Question 1

Joseph Froelicher

February 26, 2020

Part A

$$ln\left(\frac{p_i}{1-p_i}\right) = \mathbf{X}_j\beta$$

$$\hat{\beta_0} = ln\left(\frac{p(y=1|x_j=0)}{1-p(y=1|x_j=0)}\right) = ln\left(\frac{(619/2416)}{1-(619/2416)}\right) = -1.065769$$

$$\hat{\beta_1} = (\hat{\beta_0}+\hat{\beta_1}) - \hat{\beta_0} = ln\left(\frac{p(y=1|x_j=1)}{1-p(y=1|x_j=1)}\right) - \hat{\beta_0} = ln\left(\frac{(355/771)}{1-(355/771)}\right) - (-1.065769) = 0.9072015$$

$$\hat{\beta_2} = (\hat{\beta_0}+\hat{\beta_2}) - \hat{\beta_0} = ln\left(\frac{p(y=1|x_j=2)}{1-p(y=1|x_j=2)}\right) - \hat{\beta_0} = ln\left(\frac{(162/731)}{1-(162/731)}\right) - (-1.065769) = -0.1905151$$

Part B

$$L(Y_i|p) = \prod_{i=1}^{N}\binom{n_i}{Y_i}p^{Y_i}(1-p)^{n_i-Y_i}$$

$$L(Y_i|p) = \prod_{i=1}^{N}\binom{n_i}{Y_i}\left(\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{Y_i}\left(1-\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{n_i-Y_i}$$

$$L(Y_i|p) = \left(\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{\sum_{i=1}^{n}Y_i}\left(1-\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{n-\sum_{i=1}^{n}Y_i}$$

$$ln(L(Y_i|p)) = ln\left(\left(\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{\sum_{i=1}^{n}Y_i}\left(1-\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{n-\sum_{i=1}^{n}Y_i}\right)$$

$$l(Y_i|p) = \sum_{i=1}^{n}\left(ln\binom{n}{Y_i} + Y_i\,ln\left(\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right) + (n-Y_i)\,ln\left(1-\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)\right)$$

$$l(Y_i|p) = \left[ln\binom{n_1}{Y_1} + Y_1\,ln\left(\frac{e^{\hat{\beta_0}}}{1+e^{\hat{\beta_0}}}\right) + (n_1-Y_1)\,ln\left(1-\frac{e^{\hat{\beta_0}}}{1+e^{\hat{\beta_0}}}\right)\right]$$
$$+\ \left[ln\binom{n_2}{Y_2} + Y_2\,ln\left(\frac{e^{\hat{\beta_0}+X_1\hat{\beta_1}}}{1+e^{\hat{\beta_0}+X_1\hat{\beta_1}}}\right) + (n_2-Y_2)\,ln\left(1-\frac{e^{\hat{\beta_0}+X_1\hat{\beta_1}}}{1+e^{\hat{\beta_0}+X_1\hat{\beta_1}}}\right)\right]$$
$$+\ \left[ln\binom{n_3}{Y_3} + Y_3\,ln\left(\frac{e^{\hat{\beta_0}+X_2\hat{\beta_2}}}{1+e^{\hat{\beta_0}+X_2\hat{\beta_2}}}\right) + (n_3-Y_3)\,ln\left(1-\frac{e^{\hat{\beta_0}+X_2\hat{\beta_2}}}{1+e^{\beta_0+X_2\beta_2}}\right)\right]$$

By hand, the log likelihood of model 1 is -10.8699802. Using the `logLik()` function in R, we get the value of -10.8699802 for the likelihood of model 1.

Part C

$$ln(L(Y_i|p)) = ln\left(\left(\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{\sum_{i=1}^{n} Y_i}\left(1-\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)^{n-\sum_{i=1}^{n} Y_i}\right)$$

$$l(Y_i|p) = \sum_{i=1}^{n}\left(ln\binom{n}{Y_i} + Y_i\, ln\left(\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right) + (n-Y_i)\, ln\left(1-\frac{e^{\mathbf{X}\beta}}{1+e^{\mathbf{X}\beta}}\right)\right)$$

$$l(Y_i|p) = \left[ln\binom{n_1}{Y_1} + Y_1\, ln\left(\frac{e^{\hat{\beta_0}}}{1+e^{\hat{\beta_0}}}\right) + (n_1-Y_1)\, ln\left(1-\frac{e^{\hat{\beta_0}}}{1+e^{\hat{\beta_0}}}\right)\right]$$

By hand, the log likelihood of model 0 is -4.2654493. Using the `logLik()` function in R, we get the value of -4.2654493 for the likelihood of model 1.

Part D

H_0 : The null, intercept only model, sufficiently models the data. H_a : The full, covariate model, is significantly better at modeling the data.

$$T_{LR} = -2ln\left[\frac{L(p_{H_0})}{L(p_{MLE})}\right]$$

$$T_{LR} = -2ln\left[\frac{e^{-4.26545}}{e^{-10.86998}}\right]$$

$$X^2_{2,1-0.95} = 5.991$$

```
## [1] 1
```

$$P\{(T_{LR} = -13.209) > 5.991\} > 0.05$$

$$(T_{LR} = 131.0817) \not\sim X^2_{2,0.95}$$

Fail to reject the null hypothesis. The likelihood ratio test suggested that our full model, does not fit the data significantly better than the null model.

Part E

The predicted probability of a misconduct violation during the first year in prison for a prisoner with 1 strike is 0.3938046. The predicted probability of a misconduct violation during the first year in prison for a prisoner with 3 strikes is 0.4464219.

Part F

$$CI = [e^{\hat{\beta}-Z_{1-\frac{\alpha}{2}}\hat{SE}(\hat{\beta})}, e^{\hat{\beta}+Z_{1-\frac{\alpha}{2}}\hat{SE}(\hat{\beta})}]$$

CI = [0.9525757, 1.3490558]

Part G

Two reasons why model 1 is better than model 2: 1. We shouldn't try to model discrete observations as continuous variables, as we are doing in model 2. This could result in unexpected model behaviour. 2. The AIC for model 1 is significantly lower (> 10) than model 2. There is no indication that model 2 is a better fit for our data than model 1.

Appendix

```
# part b
data = data.frame(
  "strikes" = c("1", "2", "3"),
  "misconduct" = c(619, 355, 162),
  "no_misconduct" = c(1797, 416, 569)
)

# by hand
p1 = data[1, 2] / (data[1, 2] + data[1, 3])
p2 = data[2, 2] / (data[2, 2] + data[2, 3])
p3 = data[3, 2] / (data[3, 2] + data[3, 3])

n1 = data[1, 2] + data[1, 3]
n2 = data[2, 2] + data[2, 3]
n3 = data[3, 2] + data[3, 3]

c1 = data[1, 2]
c2 = data[2, 2]
c3 = data[3, 2]

w1 = data[1, 3]
w2 = data[2, 3]
w3 = data[3, 3]

ll_m1 = lchoose(n1, c1) + c1 * log(p1) + w1 * log(1 - p1)
ll_m2 = lchoose(n2, c2) + c2 * log(p2) + w2 * log(1 - p2)
ll_m3 = lchoose(n3, c3) + c3 * log(p3) + w3 * log(1 - p3)
like1_hand = sum(ll_m1, ll_m2, ll_m3)

model1 = glm(
  cbind(misconduct, no_misconduct) ~ strikes, data, family = binomial
)

like1_r = logLik(model1)

# part b
n0 = n1 + n2 + n3
c0 = c1 + c2 + c3
w0 = w1 + w2 + w3

model0 = glm(
  cbind(c0, w0) ~ 1, data, family = binomial
)

# by hand
p0 = exp(summary(model0)$coefficients[1, 1]) / (1 + exp(summary(model0)$coefficients[1, 1]))

like0_hand = lchoose(n0, c0) + c0 * log(p0) + w0 * log(1 - p0)

# check
like0_r = logLik(model0)
```

```r
# part b
t_lr = -2 * (like0_hand - like1_hand)
pchisq(t_lr, 2, lower.tail = F)

# numeric model 2
data_numeric = data.frame(
  "strikes" = c(1, 2, 3),
  "misconduct" = c(619, 355, 162),
  "no_misconduct" = c(1797, 416, 569)
)

model2 = glm(
  cbind(misconduct, no_misconduct) ~ strikes, data_numeric, family = binomial
)

summary(model2)

# part E
p_1 = summary(model2)$coefficients[1, 1] + summary(model2)$coefficients[2, 1]
p_3 = summary(model2)$coefficients[1, 1] + (3 * summary(model2)$coefficients[2, 1])

# part f
f = exp(p_3) / exp(p_1)
a = 0.05
z = qnorm(1 - (a / 2))
se = 2 * summary(model2)$coefficients[2, 2]
b = p_3 - p_1
ci = c(exp(b - (z * se)), exp(b + (z * se)))
```