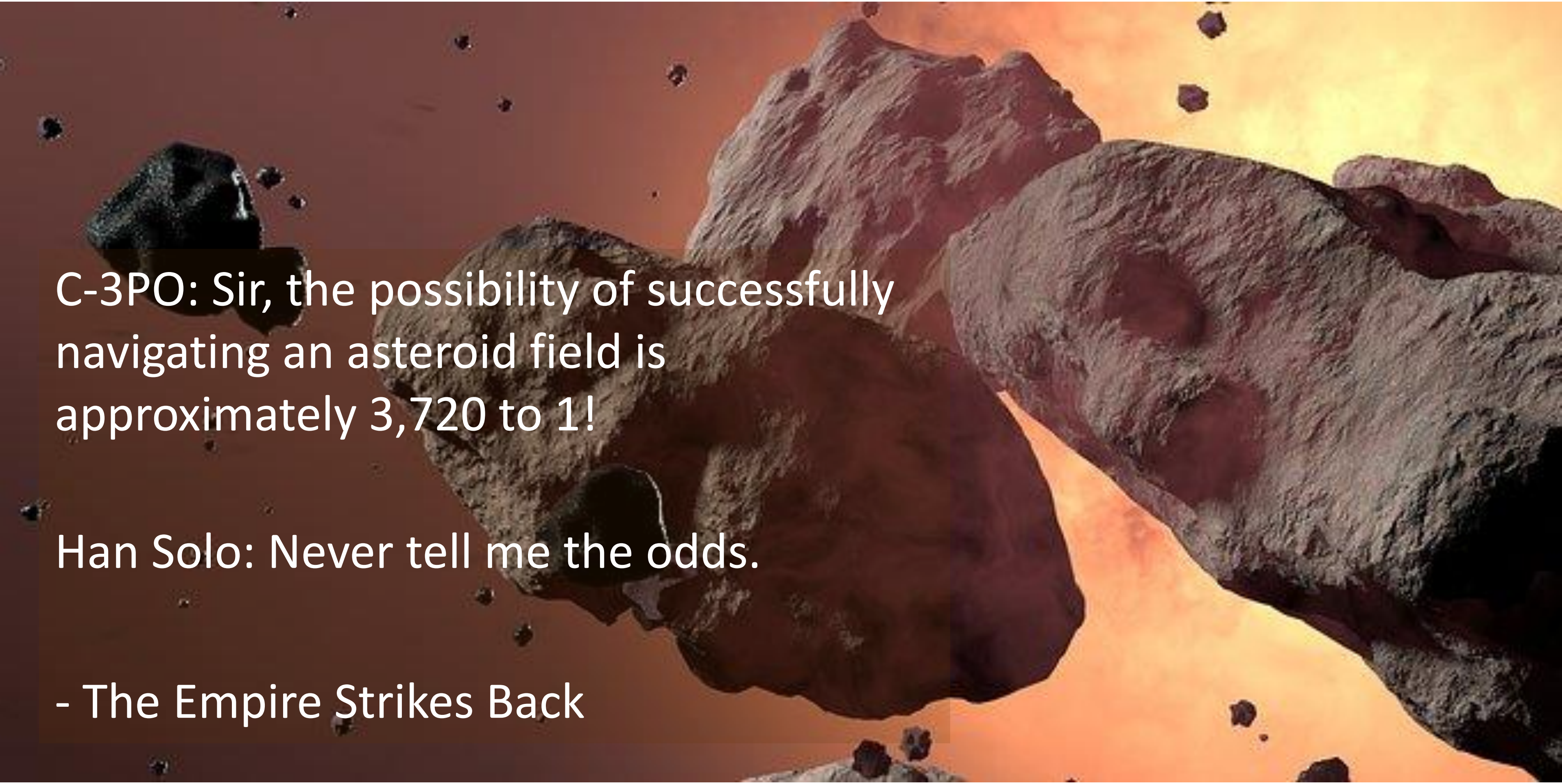
Probability, Odds, and Odds Ratios in Logistic Regression Models

Karen Grace-Martin


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

Is every ship equally likely to successfully navigate an asteroid field?

1. Collect Data

2. Run a logistic regression to see which, if any, variables predict the likelihood of successfully navigating an asteroid field

Logistic Regression Coefficients

Dependent Variable:
Successfully Navigates Asteroid Field

Variable	b	se	t	p	OR
Intercept	-8.221	2.581	3.185	.000	
Size of Ship tons	-.128	.041	3.12	.000	.88
Presence of R2 unit	.959	.327	2.93	.026	2.61 = exp(.959)

logs-odds

$e^{.959}$

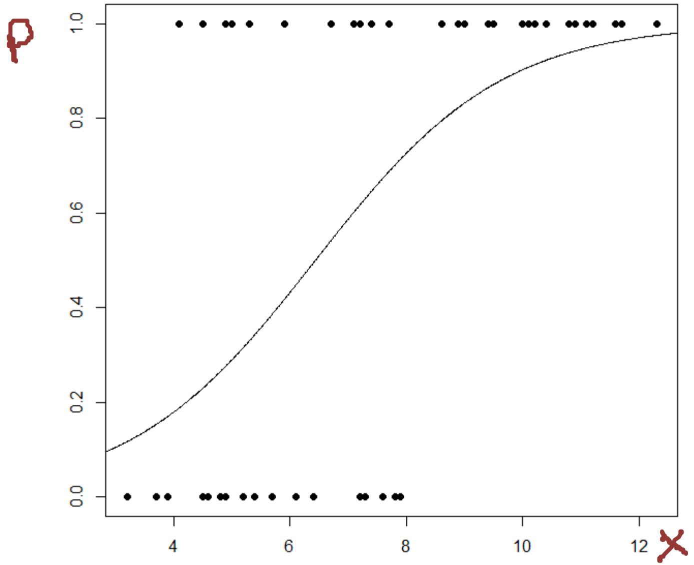
Modeling Probability

$$\underset{0-1}{\cancel{P}} = \beta_0 + \overset{b}{\beta_1} X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

Why not?

1. The right hand side of the equation ranges from - ∞ to + ∞, but the left hand side ranges from 0 to 1.

2. Relationship is not linear, but rather follows an S-shaped curve



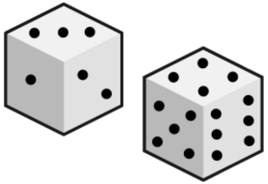
Logistic Regression

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

Using algebra, we can solve for the probability of success.

$$P=\frac{e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_k X_k}}{1+e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_k X_k}}$$

What are odds?



Odds ≠ Probability

"…probability, which is the odds of something occurring… A fair coin has 50% odds of heads coming up on any flip."

INCORRECT

- Author anonymous to protect the guilty

A Coin Experiment

Flip a coin 10,000 times and get 5,001 heads

Probability

$$P(\text{heads}) = \frac{\text{Number of heads}}{\text{Number of flips}}$$

$$P(\text{heads}) = \frac{5001}{10000}$$

$$P(\text{heads}) = .5001$$

Odds

$$\text{Odds}(\text{heads}) = \frac{\text{Number of heads}}{\text{Number of tails}}$$

$$\text{Odds}(\text{heads}) = \frac{5001}{4999}$$

$$\text{Odds}(\text{heads}) = 1.0004{:}1 = 1$$

Measure the same construct on different scales

Probability and Odds both measure likelihood *on different scales*.

.5 Odds ≠ .5 Probability

Celsius and Fahrenheit both measure temperature *on different scales.*

30°F ≠ 30°C

Probability

Heads	Tails	Total
5001	4999	10000

- Proportion of successes to the total
- Ranges from 0 to 1

P(Heads) = 5001/10000

P(Tails) = 4999/10000

P(Tails) = 1 – P(Heads)

P = .5 → Heads and tails equally likely

P > .5 → Heads is more likely than tails

P < .5 → Heads is less likely than tails

Odds

Heads	Tails	Total
5001	4999	10000

- proportion of successes to failures
- ranges from 0 to ∞

Odds(Heads) = 5001/4999

Odds(Tails) = 4999/5001

Odds(Tails) = 1/Odds(Heads)

Odds = 1 $\rightarrow$ Heads and tails equally likely

Odds > 1 $\rightarrow$ Heads is more likely than tails

Odds < 1 $\rightarrow$ Heads is less likely than tails

Converting

$$\text{Odds} = \frac{P}{1 - P}$$

P(heads) = 5001/10000 P

P(tails) = 4999/10000 = 1 – P(heads)

$$\text{Odds(heads)} = \frac{5001}{4999}$$

Num H
Num T

$$\text{Odds(heads)} = \frac{5001/\cancel{10000}}{4999/\cancel{10000}}$$

Pr(H)
Pr(T)

C-3PO:

odds un

Sir, the ~~possibility~~ of successfully navigating an asteroid field is approximately 3,720 to 1.

Odds(successful navigation) = 1/3720

Number of successful ships

Number of unsuccessful ships

Pr(successful navigation) = 1/3721

Number of successful ships

Total number of ships

Odds Ratios in Logistic Regression

Odds Ratios in Logistic Regression

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

$$\mathrm{Ln}\left(\text{odds}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

β_1 is the effect of X_1 on the ln(odds) $\beta_1 = \frac{\Delta \ln(\text{odds})}{1 \Delta X_1}$

Odds Ratios in Logistic Regression

odds = 10 tons w/ R2 unit

odds = 15 tons w/ R2 unit

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

$$.88 = \frac{\frac{\text{odds}}{\text{odds}}}{1 \text{ ton} \uparrow}$$

Logistic Regression Coefficients

$$-.128 = \frac{\Delta \ln(\text{odds})}{1 \text{ ton} \uparrow}$$

$(.88 - 1)100\%$ change $= -12\%$

Dependent Variable:
Successfully Navigates Asteroid Field

Variable		b	se	t	p	OR
Intercept		-8.221	2.581	3.185	.000	
Size of Ship	+.128	-.128	.041	3.12	.000	.88
Presence of R2 unit	e	.959	.327	2.93	.026	2.61

$\frac{1}{.88}$

$= e \sim 2.72^{-.128}$

References and Resources

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.

At The Analysis Factor:

https://www.theanalysisfactor.com/resources/by-topic/logistic-regression/

Learn More

Title: Logistic Regression for Binary, Ordinal, and Multinomial Outcomes

Instructor: Karen Grace-Martin
Stage: 3 (Extensions of Linear Models)
Software: SPSS, R, Stata, SAS

Dates: 4 modules - 2 hours First is 4/25/19

Enrollment Closes: April 18

Full price: $397 (Student $247)

https://www.theanalysisfactor.com/workshop-logistic-regression/