Flight Delay Prediction

Sirjan Kafle
sxkafle12@gmail.com
https://flightdelay.us

1 Introduction

TODO: write intro.

2 Definitions

TODO: write definitions

3 Modeling Conditional Distributions as an Exponential

3.1 Motivating Examples

TODO: show plots - full and conditioned on $D > 0$.

3.2 Proposed Model

$$f_D(\delta|\{C_i\}) = p\lambda e^{-\lambda\delta}$$

3.3 Maximum Likelihood Estimate of Parameters

Learn p, λ.

$$p = \sum_{i=1}^{n} 1_{\delta_i>0}/N$$
$$\lambda = N/\sum_{i=1}^{n} \delta_i 1_{\delta_i>0}$$

3.4 MLE Plot and Goodness of Fit

3.4.1 Metric

3.4.2 Example Fits

3.4.3 Score vs Data Size

3.4.4 Score vs Conditional Group Profile

3.4.5 Bias Variance Tradeoff in Action

3.5 Delay Survival Function

$$P(D > \delta|\{C_i\}) = pe^{-\lambda\delta}$$

4 Algorithms for Low Data Support Conditionals

4.1 Core Problem and Motivating Example

4.2 Definitions of Subset Groupings and Partition

4.3 Assumption and Setup

4.4 Parameter Estimate Derivation Given Partition

4.5 Which Partition is Best?

4.5.1 Main Idea

4.5.2 KL Divergence Derivation

4.5.3 Ranking of Partitions Algorithm

5 Final Algorithm

5.1 Model Fitting

Defer explanation of the partitions algorithm to earlier.

5.2 Inference

6 Fun Examples

Attach images from the site