RIF and Unconditional Quantile Regression

Walter Sosa Escudero
(wsosa@udesa.edu.ar)

Universidad de San Andres

September de 2006

Motivation: The effect of X on Y

x and y are RV's. We are interested in quantifying 'the effect on y of altering x a little'

What do we mean by 'altering x'? Moving its location. WLOG, consider

$$x = \mu_x + x^*$$

where μ_x is a constant and x^* is a zero mean RV. We are interested in movements that arise from altering μ_x marginally (location shifts).

The regression way

Consider the regression model

$$E(y|x) = x\beta = \mu_x\beta + x\beta$$

so

$$\beta = \frac{\partial E(y|x)}{\partial \mu_x}$$

by the Law of Iterated Expectations

$$E(y) = E[E(y|x)] = E(x\beta) = E(x)\beta = \mu_x\beta$$

so

$$\beta = \frac{\partial E(y)}{\partial \mu_x}$$

Through the LIE, β is playing the double role of capturing the effect of moving x on both, $E(y)$ and $E(y|x)$.

The quantile regression way

Consider the quantile regression model

$$Q_{y|x}(\tau) = x'\beta$$

so

$$\frac{Q_{y|x}(\tau)}{\partial \mu_x} = \beta$$

Now we are in trouble since we cannot use something like the LIE for quantiles.

We can measure how X alters conditional quantiles. What if we are interested in (unconditional) quantiles?

- The goal is to explore an estimable strategy to compute the effect of altering X on any feature of Y: mean, variance, quantiles, Gini coefficients.
- Reference: Firpo, Fortin and Lemieux (2009), *Unconditional Quantile Regressions*, Econometrica 77, 953-973.
- General approach based on influence functions.

Features as functionals

Let Y be a RV with CDF $F(y)$ and density $f(y)$. Consider its mean, μ. Then

$$\mu = \int y f(y)\, dy = \int y\, dF(y)$$

Then μ can be seen as

$$\mu = T(F) : \mathrm{Dom}(F) \rightarrow \Re : \int y\, dF(y)$$

that is, μ is a 'function' of the CDF F. 'Functions of functions' are labeled as *functionals*.

In general, features of interest of a random variable can be expressed as functionals of their CDF's. Examples

- *Variance:* $V(Y) = \int (y - mu)^2 \, dF(y)$
- *Poverty rate:* $PR(Y) = \int_0^{y_l} y \, dF(y)$, where y_l is a poverty line.
- *Quantiles:* $Q_\tau(Y) = F^{-1}(\tau)$, for invertible CDF.

It will be convenient to focus on *linear functionals* of the form

$$T(F) = \int \psi(y)\, dF(y)$$

for some function $\psi(y)$.

Example: Mean

$$T(F) = \int y\, dF(y), \qquad \psi(y) = y$$

Example: Poverty rate

$$T(F) = \int 1[y < y_l]\, dF(y), \qquad \psi(y) = 1[y < y_l]$$

Influence function

The influence function of T at F is given by

$$IF(y;F) \equiv \psi - \int \psi(y)\, dF(y)$$

Note that, rather trivially

$$E[IF(y;F)] = E[IF(y;F)] = 0$$

Examples

- Mean: $\mu = T(F) = \int x \, dF(y), \;\; \psi(y) = y$

$$IF(y;F) = y - \int y \, dF(y) = y - \mu$$

- Poverty:
$PR(Y) = T(F) = \int 1[y < y_l] \, dF(y), \;\; \psi(y) = 1[y < y_l]$

$$IF(y;F) = 1[y < y_l] - PR(Y)$$

We are still keeping some mistery about what is the interpretation of IF...

Influence and biases

Let $T(F)$ be a functional, as before.

Q: How does $\theta = T(F)$ change when observations come from some other distribution G, 'close' to F?

The problem is to compute $T(G) - T(F)$ for G close to F

For *asymptotically linear functionals*, the following 'von Mises' expansion holds

$$T(G) - T(F) = T^*(G - F) + o(d(G, F))$$

where $d(G, F)$ is a distance between G and F, and $T*$ is a linear functional, so that $T^*(G - T) = T^*(G) - T^*(F)$.

Then, for asymptotically linear functionals and when G is close to F:

$$
\begin{aligned}
T(G) - T(F) &= T^*(G - F) \\
&= \int \psi \, d(G - F) \\
&= \int \psi \, dG - \int \psi \, dF
\end{aligned}
$$

Now, from the definition of the IF, $\psi = IF - \int \psi \, dF$. Replacing

$$
\begin{aligned}
T(G) - T(F) &= \int \left[IF - \int \psi \, dF \right] dG - \int \psi \, dF \\
&= \int IF \, dG - \int \psi \, dF \int dG - \int \psi \, dF \\
&= \int IF \, dG
\end{aligned}
$$

More, formally, what we have shown is the following.
Let F and G we two CDF's, and define

$$F_t \equiv (1-t)F + tG \;=\; t(G-F) + F, \qquad 0 \leq t \leq 1$$

Then

$$\left.\frac{\partial T(F_t)}{\partial t}\right|_{t=0} = \lim_{t \downarrow 0} \frac{T(F_t) - T(F)}{t} = \int IF \; dG$$

This gives us a first interpretation for the IF...

Detour: Influence and derivative

Let $G = F_{\epsilon,z}$, with

$$F_{\epsilon,z} \equiv (1-\epsilon)F + \epsilon\,\Delta_{(z)} = F + \epsilon\,(\Delta_{(z)} - F), \qquad 0 \le \epsilon \le 1,$$

and $\Delta_{(z)}$ is a degenerate CDF with unit mass at point z.

Using our previous results

$$T(F_{\epsilon,z}) - T(F) = \int IF(y)\,dF_{\epsilon,z} + o(d(F_{\epsilon,z}, F))$$

Now

$$\begin{aligned} T(F_{\epsilon,z}) - T(F) &= \int IF(y)\,dF_{\epsilon,z} \\ &= (1-\epsilon)\int IF(z)\,dF + \epsilon\int IF(y)\,d\Delta_{(z)} \\ &= \epsilon\int IF(y)\,d\Delta_{(z)} \\ &= \epsilon\,IF(z) \end{aligned}$$

Replacing

$$T(F_{\epsilon,z}) - T(F) = \epsilon \, IF(z) + o(d(F_{\epsilon,z}, F))$$

so

$$IF(z) = \lim_{\epsilon \downarrow 0} \frac{T(F_{\epsilon,z}) - T(F)}{\epsilon}$$

- $IF(z)$ measures the effect a single point has on a functional. Recall that for the mean $IF(z) = z - \mu$.
- Influence functions have played a fundamental rol in the development of robust statistics.

Let us stop for a while.

- Recall that our goal is to measure how changes in X affect T, a functional of Y.
- We are half way. Influence functions give us a way to explore how changes in F affect T.
- The plan: changes in the distribution of X affect the distribution of Y, and this make T change.
- Idea: try to come up with something so we can use the LIE. The tool is the *recentered* influence function (RIF).

Recentered Influence Functions (RIF)

Now call F_Y the CDF of Y

$$RIF(y; F_Y) = T(F_Y) + IF(y, F_Y)$$

so, for a linear functional, $RIF(y; F_Y) = \psi(y)$.

Two results

1. $E(RIF) = T(F_Y)$
2. Let X be a RV. Using the LIE

$$\begin{aligned} T(F_Y) &= \int RIF(y, F_Y)\, dF_Y \\ &= \int \left[\int RIF(y, F_Y)\, dF_{Y|X}(y|X = x) \right] dF_X(x) \\ &= \int E[RIF(y, F_Y) \mid X = x]\, dF_X(x) \end{aligned}$$

These two results are very important

1. $T(F_Y) = E(RIF)$: any magnitude of interest can be seen as an expectation.
2. $T(F) = E[E(RIF|X)]$: we have introduced X through the LIE.

The plan

1. How small changes in X affect $E(RIF)$.
2. How small location changes in X affect $E(RIF)$.
3. The case of quantiles.
4. An estimable form.

The marginal effects of altering X

Suppose F_X changes marginally in the direction of G_X. Assume $F_{Y|X}$ stays constant.

Then

$$\left.\frac{\partial T(F_{Y,tG_Y^*})}{\partial t}\right|_{t=0} = \int E[RIF(y, F_Y) \mid X = x] \, d(G_X - F_X)(x)$$

where $F_{Y,tG_Y^*} \equiv (1-t)F_Y + t\, G_Y^*$.

Note $F_Y = F_{Y|X}F_X$, so $G_Y^* = F_{Y|X}G_X$.

Intuition

- Recall $F_Y = F_{Y|X} F_X$
- We are changing F_X by G_X.
- $G_Y^* = F_{Y|X} G_X$ would be the new CDF for Y, if we alter F_X but keep $F_{Y|X}$ constant.
- So, we are trying to measure how altering F_X towards G_X affects T.
- The whole point is that T depends on Y: X alters Y, and this alters T.

Proof

$$
\begin{aligned}
\frac{\partial T(F_{Y,tG_Y^*})}{\partial t} &= \int IF \; dG_Y^* \\
&= \int IF \; d(G_Y^* - F_Y) \\
&= \int \Big(RIF - T(.)\Big) \; d(G_Y^* - F_Y) \\
&= \int RIF \; d(G_Y^* - F_Y) - \int T(.) \; dG_Y^* - \int T(.) \; dF_Y \\
&= \int RIF \; d(G_Y^* - F_Y) \\
&= \int \left[\int RIF \; dF_{Y|X} \right] \; dG_X - \int \left[\int RIF \; dF_{Y|X} \right] \; dG_F \\
&= \int E\left[RIF \mid X = x\right] \; d(G_X - F_X)
\end{aligned}
$$

The unconditional partial effect

Let $\alpha(T)$ be the vector of partial effects on T of moving each coordinate of X separately as a location shift. Then (under some regularity)

$$\alpha(T) = \int \frac{d\, E\left[RIF \mid X = x\right]}{dx}\, dF_X(x)$$

Proof: cumbersome but easy. Some intuition. Take the case of X a scalar. Under the location shift $G_X(x) = F_X(x - \Delta)$. then

$$\int E\left[RIF \mid X = x\right]\ d(G_X - F_X) = \int E\left[RIF \mid X = x\right]\ d(F_X(x - \Delta) - F_X)$$

then take derivatives.

Unconditional Quantile Regression

Let q_τ be the $\tau-$th quantile of Y (unconditional). It can be shown that its influence function is given by

$$IF(y;F) = \frac{\tau - 1[y \leq q_\tau]}{f_Y(q_\tau)}$$

so

$$RIF(y;F) = q_\tau \ + \ \frac{\tau - 1[y \leq q_\tau]}{f_Y(q_\tau)}$$

Note: $1[y \leq q_\tau] = 1 - 1[y > q_\tau]$, so

$$
\begin{aligned}
RIF(y;F) &= q_\tau + \frac{\tau - 1}{f}\,\frac{1}{f}\,1[y > q_\tau] \\
&= c_{2,\tau} + c_{1,\tau}\,1[y > q_\tau]
\end{aligned}
$$

Taking expectations

$$E[RIF(y;F) \mid X = x] = c_{2,\tau} + c_{1,\tau}\,Pr[y > q_\tau \mid X = x]$$

Then, by our previous result, the UPE is

$$\alpha(\tau) = c_{1,\tau} \int \frac{dPr[y > q_\tau \mid X = x]}{dx}\,dF_x(x)$$

UPE for the LPM

Consider the linear probability model (LPM):

$$Pr[y > q_\tau \mid X = x] = x'\beta$$

so

$$\frac{dPr[y > q_\tau \mid X = x]}{dx} = \beta$$

Replacing above

$$\alpha(\tau) = c_{1,\tau}\beta$$

Estimation: RIF-OLS

Start with

$$1[y > q_\tau] = x'\beta + u$$

so under the LPM assumption $E(u|x) = 0$. Now

$$\begin{aligned} RIF(y;F) = 1[y > q_\tau]c_{1,\tau} + c_{2,\tau} &= c_{2,\tau} + c_{1,\tau}\, x'\beta + u \\ &= c_{2,\tau} + x'\beta^* + u \end{aligned}$$

with $\beta^* \equiv c_{1,\tau}\beta = \alpha(\tau)$

So, for the $\tau-$th quantile we have

$$RIF(y;F) = c_{2,\tau} + x'\beta^* + u$$

this is the RIF regression for the $\tau-$th quantile.

- If $RIF(y;F)$ were observable, we could regress it on x, to obtain an estimate for β^*.
- Recalling that

$$RIF(y;F) = q_\tau \; + \; \frac{\tau - 1[y \leq q_\tau]}{f_Y(q_\tau)}$$

  q_τ and $f_Y(q_\tau)$ are first estimated and then the estimated RIF is regressed on X.

- q_τ is estimated as the sample $\tau-$th quantile.
- $f_Y(q_\tau)$ is usually estimated non-parametrically (using a kernel density estimator).
- Then, for each observation we compute

$$\widehat{RIF}(y_i; F) = \hat{q}_\tau \; + \; \frac{\tau - 1[y_i \leq q_\tau]}{\hat{f}_Y(q_\tau)}$$

- Regress $\widehat{RIF}(y_i; F)$ on x_i