Supplementary Material

This is the supplement to 'Counterfactual Fairness Using Principal Stratification'. In this material, we provide the detailed proofs of theorems and propositions in the main text, as well as the implementation details.

A RULES FOR DISCRETIZATION

In Section 4, we use mat-data and por-data to test our methods. The pre-processing procedure includes discretizing continuous variables and collapsing the value space of discrete variables. The detailed rules are summarized in Table 1.

Table 1: The detailed pre-processing rules and the related variables.

Variable	Meaning	Original values	After pre-processing
AGE	student's age	from 15 to 22	from 15 to 18, or >18
TRAVELTIME	home to school travel time	1: <15 min., 2: 15 to 30 min., 3: 30 min. to 1 hour, or 4: >1 hour	1: <15 min., 2: 15 min. to 1 hour, or 3: >1 hour
MEDU	mother's education	0: none, 1: primary, 2: 5th to 9th grade, 3: secondary, or 4: higher education	1: primary or none, 2: 5th to 9th grade, 3: secondary, or 4: higher education
FEDU	father's education	0: none, 1: primary, 2: 5th to 9th grade, 3: secondary, or 4: higher education	1: primary or none, 2: 5th to 9th grade, 3: secondary, or 4: higher education
ABSENCE	number of school absences	from 0 to 93	1: <3 classes, 2: 3 to 5 classes, 3: >5 classes
G3	final grade	from 0 to 20	1: top 20% students, 0: others

B PROOFS

B.1 PROOF OF THEOREM 2

The following lemma [Shapiro, 1991] gives suffcient conditions under which the optimal solution of an estimated convex problem converges at $\sqrt{n}$ rates to the optimal solution of the target convex program.

Lemma (Shapiro, 1991) *let $\mathscr{X}$ be a compact subset of $\mathbb{R}^k$, consider the quantity $x^*(f)$ defined as the solution to the following convex optimization program:*

$$x^*(f) = \arg\min_{x\in\mathscr{X}} f(x)$$

suppose for objective function $f(x)$, there exists a positive constant α such that

$$f(x) \geqslant \inf_{x\in S} f(x) + \alpha\left[\mathrm{dist}\,(x, x^*(f))\right]^2$$

and function $\delta(x) = \xi(x) - f(x)$ is Lipschitz continuous on Θ and let W be a convex neighborhood of corresponding optimal set $x^*(f)$, if $\bar{x}(\xi) \in W$, then

$$dist(\bar{x}(\xi), x^*(f)) \leq \alpha^{-1}\kappa(\delta),$$

where

$$\kappa(\delta) = \sup\left\{\frac{|\delta(x)-\delta(y)|}{\|x-y\|} : x \in x^*(f), y \in \mathscr{X}\cap W, x \neq y\right\}$$

Notice that the linear program is equivalent to

$$\begin{aligned}
&\varepsilon^* = \textstyle\sum_{a,b}\varepsilon_{ab}\\
&+\textstyle\sum_{s_0,y_0}(\sum_{a,b} w_{s_0,a,y_0,b} - (\varepsilon_{0s_0}\hat{p}'_{(1-s_0)y_0 0} + (1-\varepsilon_{0(1-s_0)})\hat{p}'_{s_0y_0 0}))^2\\
&+\textstyle\sum_{s_1,y_1}(\sum_{a,b} w_{a,s_1,b,y_1} - (\varepsilon_{1s_1}\hat{p}'_{(1-s_1)y_1 1} + (1-\varepsilon_{1(1-s_1)})\hat{p}'_{s_1y_1 1}))^2\\
&f(w) = 0,\\
&w_{s_0s_1y_0y_1} \geq 0,\\
&\varepsilon_{ab} \geq 0\\
&\text{for } a,b,s_0,s_1,y_0,y_1 \in \{0,1\}.
\end{aligned}$$

using this lemma, we choose $f(x)$ is the objetive function, and under the assumption that $||p'_{sya} - \hat{p}'_{sya}|| = O_P(1/\sqrt{n})$, we have $\kappa(\delta) = O_p(n^{-1/2})$, conseqently, we have $\hat{x}_n$ converges to $x^*(f)$ at a rate of $O_p(n^{-1/2})$

B.2 PROOF OF THEOREM 3

For the true linear programming:

$$\begin{aligned} & \alpha^* = \min_w g(w) \\ \text{s.t.} \quad & \textstyle\sum_{a,b} w_{s_0,a,y_0,b} = \varepsilon^*_{0s_0} p'_{(1-s_0)y_0 0} + (1-\varepsilon^*_{0(1-s_0)}) p'_{s_0 y_0 0}, \\ & \textstyle\sum_{a,b} w_{a,s_1,b,y_1} = \varepsilon^*_{1s_1} p'_{(1-s_1)y_1 1} + (1-\varepsilon^*_{1(1-s_1)}) p'_{s_1 y_1 1}, \\ & w_{s_0 s_1 y_0 y_1} \geq 0, \\ & \text{for } a,b,s_0,s_1,y_0,y_1 \in \{0,1\}. \end{aligned}$$

From the definition of ε^*, it's straightforward to get $\alpha^* = 0$. When we use the estimated $\hat{\varepsilon}^*$ to replace α^*, we will have $|\hat{\alpha}^* - \alpha^*| = O_p(1/n^{-1/2})$ using the same technique of Mishler and Kennedy [2020]. As $\alpha^* = 0$, we get

$$||\hat{\alpha}^*||_1 = O_P(1/\sqrt{n}).$$

So we prove the Theorem 3.

References

Alan Mishler and Edward H Kennedy. Fairness in risk assessment instruments: Post-processing to achieve counterfactual equalized odds. *arXiv preprint arXiv:2009.02841*, 2020.

Alexander Shapiro. Asymptotic analysis of stochastic programs. *Annals of Operations Research*, 30(1):169–186, 1991.