Theorem 0.1. *For any joint distribution P of (X, S, Y), any $\sigma > 0$, as well as the base quantile estimator $\hat{q}_\alpha : \mathbb{R}^p \times [K] \to \mathbb{R}$ constructed on labeled data, the estimator $\hat{g}_\alpha$ defined in Eq. (8) satisfies*

$$(\hat{g}_\alpha(X, S) \mid S = s) \stackrel{d}{=} (\hat{g}_\alpha(X, S) \mid S = s') \quad \forall s, s' \in [K].$$

1 Appendix

1.1 Proof of Thereom 0.1

Before showing the exact DP guarantee, we utilize Lemma 1.1 (stated in Lemma 3) from Chzhen & Schreuder (2022).

Lemma 1.1. *Let $V_1, \ldots, V_n, V_{n+1}, n \geq 1$ be exchangeable real-valued random variables and U distributed uniformly on $[0, 1]$ be independent from $V_1, \ldots, V_n, V_{n+1}$, then the constructed location statistic*

$$T(V_1, \ldots, V_n, V_{n+1}, U) = \frac{1}{n+1} \sum_{i=1}^{n} \mathbb{I}\{V_i < V_{n+1}\} + \frac{1}{n+1} U \left(1 + \sum_{i=1}^{n} \mathbb{I}\{V_i = V_{n+1}\}\right)$$

is distributed uniformly on $[0, 1]$, where $\mathbb{I}$ is the indicator function.

The proof of Theorem 0.1 is a direct adaptation of quantile demographic parity guarantee from Liu et al. (2022), Chzhen & Schreuder (2022) for the Therom 2 and therom 7.2.

Proof of Theorem 0.1. To prove the claimed DP guarantee for fixed quantile level $\alpha \in \{\alpha_{\text{lo}}, \alpha_{\text{hi}}\}$, we will show that the Kolmogorov-Smirnov distance between $\nu_{\hat{g}_\alpha|s}$ and $\nu_{\hat{g}_\alpha|s'}$ equals to zero for any $s \neq s' \in [K]$. Note that, according to the formulation of $\hat{g}_\alpha$ in Eq. (12), we have for any $(x, s) \in \mathbb{R}^p \times [K]$,

$$\hat{g}_\alpha(x, s) = \sum_{s'=1}^{K} \hat{p}_{s'} \hat{F}^{-1}_{2,q_\alpha|s'} \circ \hat{F}_{1,q_\alpha|s} \circ \tilde{q}_\alpha(x, s).$$

Denote by $\hat{Q}(t) = \sum_{s'=1}^{K} \hat{p}_{s'} \hat{F}^{-1}_{2,q_\alpha|s'}$. Note that we use the training set to estimate the location statistic for the new test point, $\hat{Q}(t)$ is independent from $\hat{F}_{1,q_\alpha|s} \circ \tilde{q}_\alpha(x, s)$ for each $s \in [K]$. Since the test point belongs to group $S = s$ for some fixed $s \in [K]$ and, for all $i = 1, \ldots, |\mathcal{I}_1^s|$, set $V_i = \tilde{q}^s_{1,i}$ with $V_{N_s+1} \stackrel{d}{=} (\tilde{q}_\alpha(x, s))$ independent from $(V_i)_{i=1,\ldots,|\mathcal{I}_1^s|}$. Since the random variables $V_1, \ldots, V_{N_s}, V_{N_s+1}$ are exchangeable (more ideally, independent), Lemma 1.1 implies that for all $s \in [K]$, the location statistic $\hat{F}_{1,q_\alpha|s} \circ \tilde{q}_\alpha(x, s)$ is distributed uniformly on $[0,1]$. Thus for all $s, s' \in [K]$, we have

$$
\begin{aligned}
& \mathrm{KS}\left(\nu_{\hat{g}_\alpha \mid s}, \nu_{\hat{g}_\alpha \mid s'}\right) \\
= & \sup_{t \in \mathbb{R}} \left| P\left(\hat{g}_\alpha \leq t \mid S = s\right) - P\left(\hat{g}_\alpha \leq t \mid S = s'\right) \right| \\
= & \sup_{t \in \mathbb{R}} \left| P\left(\hat{F}_{1, q_\alpha \mid s} \circ \tilde{q}_\alpha(x, s) \leq \hat{Q}^{-1}(t) \mid S = s\right) \right. \\
- & \left. P\left(\hat{F}_{1, q_\alpha \mid s} \circ \tilde{q}_\alpha(x, s) \leq \hat{Q}^{-1}(t) \mid S = s'\right) \right| \\
= & \sup_{t \in \mathbb{R}} \left| E\left[\hat{Q}^{-1}(t) \mid S = s\right] - E\left[\hat{Q}^{-1}(t) \mid S = s'\right] \right| = 0.
\end{aligned}
$$

The first equality uses the definition of $\hat{g}_\alpha$; the second uses the fact that $\hat{Q}$ is monotone; finally since the independence of $\hat{Q}$ is independent from $\hat{F}_{1,q_\alpha|s} \circ \tilde{q}_\alpha(x,s)$ conditionally on $S = s$ for any $s \in [K]$, also $\hat{Q}$ remains independent from S. The exact DP is concluded.

References

Chzhen, E. & Schreuder, N. (2022), ‘A minimax framework for quantifying risk-fairness trade-off in regression’, *The Annals of Statistics* **50**(4), 2416–2442.

Liu, M., Ding, L., Yu, D., Liu, W., Kong, L. & Jiang, B. (2022), ‘Conformalized fairness via quantile regression’, *Advances in Neural Information Processing Systems* **35**, 11561–11572.