Discussion: Reliable inference for inequality measures with heavy-tailed distributions by Jean-Marie Dufour, Emmanuel Flachaire, Lynda Khalaf

Elise Coudin (CREST)

November 2013
Research question and approach

Issue:

- Let \((X_1, \ldots, X_n), X \sim F_x, (Y_1, \ldots, Y_m), Y \sim F_y, X_i\)s iid, \(Y_i\)s iid, \(X\) and \(Y\) independent;
- and let \(\theta(.)\) be an inequality index,
- Consider testing \(H_0 : \theta(F_x) = \theta(F_y)\).
- Both asymptotic and (classic) bootstrap inference on inequality measures perform poorly in presence of heavy-tailed distributions: income...
- Search for nonparametric methods to improve inference quality
Research question and approach

Idea/Approach:

- Permutation/randomization tests provide exact inference (in finite samples) if under the null, the two samples follow the same distribution, (and more generally when the data are generated from a distribution, which is invariant under some group of transformations.)
- Exploit results of Romano (1990) who derives conditions under which permutation tests are asymptotically valid when under the null $F_x \neq F_y$
- Expect that robustness properties of permutation tests can improve inference quality in more general setups ($F_x \neq F_y$).
Main results

Results

- Show that the equality of Generalized Entropy Indices and Gini Indices can be tested by permutation tests, if the data are well re-scaled
- Propose to use bootstrap under the null
- Provide a simulation comparison study of permutation tests, tests based on bootstrapping under the null, and usual asymptotic and bootstrap tests
Theory

Let \((X_1, \ldots, X_n), X \sim F_x, (Y_1, \ldots, Y_m), Y \sim F_y, X_i\)s iid, \(Y_i\)s iid, \(X\) and \(Y\) independent. Consider testing \(H_0 : \theta(F_x) = \theta(F_y)\).

With test statistic: \(T(X_n, Y_m) = n^{1/2}(\theta(\hat{F}_x) - \theta(\hat{F}_y))\)

Let \(Z = (X_1, \ldots, X_n, Y_1, \ldots, Y_m) = (Z_1, \ldots, Z_{n+m}), \ Z^p = (Z_{p(1)}, \ldots, Z_{p(n)}, Z_{p(n+1)}, \ldots, Z_{p(n+m)}), \) where \(p\) is a permutation of \((1, n + m)\),

Permutation statistic: \(T^p(Z_{n+m}) = n^{1/2}(\theta(\hat{F}_1) - \theta(\hat{F}_2))\)

Romano: \(T(X_n, Y_m)\) and \(T^p(Z_{n+m})\) follow asymptotically the same distribution under \(H_0\) if their asymptotic variances are the same:

\[
\text{Vas}(\theta(F_x)) + \frac{\lambda}{1 - \lambda} \text{Vas}(\theta(F_y)) = \frac{1}{1 - \lambda} \text{Vas}(\theta(\lambda F_x + (1 - \lambda)F_z))
\]

where \(\lambda = \frac{n}{n+m}\)
Theory

Dufour-Flachaire-Khalaf: A permutation test is asymptotically valid if

- under $H_0$, $\theta(F_x) = \theta(F_y) = \theta(\lambda F_x + (1 - \lambda) F_y)$
- $\theta(.)$ is linear in $F$
- and either $n/(n + m) \to \lambda = 1/2$, either
  $\text{Vas}(\theta(F_x)) = \text{Vas}(\theta(F_y))$

Discussion

- Linearity is essential. It allows to express vas of $\theta$ of the mixture distribution as a linear combination of vas of $\theta$ applied at each component.
- The method is adapted to inequality indexes derived from moments (centered or not, if the data are properly rescaled), such as GEIs, etc.. (but not a quantile ratio, and Gini?)
Simulation study

Setup

\( n = m, \lambda = .5, \) t-stat versions of \( T_n, T^* = \frac{\theta(F_x^*) - \theta(F_y^*)}{\sqrt{V(\theta(F_x^*)) - V(\theta(F_y^*))}}. \)

Compare

- **boot**: \( F_x^*, F_y^* \) obtained drawing \( n X^* \) in \( F^*_x \) and \( m Y^* \) in \( F^*_y \)
- **permut**: \( F_x^*, F_y^* \) obtained by drawing without replacement, \( n + m Z^* \) in \( F^{n+m}_Z \)
- **bootH0**: \( F_x^*, F_y^* \) obtained by drawing with replacement, \( n + m Z^* \) in \( F^{n+m}_Z \)
Simulation study

Results: both for Gini and Theil indexes

- When $F_x = F_y$: test sizes are better controlled by `permut` and `bootH0` than `boot` (very heavy upper tail) or `asymptotic` (very heavy upper tail and small $n$)

- When $F_x \neq F_y$ under $H_0$: sizes of `permut` and `bootH0` increase when the tails of $F_x$ and $F_y$ differ and $n$ is small, but less than those of `boot` or `asymptotic`.

Questions:

- Gini with $F_x = F_y$: figure 4, `boot` test size is close to .05 when $n$ small and remains constant after. Why?

- What happens when $\lambda \neq 0.5$? (and $Vas(\theta(F_x)) = Vas(\theta(F_y))$)
Figure 4: Rejection frequencies for the Gini index, with $F_x = F_y$ in the worst case ($\xi_x = \xi_y = 2.59$), as the sample size increases.
Figure 5: Rejection frequencies for the Gini index, with $F_x \neq F_y$ in the worst case ($\xi_y = 2.59$ and $\xi_x = 4.76$), as the sample size increases.
Questions/Extensions

- What about power comparison?
- May different resampling schemes allow to deal with $\lambda \neq .5$ (and $Vas(\theta(F_x)) \neq Vas(\theta(F_y))$)?
- Extend results when $X$ and $Y$ are correlated?