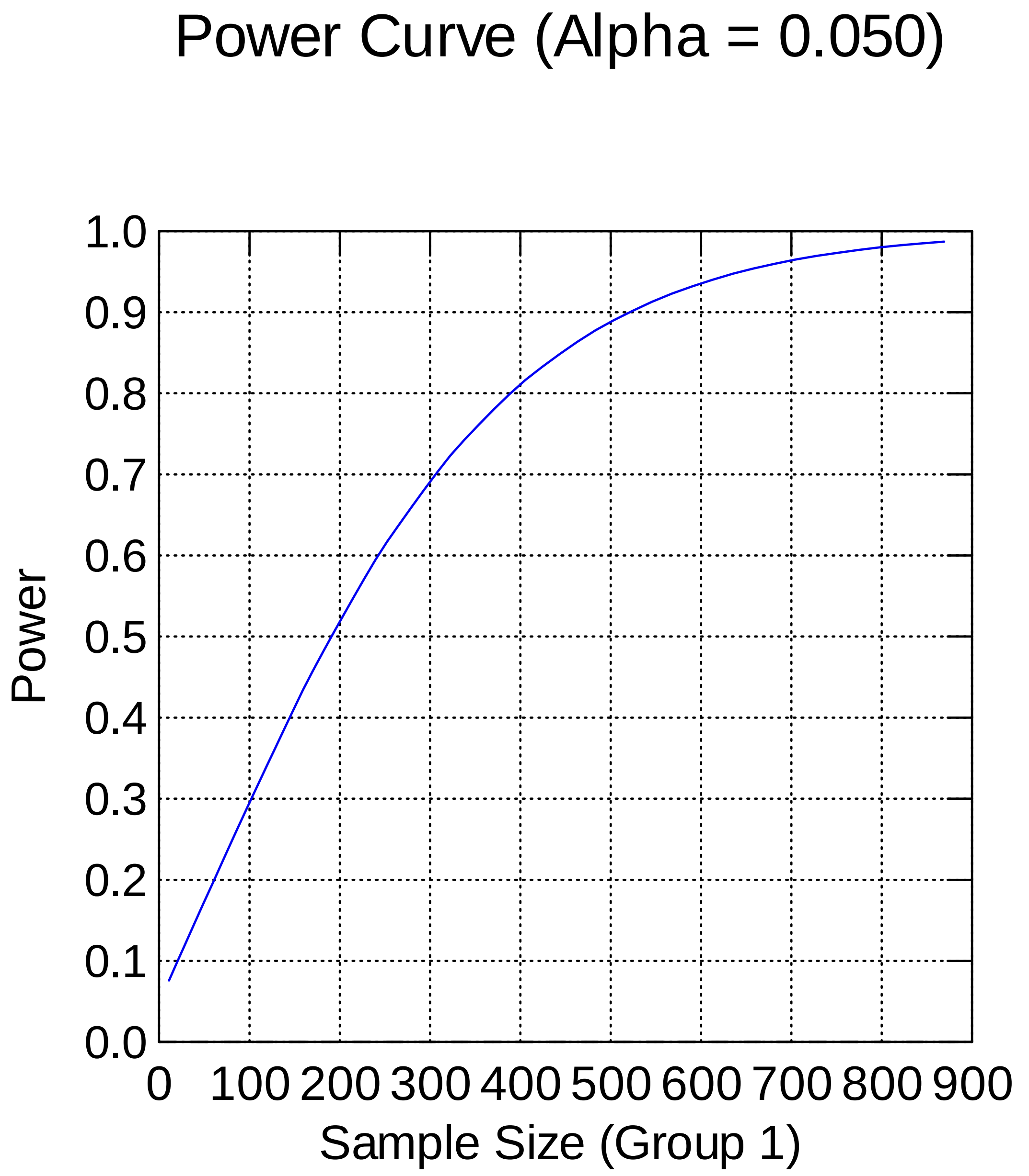


Example: Test of Equal Proportions

- Alpha = 0.050
- Power = 0.800
- Model = Equality of Proportions Alternative 'not equal'
- P1 = 0.400
- P2 = 0.500
- Ratio(2:1) = 1.000
- Effect Size = 0.201 Approximate Test
- Effect Size = 0.100 Large Sample Test

Approximate Test			Large Sample Test		
GROUP1	GROUP2	POWER	GROUP1	GROUP2	POWER
384	384	0.797	385	385	0.796
385	385	0.798	386	386	0.797
386	386	0.799	387	387	0.798
387	387	0.800	388	388	0.799
388	388	0.801	389	389	0.800

- Total Sample Size = 776 Total Sample Size = 778



Power Curve (Alpha = 0.050)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Power
0
100
200
300
400
500
600
700
800
900
Sample Size (Group 1)

Sample Size Estimation

- Equation to determine number of animals to sample to determine if disease occurs in a population

Sample Size Estimation: Disease

- $$\text{Sample Size} = [1 - (\text{alpha}^{1/D})]\left[N - \left(\frac{D-1}{2}\right)\right]$$
- Where:
- **N = Population Size**
- D = Expected number of diseased animals
- Alpha = Probability level (normally 0.05)

Example: Determining Disease Presence

-
- **Estimated population N = 400**
- **Infection Rate estimated at 2% so, 0.02*400 = 8 = D**
- **Alpha = 0.05**
-
-
- **Sample Size =** $[1 - (0.05^{1/8})]\left[400 - \left(\frac{8-1}{2}\right)\right]$
-
-
- **Sample Size = [1 – 0.687656] [396.5]**
-
- **Sample Size = 123.8 or 124 animals**
-

SAMPLE SIZE CALCULATION

-
- WHAT IS THE RELATIONSHIP BETWEEN VARIANCE AND SAMPLE SIZE???

EXAMPLE

-
- WHAT IF ONE HAD A MEAN OF 5.9 UNITS, A SD OF 0.83 AND WE WANTED TO KNOW HOW MANY OBSERVATIONS WE WOULD NEED TO OBTAIN A MEAN THAT WOULD NOT VARY FROM THE TRUE MEAN OF THE POPULATION BY MORE THAN 5% WITH 95% CONFIDENCE??

Sample Size

- FIVE PERCENT OF THE CALCULATED MEAN IS 5.9 X 0.05 = 0.3
-
- A 95% PROBABILITY IS 2 STANDARD ERRORS t = 1.96
-
- SO OUR DESIRED STANDARD ERROR BECOMES 0.3/2 = 0.15
-
- NOW WE CAN SUBSTITUTE VALUES FOR SE, AND S IN THE FORMULA
-
- SE = SD/SQUARE ROOT OF N AND SOLVE FOR N

Sample Size

- N THEN EQUALS THE NUMBER OF OBSERVATIONS NEEDED FOR A SAMPLE MEAN THAT WOULD NOT DEVIATE FROM THE TRUE MEAN BY MORE THAN 5%.
-
- SO SQUARE ROOT N = 0.83/0.15
-
- $N = (0.83/0.15)^2 = 5.53^2 = 30.62 = 31$

Sample Size

- NOW, WHAT WOULD HAPPEN IF WE INCREASED OUR STANDARD DEVIATION TO LET'S SAY 1.5 BUT THE MEAN REMAINS THE SAME
-
- SO, $(1.66/.15)^2 = (11.07)^2 = 123$
-
-
- SO, BY DOUBLING OUR STANDARD DEVIATION, WE NOW NEED 4 TIMES THE SAMPLE SIZE.