Quantile Regression

Least squares is best linear predictor of

\[ g(X) = E[Y | X] \]

Conditional expectations describes how mean of \( Y \) varies with \( X \). Many applications where one would like to know how \( X \) affects other parts of distribution of \( Y \), other than its center.

- How does education affect earnings of those who earn little? Those who earn much?
- How does income affect consumption at different consumption levels?
- How does smoking affect birthweights of small babies?

Quantile regression is designed to help answer this question.

Begin with single random variable \( Y \) to define things; the \( \tau^{th} \) quantile (percentile, fractile) is value \( q_\tau(Y) \) such that

\[ \Pr(Y \leq q_\tau(Y)) = \tau \]

In terms of the CDF \( F_Y(y) = \Pr(Y \leq y) \),

\[ F_Y(q_\tau(Y)) = \tau \quad \text{or} \quad q_\tau(Y) = F_Y^{-1}(\tau) \]

Inverse of CDF.

\[ q_\tau(Y) \]

Turns out that \( \tau^{th} \) quantile solves a minimization problem that is convenient for

[1]
estimation and extending to regression. Let

$$1(u > 0) = \begin{cases} 1, & u > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\rho_r(u) = \tau |u| 1(u > 0) + (1 - \tau) |u| 1(u < 0)$$

Then $q_Y(\tau)$ minimizes $E[\rho_r(Y - \mu)]$ over $\mu$. To explain why the quantile minimizes this objective function, we consider the first order conditions. Note that $\rho_r(u)$ is convex, and hence so is $E[\rho_r(Y - \mu)]$, so that the solution to the first-order conditions is a global minimum. Note that for $Y \neq \mu$, $\rho_r(Y - \mu)$ is differentiable, with

$$d\rho_r(Y - \mu)/d\mu = \tau 1(Y - \mu > 0) - (1 - \tau) 1(Y - \mu < 0).$$

Then assuming we can differentiate under the integral the first-order conditions are

$$0 = \tau E[1(Y - \mu > 0)] - (1 - \tau) E[1(Y < \mu)]$$

$$= \tau (1 - F_Y(\mu)) - (1 - \tau) F_Y(\mu) = \tau - F_Y(\mu)$$

First order condition solved at $\mu = q_Y(\tau)$.

Extension to regression: Replace constant $\mu$ with $X'\beta$. The population object is

$$\beta(\tau) \text{ minimizes } E[\rho_r(Y_i - X_i'\beta)].$$

An estimator of this object can be obtained as

$$\hat{\beta}(\tau) \text{ minimizes } \sum_{i=1}^{n} \rho_r(Y_i - X_i'\beta) / n.$$ 

These are called regression quantiles. Idea is that $\hat{\beta}(\tau)$ estimates effect of $X$ on $Y$ at the $\tau^{\text{th}}$ quantile for $Y$. 

[2]
**Example 1** $Y = \ln \text{earnings, } X = \text{schooling, other}$

$\hat{\beta}_1(\tau)$ effect of schooling on $\tau^{th}$ quantile for $y$;

See graphs for results from last 3 census years from Angrist, Chernozhukov, Fernandez-Val.

Note that a nonparametric approach to this problem would be to look at the conditional quantiles of $Y$ given $X$. Let $\text{Pr}(Y \leq y|X) = F_{Y|X}(y|X)$ denote the conditional CDF. Then the conditional quantile is $q_Y(\tau|X) = F_{Y|X}^{-1}(\tau|X)$. Complicated function. Difficult to estimate with high dimension $X$ (curse of dimensionality). Regression quantiles more parsimonious. Interesting question is in what sense does $X'\beta(\tau)$ approximate $q_Y(\tau|X)$. It is not minimum mean-square error predictor and it is not best predictor from the check function. See Angrist, Chernozhukov, Fernandez-Val paper discussed later.

In some cases can relate quantile regression coefficients to usual linear model and find that regression quantiles correspond to full conditional quantile.

Case A: Constant coefficient regression model.

$$Y_i = X_i'\beta_0 + u_i, \ u_i \text{ and } X_i \text{ independent.}$$

Sometimes called conditional location model, since only location varies with $X$. Note that for a constant $C$, the $\tau^{th}$ quantile of $C + Y$. Therefore, since $X_i'\beta_0$ is constant conditional on $X$,

$$q_Y(\tau|X) = X'\beta_0 + q_u(\tau|X) = X'\beta_0 + q_u(\tau),$$

where the second equality follows by independence of $u$ and $X$. Here the quantile lines are parallel. Slopes are the same for all quantiles.

Case B: Scale shift model.

$$Y_i = X_i'\beta_0 + (X_i'\gamma_0)\varepsilon_i, X_i'\gamma_0 > 0, \varepsilon_i \text{ and } X_i \text{ independent.}$$

Sometimes called conditional location scale model. Here the $\tau^{th}$ conditional quantile of $u_i = (X_i'\gamma_0)\varepsilon_i$ given $X_i$ solves

$$\tau = \text{Pr}((X'\gamma_0)\varepsilon_i \leq q_u(\tau|X)) = \text{Pr}(\varepsilon_i \leq q_u(\tau|X)/(X'\gamma_0)|X)$$

[3]
\[ F_\varepsilon(q_u(\tau|X)/(X'\gamma_0)), \]

where the last equality follows by independence of \( \varepsilon \). Inverting \( F \) and solving gives

\[ q_u(\tau|X) = (X'\gamma_0)F_\varepsilon^{-1}(\tau) = (X'\gamma_0)q_\varepsilon(\tau) \]

Then the conditional quantile function of \( Y \) given \( X \) is

\[ q_\gamma(\tau|X) = X'\beta_0 + X'\gamma_0 q_\varepsilon(\tau). \]

Here the conditional quantile function is linear, so that the regression quantile is \( \beta(\tau) = \beta_0 + \gamma_0 q_\varepsilon(\tau) \).

**Example 2** Engle curve.