Section 4K. Errors in LDA

Statistics for Data Science

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

Classification Errors: LDA Theory (2)

What are the expected value of |FP| and |FN|?

$$\mathbb{E}\left[|\mathrm{FP}|\right] = \sum_{i=1}^{N} \mathbb{E}_{XY}\left[\delta\left(C\left(X\right) - 1\right)\delta\left(Y\right)\right]$$

$$\text{(Total Expectation)} = \sum_{i=1}^{N} \mathbb{E}_{X|Y}\left[\delta\left(C\left(X\right) - 1\right)1|Y = 0\right]\pi_0 + \mathbb{E}_X\left[0|Y = 1\right]\pi_1$$

$$= \sum_{i=1}^{N} \int_{x\in\mathbb{R}} \delta\left(C\left(x\right) - 1\right) f_{X|Y}\left(x|0\right)\pi_0 dx$$

$$= N\int_{x\geq\gamma} \pi_0 f_0\left(x\right) dx$$

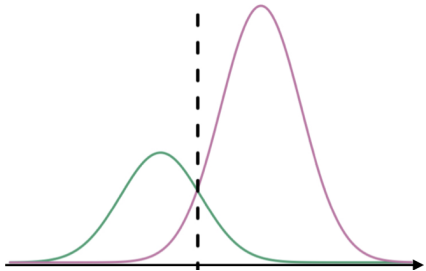
$$\mathbb{E}\left[|\mathrm{FN}|\right] = \text{(similar derivation)} = N\int_{x<\gamma} \pi_1 f_1\left(x\right) dx$$

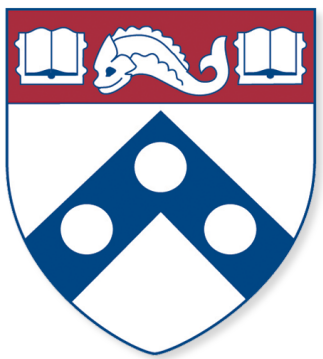
Classification Errors: LDA Theory (3)

Graphical Interpretation:

$$\mathbb{E}\left[|\mathsf{FP}|\right] = N \int_{x \geq \gamma} \pi_0 f_0\left(x\right) dx$$

$$\mathbb{E}\left[|\mathsf{FN}|\right] = N \int_{x < \gamma} \pi_1 f_1\left(x\right) dx$$





Penn
Engineering