

Logistic Regression

Objective: Determine the relationship between the probability of an event and a continuous predictor

A researcher wants to investigate the relationship between the probability of heart disease and cholesterol level. Previous studies suggest an 8% probability of heart disease for patients with an average cholesterol level (about 5 mmol/l). The investigator wants to be able to detect an increase in the probability of heart disease to at least 12% for patients with cholesterol levels 1 standard deviation above the mean. The investigator wants to detect a difference at least this large with 80% power with a significance level of 5%.

Required Information	Inputs
What is the desired power for the test?	80%
At what significance level do you want to test your hypothesis?	5%
What is the “baseline” proportion of successes?	0.08
What proportion of successes do you want to detect for a 1 standard deviation change in the predictor?	0.12
Is your hypothesis one-sided or two-sided?	Two-sided

z tests – Logistic regression
Options: Large sample z-Test, Demidenko (2007) with var corr
Analysis: A priori: Compute required sample size
Input: Tail(s) = Two
 Odds ratio = 1.5681818
 Pr(Y=1|X=1) H0 = 0.08
 α err prob = 0.05
 Power (1-β err prob) = 0.80
 R² other X = 0
 X distribution = Normal
 X parm μ = 0
 X parm σ = 1

Test family: z tests
Statistical test: Logistic regression
Type of power analysis: A priori: Compute required sample size – given α, power, and effect size

Input Parameters:
 Tail(s): Two
 Determine => Odds ratio: 1.5681818
 Pr(Y=1|X=1) H0: 0.08
 α err prob: 0.05
 Power (1-β err prob): 0.80
 R² other X: 0
 X distribution: Normal
 X parm μ: 0
 X parm σ: 1

Output Parameters:
 Critical z: 1.9599640
 Total sample size: 510
 Actual power: 0.8000301

Callout boxes:
 - "Select 'z tests' and 'Logistic regression'"
 - "Odds ratio of probability at 1 SD above mean to probability at mean or baseline. Use 'Determine =>' to calculate from 2 probabilities."
 - "'Baseline' probability of probability at mean X"
 - "0 if only one predictor"
 - "Set μ=0 and σ=1 for 1 SD change in X"

A total sample of at least 510 is needed.