Reliable inference for inequality measures with heavy-tailed distribution

Jean-Marie Dufour
McGill University

Emmanuel Flachaire
Aix-Marseille University

Lynda Khalaf
Carleton University

Motivation

In the presence of heavy-tails:

- Asymptotic and bootstrap inference perform poorly in finite sample
- Alternative methods: improvements, but inference is still unreliable

	asym	boot	varstab[1]	semip[2]	mixture[3]
Singh-Maddala					
$q = 1.7$	0.915	0.931	0.933	0.926	0.928
$q = 1.2$	0.856	0.905	0.899	0.905	0.912
$q = 0.7$	0.647	0.802	0.796	0.871	0.789

Table : Coverage of asymptotic and bootstrap confidence intervals at the 95% level for the Theil index, for several bootstrap approaches, $n = 500$.

Note that it is a large sample problem.

[1] Schluter and van Garderen (2009, JoE)

[2] Davidson and Flachaire (2007, JoE), Cowell and Flachaire (2007 JoE)

[3] Cowell and Flachaire (2013, Handbook)

Our approach

We are interested in testing

$$H_0 : \theta(F_x) = \theta(F_y) \tag{1}$$

Monte Carlo permutation tests:

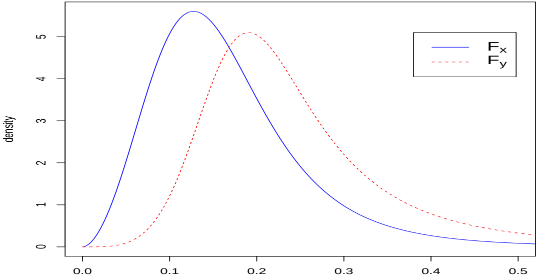
- If $F_x = F_y$, a permutation test provides exact inference.[4]
- For a nominal level α, critical value or p-value obtained from the permutation distribution would then be similar than those obtained from the true distribution.
- There is no need to obtain the permutation distribution from all the possible permutations

Problem: (1) does not guarantee that $F_x = F_y$. Two different distributions can share the same inequality index.

[4] Fisher (1935), Dwass (1957), Dufour (2006)

Our approach

The following Figure depicts Singh-Maddala distributions (Burr XII).[5] One distribution is much more heavy-tailed than the other, yet both distributions share the same value of the Theil index.



[5] with density $f(u) = aqu^{a-1}/(b^a[1 + (u/b)^a]^{1+q})$, for two choices of a, b and q: 2.8, 0.1930698, 1.7 [depicted as $\mathbf{F}_x$] and 4.8, 0.1930698, 0.6366578 [depicted as $\mathbf{F}_y$].

Our approach

We are interested in testing

$$H_0 : \theta(F_x) = \theta(F_y) \qquad (1)$$

The use of permutation tests is not justified from an exact perspective. We thus analyze the asymptotic validity of permutation tests of (1) when $F_x \neq F_y$.

- We show that permutation tests can be used reliably with the most popular inequality measures provided considered samples are recentered or rescaled adequately.
- A bootstrap method that respects the null hypothesis is also proposed.
- Simulation experiments are provided to study the finite sample properties

Outline

1. Finite and large sample theory
   - Exact inference
   - Asymptotic validity
   - Bootstrapping under the null
2. Comparing inequality measures
   - Centered and uncentered moments
   - The generalised entropy class (Theil index)
   - The Gini coefficient
3. Simulation study
   - Model design
   - Results

Permutation test

- $X = \{X_1, X_2, \ldots, X_n\} \sim F_x$ and $Y = \{Y_1, Y_2, \ldots, Y_m\} \sim F_y$. We test the null $H_0 : \theta(F_x) = \theta(F_y)$, with the statistic

$$T(X, Y) = \sqrt{n}\left[\theta(\hat{F}_x) - \theta(\hat{F}_y)\right].$$

- The permutation distribution is obtained by permuting in all possible ways the $n + m$ observations of the combined sample

$$Z = \{X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m\}.$$

It is the distribution of the permutation statistic, defined as[6]

$$T_* = \sqrt{n}\left[\theta(\hat{F}_{x_*}) - \theta(\hat{F}_{y_*})\right],$$

[6] $\hat{F}_{x_*}$ and $\hat{F}_{y*}$ are the EDF of, respectively, the first n and the remaining m observations of a permuted sample.

Asymptotic validity

Romano (1990) shows that the permutation test is asymp. valid if the asymptotic variances of the original and permutation statistics are similar, that is,

$$V[\theta(\hat{F}_w)] = (1-\lambda)V[\theta(\hat{F}_x)] + \lambda V[\theta(\hat{F}_y)]$$

where $\hat{F}_w = \lambda\hat{F}_x + (1-\lambda)\hat{F}_y$. Let $w \sim \sum_{k=1}^{K} \lambda_k F_k(w)$ and let $w_1, \ldots, w_K$ denote random variables from the K component dist.

$$\begin{aligned}
V[\theta(\hat{F}_w)] &= E\left[\left(\theta(\hat{F}_w) - E[\theta(\hat{F}_w)]\right)^2\right] \\
&= \sum_{k=1}^{K} \lambda_k E\left[\left(\theta(\hat{F}_{w_k}) - E[\theta(\hat{F}_{w_k})] + E[\theta(\hat{F}_{w_k})] - E[\hat{\theta}(F_w)]\right)^2\right]. \\
&= \sum_{k=1}^{K} \lambda_k V[\theta(\hat{F}_{w_k})] \qquad \text{if} \quad E[\theta(\hat{F}_{w_k})] = E[\theta(\hat{F}_w)], \forall k.
\end{aligned}$$

Asymptotic validity

Result

A permutation test is asymptotically valid if, under the null hypothesis, the two distributions F_x, F_y and the mixture distribution F_w share the same value of the statistic

$$\theta(F_w) = \theta(F_x) = \theta(F_y) \qquad \textit{where} \qquad F_w = \lambda F_x + (1-\lambda) F_y$$

and, either $n/(n+m) \to \lambda = 1/2$ or $V[\theta(\hat{F}_x)] = V[\theta(\hat{F}_y)]$

Permutation test is as. valid if the index is the same in F_x, F_y, F_w

The generalised entropy class of inequality measures

$$I_{\mathrm{GE}}^{\zeta}(F) = \frac{1}{\zeta^2 - \zeta}\left[\int \left[\frac{y}{\mu(F)}\right]^{\zeta} dF(y) - 1\right], \ \zeta \in \mathbb{R}, \zeta \neq 0, 1$$

$$I_{\mathrm{GE}}^{0}(F) = -\int \log\left(\frac{y}{\mu(F)}\right) dF(y)$$

$$I_{\mathrm{GE}}^{1}(F) = \int \frac{y}{\mu(F)} \log\left(\frac{y}{\mu(F)}\right) dF(y)$$

- $I_{\mathrm{GE}}^{0}(F)$ is the Mean Logarithmic Deviation index ($\zeta = 0$)
- $I_{\mathrm{GE}}^{1}(F)$ is the Theil index ($\zeta = 1$).
- The more positive ζ is, the more sensitive is the inequality measure to income differences at the top of the distribution.

A decomposable class of measures

- The GE inequality measure is decomposable by groups:

$$I^{\zeta}_{GE}(\hat{F}_w) = \sum_{k=1}^{K} I^{\zeta}_{GE}(\hat{F}_{w_k}) + I^{\zeta}_{between}$$

where $I^{\zeta}_{between} = 0$ when the groups share a common mean

- Then, permutation test is asymptotically valid if

$$\mu(F_x) = \mu(F_y)$$

This condition does not hold in general.

Rescaled samples

- The GE inequality measures are scale invariant: calculating indices from the original samples or from the *rescaled* samples

$$\left\{\frac{X_1}{\mu(F_x)},\dots,\frac{X_n}{\mu(F_x)}\right\} \quad \text{and} \quad \left\{\frac{Y_1}{\mu(F_y)},\dots,\frac{Y_m}{\mu(F_y)}\right\},$$

  gives similar results
- The rescaled samples have a common mean, equals to one.
- Permutation test is asymptotically valid, when based on

$$\left\{\frac{X_1}{\mu(F_x)},\dots,\frac{X_n}{\mu(F_x)},\frac{Y_1}{\mu(F_y)},\dots,\frac{Y_m}{\mu(F_y)}\right\}.$$

- In practice, population means are replaced by sample means

Bootstrapping under the null

$$Z_s = \left\{ \frac{X_1}{\bar{X}}, \ldots, \frac{X_n}{\bar{X}}, \frac{Y_1}{\bar{Y}}, \ldots, \frac{Y_m}{\bar{Y}} \right\}$$

Permutation approach:

- resample *without* replacement n observations in Z_s to form X_*
- the m remaining observations in Z_s are then used to form Y_*
- compute the statistic from X_* and Y_*

Bootstrap approach:

- resample *with* replacement n observations in Z_s to form $X_\flat$
- resample *with* replacement m observations in Z_s to form $Y_\flat$
- compute the statistic from $X_\flat$ and $Y_\flat$

The bootstrap respects the null (resample from same set of obs.)

Simulation: model design

We test the null $H_0 : \theta(F_x) = \theta(F_y)$ with a two-tailed t-statistic

$$T = \frac{\theta(\hat{F}_x) - \theta(\hat{F}_y)}{\sqrt{V[\theta(\hat{F}_x) - \theta(\hat{F}_y)]}}$$

We consider the following methods:

- asymptotic test
- (standard) bootstrap test
- permutation test based on the combined sample Z^s
- bootstrap under the null based on the combined sample Z^s

We compare Theil indices based on Singh-Maddala distributions

Simulation results in very small sample

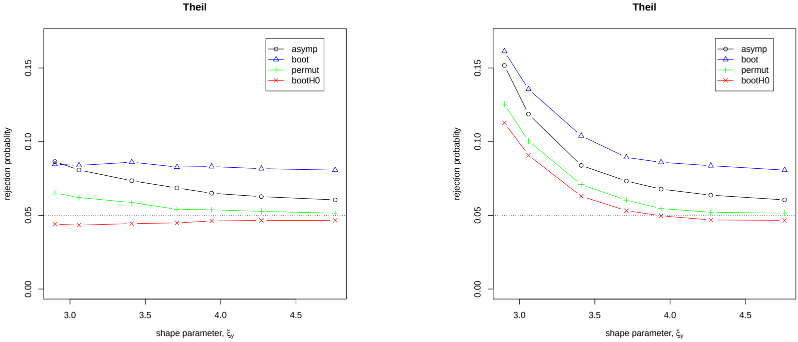


Figure : Rejection frequencies for the Theil index as the upper tail is heavier (as ξ_y decreases). Left panel: $F_x = F_y$. Right panel: $F_x \neq F_y$. $n = 20$, $B = 999$, $R = 10000$, $\alpha = 0.05$.

Simulation results as the sample size increases

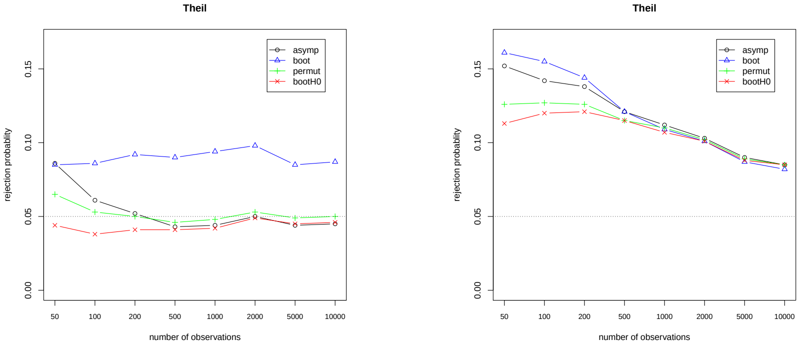


Figure : Rejection frequencies for the Theil inequality index, in the worst case, as the sample size increases. Left panel: $F_x = F_y$. Right panel: $F_x \neq F_y$. $B = 999$, $R = 10000$, $\alpha = 0.05$.

Conclusion

- Simulation results show that when the samples are drawn from two (strongly) heavy-tailed distributions which are not too different, the permutation approach and the proposed bootstrap that respects the null hypothesis perform very well in finite samples.
- When distributions differ dramatically particularly in their tails, while size distortions are not completely eradicated, our proposed methods outperform the standard asymptotic and bootstrap tests.