Calibrated Losses for the Abstain Property

Jessie Finocchiaro, Rafael Frongillo, Bo Waggoner

Boulder

Overview

Machine learning is often used to make predictions in *high-risk settings*. When given a data point $x \in \mathcal{X}$ with label $y \in \mathcal{Y}$, we want to predict the most likely outcome. However, if $Pr[Y = y \mid X = x]$ has high uncertainty, one might want their algorithm to defer to a human expert for classification. We study the **abstain property**, which yields a prediction in high-confidence settings and deferral to human expert in low-confidence settings.

Setting

$\mathcal{Y}$	Finite outcome set
$\mathcal{R} := \mathcal{Y} \cup \{\bot\}$	Report set: outcome or abstain
$\ell: \mathcal{R} \to \mathbb{R}^n_+$	Discrete loss
$L: \mathbb{R}^d \to \mathbb{R}^n_+$	Surrogate loss
$p \in \Delta_{\mathcal{Y}}$: $\langle p, L(u) \rangle$	Expected loss
$\psi: \mathbb{R}^d \to \mathcal{R}$	Link function

Properties and Calibration

Definition 1: The **abstain property**

$$\gamma(p) = \begin{cases} \operatorname{argmax}_y p_y & \max_y p_y \geq \frac{1}{2} \\ \bot & otherwise \end{cases}$$

Definition 2: We say a loss L **elicits** the property γ if, for all $p \in \Delta_{\mathcal{Y}}$, we have

$$\gamma(p) = \arg\min_{r \in \mathcal{R}} \langle p, L(r) \rangle.$$

Definition 3: Let original loss ℓ eliciting Γ, proposed surrogate L, and link function ψ be given. We say (L, ψ) is **calibrated** with respect to ℓ if, for all $p \in \Delta_{\mathcal{Y}}$,

$$\inf_{u \in \mathbb{R}^d : \psi(u) \notin \Gamma(p)} \langle p, L(u) \rangle > \inf_{u \in \mathbb{R}^d} \langle p, L(u) \rangle.$$

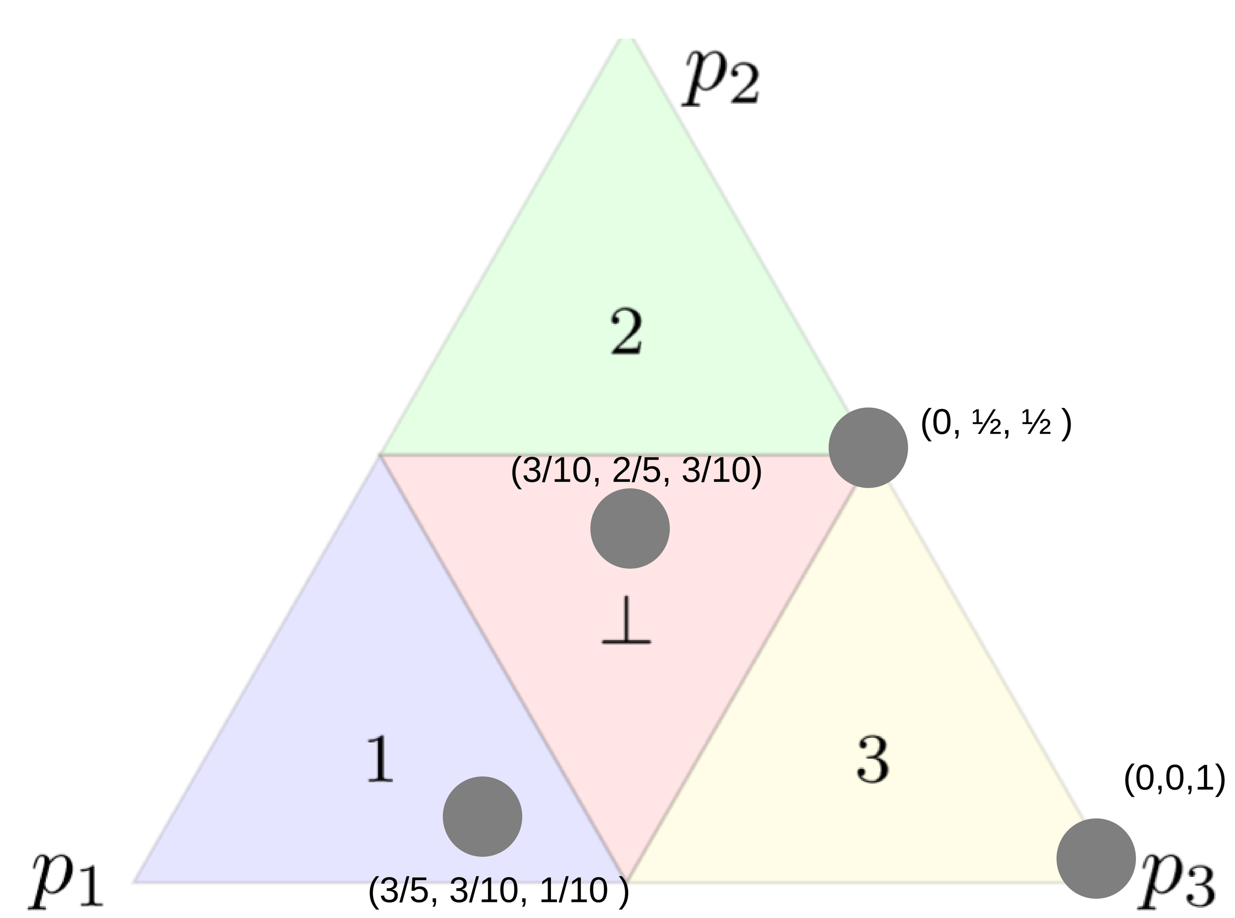
Discrete loss for abstain property

$$\ell(r, y) = \begin{cases} 0 & r = y \\ \frac{1}{2} & r = \bot \\ 1 & r \neq y, r \neq \bot \end{cases}$$

The problem

We cannot optimize ℓ. Is there an efficient surrogate with good statistical guarantees that can be linked back to the correct prediction?

Drawing the abstain property: n = 3



Main result

For any discrete loss (including abstain), there is a calibrated (surrogate, link) pair where the surrogate is convex.

Efficiency

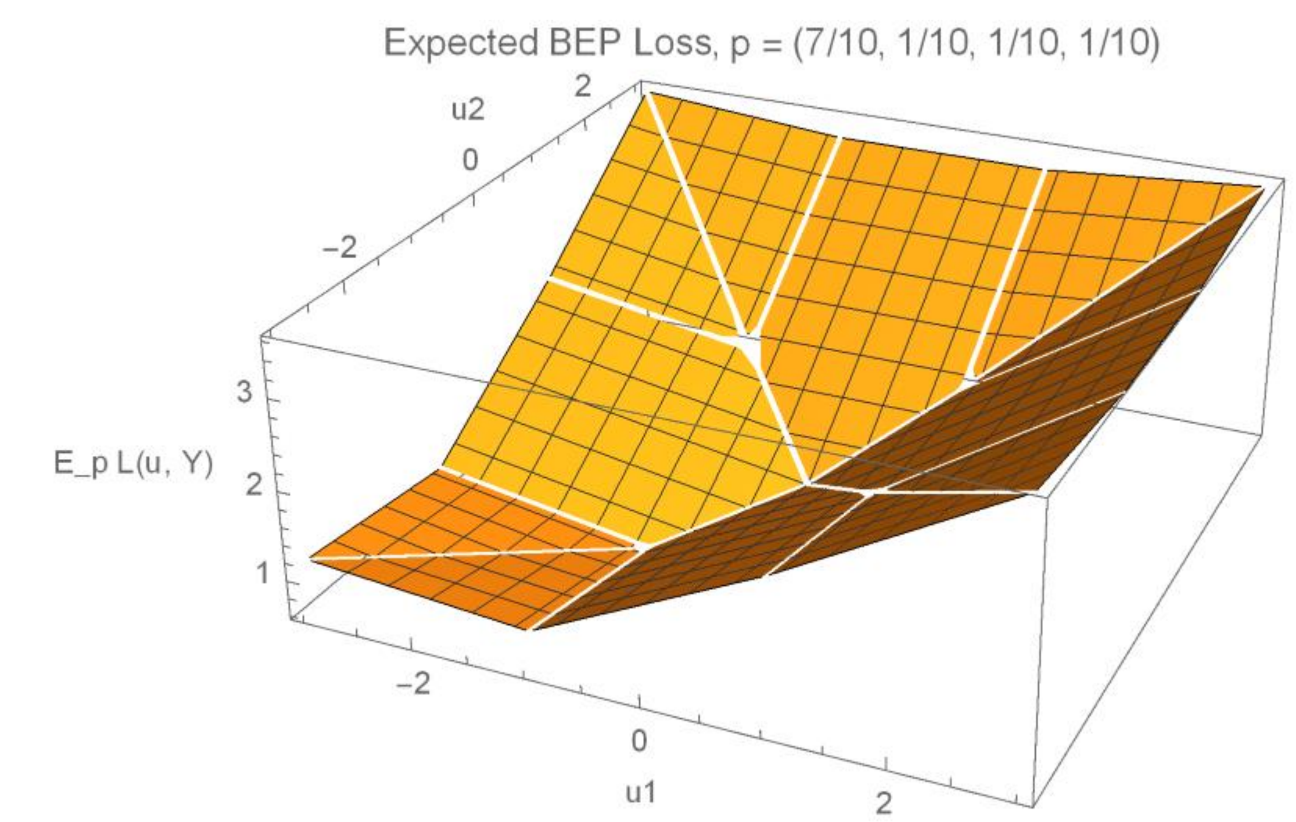
We want a surrogate $L: \mathbb{R}^d \to \mathbb{R}^n_+$ that is calibrated with respect to $\ell: \mathcal{R} \to \mathbb{R}^n_+$

Efficiency: *d* is small (relative to *n*)

Efficient surrogates reduce the dimension of the optimization problem.

Surrogate loss for abstain property: n = 4

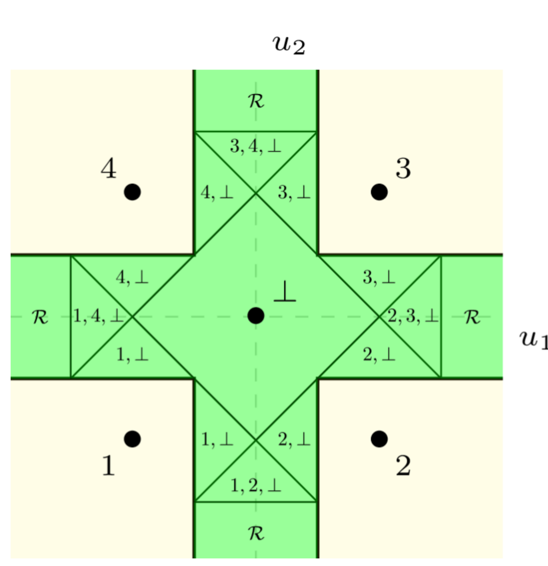
d = log(n)



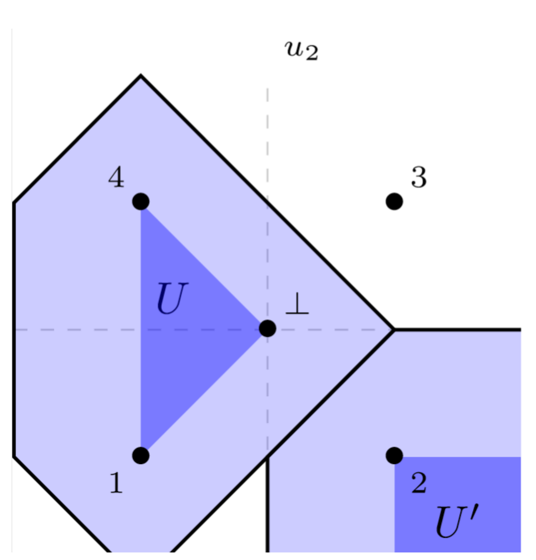
BEP Surrogate $L(u, y) = (\max_{j \in [d]} \mathrm{B_j}(y)\mathrm{u_j} + 1)_+$ is calibrated for abstain loss.

(Ramaswamy, Tewari, Agarwal. (2018.) Consistent algorithms for multiclass classification with an abstain option. In *Electronic Journal of Statistics*)

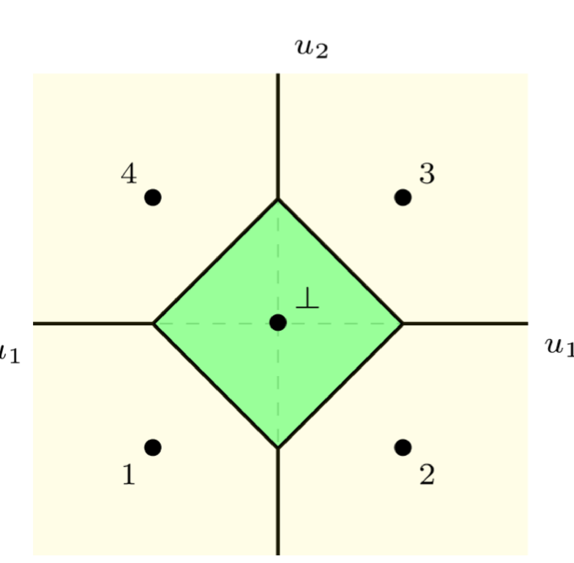
Links for abstain surrogate



Possible calibrated link values by constructing link with $\|\cdot\|_\infty$ and $\epsilon = 1/2$.



Examples of U sets that are used to calculate the calibrated link for the BEP embedding.



Calibrated link using $\|\cdot\|_1$ and $\epsilon = 1$.