HW-2: Question-1 R Notebook

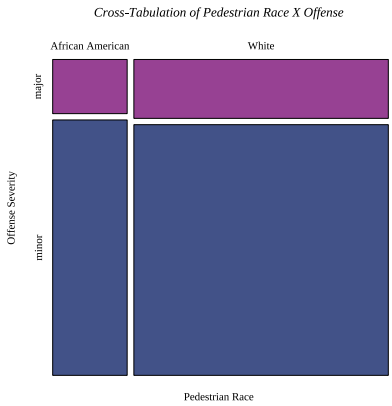
Rachel M. Smith

13 November 2016

Data Summary. Portland Police 2014 report on racial profiling in traffic stops.

Table 1: Contingency Table of Pedestrians Stopped by Portland Police: Pedestrian Race (2 Levels) x Severity of Offense (2 Levels).

	Major	Minor	π_{+j}
African American	21	78	99
White	99	332	431
π_{i+}	120	410	530



[1] *{Complete}*

1.*a*. *Loglinear Model*[1]

$$\log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$$

Loglinear Analysis Summary. A loglinear model was used to test the association between pedestrian race and whether individuals were stopped by Portland Police (*PPD*) for major versus minor offenses. 77.36% of pedestrians stopped by PPD were White, and 22.64% were African American ($N_{Total} = 530$). Among the 120 African American pedestrians, 3.96% (21) were stopped for major offenses, while 18.68% (99) were stopped for minor offenses. Among the 410 White pedestrians, 14.72% (78) were stopped for major offenses, while 62.64% (332) were stopped for minor offenses. Results from the likelihood ratio test of the $2x2$ contingency table (see above) indicated no significant differences regarding these factors, $G^2(1) = 0.14, p = 0.70$.

Call:

```
loglm(formula = ~ped.f + offense.f, data = tbl)
```

Statistics:

	X^2	df	P(> X^2)
Likelihood Ratio	0.1437	1	0.7046

	X^2	df	P(> X^2)
Pearson	0.142	1	0.7063

1.*b*. *Logistic Regression Model*[2]

[2] ~~[Complete]~~

$$\ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta X$$

$$\pi = \frac{e^{\alpha+\beta X}}{1+e^{\alpha+\beta X}}$$

Logistic Regression Analysis Summary. A logistic regression analysis was conducted to test whether pedestrian race[3] predicted whether individuals were stopped by Portland Police for *major*, versus *minor*, offenses. Results for the model indicated that the severity of offenses was not predicted according to the two levels of pedestrians' race, likelihood ratio $\chi^2(1) = 0.14$ (n.s.). The *Nagelkerke pseudo-*R^2 indicated that $< 1\%$ of the variance in offense severity was accounted for by the race of pedestrians stopped by PPD[4]. The model estimates and fit indices are summarized below.

[3] *African American* or *White*

[4] *pseudo-*$R_2 = 0.04\%$

Logistic Regression Model Summary Statistics and Fit Indices

Table 3: Logistic Regression Model Summary

	Estimate	SE	Z	p.value
(Intercept)	1.58	0.30	5.22	0.00
ped	-0.03	0.07	-0.38	0.71

Table 4: Likelihood Ratio χ^2

Log Likelihood	*df*	χ^2	*p*
-255	1	0.14	0.7

Confidence Intervals (CI) & Odds Ratios (OR)

Table 5: Logistic Regression Coefficients (β) & Coresponding Confidence Intervals (*CI*)

		CI_β	
	β	2.5 %	97.5 %
(Intercept)	1.5761	1.0111	2.2009
ped	-0.0255	-0.1634	0.1035

Table 6: Logistic Regression Odds Ratios (Φ) & Coresponding Confidence Intervals (*CI*) [1]

		CI_{Φ}	
	Φ	2.5 %	97.5 %
(Intercept)	4.8363	2.7485	9.0333
ped	0.9748	0.8492	1.1091

Note:
[1] Confidence intervals are based on the logistic regression model's profiled log-likelihood function, rather than the standard errors

Additional Model Fit Indices

Table 7: Logistic Regression Model Fit Statistics

	Estimate	Degrees of Freedom
Null Deviance	510.43	529
Residual Deviance	510.29	528
AIC	514.29	

1.*c*. *Previous Analysis & Results*[5]

[5] **[ToDo]**

Look back at your findings from the analysis of the contingency table in **HW-1 Problem-3c**. Briefly compare your results from HW-1 Problem-3c to the results from the loglinear analysis and the logistic regression model[6]. Refer to specific values when making your comparisons.

[6] *no more than 2-3 sentences is needed*

Contingency Table & Group Comparisons via Pearson χ^2[7]

[7] Original analysis conducted for ***Homework-1, Question-3c***