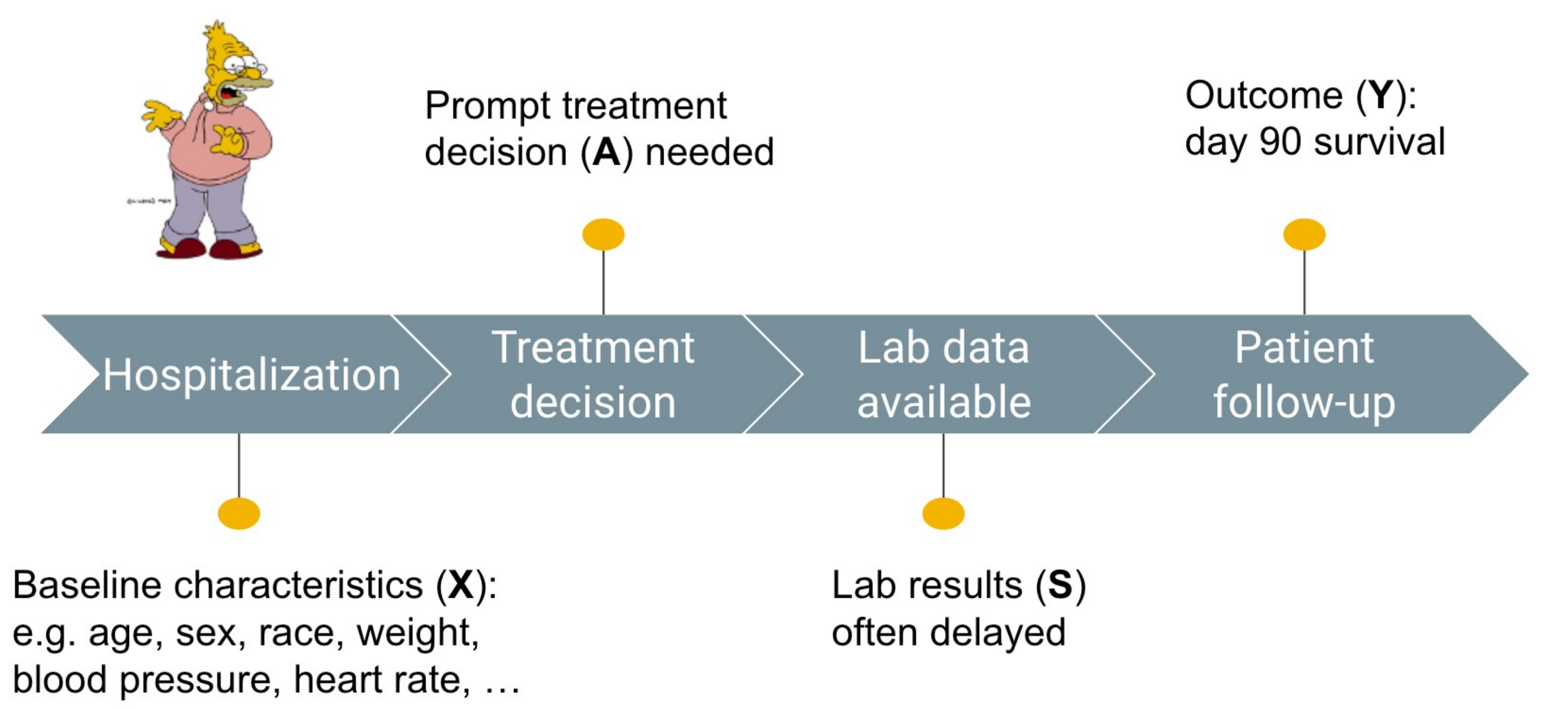
RISE: Robust Individualized Decision Learning with Sensitive Variables

Xiaoqing Tan (xit31@pitt.edu)
Zhengling Qi @ GWU
Christopher W. Seymour
Lu Tang
University of Pittsburgh

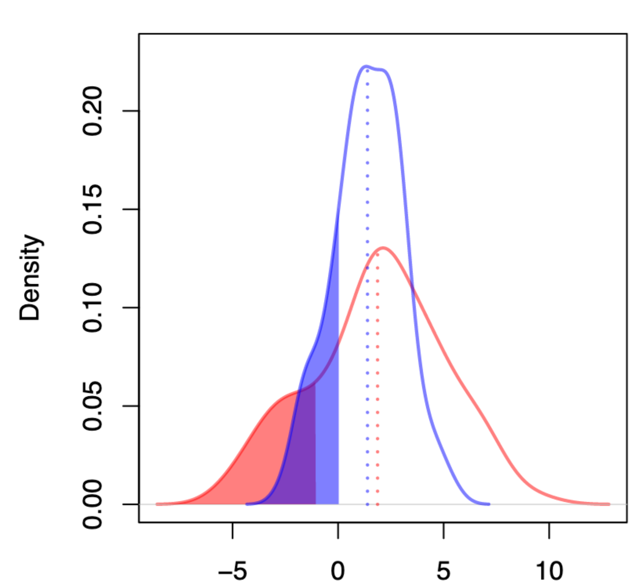
Motivation



- When deriving individualized decision rules (IDRs) in an *offline* environment, some variables are important to the intervention, while their inclusion in the decision rule is prohibited
  - **Scenario 1 – Delayed availability:** lab results for patients in life-threatening conditions
  - **Scenario 2 – Fairness concerns:** sensitive characteristics of subjects e.g., income, sex, and race
- We define ***sensitive variables*** as variables whose inclusion into decision rules is prohibited

Methods to deal with sensitive variables

- **Naive approaches:** omit sensitive variables S, introduce bias in finding optimal decision rule from the causal perspective
  - $\underset{d\in\mathcal{D}}{\text{maximize}}\, E_X$ {outcome under d}
- **Mean-optimal approaches:** under value function framework (Manski, 2004; Qian & Murphy, 2011)
  - $\underset{d\in\mathcal{D}}{\text{maximize}}\, E_X$ {expected outcome caused by S}
- **RISE (proposed):** robust quantile/infimum-optimal objective, consider worst-case outcomes of individuals
  - $\underset{d\in\mathcal{D}}{\text{maximize}}\, E_X$ {worst−case outcome caused by S}



Density of outcome given X
Red: mean-optimal rule
Blue: RISE

An illustrative example

Binary actions $A \in \{-1,1\}$. Note that the decision can only be made based on baseline covariate X whereas S is a sensitive variable.

Table 1: Toy example setup.

	$X \leq 0.5$		$X > 0.5$	
$E(Y\|X,S,A)$	$S=0$	$S=1$	$S=0$	$S=1$
$A=-1$	11	13	5	27
$A=1$	30	0	15	13

Table 2: Toy example results.

	Average reward	
	Overall	Vulnerable
Mean-optimal rule	**14.4**	7.1
RISE	13	**14**

Red: vulnerable subjects (those with low outcome values given X)
Blue: The worst-case outcomes of the rule by the proposed RISE

For X ≤ 0.5,

- Mean-optimal rule gives A = 1, as it achieves the largest average reward across S = 0 and S = 1.
- However, this greatly harms subjects with S = 1 as they could get the worst expected outcome of 0.
- RISE gives A = -1, which improves the worst-case outcome of these vulnerable subjects

RISE achieves a larger reward among vulnerable subjects while maintaining a comparable overall expected reward

Robust Optimality with Sensitive Variables

The proposed RISE estimates the following IDR

$$d^* \in \arg\max_{d\in\mathbb{D}} E_X\left[G_{S|X}\{E(Y|X,S,A=d(X))\}\right]$$

where $G_{S|X}(\cdot)$ could be chosen as some risk measure for evaluating $E(Y|X,S,A=d(X))$ for each $S \in \mathbb{S}$, e.g., variance, quantiles

For discrete *S* (infimum): find A with the best worst-case scenario among all possible values of S for every $X \in \mathbb{X}$

$$d^* \in \text{argmax}_{\mathbb{D}} E_X\left[\inf_{s\in\mathbb{S}}\{E(Y|X,S=s,A=d(X))\}\right]$$

$$d^*(X) \in \text{sign}(\inf_{s\in\mathbb{S}}\{E(Y|X,S=s,A=1) - \inf_{s\in\mathbb{S}}\{E(Y|X,S=s,A=-1)\})$$

For continuous *S* (conditional quantile): find A with the largest τ-th quantile of the outcome over the distribution related to S

$$d^* \in \text{argmax}_{\mathbb{D}} E_X\left[Q^{\tau}_{S|X}\{E(Y|X,S,A=d(X))\}\right]$$

$$d^*(X) \in \text{sign}(\{Q^{\tau}_{S|X}\{E(Y|X,S,A=1)\} - Q^{\tau}_{S|X}\{E(Y|X,S,A=-1)\})$$

Estimation and Algorithm

Estimate $d^*(X)$ with a classification-based optimization framework

Algorithm 1 RISE (Robust individualized decision learning with sensitive variables)

Input Training data $\mathcal{D}_n = \{Y_i, A_i, X_i, S_i\}_{i=1}^n$
Output Estimated decision rule $\hat{d}$

1: $\hat{Y}_i(x_i, s_i, a_i) \leftarrow$ Model $E(Y|X,S,A=a)$ using $\mathcal{D}_n$ with $a=1$ and $a=-1$, respectively.
2: **if** S is continuous **then**
3: $g_1(x_i) \leftarrow$ Model $Q_{S|X,A}\{E(Y|X,S,A=a)\}$ via quantile regressions of $\hat{Y}_i(x_i,s_i,a_i)$ on x_i, for $\mathcal{D}_n$ with $a=1$.
4: $g_2(x_i) \leftarrow$ Model $Q_{S|X,A}\{E(Y|X,S,A=a)\}$ via quantile regressions of $\hat{Y}_i(x_i,s_i,a_i)$ on x_i, for $\mathcal{D}_n$ with $a=-1$.
5: **if** S is discrete **then**
6: $g_1(x_i) \leftarrow$ Compute $\inf_{s\in\mathbb{S}}\{\hat{Y}_i(x_i,s,a_i=1)\}, \forall i$.
7: $g_2(x_i) \leftarrow$ Compute $\inf_{s\in\mathbb{S}}\{\hat{Y}_i(x_i,s,a_i=-1)\}, \forall i$.
8: $\hat{d} \leftarrow$ Build a weighted classification model with features x_i, label $\text{sgn}\{g_1(x_i) - g_2(x_i)\}$, and sample weight $|g_1(x_i) - g_2(x_i)|$.
9: **Return** $\hat{d}$

Real-data Applications

Three real-world examples from fairness and safety perspectives
(a) Fairness in a job training program (LaLonde, 1986)
(b) Improvement of HIV treatment (Hammer et al., 1996)
(c) Safe resuscitation for patients with sepsis (Seymour et al., 2016

Dataset	IDR	Obj. (all)	Obj. (vulnerable)	Value (all)	Value (vulnerable)
NSW *log(income+1)*	Base	5.26 (0.04)	5.28 (0.05)	6.32 (0.05)	6.33 (0.07)
	Exp	5.22 (0.04)	5.24 (0.05)	6.37 (0.05)	6.37 (0.07)
	RISE	**5.43** (0.04)	**5.44** (0.04)	**6.42** (0.04)	**6.42** (0.06)
ACTG175 *CD4 T-cell amount*	Base	336.9 (1.65)	338.1 (2.23)	350.5 (1.86)	357.5 (2.24)
	Exp	337.5 (1.65)	338.9 (1.80)	**351.9** (1.95)	359.1 (2.21)
	RISE	**351.5** (1.67)	**351.2** (1.80)	351.8 (1.88)	**363.1** (2.19)
Sepsis *survival rate*	Base	0.752 (0.001)	0.721 (0.001)	0.965 (0.001)	0.905 (0.002)
	Exp	0.752 (0.001)	0.721 (0.002)	0.966 (0.001)	0.908 (0.002)
	RISE	**0.771** (0.001)	**0.735** (0.001)	**0.972** (0.001)	**0.923** (0.002)

Main Contributions

- We propose a novel framework to handle sensitive variables in **causality-driven decision making**. Robustness is introduced to improve the worst-case outcome caused by sensitive variables
- To the best of our knowledge, we are among the first to propose a **robust-type fairness criterion under causal inference**
- We introduce a flexible **classification-based optimization framework** that can easily leverage most existing classification tools
- We illustrate the robustness of RISE using **three real-world examples from fairness and safety perspectives**