Quantile Regression (Theory):

Let Y be a random variable. The τ-quantile of Y is a number $Q_Y(\tau)$ s.t.

$Pr[Y \leq Q_Y(\tau)] = \tau$ (continuous Y case). More generally, τ-quantile is defined as

$$Q_Y(\tau) = \inf\{y : Pr[Y \leq y] \geq \tau\}$$

Clearly, $Q_Y(\tau)$ is the inverse of $F_Y(y) = Pr[Y \leq y]$

Skorohod representation (Monte-Carlo) representation of random variable Y:

$Y = Q_Y(U)$ where $U \sim \underbrace{U}_{\text{uniform}}[0,1]$

$$\begin{aligned} Pr[Y \leq y] &= Pr[Q_Y(U) \leq y] \\ &= Pr[U \leq Q_Y^{-1}(y)] \\ &= Pr[U \leq F_Y(y)] = F_Y(y) \end{aligned}$$

Conditional (regression) case:

$Q_Y(\tau|X)$ solves $Pr[Y \leq Q_Y(\tau|X) | X] = \tau$ so that $Q_Y(\tau|X)$ can be seen as the inverse of the conditional distribution function

$$F_Y(y|X) = Pr[Y \leq y | X]$$

- Skorohod representation in the conditional case:

$$Y = Q_Y(U|X) \quad \text{where} \quad U|X \sim U[0,1]$$

- We will use approximation

$$Q_Y(\tau|X) \cong X'\beta(\tau)$$

The corresponding regression equation is

$$Y = X'\beta(U), \quad U|X \sim U[0,1]$$

- random coefficient model. (very nonlinear)
- People differ in their propensity to consume

How does quantile regression compare to conventional regression?

QR versus Conventional Regression

Location model:

$$Y_i = X_i'\beta_0 + \varepsilon_i, \ \varepsilon_i \perp\!\!\!\perp X_i \quad \text{traditional model}$$

$$= X_i'\beta_0 + Q_{\varepsilon_i}(U_i) \qquad U_i \sim U[0,1]$$

- slopes are constant, intercepts are random
  - This is a special case of the above random coefficient model.

Has conditional quantile function

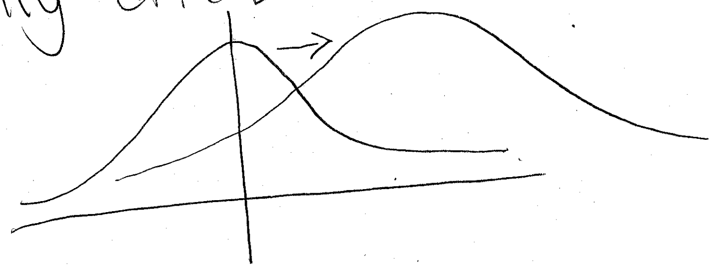
$$Q_Y(\tau|X) = X'\beta_0 + Q_\varepsilon(\tau)$$

$$= \underline{X}'\beta(\tau), \text{ where } \beta(\tau) = \begin{bmatrix} \beta_{01} + Q_\varepsilon(\tau) \\ \beta_{02} \\ \vdots \\ \beta_{0K} \end{bmatrix}$$

Location-Scale Model

$$\underline{Y}_i = \underline{X}_i'\beta_0 + \underbrace{(\underline{X}_i'\gamma_0)}_{>0}\cdot\varepsilon_i \qquad \varepsilon_i \perp\!\!\!\perp \underline{X}_i$$

- only affect scale - not skewness / kurtosis

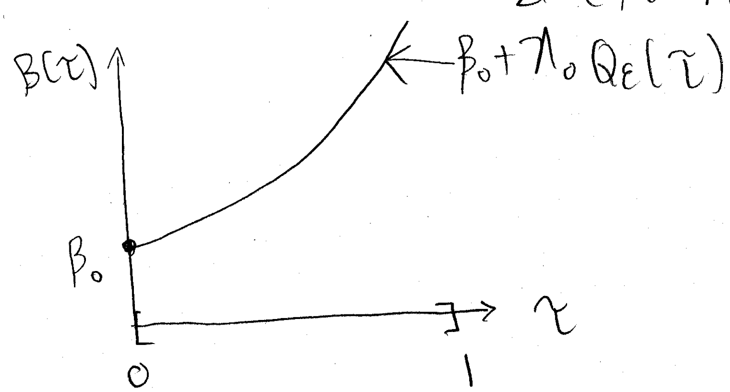


$$\Rightarrow \underline{Y}_i = \underline{X}_i'\beta_0 + (\underline{X}_i'\gamma_0)\cdot Q_\varepsilon(U_i)$$

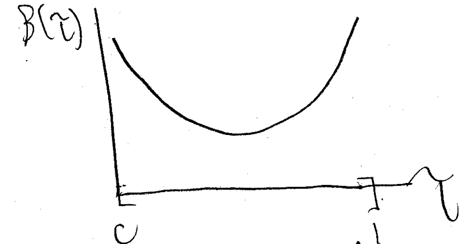
$$\Rightarrow Q_{\underline{Y}}(\tau|\underline{X}) = \underline{X}'\beta_0 + (\underline{X}'\gamma_0)Q_\varepsilon(\tau)$$

$$= \underline{X}'\beta(\tau) \quad \text{where } \beta(\tau) = \begin{bmatrix} \beta_{01} + \gamma_{01}Q_\varepsilon(\tau) \\ \vdots \\ \beta_{0K} + \gamma_{0K}Q_\varepsilon(\tau) \end{bmatrix}$$

$$= \underline{X}'(\beta_0 + \gamma_0 Q_\varepsilon(\tau))$$



- This is restrictive, since it imposes linear structure and hence monotonicity. We reject the location-scale model if we have:



- This is inconsistent with the location-scale model.

$$\left[\begin{array}{l} \hat{\varepsilon}_i = \dfrac{Y_i - X_i'\hat{\beta}}{X_i'\hat{\gamma}} \\ \hat{Q}_Y(\tau|X) = X'\hat{\beta}_0 + (X'\hat{\gamma})\hat{Q}_{\hat{\varepsilon}}(\tau) \end{array}\right] \leftarrow \text{aside}$$

<u>Doksum</u> (1974)

- Annals of Statistics
- Biometrika

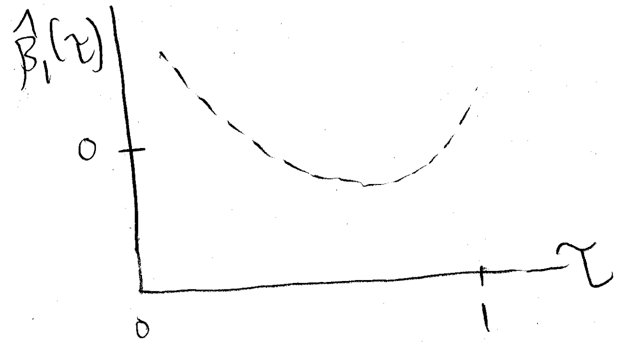
Y_i = duration of survival of guinea pigs

$X_i = \begin{bmatrix} 1 \\ D_i \end{bmatrix}$, D_i = indicator of vaccine treatments

$$Q_Y(\tau|X) = \beta_0(\tau) + \beta_1(\tau) \cdot D$$

τ - "proneness"

- unobserved health or immunity.



$$Q_Y(\tau|\text{control}) = \beta_0(\tau)$$

$$Q_Y(\tau|\text{treatment}) = \beta_0(\tau) + \beta_1(\tau)$$

$$\Delta = \beta_1(\tau)$$