Chapter 1

Quantile Estimation - Exercises

Shu Shen

Exercise 24.3 Define $\psi(x) = \tau - \mathbb{I}(x < 0)$. Let θ satisfy $E[\psi(Y-\theta)] = 0$. Is θ a quantile of the distribution of Y ?

Solution. 24.3

$$\begin{aligned} 0 = E[\psi(Y-\theta)] &= E[\tau - \mathbb{I}(Y-\theta<0)] \\ &= E[\tau - \mathbb{I}(Y<\theta)] \\ &= \tau - E[\mathbb{I}(Y<\theta)] \\ &= \tau - P(Y<\theta) \end{aligned}$$

Since $P(Y<\theta) = \tau$, the τ-quantile of the distribution of Y is θ.

Exercise 24.4 Take the model $Y = X'\beta + e$ where the distribution of e given X is symmetric about zero.

(a) Find $E[Y \mid X]$ and $\operatorname{med}[Y \mid X]$.

(b) Do OLS and LAD estimate the same coefficient β or different coefficients?

(c) Under which circumstances would you prefer LAD over OLS? Under which circumstances would you prefer OLS over LAD? Explain.

Solution. 24.4

(a) The conditional mean is $E[Y \mid X] = E[X'\beta + e \mid X] = X'\beta + E[e \mid X] = X'\beta$. Regarding the conditional media, we have

$$\begin{aligned} 0.5 = P(Y \le \operatorname{med}[Y \mid X] \mid X) = P\left(X'\beta + e \le \operatorname{med}[Y \mid X] \mid X\right) \\ = P\left(e \le \operatorname{med}[Y \mid X] - X'\beta \mid X\right) \end{aligned}$$

Since the distribution of e given X is symmetric about zero, we also have $P(e \leq 0 \mid X) = 0.5$. We conclude that $\text{med}[Y \mid X] = X'\beta$.

(b) OLS estimates the population coefficient $\beta_0 = \arg\min_\beta E\left[(Y_i - X_i'\beta)^2\right]$. LAD estimates the population coefficient

$\beta_{0.5} = \arg\min_\beta E[|Y_i - X_i'\beta|]$. Since $f(e \mid X) = f(Y - X'\beta \mid X)$ is symmetric about zero, $\beta_0 = \beta_{0.5}$. That is, the two estimation methods estimate the same population coefficient.

(c) When the errors have a thin-tailed distribution, the OLS estimator has some desirable characteristics, including being the maximum likelihood estimator and having the smallest variance among all linear unbiased estimators. Therefore, OLS is usually preferred in this case. If the data contains outliers or influential observations, then LAD may be preferred over OLS. This is because the LAD estimator is less sensitive to the influence of extreme values in the data.