A Friendly Amendment of the Theil Index

INFORMS October 2009

Kobi Abayomi [1] William Darity Jr. [2]

[1] Asst. Professor
Industrial Engineering - Statistics Group
Georgia Institute of Technology

[2] Department of Public Policy
Duke University and
Department of Economics
University of North Carolina

October 2009

Theil's Index

Theil's Index, a version of Shannon's Entropy, is introduced in econometrics as a measure for inequality. It is improperly specified for statistical use. We explore an adjustment of the Theil index by considering...

- Shannon's original Axiomatization
- Theil's (Mis)-Specification via Shannon
- Adjustment and Re-Specification

Theil's Index

Theil's Index, a version of Shannon's Entropy, is introduced in econometrics as a measure for inequality. It is improperly specified for statistical use. We explore an adjustment of the Theil index by considering...

- Shannon's original Axiomatization
- Theil's (Mis)-Specification via Shannon
- Adjustment and Re-Specification

Theil's Index

Theil's Index, a version of Shannon's Entropy, is introduced in econometrics as a measure for inequality. It is improperly specified for statistical use. We explore an adjustment of the Theil index by considering...

- Shannon's original Axiomatization
- Theil's (Mis)-Specification via Shannon
- Adjustment and Re-Specification

Outline

Shannon's Axioms

Shannon represents an 'Information Source' - a random process on a discrete space - as a Markov process.
He supposes a "measure" *H* should have these qualities:

- *H* should be continuous in p_i
- For $p_i = p, \; \forall \; i$, H should monotonically increase
- "If a choice be broken down, the original [*H*] should be the weighted sum"

Shannon's Axioms

Shannon represents an 'Information Source' - a random process on a discrete space - as a Markov process.
He supposes a "measure" H should have these qualities:

- H should be continuous in p_i
- For $p_i = p, \ \forall \ i$, H should monotonically increase
- "If a choice be broken down, the original [H] should be the weighted sum"

Shannon's Axioms

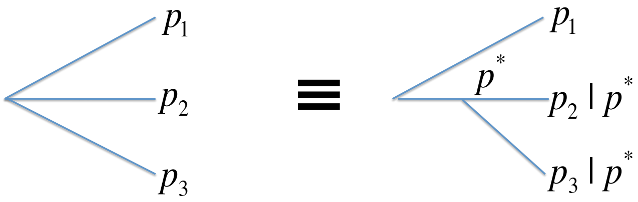
Shannon represents an 'Information Source' - a random process on a discrete space - as a Markov process.
He supposes a "measure" H should have these qualities:

- H should be continuous in p_i
- For $p_i = p, \; \forall \; i$, H should monotonically increase
- "If a choice be broken down, the original [H] should be the weighted sum"

"If a choice be broken down"

Consistency over conditioning...

$$H(p_1, p_2, p_3) \equiv H(p_1, p^*) + p^* H(p_2|p^*, p_3|p^*)$$



Shannon's Theorem

The only H satisfying the three axioms is of the form (1949)

$$H = -K \sum_{i=1}^{n} p_i \, log \, p_i$$

Other 'Entropies'

- Gibbs Entropy: $S = -k_B \sum p_i \, log \, p_i$; (1872, 1878)
- von Neumann Entropy: $S - k_B Tr[\rho \, log_e \, \rho]$

Other 'Entropies'

- Gibbs Entropy: $S = -k_B \sum p_i \, log \, p_i$; (1872, 1878)
- von Neumann Entropy: $S - k_B Tr[\rho \; log_e \; \rho]$

Entropy's Career

Physics	Electrical Engineering → Computing → Computer Science
Statistics	Frechet → Ash → Kullback → Rissanen
Humanities	Theil: Econometrics

Entropy's Career

Physics	Electrical Engineering → Computing → Computer Science
Statistics	Frechet → Ash → Kullback → Rissanen
Humanities	Theil: Econometrics

Entropy's Career

Physics	Electrical Engineering → Computing → Computer Science
Statistics	Frechet → Ash → Kullback → Rissanen
Humanities	Theil: Econometrics

Outline

Theil's Version

A version of Theil's index is

$$T = n^{-1} \sum \frac{x_i}{n^{-1} \sum_j x_j} \, log \, \frac{x_i}{n^{-1} \sum_j x_j}$$

The probability of a particular event/realization is replaced with the income share for a particular element.

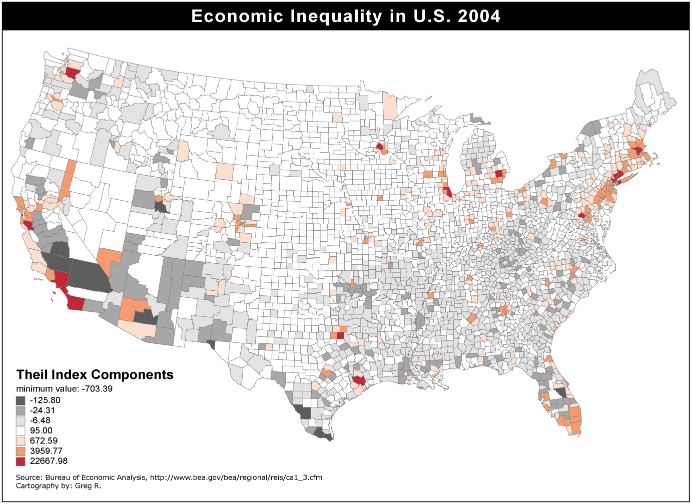
Theil's Decomposition

When collection of elements can be divided into m groups, $g_1, ..., g_m$ each with n_j elements ($=$ individuals); $G = \bigcup g_j$, $g_j \cap g_j^* = \emptyset$, $\forall j \neq j^*$, $\sum_j n_j = n$.

$$T = \sum_G \frac{n_j}{n} \frac{n_j^{-1} \sum_{g_j} x_j}{n^{-1} \sum_G x_j} log \frac{n_j^{-1} \sum_{g_j} x_j}{n^{-1} \sum_G x_j} +$$

$$\sum_G \frac{n_j^{-1} \sum_{g_j} x_j}{n^{-1} \sum_G x_j} \cdot n_j^{-1} \sum_{g_j} \frac{x_j}{n^{-1} \sum_G x_j} log \frac{x_j}{n^{-1} \sum_G x_j}$$

Illustration



Comments

- The probability of a particular event/realization is replaced with the income share for a particular individual.
- We were measuring prob mass of events, now we are measuring the 'size' of individual
- In this sense: The individual is the event, and the income share is the prob mass

Comments

- The probability of a particular event/realization is replaced with the income share for a particular individual.
- We were measuring prob mass of events, now we are measuring the ‘size’ of individual
- In this sense: The individual is the event, and the income share is the prob mass

Comments

- The probability of a particular event/realization is replaced with the income share for a particular individual.
- We were measuring prob mass of events, now we are measuring the 'size' of individual
- In this sense: The individual is the event, and the income share is the prob mass

Comments

- The lack of a true event space (σ-algebra) yields a degenerate probability model...
- Though heuristic is consistent: n-dimensional simplex / Liouville family of distributions.
- But this belies a "deeper confusion" about Entropy and Probability [Jaynes (1965), *American Journal of Physics*].

Comments

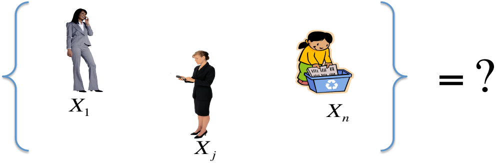
- The lack of a true event space (σ-algebra) yields a degenerate probability model...
- Though heuristic is consistent: n-dimensional simplex / Liouville family of distributions.
- But this belies a "deeper confusion" about Entropy and Probability [Jaynes (1965), *American Journal of Physics*].

Comments

- The lack of a true event space (σ-algebra) yields a degenerate probability model...
- Though heuristic is consistent: *n*-dimensional simplex / Liouville family of distributions.
- But this belies a "deeper confusion" about Entropy and Probability [Jaynes (1965), *American Journal of Physics*].

Illustration

$$\Omega = \left\{ E_1 \quad E_j \quad E_n \right\}$$



Shannon's Axiomatization

Notice:

$$T = log(n) - H$$

and that Shannon's original proof specified

$$H = -K \sum_{i=1}^{n} p_i \; log \; p_i$$

- Shannon's advice: K amounts to a choice of unit of measure (σ-algebra)
- Shannon chose the binary log (base $b = 2$)

Shannon's Axiomatization

Notice:

$$T = log(n) - H$$

and that Shannon's original proof specified

$$H = -K \sum_{i=1}^{n} p_i \; log \; p_i$$

- Shannon's advice: K amounts to a choice of unit of measure (σ-algebra)
- Shannon chose the binary log (base $b = 2$)

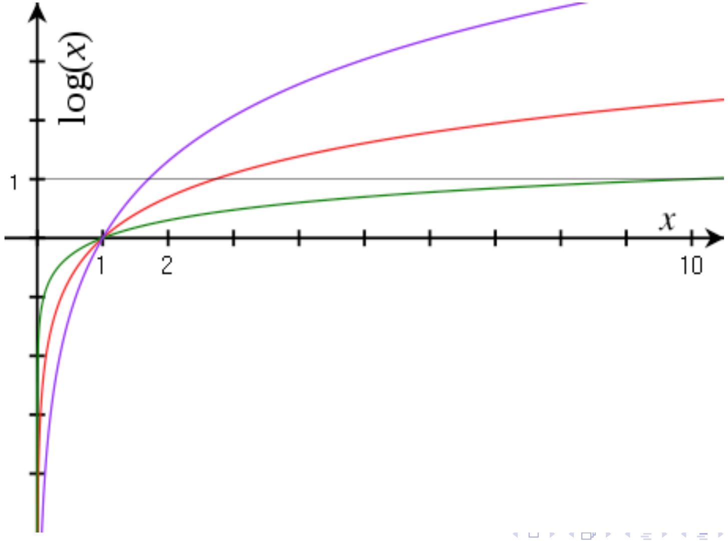
Skew-Sensitivity

One can arbitrarily choose partitions that guarantee across group sum dominates within groups sum.
Merely choose a collection G^* such that $x_j > n_j^{-1} \sum_{g_j} x_j$ for groups $g_1, \ldots, g_{[n/m]}$.
On this collection

$$A \equiv log \frac{n_j^{-1} \sum_{g_j} x_j}{n^{-1} \sum_G x_j} < log \frac{x_j}{n^{-1} \sum_G x_j} \equiv B$$

obscuring within group effects.

Illustration



Log base characterizes 'Granularity'

As well, there exist collections G^* for every choice of b such that on G^* / g^*

$$A, B > 1$$

but

$$A, B < 1$$

on g^*

The Amnesia of (Probability) Measure

- The problems don't arise on Shannon's specification via the unit simplex
- Theil Index is on simplex of arbitrary sum
- Theil Index is an *ex parte* function on probability measures

The Amnesia of (Probability) Measure

- The problems don't arise on Shannon's specification via the unit simplex
- Theil Index is on simplex of arbitrary sum
- Theil Index is an *ex parte* function on probability measures

The Amnesia of (Probability) Measure

- The problems don't arise on Shannon's specification via the unit simplex
- Theil Index is on simplex of arbitrary sum
- Theil Index is an *ex parte* function on probability measures

Outline

Suggestions

Adjust b Rocke and Durbin Log-Linear hybridization; Box-Cox, Tukey families of transforms.

Rank Transform Estimate empirical CDF across and within groups and substitute for raw income shares

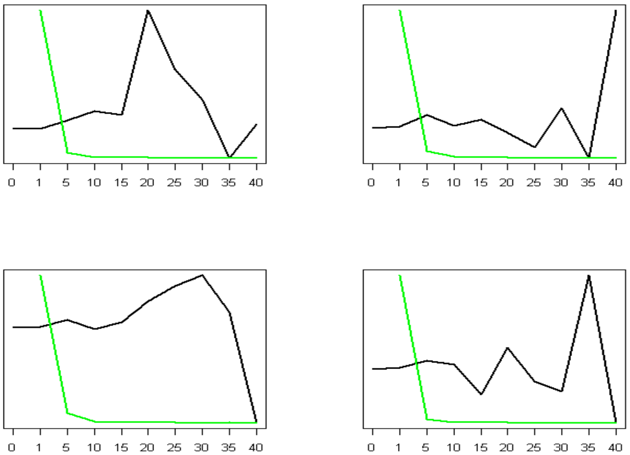
Suggestions

Adjust b Rocke and Durbin Log-Linear hybridization; Box-Cox, Tukey families of transforms.

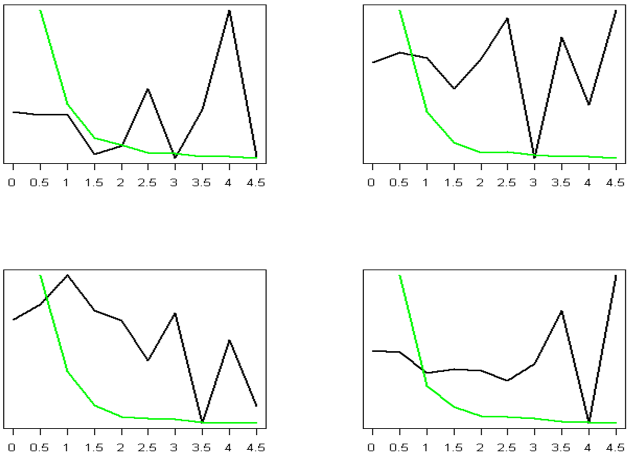
Rank Transform Estimate empirical CDF across and within groups and substitute for raw income shares

Outline

Groups of Equal Mean Uniform Discrete Random Variables



Groups of Equal Mean Uniform Discrete Random Variables



Groups of Equal Mean Uniform Discrete Random Variables, *b* is min.