Combined Data Analysis for Spring 2016 and Fall 2016

Prepared by Ramjee Sharma

# Logistic Regression Analysis

In this analysis, we only consider students who either over estimate or under estimate. Dependent Variable=Behavior(overestimate=1, underestimate=0)-Binomial Variable

Independent variable/predictor variables:

Semester (Spring=1, Fall=2)-Categorical

CourseLevel(lower level courses=1, upper level courses=2)-Categorical

Gender (Female=1, Male=2)-Categorical

InstructorsGrade (Continuous Variable)-Continuous

## Analysis Result

Deviance Residuals:

Min 1Q Median 3Q Max

-2.966 -0.861 0.261 0.868 1.917

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 8.60209 0.52607 16.35 < 0.0000000000000002 \*\*\*

Gender1 0.14224 0.10421 1.36 0.17

Semester2 0.81653 0.17775 4.59 0.0000044 \*\*\*

Grade -0.11339 0.00736 -15.40 < 0.0000000000000002 \*\*\*

CourseLevel2 0.43345 0.10818 4.01 0.0000616 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2998.5 on 2235 degrees of freedom

Residual deviance: 2243.6 on 2231 degrees of freedom

AIC: 2254

Number of Fisher Scoring iterations: 5

## Interpretation:

When the gender2 (female) changes to gender1(male), the log(probability of under estimate) increases by 0.14, but this is not significant as the p value is more than 0.05. This means, the gender analysis is not suitable for the logistic regression. When the grade increases by 1, the log(probability of under estimate) decreases by 0.11. When the course level2 (higher level) changes into course level1 (lover level), the log(probability of under estimate) increases by 0.4. When the semester2(Fall) changes to semester1(spring), the log(probability of under estimate) increases by 0.8.

# Logistic Regression Analysis for Fall 2016 Calculus II

We have used the following quantification for the categories:

Binomial variable= (over estimate=1, under estimate=0)

Continuous variable=Instructor’s Grade

Categorical Variables:

Gender= (Female=0, Male=1)

Category =(KK=1,NKK=2, KNK=3, NKNK=4, KBNKW=5, Others=6)

Required= (Yes(1),No(0))

First time= (Yes(1), No(0))

Preparation= (1,2,3,4,5)

Perceived difficulty level= (1,2,3,4,5)

Expected grade= (A=1, B=2, C=3, D=4, F=5)

Taken upper level courses =Yes(1), No(0)

Actual Grades= (A=1, B=2, C=3, D=4, F=5)

This analysis finds the log(probability of over estimate) using the logistic regression model.

Call:

glm(formula = Behavior ~ Grade + Gender + Category + Required +

FirstTime + Preparation + PerceivedDiffLevel + ExpGrade +

TakeUpDiv + ActualCourseGrade, family = "binomial")

Deviance Residuals:

Min 1Q Median 3Q Max

-2.5866 -0.3581 -0.0002 0.3111 2.4491

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 40.8035 2694.6978 0.02 0.98792

Grade -0.1119 0.0235 -4.77 0.0000019 \*\*\*

Gender1 1.9126 0.5968 3.20 0.00135 \*\*

Category2 -3.7246 1.1780 -3.16 0.00157 \*\*

Category3 -4.0944 1.1821 -3.46 0.00053 \*\*\*

Category4 -0.5375 0.8079 -0.67 0.50583

Category5 -4.1313 1.1598 -3.56 0.00037 \*\*\*

Category6 -17.6711 2781.9914 -0.01 0.99493

Required1 -14.5990 1960.9897 -0.01 0.99406

FirstTime1 -1.2223 0.8454 -1.45 0.14821

Preparation3 1.4907 2.0980 0.71 0.47736

Preparation4 0.5428 1.9490 0.28 0.78063

Preparation5 1.0955 2.2771 0.48 0.63046

PerceivedDiffLevel3 -16.5834 1848.2174 -0.01 0.99284

PerceivedDiffLevel4 -15.9075 1848.2171 -0.01 0.99313

PerceivedDiffLevel5 -16.4217 1848.2172 -0.01 0.99291

ExpGrade2 -0.7262 0.7001 -1.04 0.29963

ExpGrade3 -3.4593 1.5797 -2.19 0.02853 \*

TakeUpDiv1 -0.5957 0.6843 -0.87 0.38399

ActualCourseGrade2 0.2495 0.8437 0.30 0.76745

ActualCourseGrade3 -0.0550 0.6740 -0.08 0.93497

ActualCourseGrade4 0.3389 1.3010 0.26 0.79447

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 293.88 on 211 degrees of freedom

Residual deviance: 113.13 on 190 degrees of freedom

(60 observations deleted due to missingness (Difference=0)

AIC: 157.1

Number of Fisher Scoring iterations: 16