



IIT ROORKEE



NPTEL ONLINE
CERTIFICATION COURSE

LOGISTIC REGRESSION PART-7

LECTURE 52

DR. GAURAV DIXIT

DEPARTMENT OF MANAGEMENT STUDIES



LOGISTIC REGRESSION

- Linear Regression for a categorical outcome variable?
 - Can be done by treating the outcome variable as continuous and coding it numerically
 - However, anomalies will lead to spurious modeling
 - Predictions can take any value, not just dummy values $\{0,1\}$
 - Outcome variable or residuals don't follow normal distribution
 - binomial distribution
 - Variance of outcome variable is not constant across all records (violation of homoscedasticity)
 - $np(1-p)$

LOGISTIC REGRESSION

- Logistic Regression for Profiling Task
 - Apart from model performance on validation partition
 - Model's fit to data is assessed on training partition
 - However, still avoid overfitting
 - Usefulness of predictors is examined
 - Goodness of fit metrics
 - Overall fit of the model
 - Deviance (equivalent to SSE in linear regression)
 - 1 - Deviance/Null Deviance (equivalent to multiple R^2 in linear regression)
 - Single predictors



LOGISTIC REGRESSION

- Outcome variable with m classes ($m > 2$)
 - Multinomial logistic regression
 - Separate binary logistic regression model for $m-1$ classes (one class is treated as reference class)
 - Ordinal logistic regression
 - Large no. of ordinal classes: treat ordinal variable as continuous variable and apply multiple linear regression



LOGISTIC REGRESSION

- Outcome variable with m classes ($m > 2$)
 - Ordinal logistic regression
 - Small no. of ordinal classes: Proportional odds or cumulative logit method
 - Separate binary logistic regression model for m-1 cumulative probabilities

For a three class case: C1, C2, and C3 and a single predictor x_1

$$\begin{aligned} \text{logit}(C1) &= \alpha_0 + \beta_1 x_1 \\ \text{logit}(C1 \text{ or } C2) &= \beta_0 + \beta_1 x_1 \end{aligned}$$

- RStudio



Key References

- Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data by EMC Education Services (2015)
- Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner by Shmueli, G., Patel, N. R., & Bruce, P. C. (2010)

Thanks...

