**Week 8, Day 2**

**Logistic Regression**

* Odds Ratios -- Situation 2: **Numerical** explanatory var.
  + Example: math & CS grads by year

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Graduating Year | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
| Math seniors | 22 | 24 | 17 | 28 | 27 | 26 |
| CS seniors | 13 | 5 | 16 | 14 | 15 | 20 |
| Total seniors | 35 | 29 | 33 | 42 | 42 | 46 |
| Prop. of dept in math | 0.629 | 0.828 | 0.515 | 0.667 | 0.643 | 0.565 |
| Odds of math | 1.692 | 4.8 | 1.063 | 2.0 | 1.8 | 1.3 |

Odds ratios: 2.837 0.221 1.881 0.9 0.722

(next year/previous year)

Interpretations: In the department, the odds of choosing a math major (rather than a CS major) in 2015 is 2.8x the odds for 2014 grads. The odds of choosing a math major over a CS major in 2016 is 22% the odds in 2015.

* + Another Example: Trashball by distance

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Distance | 4 feet | 8 feet | 12 feet | 16 feet | 20 feet | 24 feet |
| Shots made | 24 | 21 | 12 | 7 | 2 | 6 |
| Shots missed | 0 | 3 | 12 | 17 | 22 | 18 |
| Odds of success | 24/0? | 21/3 | 12/12 | 7/17 | 2/22 | 6/18 |

Odds ratios: ? 7 2.43 4.53 0.27

**Trick** for division by 0: Include a hypothetical shot that was “half-made”.   
Then odds(4 feet) = 24.5/0.5, and OR(4 to 8) = (24.5/0.5)/(21.3) = 7.

(R knows how to deal with 0’s; this is just for us to understand what’s happening and be able to *visualize*.)

* + Activity: Odds & ORs, Part 2
* The linearity condition: What is it and how do we check it?
  + When explanatory variable is categorical
  + When explanatory variable is numerical