The Craft of Statistical Analysis Webinars

Probability, Odds, and Odds Ratios in Logistic Regression Models

Karen Grace-Martin

THE ANALYSIS FACTOR

C-3PO: Sir, the possibility of successfully navigating an asteroid field is approximately 3,720 to 1.

Han Solo: Never tell me the odds.

- The Empire Strikes Back

Want to predict for a set of 90 students whether a student goes on Academic Warning after the first semester (1st semester GPA < 2.0.

Response:

Y = 1 if a student is on academic warning

Y = 0 if a student is not on academic warning

Predictors:

- High school GPA
- Verbal SAT score
- Math SAT score
- Gender
- Involvement in high school sports

Of the 90 students in our sample, 23 were put on academic warning and 67 passed their classes successfully.

Passed	Failed	Total
67	23	90

Of the 90 students in our sample, 23 were put on academic warning and 67 passed their classes successfully.

Passed	Failed	Total
67	23	90

Probability

Probability is proportion of passes to the total

Probability ranges from 0 to 1

Pr(Failing) = 23/90 = .256

Pr(Passing) = 67/90 = .744 = 1 - .256

Pr(Failing) = 1 - Pr(Passing)

Pr = .5 → Pass and Fail equally likely

Pr > .5 → Failing is more likely than Passing

Pr < .5 → Failing is less likely than Passing

Of the 90 students in our sample, 23 were put on academic warning and 67 passed their classes successfully.

Passed	Failed	Total
67	23	90

Odds

Odds are the proportion of passes to failures

Odds ranges from 0 to ∞

Odds(Failing) = 23/67 = .34

Odds(Passing) = 67/23 = 2.9 = 1/.34

Odds(Failing) = 1/Odds(Passing)

Odds = 1 → Failing and passing equally likely

Odds > 1 → Failing is more likely than Passing

Odds < 1 → Failing is less likely than Passing

$$\text{Odds(failing)} = \frac{\Pr(\text{failing})}{\Pr(\text{passing})} = \frac{P}{1-P} = \frac{Failures/Total}{Passes/Total}$$

$$\text{Odds(failing)} = \frac{\Pr(\text{failing})}{\Pr(\text{passing})} = \frac{.256}{.744} = .34$$

C-3PO: Sir, the ~~possibility~~ odds of successfully navigating an asteroid field is approximately 3,720 to 1.

Pr(unsuccessful navigation) = 1/3721

= .000269

Odds Ratios in Logistic Regression

$$Ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	HSGPA	-1.161	.922	1.586	1	.208	.313	.051	1.908
	SATV	-.011	.009	1.606	1	.205	.989	.972	1.006
	SATM	-.021	.008	6.386	1	.012	.979	.964	.995
	SPORTS(1)	-.408	.848	.232	1	.630	.665	.126	3.506
	GENDER	1.645	.797	4.259	1	.039	5.179	1.086	24.694
	Constant	15.591	5.231	8.883	1	.003	5905214		

a. Variable(s) entered on step 1: HSGPA, SATV, SATM, SPORTS, GENDER.

Gender (0 = Female, 1 = Male) OR = 5.2

Involvement in high school sports (0 = No, 1 = Yes) OR = .67

References

DeMaris, Alfred. (1995). A Tutorial in Logistic Regression. Journal of Marriage and Family, 57, 956-968.

Menard, Scott. (1995). Applied Logistic Regression Analysis. Sage Publications, Thousand Oaks, CA.

Morgan, S. Philip & Teachman, Jay D. (1988). Logistic Regression: Description, Examples, and Comparisons. Journal of Marriage and the Family, 50, 929-936.

At The Analysis Factor:

http://www.theanalysisfactor.com/category/other-regression-models/logistic-regression/