Introduction and Motation

Machine learning methods have proven useful in the field of fraud detection (reference?). It is crucial that the man and machine collaboration is considered in conjunction - a computer cannot determine if a certain event was fraudulent or not, it can only suggest which events to further inspect. The need for human interaction limits the capability of the algorithm, and so the events that should be inspected need to be chosen optimally. We propose a definition of the strategy a company should adopt in order to maximize their fraud detection, and how this strategy is influenced by changes in the underlying fraud drivers.

Set-up

$\mathcal{I}$ set of events. Denote by X_i an indicator describing whether the i^{th} event is fraudulent - could be insurance claims, credit card transactions, stock tradings or any other single event that can be fraudulent or not. Discrete time-points τ, should be thought of as the smallest of the time it takes to evaluate k policies, and the time it takes for $\mathcal{I}$ to be updated.

Example 0.1.
Let the number of events be $\#\mathcal{I} = 2n$, whereof n events are of type 1 and the rest are of type 2. $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$. Assume that

$$P(X_i = 1) = \begin{cases} p_1 & \text{for} \quad i \in \mathcal{I}_1 \\ p_2 & \text{for} \quad i \in \mathcal{I}_2 \end{cases}$$

$\triangle$

Definition 0.1.
Strategy $\mathcal{K}$ is optimal if

$$\mathcal{K}_\tau = \arg\sup_K \mathrm{E}\left[\sum_{i \in K} X_\tau^i\right]$$