AGRODEP Training Session

"Poverty measurement and analysis"

Anne-Claire Thomas & Philippe Van Kerm
Université Catholique de Louvain (Belgium)
& CEPS/INSTEAD (Luxembourg)

International Food Policy Research Institute (IFPRI), Dakar
April 24–27 2012

Session 7

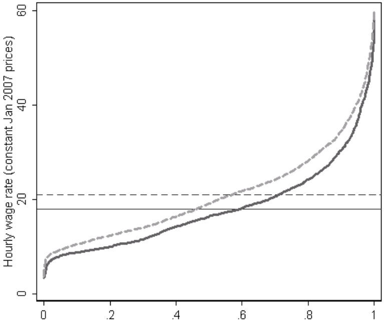
Accounting for distribution change

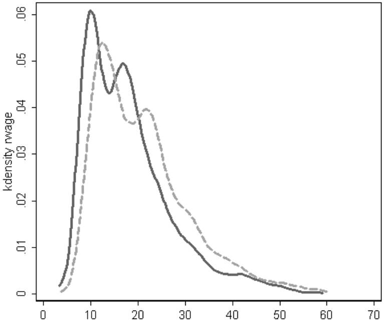
Content

[outline]

Assessing distribution differences

Comparison of the income distribution at year 1 and at year 2 (or for women versus distribution for men, or immigrants vs. natives, or urban vs. rural, etc.)





Because of change in population composition/characteristics? Or because of different return to characteristics?

Assessing distribution differences

Objective is to simulate counterfactual distributions that would be observed with fixed coefficients or controlling for differences in population characteristics

With simulated distributions, can assess differences in

- any poverty indicator
- inequality measures
- pro-poor growth indices
- ...

net of the impact of differences in population composition

The analysis of **mean** differences first...

Law of iterated expectations:

$$E(y) = E_x(E(y|x)) = \int_{\Omega_x} E(y|x)\, h(x)\, dx = \int_{\Omega_x} \int y\, f(y|x)\, h(x)\, dydx$$

$\Longrightarrow$ Decompose the overall mean ($E(y)$) as weighted sum of conditional means ($E(y|x)$) with weights being proportion of population with x ($h(x)$)

The analysis of mean differences

Easy to decompose differences in two overall means as differences attributed to $E(y|x)$ ('structural') and $h(x)$ ('compositional').

Create a counterfactual mean by keeping $h(x)$ identical in two groups and see impact of differences in $E(y|x)$ on overall mean:

$$E^*(y) = \int_{\Omega_x} E^m(y|x)\, h^f(x)\, dx$$

and

$$(E^f(y) - E^m(y)) = (E^f(y) - E^*(y)) + (E^*(y) - E^m(y))$$

'Oaxaca-Blinder' decomposition

In a regression model one sets $E(y|x) = x\beta$ and the decomposition is into 'coefficient effect' (β's) vs. composition.
Sample analog to iterated expectations expression:

$$\hat{\mu}^f = \frac{1}{N^f}\sum_{i=1}^{N^f}\hat{\mu}^f(x_i) = \frac{1}{N^f}\sum_{i=1}^{N^f}x_i\hat{\beta}^f$$

and

$$\hat{\mu}^* = \frac{1}{N^f}\sum_{i=1}^{N^f}\hat{\mu}^m(x_i) = \frac{1}{N^f}\sum_{i=1}^{N^f}x_i\hat{\beta}^m$$

So

$$\hat{\mu}^f - \hat{\mu}^* = \frac{1}{N^f}\sum_{i=1}^{N^f}x_i(\hat{\beta}^f - \hat{\beta}^m) = (\hat{\beta}^f - \hat{\beta}^m)\left(\frac{1}{N^f}\sum_{i=1}^{N^f}x_i\right)$$

[outline]

Mean as least squares minimization

The mean can be expressed as $\hat{\mu} = \arg\min_{\mu} \sum_i (y_i - \mu)^2$

In the presence of covariates interest is on *conditional* mean
$\mu(x) = E(y|x)$:
$\hat{\mu}(x) = \arg\min_{\mu} \sum_i \mathbf{1}(x_i = x)\ (y_i - \mu)^2$

Of course only possible with large samples and discrete x.

OLS regression – the linear regression model

Assume a particular relationship (linear) between conditional mean and x:

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

Or equivalently $y_i = x_i\beta + e_i$ with $E(e_i|x_i) = 0$

$$\hat{\beta} = \arg\min_{\beta} \sum_i (y_i - x_i\beta)^2$$

and $\hat{\mu}(x) = x\hat{\beta}$

Quantiles as check function minimzer

QR is fully analogous to mean regression

The trick is to express Q_τ as similar minimization problem:

$$\hat{Q}_\tau = \arg\min_{\xi} \sum_i \rho_\tau(y_i - \xi)$$

where

$$\rho_\tau(u) = u(\tau - \mathbf{1}(u < 0))$$

Quantiles as check function minimzer

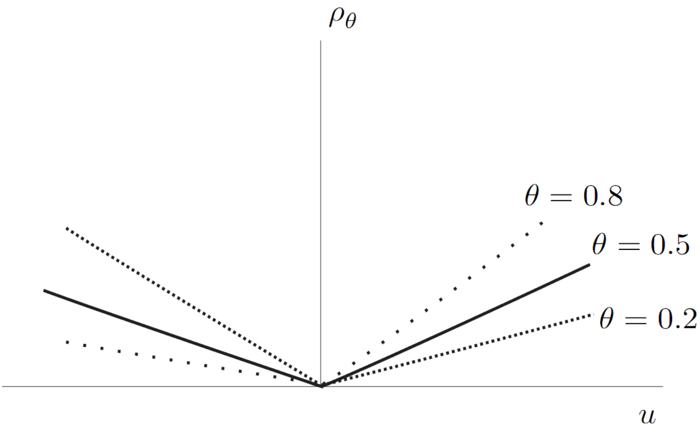


Figure: Check function $\rho_\theta(u) = (\theta - 1_{\{u<0\}})u$.

Linear quantile regression model

Focus on conditional quantile now and assume a particular relationship (linear) between conditional quantile and x:

$$Q_\tau(y|x) = x\beta_\tau$$

(Or equivalently $y_i = x_i\beta_\tau + u_i$ where $F^{-1}_{u_i|x_i}(\tau) = 0$)

$$\hat{\beta}_\tau = \arg\min_\beta \sum_i \rho_\tau(y_i - x_i\beta)$$

Estimate of the conditional quantile (given linear model):

$$\hat{Q}_\tau(y|x) = x\hat{\beta}_\tau$$

Interpretation of linear QR

- Estimation of $\hat{Q}_\tau(y|x)$ for a range of τ in $(0, 1)$ provides a linear approximation for the entire conditional distribution of Y given X
- $\hat{\beta}_\tau$ can be interpreted as the marginal change in the τ conditional quantile for a marginal change in x ...
- ... However, this does not imply that it is really the effect of a change in x for a person at the τ quantile of the distribution! This would require a “rank-preservation” assumption.

Advantages over OLS?

- Provides description of full conditional distribution (possibly revealing relevant heteroscedascity, location and scale variations with x)
- Median regression as a linear measure of central tendency of the data very similar to OLS regression if error symmetric, yet
  - robust to outliers
  - and may be more efficient if data heteroscedastic (violation of Gauss-Markov assumptions)
  - invariant to monotonic transformations: $Q(h(y)|x) = h(Q(y|x))$

Estimation

Seemingly simple minimization problem ...

- ... but target function is not differentiable (kink at 0) ...
- ... so standard optimization algorithm (e.g. Newton-Raphson) not applicable

Solution: Linear programming algorithm

Inference

MM framework used for deriving asymptotic properties.

$$\sqrt{n}(\hat{\beta}_\tau - \beta_\tau) \to N(0, A^{-1}BA^{-1})$$

where

$$A = E(f_{u|X}(0)X'X)$$

$$B = \tau(1-\tau)E(X'X)$$

Difficulty is estimation of $f_{u|x}(0)$ (if $\dim(X)$ is large).
Common to assume $f_{u|x}(0) = f_u(0)$ but very unattractive assumption!

Alternative to analytic SE's: Bootstrap
(simple paired bootstrap shown to perform well)

Decomposition of ‘quantile’ differences

Unfortunately, there is no law of ‘iterated quantiles’! We can not derive the unconditional quantiles from conditional quantiles:

$$Q_\tau(y) \neq \int_{\Omega_x} Q_\tau(y|x)\, h(x)\, dx$$

Decomposition of quantile differences

Solution is to use some form of 'simulation'.

- Original proposal by Machado-Mata (JAE 2004)
- Simplification (see Melly (LabEco 2005) for case of quantile):
  - uniform (equaly spaced) sequence of quantile predictions is a quasi-random sample from the conditional distribution (cf. earlier Stata example)
  - so can construct simulated samples from the unconditional distribution by pooling sequence of predictions for different x
  - can be used to estimate unconditional quantiles (and many other stats)

Decomposition of quantile differences (ctd.)

Simulation:

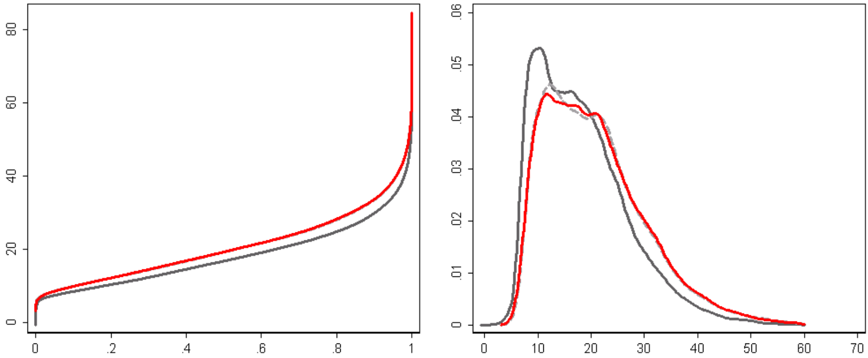
- Estimate a large number K of equally spaced quantile regressions from group of interest (say women) and make K predictions for all individuals
- Do predictions both in- and out-of-sample! Estimate β^f_τ from N^f women but predict $q^f_\tau(x_i) = x_i\hat{\beta}^f_\tau$ for all $N^f + N^m$ men and women (gives quantile 'as if paid like women' – for both men and women)
- Do same for men: estimate QR and predict for both men and women

Decomposition of quantile differences (ctd.)

Let $\{x_i\hat{\beta}^f_\theta\}^f_{\theta\in(0,1)}$ be the $K \times N^f$ predictions from women model
Let $\{x_i\hat{\beta}^m_\theta\}^f_{\theta\in(0,1)}$ be the $K \times N^f$ predictions from men model
And similarly $\{x_i\hat{\beta}^f_\theta\}^m_{\theta\in(0,1)}$ and $\{x_i\hat{\beta}^m_\theta\}^m_{\theta\in(0,1)}$ be the $K \times N^m$ predictions on the men sample

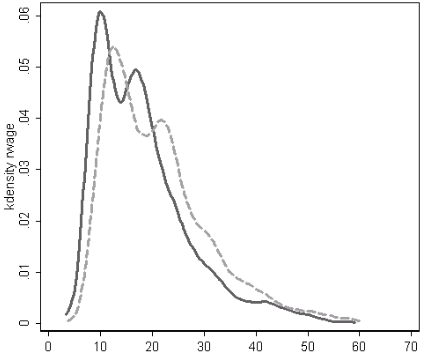
$$
\begin{aligned}
\hat{q}^f_\tau - \hat{q}^m_\tau &= q_\tau\left(\{x_i\hat{\beta}^f_\theta\}^f_{\theta\in(0,1)}\right) - q_\tau\left(\{x_i\hat{\beta}^m_\theta\}^m_{\theta\in(0,1)}\right) \\
&= (\hat{q}^f_\tau - \hat{q}^*_\tau) + (\hat{q}^*_\tau - \hat{q}^m_\tau) \\
&= \left(q_\tau\left(\{x_i\hat{\beta}^f_\theta\}^f_{\theta\in(0,1)}\right) - q_\tau\left(\{x_i\hat{\beta}^m_\theta\}^f_{\theta\in(0,1)}\right)\right) \\
&- \left(q_\tau\left(\{x_i\hat{\beta}^m_\theta\}^f_{\theta\in(0,1)}\right) - q_\tau\left(\{x_i\hat{\beta}^m_\theta\}^m_{\theta\in(0,1)}\right)\right)
\end{aligned}
$$

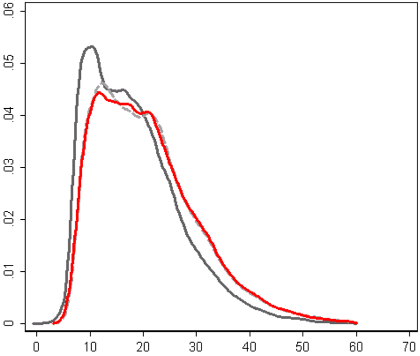
Illustration



It's all in the coefficients!

Illustration





(a word of caution: difference between true and simulated densities: check modelling assumptions!)