Math 327

Section 9.2
Logistic Regression and Odds Ratios

Remember…

- Probability, p
  - Scale is 0 to 1

- Odds $= \frac{p}{1-p}$
  - Scale is 0 to ∞

- Log (Odds) $= \ln(Odds) = \ln\left(\frac{p}{1-p}\right)$
  - Scale is $-\infty$ to ∞

Pg. 424

Odds Ratio

- In this section, we focus on a binary predictor variable
  - Remember, the response variable is also binary
- Odds Ratio
  - A way to compare two odds values
  - Remember: Odds is a ratio, so the Odds Ratio is a ratio of ratios
- Logit form of the model
  - $\log(odds) = \beta_0 + \beta_1 X$
  - So $odds = e^{\beta_0 + \beta_1 X}$
  - And $probability = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$ since $p = \frac{Odds}{1 + Odds}$

Pg. 424-5

Example 9.6

- Explanatory variable: Treatment group, TMS or Placebo
- Response variable: Pain free after 2 hours, Yes or No
- Odds of Yes
  - Odds(Yes) for TMS: 39/61 = 0.639
  - Odds(Yes) for Placebo: 22/78 = 0.282
  - Odds Ratio, TMS vs. Placebo, $OR = \frac{0.639}{0.282} = 2.27$
  - Odds of <u>being pain free</u> after 2 hours is 2.27 times higher taking <u>TMS vs. Placebo</u>
- Odds of No
  - Odds(No) for TMS: 61/39 = 1.564
  - Odds(No) for Placebo: 78/22 = 3.545
  - Odds Ratio, Placebo vs. TMS, $OR = \frac{3.545}{1.564} = 2.27$
  - Odds of <u>not being pain free</u> after 2 hours is 2.27 times higher taking <u>Placebo vs. TMS</u>

	TMS	Placebo	Total
Pain-free 2 hrs later	39	22	61
Not pain-free 2 hrs. later	61	78	139
Total	100	100	200

Pg. 425

Example 9.7

- Letrozole therapy
  - Treatment for post-menopausal women with breast cancer who have completed 5 years of tamoxifen therapy
  - 5000 women randomized to one of two groups: letrozole or placebo
  - Response variable: Disease-free survival (Yes or No)
  - Results
    - Letrozole: $Odds(DFS) = \frac{2390}{185} = 12.9$
    - Placebo: $Odds(DFS) = \frac{2241}{341} = 6.57$
    - Odds Ratio: $\mathrm{OR} = \frac{12.9}{6.57} = 1.97$
  - Odds of disease-free survival is about twice as high for letrozole as for placebo

Disease-free survival	Letrozole	Placebo	Total
Yes	2390	2241	4631
No	185 (7.2%)	341 (13.2%)	566
Total	2575	2582	5157

Pg. 425-6

Example 9.8

- Transition to marriage
  - Does the chance of a single mother getting married depend on the gender of their child?
  - If child is a boy:
    - $Odds(married) = \frac{176}{134} = 1.31$
  - If child is a girl:
    - $Odds(married) = \frac{148}{142} = 1.04$
  - Odds ratio:
    - $OR = \frac{1.31}{1.04} = 1.26$
  - The odds of getting married if the child is a boy is 26% higher (1.26 times higher) than if the child is a girl

Mother married	Boy	Girl	Total
Yes	176	148	324
No	134	142	276
Total	310	290	600

Odds Ratio and Slope

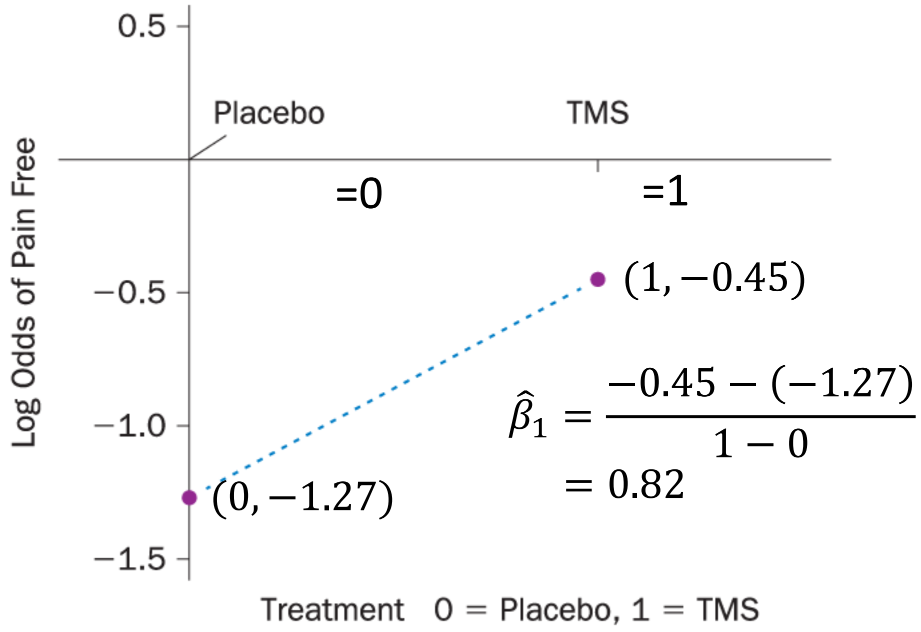
- In logistic regression, there is a constant slope on the logit scale
- Use empirical logits to help visualize the slope
  - Empirical logit $= logit(\hat{p}) = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \log\left(\frac{\#Yes}{\#No}\right)$
- Plot
  - Empirical logit vs. the 0, 1 predictor
  - Graphical summary of the 2x2 table

Pg. 427

Example 9.9

- Migraines: TMS treatment vs Placebo

Pain-free?	TMS	Placebo	TMS vs. Placebo
Yes	39	22	
No	61	78	
Total	100	100	
Odds (Pain-free)	0.64	0.28	$OR = \frac{0.64}{0.28} = 2.27$
Log(Odds)	-0.45	-1.27	$Slope$ $= -0.45 - (-1.27)$ $= 0.82$



Cannon, et. al., *STAT2: Modeling with Regression and ANOVA*, 2e, © 2019 W. H. Freeman and Company

Note: $e^{Slope} = OR$

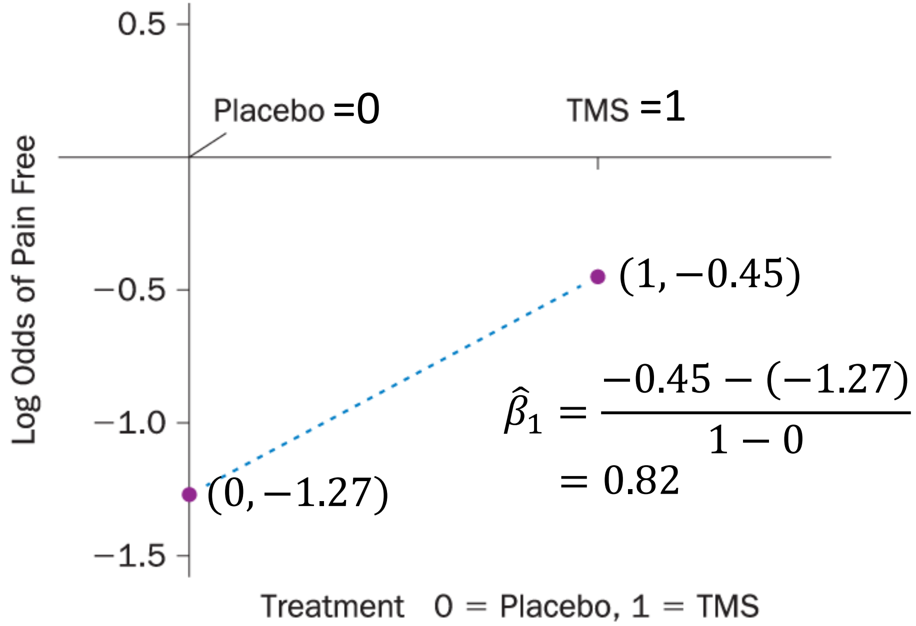
$e^{0.82} = 2.27$

Derivation:

$$e^{\hat{\beta}_1} = e^{Odds_2 - Odds_1} = \frac{e^{Odds_2}}{e^{Odds_1}} = \frac{Odds_2}{Odds_1} = OR$$

Fitted Slope and Odds Ratio

- In general the fitted slope is
  - $Slope = \frac{rise}{run} = \frac{change\ in\ \log odds}{1-0} = \log(OR)$
- Thus,
  - $OR = e^{slope} = e^{\hat{\beta}_1}$

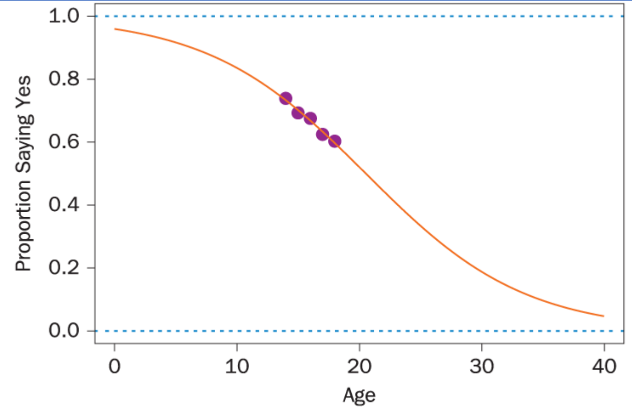


Cannon, et. al., *STAT2: Modeling with Regression and ANOVA,* 2e, © 2019 W. H. Freeman and Company

Slope and Odds Ratio when Predictor is Quantitative

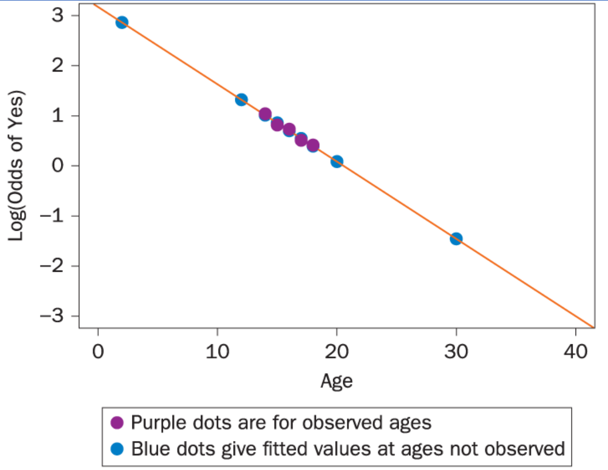
Pg. 428

- Probability scale for the response
  - The slope changes with X (logistic curve)



Cannon, et. al., *STAT2: Modeling with Regression and ANOVA*, 2e, © 2019 W. H. Freeman and Company

- Log(Odds) scale for the response
  - The slope is constant, but it's "meaning keeps changing"
    - Explanation of "meaning keeps changing:" The impact of a one-unit change in X on the probability scale is different for different values of X. The impact is constant on the log(Odds) scale
  - The Odds Ratio is constant



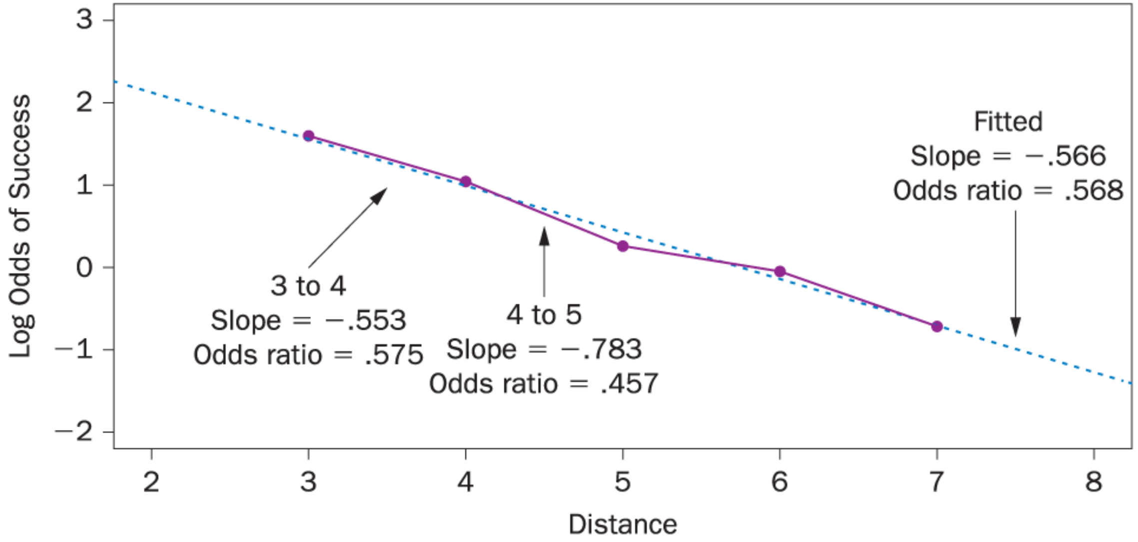
Cannon, et. al., *STAT2: Modeling with Regression and ANOVA*, 2e, © 2019 W. H. Freeman and Company

Pg. 428-9

Example 9.10

- Putting success in golf
  - Response: Putt is successful
  - Predictor: Length of the putt (feet)

Length of putt (in feet)	3	4	5	6	7
Number of successes	84	88	61	61	44
Number of failures	17	31	47	64	90
Total number of putts	101	119	108	125	134
Proportion of successes	**0.832**	**0.739**	**0.565**	**0.488**	**0.328**
Odds of success	4.941	2.839	1.298	0.953	0.489
Empirical logit	**1.60**	**1.04**	**0.26**	**−0.05**	**−0.72**
Fitted logit value	1.558	0.992	0.426	−0.140	−0.706
Change in fitted logit value		−0.566	−0.566	−0.566	−0.566
Fitted probability	0.826	0.730	0.605	0.465	0.330
Change in fitted probability per foot of putt length		−0.097	−0.125	−0.140	−0.135



Cannon, et. al., *STAT2: Modeling with Regression and ANOVA*, 2e, © 2019 W. H. Freeman and Company

Length of putt (feet):	4 to 3	5 to 4	6 to 5	7 to 6
Empirical Odds Ratio	0.575	0.457	0.734	0.513
Fitted logistic Odds Ratio	0.568	0.568	0.568	0.568

Pg. 429-30

Example 9.11

- Medical school admissions: Interpreting slope

```
> fit.med1 = glm (Acceptance ~ GPA, data=MedGPA, family=binomial)
> summary (fit.med1)

Call:
glm(formula = Acceptance ~ GPA, family = binomial, data = MedGPA)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7805  -0.8522   0.4407   0.7819   2.0967

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -19.207      5.629  -3.412 0.000644 ***
GPA            5.454      1.579   3.454 0.000553 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- $\hat{\beta}_1 = 5.45$ logit units per GPA unit
- $Odds_{GPA} = e^{-19.21+5.45 \cdot GPA}$
- $Odds_{GPA+1} = e^{-19.21+5.45(GPA+1)}$
- Describe one-unit increase in GPA in terms of the odds ratio:
  - $\dfrac{Odds_{GPA+1}}{Odds_{GPA}} = \dfrac{e^{-19.21+5.45(GPA+1)}}{e^{-19.21+5.45 \cdot GPA}} = e^{5.45}$

- A one-unit increase in GPA is associated with an $e^{5.45}$, or 233.7-fold, increase in the odds of acceptance
- Be aware of units: A 0.1-unit increase in GPA is associated with an $e^{0.545}$, or 1.72-fold, increase in the odds of acceptance, or not quite doubled
- Note: $1.72^{10} = 233.7$. If the odds of admission increases by a factor of 1.72 per 0.1 GPA units, then multiplying times 1.72 ten times is the same as multiplying by $1.72^{10} = 233.7$.

Example 9.11

- Confidence intervals for logistic regression parameters
  - We usually do not interpret the intercept, but $e^{\hat{\beta}_0}$ is the predicted odds of acceptance when $X = 0 \; (GPA = 0)$
    - In this case, it's nearly zero (4.557e-09)
  - With 95% confidence, the odds ratio for GPA is between 14.8 and 7,829
    - For every <u>one unit</u> increase in GPA, the odds of acceptance is between 14.8 and 7,829 times higher
  - With 95% confidence, the odds ratio for GPA/10 is between 1.31 and 2.45
    - For every <u>one-tenth</u> unit increase in GPA, the odds of acceptance increases between 1.31 and 2.45 times, or between 31% and 145% higher

```
> exp.betahat = exp (fit.med1$coefficients)
> ci.exp.beta = exp (confint (fit.med1))
Waiting for profiling to be done...
>
> cbind.data.frame (exp.betahat, ci.exp.beta)
               exp.betahat        2.5 %       97.5 %
(Intercept) 4.557449e-09 1.686955e-14 8.472476e-05
GPA         2.337298e+02 1.482501e+01 7.829246e+03
```

```
> exp.betahat0.1 = exp (fit.med1$coefficients/10)
> ci.exp.beta0.1 = exp (confint (fit.med1)/10)
Waiting for profiling to be done...
> cbind.data.frame (exp.betahat0.1, ci.exp.beta0.1)
            exp.betahat0.1      2.5 %    97.5 %
(Intercept)      0.1465117 0.04194791 0.3915624
GPA              1.7253270 1.30948191 2.4511619
```