

Section 4F. Logistic Regression with More Classes

Statistics for Data Science

Victor M. Preciado, PhD MIT EECS
Dept of Electrical & Systems Engineering
University of Pennsylvania
preciado@seas.upenn.edu

Logistic Regression: More than two classes

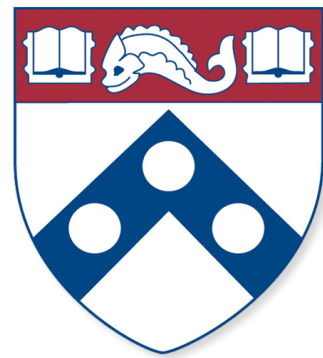
So far, we have considered the case when the output classes are $\mathcal{C} = \{0, 1\}$. How about if we have K classes?

- ▶ We can use a multiclass logistic regression

$$\Pr(Y = k|X) = \frac{e^{\beta_{0k} + \beta_{1k}X_1 + \dots + \beta_{pk}X_p}}{\sum_{\ell=1}^K 1 + e^{\beta_{0\ell} + \beta_{1\ell}X_1 + \dots + \beta_{p\ell}X_p}}$$

- ▶ We can find estimates $\hat{\beta}_{0\ell}, \hat{\beta}_{1\ell}, \dots, \hat{\beta}_{p\ell}$ for $\ell = 1, \dots, K$ using the Maximum Likelihood Criterion
- ▶ We can make predictions for an input \mathbf{x} using the functions

$$\hat{p}_k(\mathbf{x}) = \frac{e^{\hat{\beta}_{0k} + \hat{\beta}_{1k}x_1 + \dots + \hat{\beta}_{pk}x_p}}{\sum_{\ell=1}^K 1 + e^{\hat{\beta}_{0\ell} + \hat{\beta}_{1\ell}x_1 + \dots + \hat{\beta}_{p\ell}x_p}}$$



Penn Engineering

Copyright 2020 University of Pennsylvania
No reproduction or distribution without permission.