# Homework 07

Spencer Pease

3/06/2020

## *(Q1)*

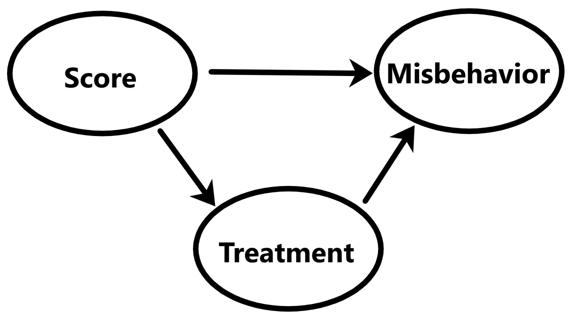


Figure 1: Casual diagram relating score, treatment, and behavior

## *(Q2)*

$$\text{logit}(P(response = 1 \mid score)) = \beta_0 + \beta_1 \cdot score$$

Table 1: Inferential statistics for the logistic regression of behavior response on score

| Parameter | Estimate | Robust SE | e(Est) | e(95%L) | e(95%H) | Pr(>\|z\|) |
|---|---|---|---|---|---|---|
| (Intercept) | -1.289 | 0.062 | 0.276 | 0.244 | 0.311 | 4.90e-90 |
| score | 0.012 | 0.002 | 1.013 | 1.009 | 1.016 | 1.15e-14 |

The coefficient of the score term (slope) is the log-odds ratio of the associative relationship between score and misbehavior. On average, between two groups with a one-unit difference in assigned score, the odds ($e^{\beta_1}$) of a misbehavior event are 1.013 (95% CI: 1.009, 1.016) times higher in the higher-scored group. A $P$-value of 1.15e-14 suggests we can reject the null hypothesis that there is no association between odds of misbehavior and assigned score in favor of there being some statistically significant difference.

## *(Q3)*

Table 2: Predicted probability of misconduct based on score

| score | prob. misconduct |
|---|---|
| 25 | 0.274 |
| 50 | 0.340 |
| 75 | 0.413 |

## *(Q4)*

Yes, since treatment is known to affect behavior

## *(Q5)*

$$\text{logit}(P(response = 1 \mid score, treat)) = \beta_0 + \beta_1 \cdot score + \beta_2 \cdot treat$$

### *(Q5.a)*

Table 3: Inferential statistics for the logistic regression of behavior response on score, adjusted for treatment

| Parameter | Estimate | Robust SE | e(Est) | e(95%L) | e(95%H) | Pr(>\|z\|) |
|---|---|---|---|---|---|---|
| (Intercept) | -1.504 | 0.081 | 0.222 | 0.190 | 0.260 | 2.64e-74 |
| score | 0.024 | 0.003 | 1.024 | 1.018 | 1.030 | 2.92e-15 |
| treat | -0.732 | 0.161 | 0.481 | 0.351 | 0.659 | 5.32e-06 |

For two groups of the population with the same treatment status, a one-unit difference in assigned score is associated with a 1.024 (95% CI: 1.018, 1.03) times change in the odds of a misbehavior event ($e_1^{\beta}$) in the higher-scored group, on average.

For two groups of the population with the same score, being assigned to a maximum security prison is associated with a 0.481 (95% CI: 0.351, 0.659) times change in odds of a misbehavior event ($e_2^{\beta}$), on average.

In both cases, the estimated coefficients are significant ($P$-values 2.92e-15 and 5.32e-06, respectively) and not equal to zero, so we can separately reject the null hypotheses that each included variable is not significantly associated with misbehavior.

### *(Q5.b)*

Table 4: Predicted probability of misconduct based on score and treatment

| treat | score | prob. misconduct |
|---|---|---|
| 0 | 25 | 0.288 |
| 1 | 25 | 0.388 |
| 0 | 50 | 0.462 |
| 1 | 50 | 0.355 |
| 0 | 75 | 0.645 |

| treat | score | prob. misconduct |
|---|---|---|
| 1 | 75 | 0.324 |

## *(Q6)*

*(Optional)*

## *(Q7)*

$$\text{logit}(P(response = 1 \mid score, treat)) = \beta_0 + \beta_1 \cdot score + \beta_2 \cdot treat + \beta_3(score \cdot treat)$$

Table 5: Inferential statistics for the logistic regression of behavior response on score, adjusted for treatment (including interaction term)

| Parameter | Estimate | Robust SE | e(Est) | e(95%L) | e(95%H) | Pr(>\|z\|) |
|---|---|---|---|---|---|---|
| (Intercept) | -1.654 | 0.088 | 0.191 | 0.161 | 0.227 | 9.19e-76 |
| score | 0.030 | 0.003 | 1.030 | 1.024 | 1.037 | 8.31e-20 |
| treat | 1.336 | 0.436 | 3.803 | 1.616 | 8.949 | 2.23e-03 |
| score:treat | -0.036 | 0.007 | 0.965 | 0.951 | 0.979 | 1.11e-06 |

From the above logistic regression, we estimate that for two groups of the population both assigned to maximum security, a one-unit increase in assigned score modifies the association of the log-odds ratio of misbehavior by -0.036. A $P$-value of 1.11e-06 suggests we reject the null hypothesis that there is no association between the interaction of assigned score and treatment.

## *(Q8)*

### *(Q8.a,b)*

- Odds ratio characterizing the association between classification score and misconduct for prisoners **not** assigned to maximum security:
- Odds ratio characterizing the association between classification score and misconduct for prisoners assigned to maximum security:

### *(Q8.c)*

## *(Q9)*

### *(Q9.a,b)*

For the question of interest, **mother's education** plays the role of a *confounder*, while **baby's sex** plays is a *precision* variable.

### *(Q9.c)*

Table 6: Inferential statistics for the logistic regression of low birth weight by age, adjusted for race, education, and sex

| Parameter | Estimate | Robust SE | e(Est) | e(95%L) | e(95%H) | Pr(>\|z\|) |
|---|---|---|---|---|---|---|
| (Intercept) | -1.291 | 1.277 | 0.275 | 0.022 | 3.369 | 0.3122 |
| age | 0.022 | 0.025 | 1.022 | 0.974 | 1.073 | 0.3714 |
| factor(race)White | -0.512 | 1.069 | 0.599 | 0.074 | 4.884 | 0.6320 |
| factor(race)Black | -0.165 | 1.154 | 0.848 | 0.088 | 8.165 | 0.8863 |
| factor(race)Asian | -0.064 | 1.084 | 0.938 | 0.112 | 7.874 | 0.9533 |
| factor(race)Hispanic | -2.582 | 1.481 | 0.076 | 0.004 | 1.382 | 0.0815 |
| education | -0.114 | 0.050 | 0.892 | 0.809 | 0.983 | 0.0217 |
| sex | 0.078 | 0.257 | 1.081 | 0.653 | 1.789 | 0.7617 |

From our logistic regression model fit with robust standard error estimates, we estimate that for two subgroups of the population with the same race, education level, and baby's sex, a one year difference in age between the groups is associated with a 1.022 (95% CI: 0.974, 1.073) times change in the odds of low birthweight. From the table above, we see this estimate is not statistically significant enough to reject the hypothesis that there is no association between lowe birthweight and age. The only covariate that was statistically significant was education, which decreased the odds of having a low birthweight child between groups (Odds 0.892, 95% CI: 0.809, 0.983).