Statistical Approaches for Constrained/Limited Sums. Some topics in (Environmental) Economics

January 2012

Kobi Abayomi: Valerie Thomas, Dexin Luo, William Darity, Tracy Yandle

January 2012

Outline

Setup

We investigate the effect of biofuels on land use change via U.S. corn production data.

- Agricultural models are used to estimate the effect of biofuel production on crop production $\Rightarrow$ land use change
- Statistical determination of the influence of biofuel production is non-standard $\Rightarrow$ total crop use is always the sum of the different uses
- These (induced) distributions are *Compositional Distributions* $\Rightarrow$ Methodology for Dependency/Competition among Compositional Distributions

Setup

We investigate the effect of biofuels on land use change via U.S. corn production data.

- Agricultural models are used to estimate the effect of biofuel production on crop production $\Rightarrow$ land use change
- Statistical determination of the influence of biofuel production is non-standard $\Rightarrow$ total crop use is always the sum of the different uses
- These (induced) distributions are *Compositional Distributions* $\Rightarrow$ Methodology for Dependency/Competition among Compositional Distributions

Setup

We investigate the effect of biofuels on land use change via U.S. corn production data.

- Agricultural models are used to estimate the effect of biofuel production on crop production $\Rightarrow$ land use change
- Statistical determination of the influence of biofuel production is non-standard $\Rightarrow$ total crop use is always the sum of the different uses
- These (induced) distributions are *Compositional Distributions* $\Rightarrow$ Methodology for Dependency/Competition among Compositional Distributions

Energy Independence and Security Act

- EISA-2007 mandates an increase in ethanol production to 36 billion gallons per year by 2022.
- Ethanol production has increased more than 5000 percent since 1980
- At the same time, the total U.S. corn yield has less than doubled

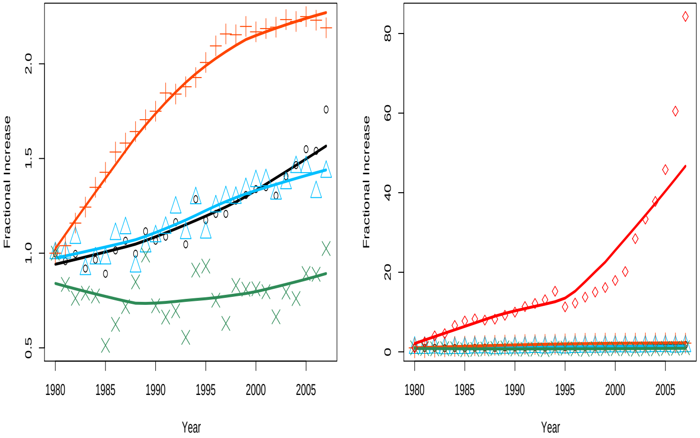
Energy Independence and Security Act

- EISA-2007 mandates an increase in ethanol production to 36 billion gallons per year by 2022.
- Ethanol production has increased more than 5000 percent since 1980
- At the same time, the total U.S. corn yield has less than doubled

Energy Independence and Security Act

- EISA-2007 mandates an increase in ethanol production to 36 billion gallons per year by 2022.
- Ethanol production has increased more than 5000 percent since 1980
- At the same time, the total U.S. corn yield has less than doubled

Fractional Increase



Environmental Impacts

- Net energy budget
- Competition with corn-based commodities
- Greenhouse emissions
- $\Rightarrow$ Competition within distribution of Corn Yield constitutents

Environmental Impacts

- Net energy budget
- Competition with corn-based commodities
- Greenhouse emissions
- $\Rightarrow$ Competition within distribution of Corn Yield constitutents

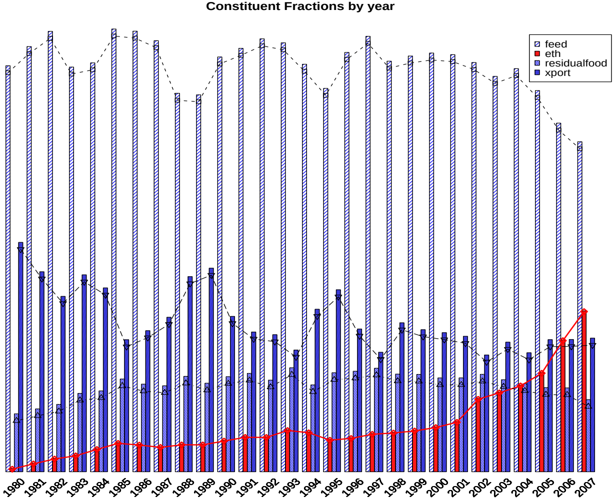
Environmental Impacts

- Net energy budget
- Competition with corn-based commodities
- Greenhouse emissions
- $\Rightarrow$ Competition within distribution of Corn Yield constitutents

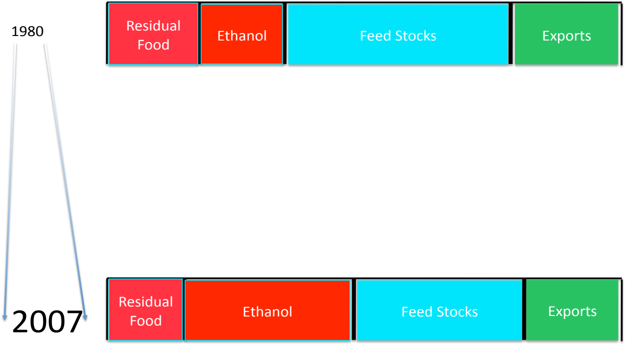
Environmental Impacts

- Net energy budget
- Competition with corn-based commodities
- Greenhouse emissions
- $\Rightarrow$ Competition within distribution of Corn Yield constitutents

Constituents of Corn Yield



Compositional Distribution



Aitchison Notation

Let

$$\mathbf{x} = (x_1, ..., x_k) \tag{1}$$

be a basis or open vector of positive quantities
In this example

$$\mathbf{x} = (x_{eth}, x_{rfood}, x_{feed}, x_{xport})$$

(in bushels) corn of: ethanol production, residual food stock, feed stock, and exports.

Aitchison Notation

Let

$$y_j = x_j / \sum_j^k x_j \tag{2}$$

$\mathbf{y} = (y_1, ..., y_k)$ the vector of fractions.
Aitchison defines $y_{k+1} = 1 - \sum_j^k y_j$;
Here $\sum_j^k y_j = 1$

Aitchison Notation

A (log-ratio) transformation sets

$$v_{\substack{j \\ m}} = log(\frac{y_j}{y_m}) = log\, y_j - log\, y_m \tag{3}$$

in a slight modification of Aitchison's notation
(where $v_j = log(y_j/y_{k+1})$).

Modifying Aitchison Notation

- The total is fixed and known $\Rightarrow$ the residual is y_{k+1}=0 and Aitchison's v_j is undefined
- In the original notation v_j is the log of the relative fraction of constituent j to the residual component of the basis
- $v_{\underset{m}{\cdot}}$, maps $\mathbb{S}^{m,k} = \{\mathbf{y}_{-m}, y_m\}$ to $\mathbb{R}^{k-|m|}$

Modifying Aitchison Notation

- The total is fixed and known $\Rightarrow$ the residual is y_{k+1}=0 and Aitchison's v_j is undefined
- In the original notation v_j is the log of the relative fraction of constituent j to the residual component of the basis
- $v_{\underset{m}{\cdot}}$, maps $\mathbb{S}^{m,k} = \{\mathbf{y}_{-m}, y_m\}$ to $\mathbb{R}^{k-|m|}$

Modifying Aitchison Notation

- The total is fixed and known $\Rightarrow$ the residual is y_{k+1}=0 and Aitchison's v_j is undefined
- In the original notation v_j is the log of the relative fraction of constituent j to the residual component of the basis
- $v_{\underset{m}{\cdot}}$, maps $\mathbb{S}^{m,k} = \{\mathbf{y}_{-m}, y_m\}$ to $\mathbb{R}^{k-|m|}$

Why Transform?

(Natural) Dirichlet model for $(y_1, ..., y_k)$; $\sum_j y_j = 1$; $y_j > 0 \; \forall j$ is:

$$dF(\mathbf{y}) \propto (1 - \sum_j y_j)^{\alpha_{k+1}-1} \cdot \prod_j y_j^{\alpha_j - 1} \tag{4}$$

with parameters $\alpha = (\alpha_1, ..., \alpha_{k+1})$
Insufficient for non-*neutral* proportions

Why Transform?

(Generalization) Liouville distribution is:

$$dF \propto h(\sum_j y_j) \prod y_j^{\alpha_j - 1} \tag{5}$$

with $\alpha_j > 0$ (as before) and g some function.
Note that when $h(t) = 1 - t$ the Liouville distribution is the special case Dirichlet distribution with $\alpha_{k+1} = 1$
Thus h is an additional parameter of interest for estimation

Aitchison Approach

- Aitchison fits a log-normal distribution on $\mathbf{v}$
- Σ is sufficient for dependency in the log-normal distribution

Aitchison Approach

- Aitchison fits a log-normal distribution on $\mathbf{v}$
- Σ is sufficient for dependency in the log-normal distribution

Aitchison Approach

- Under a composition

$$\Sigma_{\mathbf{v}} \propto diag(\omega_1, ..., \omega_k) + \omega_{k+1}, \tag{6}$$

- $\Sigma_{\mathbf{v}}$ is constrained to the positive orthant and proportional to the units of the residual component

Aitchison Approach

- Under a composition

$$\Sigma_{\mathbf{v}} \propto diag(\omega_1, ..., \omega_k) + \omega_{k+1}, \tag{6}$$

- $\Sigma_{\mathbf{v}}$ is constrained to the positive orthant and proportional to the units of the residual component

Aitchison Approach

- Aitchison uses a likelihood ratio test
- Test statistic is iteratively estimated due to the constraints on the support of the parameter space
- $\Sigma_{\mathbf{v}} \geq \mathbf{0}$ and proportional to the choice of y_{k+1}

Aitchison Approach

- Aitchison uses a likelihood ratio test
- Test statistic is iteratively estimated due to the constraints on the support of the parameter space
- $\Sigma_{\mathbf{v}} \geq \mathbf{0}$ and proportional to the choice of y_{k+1}

Aitchison Approach

- Aitchison uses a likelihood ratio test
- Test statistic is iteratively estimated due to the constraints on the support of the parameter space
- $\Sigma_{\mathbf{v}} \geq \mathbf{0}$ and proportional to the choice of y_{k+1}

Our Approach

Multivariate Version of KS distance:

$$D_{n,k} = \sup_{\mathbf{t}} |F_n(\mathbf{t}) - F(\mathbf{t})| \tag{7}$$

For $\mathbf{t} = (t_1, t_2, ...)$ the distance is a probability measure on Kendall's distributions...chi-square convergence does not hold (Nelsen 2003).

Our Approach

A Kolmogorov-Smirnov statistic (distance) for multivariate independence can be written:

$$D_{n,k}^{\Pi} = \sup_{\mathbf{t}} |F_n(\mathbf{t}) - \prod_j F_j(t_j)|. \tag{8}$$

Under independence the distance converges to zero (via Glivenko-Cantelli), but distributionally for $k \geq 2$?

Our Approach

- Let $\mathbf{u} = (u_1, ..., u_k)$, where each $u_j = F_j(v_j)$
- Let the joint distribution for $\mathbf{v}$ be $F(\mathbf{v})$
- The *copula* for $\mathbf{u}$ is

$$C(\mathbf{u}) = F(F_1(v_1), ..., F_k(v_k)) \tag{9}$$

Our Approach

- Let $\mathbf{u} = (u_1, ..., u_k)$, where each $u_j = F_j(v_j)$
- Let the joint distribution for $\mathbf{v}$ be $F(\mathbf{v})$
- The *copula* for $\mathbf{u}$ is

$$C(\mathbf{u}) = F(F_1(v_1), ..., F_k(v_k)) \tag{9}$$

Our Approach

- Let $\mathbf{u} = (u_1, ..., u_k)$, where each $u_j = F_j(v_j)$
- Let the joint distribution for $\mathbf{v}$ be $F(\mathbf{v})$
- The *copula* for $\mathbf{u}$ is

$$C(\mathbf{u}) = F(F_1(v_1), ..., F_k(v_k)) \tag{9}$$

Our Approach

With $C_n(\cdot)$ a multivariate version of the *empirical copula*:

$$C_n(\mathbf{u}) = \frac{\#\{\mathbf{t} \mid t_1 \leq F_1^{-1}(u_1), ..., t_k \leq F_k^{-1}(u_k)\}}{n} \tag{10}$$

Our Approach

- Fit a *Dirichlet* distribution (i.e. estimate $\hat{\alpha} = (\hat{\alpha}_1, ..., \hat{\alpha}_k)$ for $\alpha = (\alpha_1, ..., \alpha_k)$) to the composition data $\mathbf{y}$.
- Generate T *Dirichlet* replicates, parameter $\hat{\alpha}$, each of dimension $n \times k$: $(\mathbf{y}^{\hat{\alpha},1}, ..., \mathbf{y}^{\alpha,T})$.
- Compute $m = 1...k$ versions of Aitchison's log-ratios on the replicates: $\mathbf{v}_m^{\hat{\alpha},1} ..., \mathbf{v}_m^{\hat{\alpha},T}$
- For $m = 1..k$ compute $D_{\substack{n,k \\ m}}^{\Pi,1}, ..., D_{\substack{n,k \\ m}}^{\Pi,T}$ of

$$D_{n,k}^{\Pi} = \sup_{\mathbf{u}} |C_n(\mathbf{u}^{\hat{\alpha}}) - \prod_j u_j^{\hat{\alpha}_j}|. \tag{11}$$

Our Approach

- Fit a *Dirichlet* distribution (i.e. estimate $\hat{\alpha} = (\hat{\alpha}_1, ..., \hat{\alpha}_k)$ for $\alpha = (\alpha_1, ..., \alpha_k)$) to the composition data $\mathbf{y}$.
- Generate T *Dirichlet* replicates, parameter $\hat{\alpha}$, each of dimension $n \times k$: $(\mathbf{y}^{\hat{\alpha},1}, ..., \mathbf{y}^{\alpha,T})$.
- Compute $m = 1...k$ versions of Aitchison's log-ratios on the replicates: $\mathbf{v}_m^{\hat{\alpha},1} ..., \mathbf{v}_m^{\hat{\alpha},T}$
- For $m = 1..k$ compute $D_{\substack{n,k \\ m}}^{\Pi,1}, ..., D_{\substack{n,k \\ m}}^{\Pi,T}$ of

$$D_{n,k}^{\Pi} = \sup_{\mathbf{u}} |C_n(\mathbf{u}^{\hat{\alpha}}) - \prod_j u_j^{\hat{\alpha}_j}|. \tag{11}$$

Our Approach

- Fit a *Dirichlet* distribution (i.e. estimate $\hat{\alpha} = (\hat{\alpha}_1, ..., \hat{\alpha}_k)$ for $\alpha = (\alpha_1, ..., \alpha_k)$) to the composition data $\mathbf{y}$.
- Generate T *Dirichlet* replicates, parameter $\hat{\alpha}$, each of dimension $n \times k$: $(\mathbf{y}^{\hat{\alpha},1}, ..., \mathbf{y}^{\alpha,T})$.
- Compute $m = 1...k$ versions of Aitchison's log-ratios on the replicates: $\mathbf{v}_m^{\hat{\alpha},1} ..., \mathbf{v}_m^{\hat{\alpha},T}$
- For $m = 1..k$ compute $D_{\substack{n,k\\m}}^{\Pi,1}, ..., D_{\substack{n,k\\m}}^{\Pi,T}$ of

$$D_{n,k}^{\Pi} = \sup_{\mathbf{u}} |C_n(\mathbf{u}^{\hat{\alpha}}) - \prod_j u_j^{\hat{\alpha}_j}|. \tag{11}$$

Our Approach

- Fit a *Dirichlet* distribution (i.e. estimate $\hat{\alpha} = (\hat{\alpha}_1, ..., \hat{\alpha}_k)$ for $\alpha = (\alpha_1, ..., \alpha_k)$) to the composition data $\mathbf{y}$.
- Generate T *Dirichlet* replicates, parameter $\hat{\alpha}$, each of dimension $n \times k$: $(\mathbf{y}^{\hat{\alpha},1}, ..., \mathbf{y}^{\alpha,T})$.
- Compute $m = 1...k$ versions of Aitchison's log-ratios on the replicates: $\mathbf{v}_m^{\hat{\alpha},1} ..., \mathbf{v}_m^{\hat{\alpha},T}$
- For $m = 1..k$ compute $D_{\substack{n,k\\m}}^{\Pi,1}, ..., D_{\substack{n,k\\m}}^{\Pi,T}$ of

$$D_{n,k}^{\Pi} = \sup_{\mathbf{u}} |C_n(\mathbf{u}^{\hat{\alpha}}) - \prod_j u_j^{\hat{\alpha}_j}|. \tag{11}$$

Our Approach

- Yields a distribution for the statistic under an independence hypothesis among the compositions...
- m versions of $D^{\Pi,1}_{n,k \atop m}, ..., D^{\Pi,T}_{n,k \atop m}$ are proxies for tests of *complete subcompositional independence...*
- Calculating on the log-ratios ($\mathbf{v}$) of the replicates, and not the *Dirichlet* draws picks each of m components to serve as 'residual' (via the basis $\mathbf{x}$ or composition $\mathbf{y}$) without requiring m estimates of α and m-fold random draws.

Our Approach

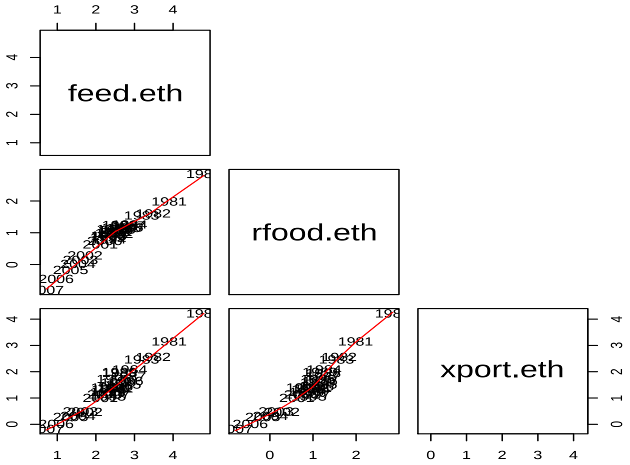
- Yields a distribution for the statistic under an independence hypothesis among the compositions...
- m versions of $D^{\Pi,1}_{\underset{m}{n,k}}, \ldots, D^{\Pi,T}_{\underset{m}{n,k}}$ are proxies for tests of *complete subcompositional independence...*
- Calculating on the log-ratios ($\mathbf{v}$) of the replicates, and not the *Dirichlet* draws picks each of m components to serve as 'residual' (via the basis $\mathbf{x}$ or composition $\mathbf{y}$) without requiring m estimates of α and m-fold random draws.

Our Approach

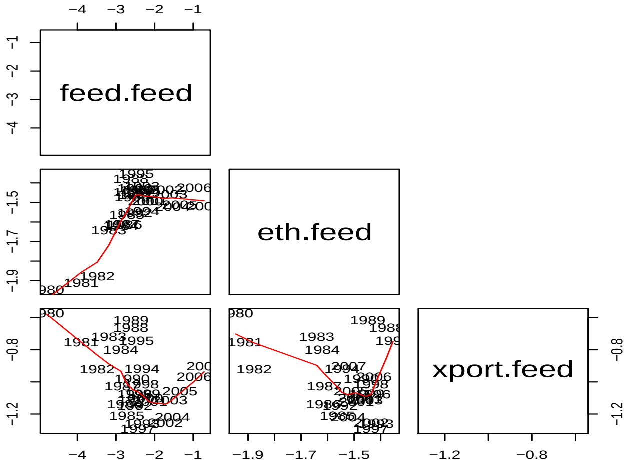
- Yields a distribution for the statistic under an independence hypothesis among the compositions...
- m versions of $D^{\Pi,1}_{\frac{n,k}{m}}, \ldots, D^{\Pi,T}_{\frac{n,k}{m}}$ are proxies for tests of *complete subcompositional independence...*
- Calculating on the log-ratios (**v**) of the replicates, and not the *Dirichlet* draws picks each of m components to serve as 'residual' (via the basis **x** or composition **y**) without requiring m estimates of α and m-fold random draws.

Compositional Distribution

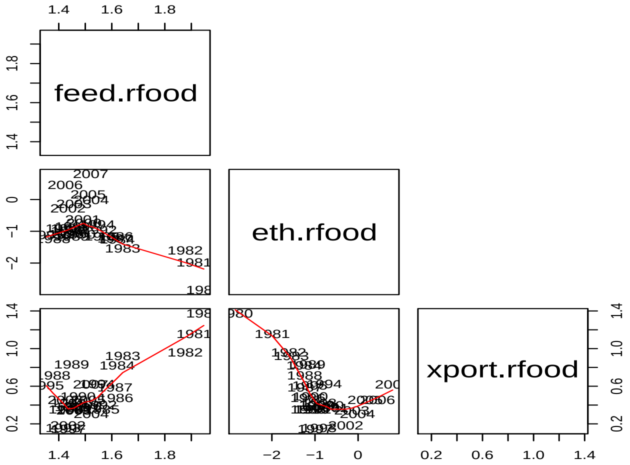
W.R.T. Ethanol


feed.eth
rfood.eth
xport.eth
1981
1982
1983
2002
2004
2005
2006

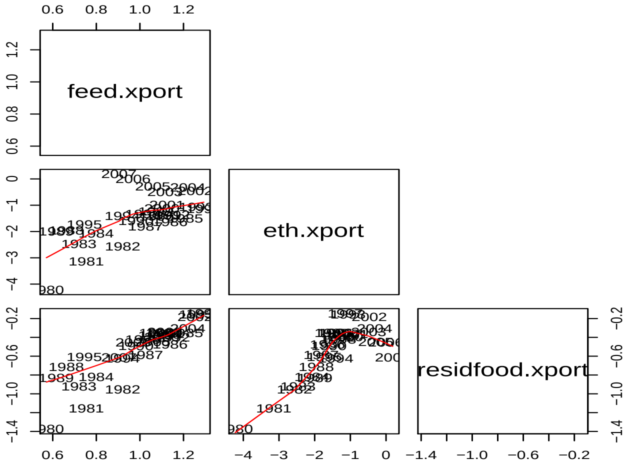
W.R.T. Feed



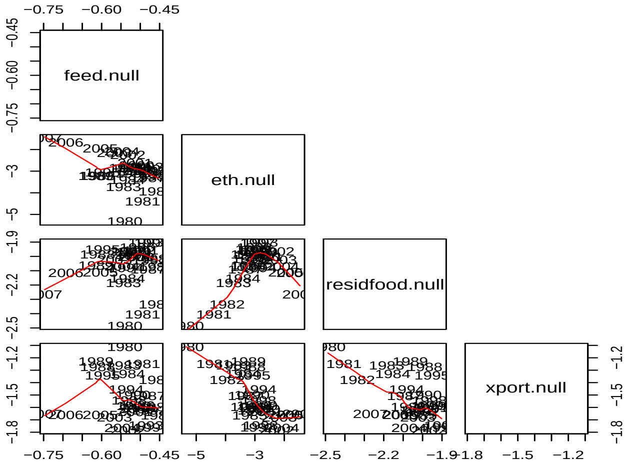
W.R.T. Food



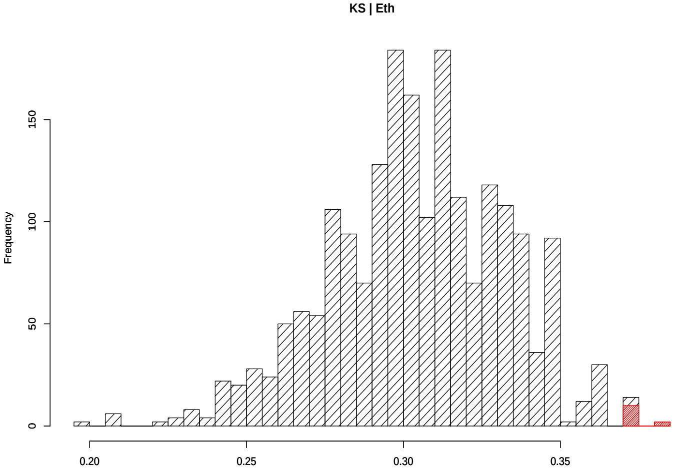
W.R.T. Export


feed.xport
eth.xport
residfood.xport
0.6
0.8
1.0
1.2
0
−1
−2
−3
−4
−0.2
−0.6
−1.0
−1.4

W.R.T. Null

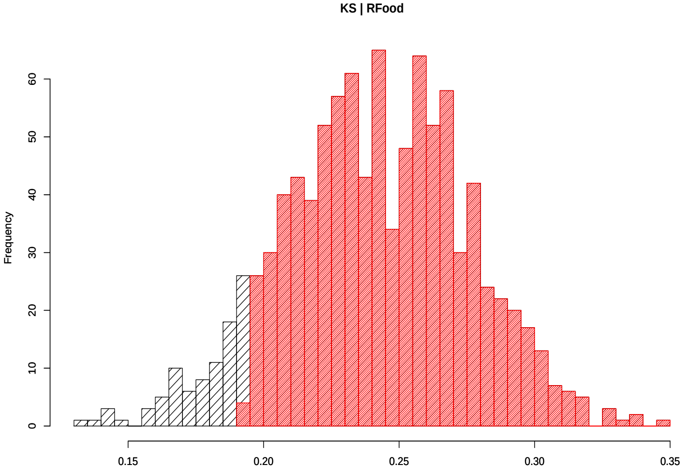

feed.null
eth.null
residfood.null
xport.null
−0.75
−0.60
−0.45
−5
−3
−2.5
−2.2
−1.9
−1.8
−1.5
−1.2

W.R.T. Ethanol

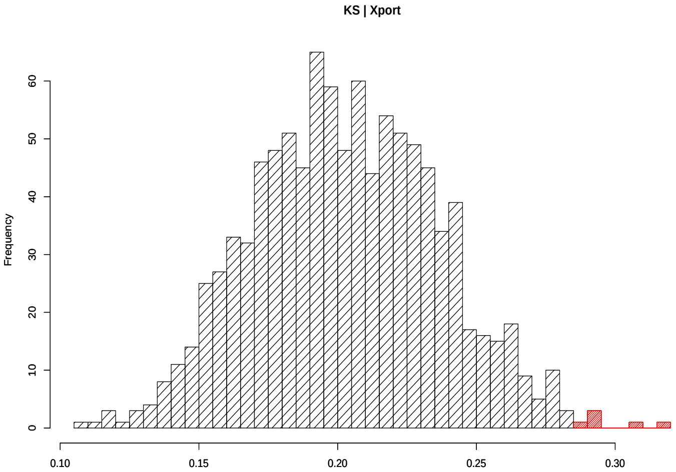


W.R.T. Feed

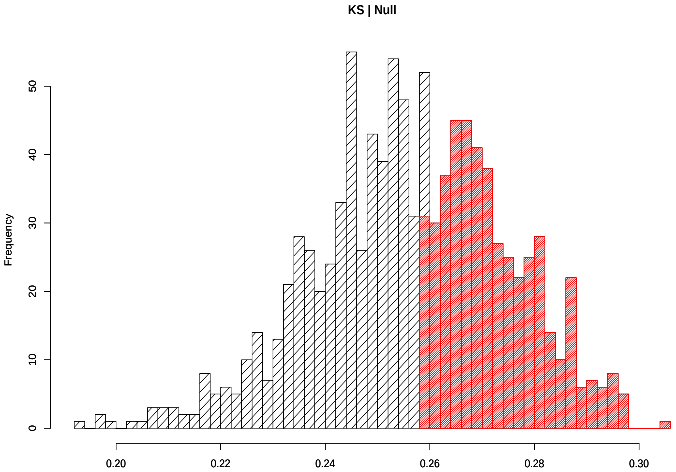
W.R.T. Residual Food



W.R.T. Export



W.R.T. Null



Comments

- Distance is $\mathcal{L}_\infty$ norm, dominates $\mathcal{L}_2$ - Cramer von-Mises distance
- Empirical Prob Integral Transform : : Order statistics $\Rightarrow$ Invariant to increasing (log-ratio) transform.
- In high dimensions Aitchison's method 'tail-migrates'
- Distance is Euclidean (not on Aitchison Geometry!) on prob measure space.

Comments

- Distance is $\mathcal{L}_\infty$ norm, dominates $\mathcal{L}_2$ - Cramer von-Mises distance
- Empirical Prob Integral Transform : : Order statistics $\Rightarrow$ Invariant to increasing (log-ratio) transform.
- In high dimensions Aitchison's method 'tail-migrates'
- Distance is Euclidean (not on Aitchison Geometry!) on prob measure space.

Comments

- Distance is $\mathcal{L}_\infty$ norm, dominates $\mathcal{L}_2$ - Cramer von-Mises distance
- Empirical Prob Integral Transform : : Order statistics $\Rightarrow$ Invariant to increasing (log-ratio) transform.
- In high dimensions Aitchison's method 'tail-migrates'
- Distance is Euclidean (not on Aitchison Geometry!) on prob measure space.

Comments

- Distance is $\mathcal{L}_\infty$ norm, dominates $\mathcal{L}_2$ - Cramer von-Mises distance
- Empirical Prob Integral Transform : : Order statistics $\Rightarrow$ Invariant to increasing (log-ratio) transform.
- In high dimensions Aitchison's method 'tail-migrates'
- Distance is Euclidean (not on Aitchison Geometry!) on prob measure space.

Deeper Issues

- "Richness" of *Dirichlet* replicates w.r.t generalized Liouville?
- Bayesian ‘prior’ for $\alpha \Rightarrow$ posterior distribution for D
- Time dimension?

Deeper Issues

- "Richness" of *Dirichlet* replicates w.r.t generalized Liouville?
- Bayesian 'prior' for $\alpha \Rightarrow$ posterior distribution for *D*
- Time dimension?

Deeper Issues

- "Richness" of *Dirichlet* replicates w.r.t generalized Liouville?
- Bayesian ‘prior’ for $\alpha \Rightarrow$ posterior distribution for *D*
- Time dimension?

Outline

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

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

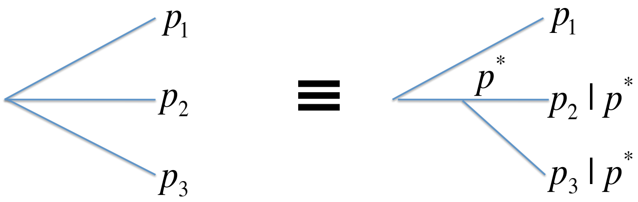
Shannon represents an ‘Information Source’ - a random process on a discrete space - as a Markov process.
He supposes a "measure" H should have these qualities:

- 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 $\rightarrow$ Computing $\rightarrow$ Computer Science
Statistics	Frechet $\rightarrow$ Ash $\rightarrow$ Kullback $\rightarrow$ Rissanen
Humanities	Theil: Econometrics

Entropy's Career

Physics	Electrical Engineering $\rightarrow$ Computing $\rightarrow$ Computer Science
Statistics	Frechet $\rightarrow$ Ash $\rightarrow$ Kullback $\rightarrow$ 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.

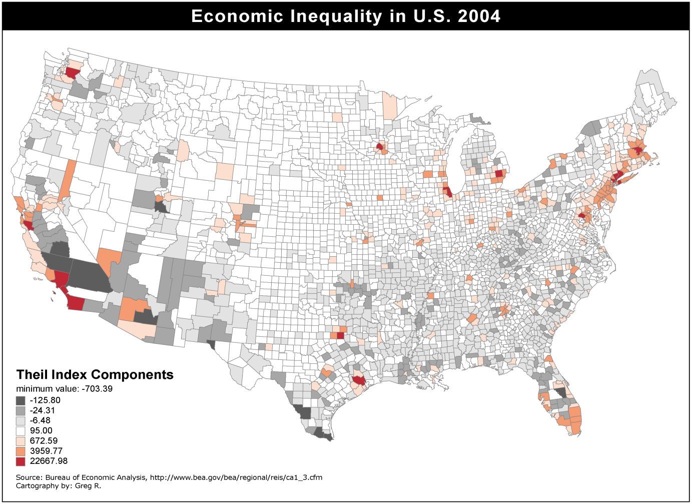
Theil's Decomposition

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

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 $b = 2$)

Partitioning T

Consider this rewrite

$$T = \sum_{G} \frac{n_g}{n} \frac{\overline{X}_g}{\overline{\overline{X}}} log_b \frac{\overline{X}_g}{\overline{\overline{X}}} + \sum_{G} \frac{\overline{X}_g}{\overline{\overline{X}}} \frac{1}{n_g} \sum_{g} \frac{X_{ig}}{\overline{X}_g} log_b \frac{X_{ig}}{\overline{X}_g} \tag{12}$$

Decomposition

The first term on the rhs

$$Across = \sum_{G} \frac{n_j}{n} \frac{\overline{X}_g}{\overline{\overline{X}}} \, log_b \frac{\overline{X}_g}{\overline{\overline{X}}} \tag{13}$$

is the measure of the *across* or *between* group inequality; the second term

$$Within = \sum_{G} \frac{\overline{X}_g}{\overline{\overline{X}}} \frac{1}{n_g} \sum_{g} \frac{X_{ig}}{\overline{\overline{X}}_g} log_b \frac{X_{ig}}{\overline{\overline{X}}_g} \tag{14}$$

the *within* group inequality.

Theil's Version

Since

$$\frac{log_{b_0}\ t}{log_{b_1}\ t} = K,\ \forall t \tag{15}$$

K can serve as a conversion between information units b_0 and b_1. **This feature is elided from Theil's construction** with the loss of Shannon's constant.

Misspecification

Examine the expectation of the *within* term

Theorem

$E[Within] \geq 0$

Misspecification

Proof.

$$E[Within] = E[E[\sum_G \frac{\overline{X}_g}{\overline{X}} \frac{1}{n_g} \sum_g \frac{X_{ig}}{\overline{X}_g} log_b \frac{X_{ig}}{\overline{X}_g} | G = g]] \qquad (16)$$

which yields

$$E[Within] = E[\sum_G \frac{1}{n_g} \frac{\overline{X}_g}{\overline{X}_g} E[\frac{1}{\overline{X}} \sum_g X_{ig} log_b \frac{X_{ig}}{\overline{X}_g} | G = g]] \qquad (17)$$

and then, by conditioning on $G = g$

$$E[Within] = E[\sum_G \frac{1}{\overline{X} n_g} \sum_g E[X_{ig} log_b \frac{X_{ig}}{\overline{X}_g} | G = g]] \qquad (18)$$

$$= E[\sum_G \frac{1}{\overline{X} n_g} E[\sum_g X_{ig} log_b \frac{X_{ig}}{\overline{X}_g}]] \qquad (19)$$

Misspecification

$$E[Within] = E[\sum_G \frac{1}{\overline{X} n_g} \sum_g E[X_{ig} log_b \frac{X_{ig}}{\overline{X}_g} | G = g]] \tag{22}$$

$$= E[\sum_G \frac{1}{\overline{X} n_g} E[\sum_g X_{ig} log_b \frac{X_{ig}}{\overline{X}_g}]] \tag{23}$$

$$\geq E[\sum_G \frac{1}{\overline{X} n_g} E[n_g \overline{X}_g log_b \frac{n_g \overline{X}_g}{n_g \overline{X}_g}]] \tag{24}$$

$$\geq E[\sum_G \frac{1}{\overline{X} n_g} E[0]] = 0 \tag{25}$$

Misspecification

Now the across term

Theorem

$$E[Across] \leq E[\sum_{G} \beta_g^b]$$

when

$$\sum_{g_i \neq g_j} \alpha_{g_i} \beta_{g_j}^b > 0 \tag{26}$$

Misspecification

Proof.

With $\sum_G \alpha_g = 1$, and $\alpha_g \geq 0$, by assumption, $\forall g$. Then:

$$E[Across] = E[\sum_G \frac{n_g}{n} \frac{\overline{X}_g}{\overline{\overline{X}}} \, log_b \frac{\overline{X}_g}{\overline{\overline{X}}}] = E[\sum_G \alpha_g \, \beta_g^b] \tag{27}$$

$$E[Across] = E[\sum_G \alpha_g \, \beta_g^b] \tag{28}$$

$$= E[\sum_G \alpha_g \, \sum_G \beta_g^b - \sum_{g_i \neq g_j} \alpha_{g_i} \beta_{g_j}^b] \tag{29}$$

$$\leq E[\sum_G \alpha_g \, \sum_G \beta_g^b] = E[1 \cdot \sum_G \beta_g^b] \tag{30}$$

$$= E[\sum_G \beta_g^b] \tag{31}$$

Misspecification

Consider these propositions:

Theorem

(I)

$$b \leq \min_G(\overline{X}_g/\overline{X}) \implies E(\sum_G \beta_g^b) \geq 0 \quad (32)$$

(II)

$$b \geq \max_G(\overline{X}_g/\overline{X}) \implies E(\sum_G \beta_g^b) \leq 0 \quad (33)$$

Misspecification

Consider these propositions:

Theorem

(I)

$$b \leq \min_G(\overline{X}_g/\overline{X}) \implies E(\sum_G \beta_g^b) \geq 0 \tag{32}$$

(II)

$$b \geq \max_G(\overline{X}_g/\overline{X}) \implies E(\sum_G \beta_g^b) \leq 0 \tag{33}$$

Misspecification

Consider these propositions:

Theorem

(I)

$$b \leq \min_G(\overline{X}_g/\overline{X}) \implies E(\sum_G \beta_g^b) \geq 0 \tag{32}$$

(II)

$$b \geq \max_G(\overline{X}_g/\overline{X}) \implies E(\sum_G \beta_g^b) \leq 0 \tag{33}$$

Proof Illustration

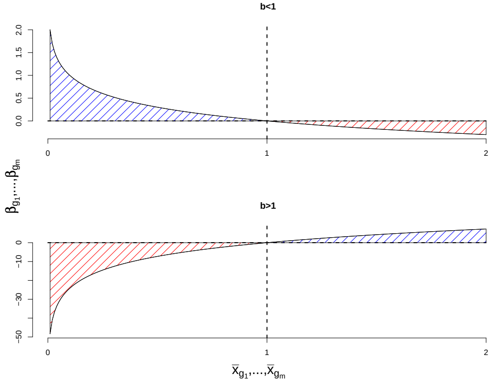


Figure: Illustration of β_g vs. $\overline{x}_g$ for log base $b < 1$ and $b > 1$. The contribution to the across term of Theil's index - T, equation (13), is concave up or down by choice of b. See paper.

Misspecification

- Choosing b is like choosing ‘natural’ ratio of group to overall inequality
- The effect is inflated in data where few, small groups n_g have high, or low, income relative to population size
- The effect in b is non-linear, while the contribution n_g is only linear.

Misspecification

- Choosing b is like choosing 'natural' ratio of group to overall inequality
- The effect is inflated in data where few, small groups n_g have high, or low, income relative to population size
- The effect in b is non-linear, while the contribution n_g is only linear.

Misspecification

- Choosing b is like choosing 'natural' ratio of group to overall inequality
- The effect is inflated in data where few, small groups n_g have high, or low, income relative to population size
- The effect in b is non-linear, while the contribution n_g is only linear.

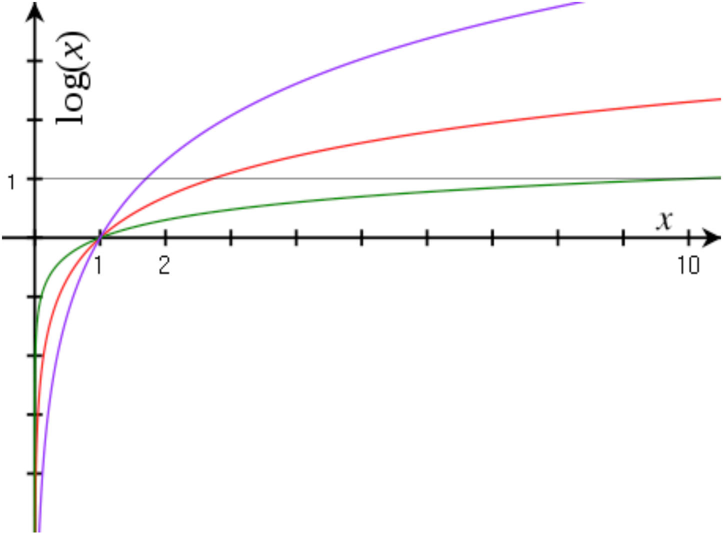
Misspecification

- Choosing b is like choosing 'natural' ratio of group to overall inequality
- The effect is inflated in data where few, small groups n_g have high, or low, income relative to population size
- The effect in b is non-linear, while the contribution n_g is only linear.

Misspecification

- Choosing b is like choosing 'natural' ratio of group to overall inequality
- The effect is inflated in data where few, small groups n_g have high, or low, income relative to population size
- The effect in b is non-linear, while the contribution n_g is only linear.

Illustration



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

Technical Comments

- The lack of a true event space yields a degenerate probability model (σ algebra)
- Though the heuristic is consistent: n dimensional simplex / Liouville family of distributions.

Technical Comments

- The lack of a true event space yields a degenerate probability model (σ algebra)
- Though the heuristic is consistent: *n* dimensional simplex / Liouville family of distributions.

Technical Comments

- The lack of a true event space yields a degenerate probability model (σ algebra)
- Though the heuristic is consistent: *n* dimensional simplex / Liouville family of distributions.

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

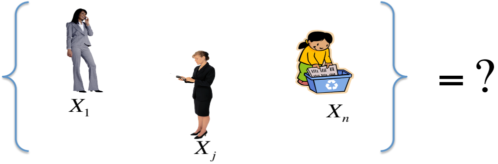
- The lack of a true event space (σ-algebra) yields a degenerate probability model...
- Though heuristic is consistent: n-dimensional simplex / Liouville family of distributions.
- But this belies a "deeper confusion" about Entropy and Probability [Jaynes (1965), *American Journal of Physics*].

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

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



A T-test for T

$$\begin{aligned} H_0 &: \mathcal{T} = 0 \\ vs. \\ H_a &: \mathcal{T} > 0 \end{aligned} \tag{34}$$

$$p - value = \mathbb{P}_{H_0:\mathcal{T}=0,b}\left(Z > \frac{T}{s.e.(T)}\right) \tag{35}$$

U Mich HRS Data

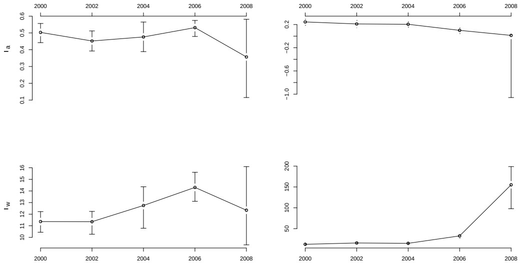


Figure: Illustration of Theil's index calculated on wealth - left hand column - and income - right hand column - using the University of Michigan's Health and Retirement Survey (HRS) data: 2000, 2002, 2004, 2006, and 2008. The upper row is the across term, the lower row is the within term. Both terms are fixed by log base $b = \min(\overline{x}_g/\overline{x})$ the ratio of the poorer (black) group sample mean to the overall mean.

Outline

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Or inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

Motivation

Consolidation

Market Consolidation in ITQs

- Individual Transferable Quotas (ITQ); Fishery Harvesting Permits
  - Mechanism for control of fishery management practices
- Consolidation and aggregation of catching rights are well-documented issues in ITQs.
  - Is consolidation an indicator of efficiency?
  - Of inequality?

GOAL: Investigate distribution of ITQs as measurements of market consolidation/inequality

New Zealand Fisheries

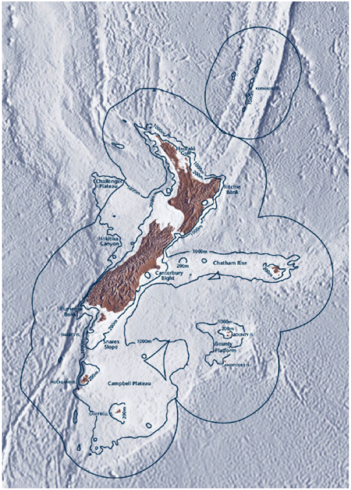
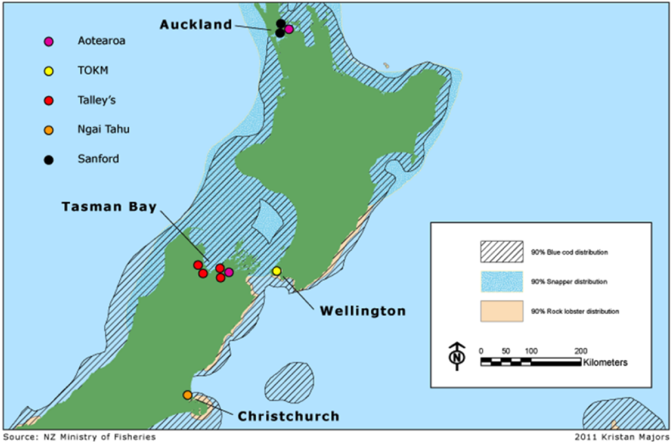
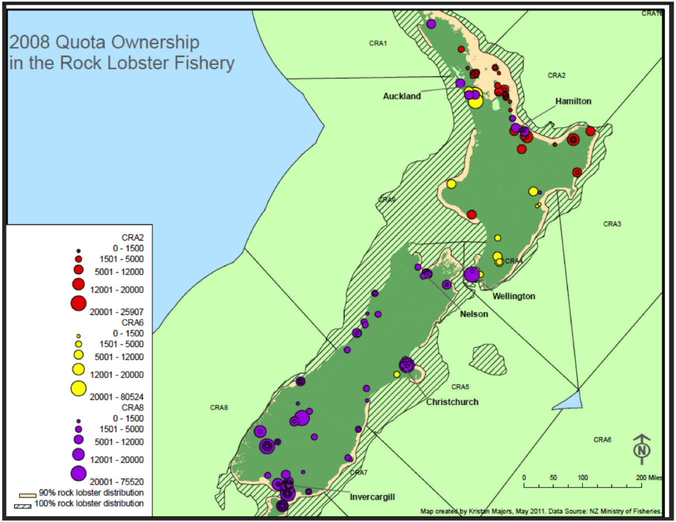


Figure: EEZ = 15 times land mass. NZ $4.0 billion; Seafood exports NZ $ 1.4 billion. Deep water fisheries: commercial, vertically integrated. Inshore

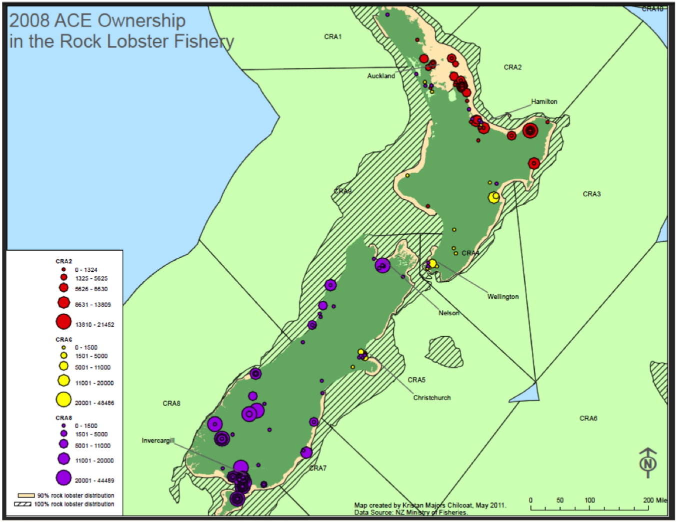
Large Fishing Rights Holders



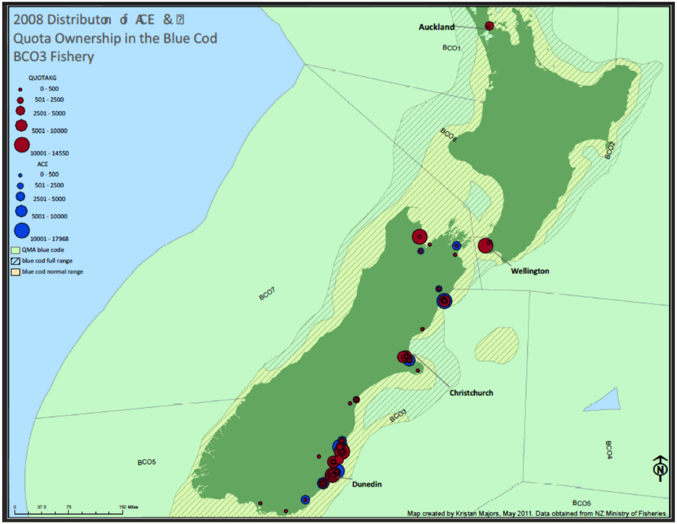
Rock Lobster ITQs, Permanent Catching Rights



Rock Lobster ACEs, Leased Catching Rights



Blue Cod ITQ and ACE



Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some 'random' process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some 'random' process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some 'random' process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some 'random' process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some 'random' process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some 'random' process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some ‘random’ process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some ‘random’ process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Motivation

ITQs: Constrained Sum Data

Quantifying Market Concentration in the Presence of Covariates

- Partition concentration
  - Group-wise
  - Contribution-wise
- Statistically specify concentration
  - As from data...
  - ...from some 'random' process
  - with distribution, thus tests of significant difference

GOAL: Straightforward (Easy) Conditional/Groupwise Estimates of Concentration, with Probability Intervals

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \le y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3$, $y_2 = 7$ then $\mathbb{1}_{[y_1 \le y]} = 1$ but $\mathbb{1}_{[y_2 \le y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \le y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \leq y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3$, $y_2 = 7$ then $\mathbb{1}_{[y_1 \leq y]} = 1$ but $\mathbb{1}_{[y_2 \leq y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \leq y]} \quad (36)$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \leq y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3$, $y_2 = 7$ then $\mathbb{1}_{[y_1 \leq y]} = 1$ but $\mathbb{1}_{[y_2 \leq y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \leq y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \le y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3,\ y_2 = 7$ then $\mathbb{1}_{[y_1 \le y]} = 1$ but $\mathbb{1}_{[y_2 \le y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \le y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \leq y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3,\ y_2 = 7$ then $\mathbb{1}_{[y_1 \leq y]} = 1$ but $\mathbb{1}_{[y_2 \leq y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \leq y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \le y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3,\ y_2 = 7$ then $\mathbb{1}_{[y_1 \le y]} = 1$ but $\mathbb{1}_{[y_2 \le y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \le y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \leq y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3, \; y_2 = 7$ then $\mathbb{1}_{[y_1 \leq y]} = 1$ but $\mathbb{1}_{[y_2 \leq y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \leq y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \leq y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3,\ y_2 = 7$ then $\mathbb{1}_{[y_1 \leq y]} = 1$ but $\mathbb{1}_{[y_2 \leq y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \leq y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \leq y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3, \; y_2 = 7$ then $\mathbb{1}_{[y_1 \leq y]} = 1$ but $\mathbb{1}_{[y_2 \leq y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \leq y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Just a little notation

Brief Notation

- $\mathbf{y} = (y_1, ..., y_n) \leftarrow$ (non-negative) quota shares for $i = 1, ..., n$ shareholders
- $\mathbb{1}_{[y_i \leq y]} \leftarrow$ Indicator function. Say $y = 5$ and $y_1 = 3,\ y_2 = 7$ then $\mathbb{1}_{[y_1 \leq y]} = 1$ but $\mathbb{1}_{[y_2 \leq y]} = 0$
- $\sum_{i=1}^{n} apple_i \leftarrow$ add up apples 1 through N.
- Empirical distribution function (ecdf)

$$F_Y^n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{[y_i \leq y]} \tag{36}$$

The ecdf in this context is just the proportion of people with a less or equal share y of the quota

Categorical Variables

Variable	Category	Description
Location		
	Inshore	Close to shore
	Deepwater	Offshore
	HMS	Highly Migratory Species
Market		
	Top Export	Rock Lobster, Hoki, Squid, Orange Roughy, Jack Mackerel,
	Not Top Export	All other species
Fishery		
	SNA	Snapper
	BCO	Blue Cod
	ORH	Orange Roughy
	CRA	Rock Lobster

ITQ data by Fishery via ecdf

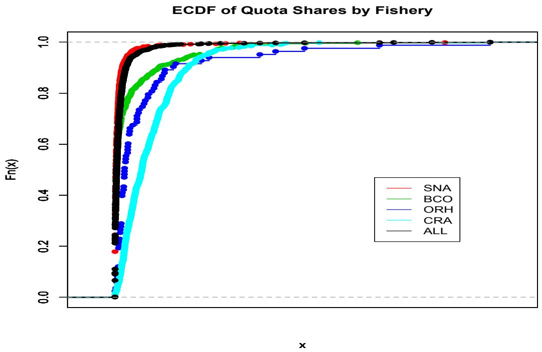


Figure: Graph of empirical cumulative distribution function (ecdf) of Quota Shares by Fishery Type

Measuring Inequality

Essentially all functions of ecdf

Via Quantiles

Notice that the ecdf in equation (45) generates, at least, n quantiles

$$F_n{}^{-1}(p) = y_{(\lfloor p \cdot n \rfloor)} \tag{37}$$

Lorenz Curve

The beautiful Lorenz Curve

The Lorenz curve is just a list of population proportions — numbers between 0 and 1 — joined to the list of 'good' proportions,

$$L_n(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} y_{(i)} \tag{38}$$

also numbers between 0 and 1.

Lorenz Curve

The beautiful Lorenz Curve

The Lorenz curve is just a list of population proportions — numbers between 0 and 1 — joined to the list of ‘good’ proportions,

$$L_n(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} y_{(i)} \tag{38}$$

also numbers between 0 and 1.

Lorenz Curve

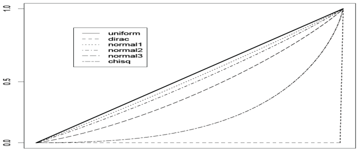
The beautiful Lorenz Curve

The Lorenz curve is just a list of population proportions — numbers between 0 and 1 — joined to the list of 'good' proportions,

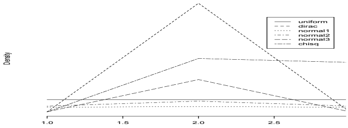
$$L_n(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} y_{(i)} \tag{38}$$

also numbers between 0 and 1.

Lorenz Curves



(a) Lorenz Curves



(b) Density Curves

Figure: Illustrations of Lorenz curves on parametric distributional models. The 45° line is the Lorenz curve on a uniform distribution, the right angle is the dirac distribution, completely concentrated at one point. Example distributions listed in the legend are in order of distributional 'inequality': the uniform distribution is perfectly equal, the dirac perfectly inequal, the normal distributions in order of increasing variance, the chi-squared distribution is right-skewed. of The Gini index is the area between the 45° line — the Lorenz curve for an equal distribution — and the particular Lorenz curve divided by $1/2$, the max area of concentration.

Measuring Concentration

Essentially all functions of ecdf

'Lorenz via ecdf'

$$L_n(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor Np \rfloor} F_n^{-1}(i/n) \tag{39}$$

Measuring Concentration

'Mean Absolute Deviation'
Gini Index:

$$G = \binom{n}{2}^{-1} \sum_{i<j} |y_i - y_j| \tag{40}$$

$$G_n = \frac{\frac{1}{2} - \sum_{p=1/n}^{n} \frac{1}{n} L_n(p)}{1/2} = 1 - 2\frac{1}{n} \sum_{p=1/n}^{n} L_n(p) \tag{41}$$

$$= 1 - 2\frac{1}{n} \sum_{p=1/n}^{n} (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor n \cdot p \rfloor} F_n^{-1}(i/n) \tag{42}$$

Measuring Concentration

'Mean Absolute Deviation'
Gini Index:

$$G = \binom{n}{2}^{-1} \sum_{i<j} |y_i - y_j| \tag{40}$$

$$G_n = \frac{\frac{1}{2} - \sum_{p=1/n}^{n} \frac{1}{n} L_n(p)}{1/2} = 1 - 2\,\frac{1}{n} \sum_{p=1/n}^{n} L_n(p) \tag{41}$$

$$= 1 - 2\,\frac{1}{n} \sum_{p=1/n}^{n} (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor n \cdot p \rfloor} F_n^{-1}(i/n) \tag{42}$$

Measuring Concentration

'Herfindahl/Hirschman Index'
HHI:

$$H_n = \sum_{i=1}^{n^*} \left[\frac{y_{(i)}}{n^* \cdot \overline{y}} \right]^2 \tag{43}$$

since it doesn't matter whether the data are sorted, and then — again immediately,

$$H_n = \sum_{i=1}^{n^*} \left[\frac{F_n^{-1}(i/n)}{n^* \cdot \overline{y}} \right]^2 = [(n^*) \cdot \overline{y}]^{-2} \sum_{i=1}^{n} \left[F_n^{-1}(i/n) \right]^2 \tag{44}$$

where $n^* = n \wedge 50$ is the minimum of the sample size and fifty.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, \ldots, y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \qquad (45)$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\bar{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, \ldots, y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \bar{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\bar{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, \ldots, y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \bar{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, \ldots, y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, ..., y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that $p\%$ of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \qquad (45)$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, \ldots, y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, ..., y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, \ldots, y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that $p\%$ of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, \ldots, y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Just a little more notation

Brief Notation

- $\overline{y} \leftarrow$ the observed mean
- $\mathbf{y}_{()} = (y_{(1)}, ..., y_{(n)}) \leftarrow$ the sorted list
- $F_n^{-1}(p) \leftarrow$ the observed *pth* quantile, the magnitude of share that p% of the people have less than (or equal to).
- Lorenz Curve

$$L(p) = (n \cdot \overline{y})^{-1} \sum_{i=1}^{\lfloor np \rfloor} F_n^{-1}(i/n) \tag{45}$$

The Lorenz curve is just the sorted, cumulative list of shares by population proportion.

Lorenz → Gini

The Gini coefficient is a function of the Lorenz curve...

$$G = \frac{\frac{1}{2} - \sum_{p=1/n}^{n} \frac{1}{n} L(p)}{1/2} = 1 - 2\,\frac{1}{n} \sum_{p=1/n}^{n} L(p) \tag{46}$$

...the scaled difference between the area under the observed Lorenz and equality

Lorenz $\rightarrow$ Gini

The Gini coefficient is a function of the Lorenz curve...

$$G = \frac{\frac{1}{2} - \sum_{p=1/n}^{n} \frac{1}{n} L(p)}{1/2} = 1 - 2\,\frac{1}{n} \sum_{p=1/n}^{n} L(p) \tag{46}$$

...the scaled difference between the area under the observed Lorenz and equality

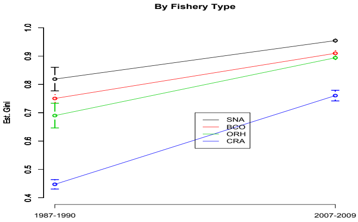
Lorenz → Gini

The Gini coefficient is a function of the Lorenz curve...

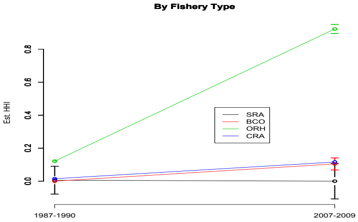
$$G = \frac{\frac{1}{2} - \sum_{p=1/n}^{n} \frac{1}{n} L(p)}{1/2} = 1 - 2\,\frac{1}{n} \sum_{p=1/n}^{n} L(p) \tag{46}$$

...the scaled difference between the area under the observed Lorenz and equality

Lorenz → Gini



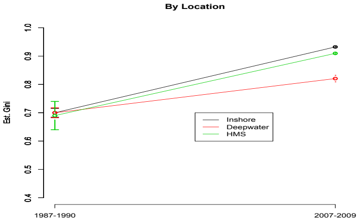
(a) Gini Index



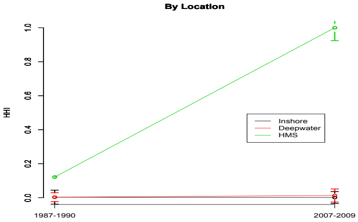
(b) HHI Index

Figure: (a): Gini indices, past and recent - by fishery type - with 95% confidence bars calculated by bootstrap. There is statistically significant evidence of an increase in concentration of quota shares. (b) HHI indices, past and recent - by fishery type - with 95% confidence bars calculated by bootstrap. The HHI indices generally have wider confidence intervals; the HHI by definition is defined on a maximum of 50 observations. Notice the difference in ranges (y-axis) for Gini and HHI plots: on data with many observations the HHI is often smaller than the Gini.

Lorenz → Gini



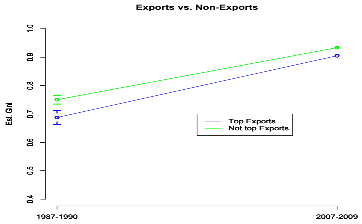
(a) Gini Index



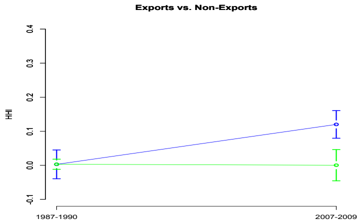
(b) HHI Index

Figure: (a): Gini indices, past and recent - by location - with 95% confidence bars calculated by bootstrap. There is statistically significant evidence of an increase in concentration of quota shares. (b) HHI indices, past and recent - by location - with 95% confidence bars calculated by bootstrap. There is no significant increase in measured concentration via HHI for inshore and deepwater fish species; in general the confidence intervals for HHI are wider than the Gini, as it is defined on less data. The observed HHI for Highly Migratory Species (HMS) is nearly maximal. Notice the difference in ranges

Lorenz → Gini



(a) Gini Index



(b) HHI Index

Figure: (a): Gini indices, past and recent - by export type - with 95% confidence bars calculated by bootstrap. There is statistically significant evidence of an increase in concentration of quota shares. (b) HHI indices, past and recent - by export type - with 95% confidence bars calculated by bootstrap. Notice the difference in ranges (y-axis) for Gini and HHI plots: on data with many observations the HHI is often smaller than the Gini.

Conditional Lorenz Curve

The goal is to represent overall inequality via contribution from conditional covariates.

The trick is to see covariates as 'conditional information'

Aaberge et al define pseudo-Lorenz regression curve as a function, in the presence of covariates $\mathbf{x}$ for y, such that

$$E[\Lambda(p|\mathbf{x})] = L(p) \tag{47}$$

e.g. that the conditional curves should 'sum' to the original curve

Conditional Lorenz Curve

The goal is to represent overall inequality via contribution from conditional covariates.

The trick is to see covariates as 'conditional information'

Aaberge et al define pseudo-Lorenz regression curve as a function, in the presence of covariates $\mathbf{x}$ for y, such that

$$E[\Lambda(p|\mathbf{x})] = L(p) \tag{47}$$

e.g. that the conditional curves should 'sum' to the original curve

Conditional Lorenz Curve

This is just the law of iterated expectation...

for discrete, i.e. categorical, covariates, this is easy

$$L(p) = \sum_{j=1}^{m} \pi_j \, \Lambda(p|\mathbf{x} \in C_j) \tag{48}$$

and setting

$$\Lambda(p|C_j) = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \, L(F_j(F^{-1}(p))|C_j) \tag{49}$$

guarantees that the overall Lorenz curve will be the weighted sum of conditional ‘pseudo’-Lorenz curves.

Conditional Lorenz Curve

This is just the law of iterated expectation...

for discrete, i.e. categorical, covariates, this is easy

$$L(p) = \sum_{j=1}^{m} \pi_j \, \Lambda(p|\mathbf{x} \in C_j) \tag{48}$$

and setting

$$\Lambda(p|C_j) = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \, L(F_j(F^{-1}(p))|C_j) \tag{49}$$

guarantees that the overall Lorenz curve will be the weighted sum of conditional ‘pseudo’-Lorenz curves.

Conditional Lorenz Curve

This is just the law of iterated expectation...

for discrete, i.e. categorical, covariates, this is easy

$$L(p) = \sum_{j=1}^{m} \pi_j \, \Lambda(p|\mathbf{x} \in C_j) \tag{48}$$

and setting

$$\Lambda(p|C_j) = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \, L(F_j(F^{-1}(p))|C_j) \tag{49}$$

guarantees that the overall Lorenz curve will be the weighted sum of conditional 'pseudo'-Lorenz curves.

Conditional Lorenz Curve

This is just the law of iterated expectation...

for discrete, i.e. categorical, covariates, this is easy

$$L(p) = \sum_{j=1}^{m} \pi_j \, \Lambda(p|\mathbf{x} \in C_j) \tag{48}$$

and setting

$$\Lambda(p|C_j) = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \, L(F_j(F^{-1}(p))|C_j) \tag{49}$$

guarantees that the overall Lorenz curve will be the weighted sum of conditional 'pseudo'-Lorenz curves.

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\bar{y}_j}{\bar{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* **y**
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* y
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* y
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* y
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* y
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* **y**
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group *j*
- *p* the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* **y**
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group *j* at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group *j* on the observed proportion of population in group *j* at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* $\mathbf{y}$
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* $\mathbf{y}$
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Just a little more notation

More Brief Notation

- $\pi_j = \frac{\overline{y}_j}{\overline{y}} \cdot n_j \leftarrow$ the proportional size of group j
- p the proportion of the population
- $F_N^{-1}(p) \leftarrow$ the observed *pth* quantile *of overall* **y**
- $F_j(F^{-1}(p)) \leftarrow$ the observed proportion of population in group j at the *pth* quantile of the *overall* distribution
- $L(F_j(F^{-1}(p))|C_j) \leftarrow$ the Lorenz curve of group j on the observed proportion of population in group j at the *pth* quantile of the *overall* distribution

In Layman's terms...

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution.*$\rightarrow F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow \frac{\bar{y}_j}{\bar{y}} \cdot n_j\, L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution*.$\rightarrow F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow \frac{\bar{y}_j}{\bar{y}} \cdot n_j\ L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution.*$\rightarrow F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow \frac{\bar{y}_j}{\bar{y}} \cdot n_j \, L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution*.$\rightarrow F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow \frac{\overline{y}_j}{\overline{y}} \cdot n_j\, L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution*.$\rightarrow F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow \frac{\bar{y}_j}{\bar{y}} \cdot n_j\, L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution*.$\rightarrow F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow \frac{\bar{y}_j}{\bar{y}} \cdot n_j\ L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution.*$\rightarrow F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow \frac{\overline{y}_j}{\overline{y}} \cdot n_j \; L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow$ $F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution.*$\rightarrow$ $F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow$ $L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow$ $\frac{\overline{y}_j}{\overline{y}} \cdot n_j \; L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

A simple algorithm

1. Sort all the data; Generate the *pth* quantiles of the unconditioned distribution.$\rightarrow$ $F_N, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the *pth* quantiles, *of the original distribution*.$\rightarrow$ $F_j(F^{-1}(p))$
3. Join the *pth* proportions for each group F_j with the cumulative proportion of income at each group.$\rightarrow$ $L(F_j(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each *pth* proportion.$\rightarrow$ $\frac{\overline{y}_j}{\overline{y}} \cdot n_j \; L(F_j(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

Algorithm

1. Sort all the data; Generate the pth quantiles of the unconditioned distribution. In the notation: $F_n, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the pth quantiles of the original distribution. These are $F_{n,j}(F_n^{-1}(p))$
3. Join the *pth* proportions for each group $F_{n,j}$ with the cumulative proportion of income at each group. This is $L(F_{n,j}(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each pth proportion: $\frac{\overline{y}_j}{\overline{y}} \cdot n_j \, L(F_{n,j}(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

Algorithm

1. Sort all the data; Generate the pth quantiles of the unconditioned distribution. In the notation: $F_n, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the pth quantiles of the original distribution. These are $F_{n,j}(F_n^{-1}(p))$
3. Join the *pth* proportions for each group $F_{n,j}$ with the cumulative proportion of income at each group. This is $L(F_{n,j}(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each pth proportion: $\frac{\bar{y}_j}{\bar{y}} \cdot n_j \, L(F_{n,j}(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

Algorithm

1. Sort all the data; Generate the pth quantiles of the unconditioned distribution. In the notation: $F_n, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the pth quantiles of the original distribution. These are $F_{n,j}(F_n^{-1}(p))$
3. Join the *pth* proportions for each group $F_{n,j}$ with the cumulative proportion of income at each group. This is $L(F_{n,j}(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each pth proportion: $\frac{\overline{y}_j}{\overline{y}} \cdot n_j\, L(F_{n,j}(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

Algorithm

1. Sort all the data; Generate the pth quantiles of the unconditioned distribution. In the notation: $F_n, F^{-1}(p)$
2. Sort the data within each group; Generate the ecdf for each group (conditional distribution) at the pth quantiles of the original distribution. These are $F_{n,j}(F_n^{-1}(p))$
3. Join the *pth* proportions for each group $F_{n,j}$ with the cumulative proportion of income at each group. This is $L(F_{n,j}(F^{-1}(p))|C_j)$
4. Compute the contribution to the overall Lorenz curve, at each pth proportion: $\frac{\overline{y}_j}{\overline{y}} \cdot n_j \, L(F_{n,j}(F^{-1}(p))|C_j)$

Conditional Lorenz Curve

Example

Consider this data

```
g1<-c(1,5,5,1) g2<-c(3,3,3,3) g3<-c(1,1,1,9)
```

Sort all the data

```
sort(c(g1,g2,g3))
1 1 1 1 1 3 3 3 3 5 5 9
```

Conditional Lorenz Curve

Example

Consider this data

```
g1<-c(1,5,5,1) g2<-c(3,3,3,3) g3<-c(1,1,1,9)
```

Sort all the data

```
sort(c(g1,g2,g3))
1 1 1 1 1 3 3 3 3 5 5 9
```

Conditional Lorenz Curve

Example

Consider this data

```
g1<-c(1,5,5,1) g2<-c(3,3,3,3) g3<-c(1,1,1,9)
```

Sort all the data

```
sort(c(g1,g2,g3))
1 1 1 1 1 3 3 3 3 5 5 9
```

Simple Example

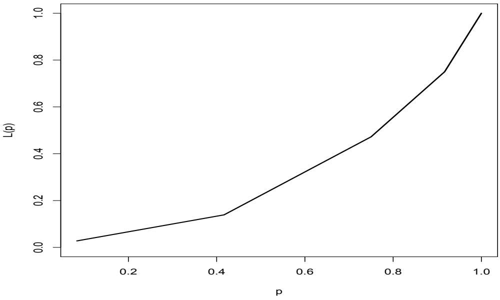


Figure: Overall curve

Conditional Lorenz Curve

Example

Illustrate the conditional lorenz curves for each group

```
lnew1<-lorenz(g1); lnew2<-lorenz(g2);
lnew3<-lorenz(g3)
```

The function to compute the lorenz curve is sooooo easy

```
lorenz function(x)
y<-sort(x)
m<-mean(y); s<-sum(y)
l<-cumsum(y)/s
l
```

Conditional Lorenz Curve

Example

Illustrate the conditional lorenz curves for each group

```
lnew1<-lorenz(g1); lnew2<-lorenz(g2);
lnew3<-lorenz(g3)
```

The function to compute the lorenz curve is sooooo easy

```
lorenz function(x)
y<-sort(x)
m<-mean(y); s<-sum(y)
l<-cumsum(y)/s
l
```

Conditional Lorenz Curve

Example

Illustrate the conditional lorenz curves for each group

```
lnew1<-lorenz(g1); lnew2<-lorenz(g2);
lnew3<-lorenz(g3)
```

The function to compute the lorenz curve is sooooo easy

```
lorenz function(x)
y<-sort(x)
m<-mean(y); s<-sum(y)
l<-cumsum(y)/s
l
```

Simple Example

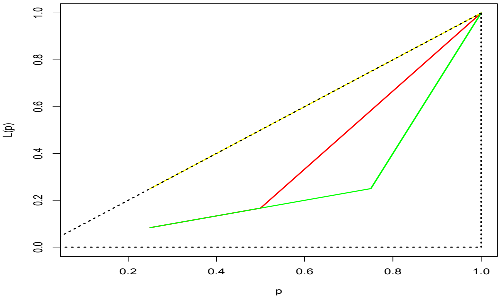


Figure: Conditional curves.

Conditional Lorenz Curve

Example

Generally the ‘resolution’ can be set ‘arbitrarily’. (but it’s easy to set it at fewest group)

```
> lresg [1] 4
```

And compute the multipliers for each of the groups

```
meanratiosg<-c(mg1,mg2,mg3)/mgall
[1] 0.6859177 0.8883197 5.2228916 0.8163842
groupsizesg<-c(4,4,4)/12 [1] 0.3333333 0.3333333
0.3333333
```

Conditional Lorenz Curve

Example

Generally the ‘resolution’ can be set ‘arbitrarily’. (but it’s easy to set it at fewest group)

```
> lresg [1] 4
```

And compute the multipliers for each of the groups

```
meanratiosg<-c(mg1,mg2,mg3)/mgall
[1] 0.6859177 0.8883197 5.2228916 0.8163842
groupsizesg<-c(4,4,4)/12 [1] 0.3333333 0.3333333
0.3333333
```

Conditional Lorenz Curve

Example

Generally the ‘resolution’ can be set ‘arbitrarily’. (but it’s easy to set it at fewest group)

```
> lresg [1] 4
```

And compute the multipliers for each of the groups

```
meanratiosg<-c(mg1,mg2,mg3)/mgall
[1] 0.6859177 0.8883197 5.2228916 0.8163842
groupsizesg<-c(4,4,4)/12 [1] 0.3333333 0.3333333
0.3333333
```

Conditional Lorenz Curve

Example

Generally the ‘resolution’ can be set ‘arbitrarily’. (but it’s easy to set it at fewest group)

```
> lresg [1] 4
```

And compute the multipliers for each of the groups

```
meanratiosg<-c(mg1,mg2,mg3)/mgall
[1] 0.6859177 0.8883197 5.2228916 0.8163842
groupsizesg<-c(4,4,4)/12 [1] 0.3333333 0.3333333
0.3333333
```

Conditional Lorenz Curve

Example

Generally the ‘resolution’ can be set ‘arbitrarily’. (but it’s easy to set it at fewest group)

```
> lresg [1] 4
```

And compute the multipliers for each of the groups

```
meanratiosg<-c(mg1,mg2,mg3)/mgall
[1] 0.6859177 0.8883197 5.2228916 0.8163842
groupsizesg<-c(4,4,4)/12 [1] 0.3333333 0.3333333
0.3333333
```

Conditional Lorenz Curve

Example

Essentially the contribution to the overall lorenz curve is calculated pointwise

```
for(pg in uppsg)
lpg<-c(lnew1[pg],lnew2[pg],lnew3[pg])
conditionallorenzedg[pg]<-
as.double(sum(meanratiosg*lpg*groupsizesg))
conditionallorenzedg
[1] 0.1388889 0.2777778 0.5277778 1.0000000
```

For instance at $p = .50$, the conditional lorenz curves are

```
[1] 0.1666667 0.5000000 0.1666667
```

And their contributions to the overall lorenz curve are

```
[1] 0.3333333 0.3333333 0.3333333
```

Conditional Lorenz Curve

Example

Essentially the contribution to the overall lorenz curve is calculated pointwise

```
for(pg in uppsg)
lpg<-c(lnew1[pg],lnew2[pg],lnew3[pg])
conditionallorenzedg[pg]<-
as.double(sum(meanratiosg*lpg*groupsizesg))
conditionallorenzedg
[1] 0.1388889 0.2777778 0.5277778 1.0000000
```

For instance at $p = .50$, the conditional lorenz curves are

```
[1] 0.1666667 0.5000000 0.1666667
```

And their contributions to the overall lorenz curve are

```
[1] 0.3333333 0.3333333 0.3333333
```

Conditional Lorenz Curve

Example

Essentially the contribution to the overall lorenz curve is calculated pointwise

```
for(pg in uppsg)
lpg<-c(lnew1[pg],lnew2[pg],lnew3[pg])
conditionallorenzedg[pg]<-
as.double(sum(meanratiosg*lpg*groupsizesg))
conditionallorenzedg
[1] 0.1388889 0.2777778 0.5277778 1.0000000
```

For instance at $p = .50$, the conditional lorenz curves are

```
[1] 0.1666667 0.5000000 0.1666667
```

And their contributions to the overall lorenz curve are

```
[1] 0.3333333 0.3333333 0.3333333
```

Conditional Lorenz Curve

Example

Essentially the contribution to the overall lorenz curve is calculated pointwise

```
for(pg in uppsg)
lpg<-c(lnew1[pg],lnew2[pg],lnew3[pg])
conditionallorenzedg[pg]<-
as.double(sum(meanratiosg*lpg*groupsizesg))
conditionallorenzedg
[1] 0.1388889 0.2777778 0.5277778 1.0000000
```

For instance at $p = .50$, the conditional lorenz curves are

```
[1] 0.1666667 0.5000000 0.1666667
```

And their contributions to the overall lorenz curve are

```
[1] 0.3333333 0.3333333 0.3333333
```

Simple Example

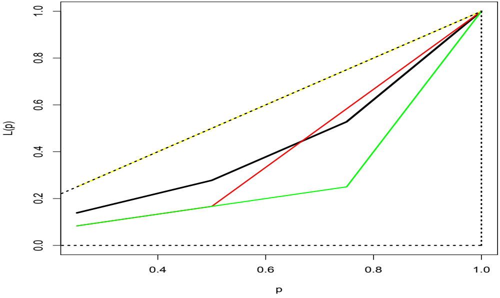


Figure: Conditional curves, overall in black.

Conditional Lorenz Curve

Example

And the Gini's are easy to compute

```
ginioverall<-1-2*sum(conditionallorenzedg)/4
0.02777778
```

Conditional Lorenz Curve

Example

And the Gini's are easy to compute

```
ginioverall<-1-2*sum(conditionallorenzedg)/4
0.02777778
```

Simple Example

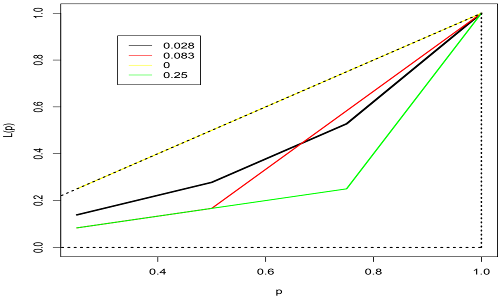


Figure: Conditional curves, with Ginis: overall in black.

Significant Differences

Effect of group membership on overall inequality

Like in Linear Regression we want effect of covariate (here c_j) on response (here Lorenz/Gini)

Mathematically this is

$$\frac{\partial L(p)}{\partial C}\bigg|_{C=c_j} [\sum_{j=1}^{m} \pi_j \frac{\bar{y}_j}{\bar{y}} \cdot n_j \, L(F_j(F^{-1}(p))|C_j)] \tag{50}$$

But if we remember the definition of the derivative, and that the categorical covariate is 'singular', this is just

$$L(p)\bigg|_{C_{-j}} - L(p)\bigg|_{C} \tag{51}$$

Just the difference between the overall (conditionally defined) lorenz curve without and with the *jth* group.

Significant Differences

Effect of group membership on overall inequality

Like in Linear Regression we want effect of covariate (here c_j) on response (here Lorenz/Gini)
Mathematically this is

$$\frac{\partial L(p)}{\partial C}\bigg|_{C=c_j}[\sum_{j=1}^{m}\pi_j \frac{\overline{y}_j}{\overline{y}} \cdot n_j \ L(F_j(F^{-1}(p))|C_j)] \tag{50}$$

But if we remember the definition of the derivative, and that the categorical covariate is ‘singular’, this is just

$$L(p)\Big|_{C_{-j}} - L(p)\Big|_{C} \tag{51}$$

Just the difference between the overall (conditionally defined) lorenz curve without and with the *jth* group.

Significant Differences

Effect of group membership on overall inequality

Like in Linear Regression we want effect of covariate (here c_j) on response (here Lorenz/Gini)
Mathematically this is

$$\frac{\partial L(p)}{\partial C}\bigg|_{C=c_j} [\sum_{j=1}^{m} \pi_j \frac{\overline{y}_j}{\overline{y}} \cdot n_j \, L(F_j(F^{-1}(p))|C_j)] \tag{50}$$

But if we remember the definition of the derivative, and that the categorical covariate is 'singular', this is just

$$L(p)\bigg|_{C_{-j}} - L(p)\bigg|_{C} \tag{51}$$

Just the difference between the overall (conditionally defined) lorenz curve without and with the *jth* group.

Significant Differences

Statistical significance

We can test for statistical significance using exploiting the duality between the Lorenz curve and the ecdf

since

$$F_n(t) \sim N(F(t), F(t)[1 - F(t)]) \tag{52}$$

Then

$$L_n(p) \sim N\left(L(p), \frac{L(p)[1 - L(p)]}{n}\right) \tag{53}$$

and we can use normal confidence bounds (pointwise), or at least the Kolomorogov-Smirnov (KS) test for differences in distributions to test for significant effects. See Abayomi, Yandle 2011.

> *We must be careful not to confuse data with the abstractions we use to analyze them.*

-William James

Significant Differences

Statistical significance

We can test for statistical significance using exploiting the duality between the Lorenz curve and the ecdf
since

$$F_n(t) \sim N(F(t), F(t)[1 - F(t)]) \tag{52}$$

Then

$$L_n(p) \sim N\left(L(p), \frac{L(p)[1 - L(p)]}{n}\right) \tag{53}$$

and we can use normal confidence bounds (pointwise), or at least the Kolomorogov-Smirnov (KS) test for differences in distributions to test for significant effects. See Abayomi, Yandle 2011.

> *We must be careful not to confuse data with the abstractions we use to analyze them.*
>
> -William James

Significant Differences

Statistical significance

We can test for statistical significance using exploiting the duality between the Lorenz curve and the ecdf
since

$$F_n(t) \sim N(F(t), F(t)[1 - F(t)]) \tag{52}$$

Then

$$L_n(p) \sim N\left(L(p), \frac{L(p)[1 - L(p)]}{n}\right) \tag{53}$$

and we can use normal confidence bounds (pointwise), or at least the Kolomorogov-Smirnov (KS) test for differences in distributions to test for significant effects. See Abayomi, Yandle 2011.

> *We must be careful not to confuse data with the abstractions we use to analyze them.*

-William James

Results

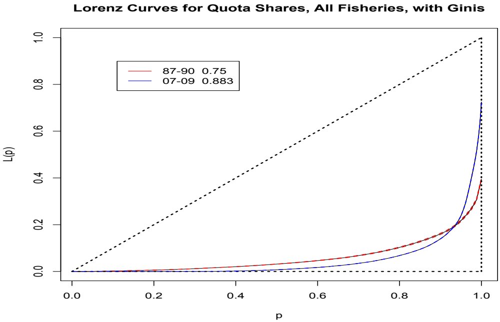


Figure: Lorenz Curves - over all locations - with 95% confidence bars on quota shares for SNA, BCO, ORH and CRA.

Results

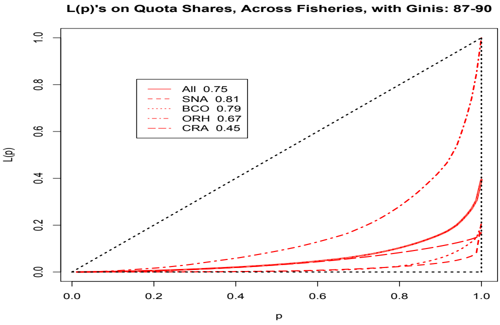


Figure: Lorenz Curves - over all locations - with 95% confidence bars on quota shares for SNA, BCO, ORH and CRA. The curves are significantly different at $\alpha = .05$ across fisheries.

Results

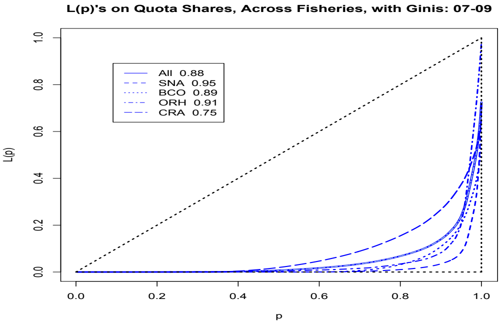


Figure: Lorenz Curves - over all locations - with 95% confidence bars on quota shares for SNA, BCO, ORH and CRA. The curves are significantly different at $\alpha = .05$ across fisheries.

Outline

Sustainability Characterized

Chichilnisky (1997) introduces an axiomatization for intergenerational equity as a criteria for sustainability:

Sustainability Characterized

Letting $\mathbf{u} = \{u_t\}_{t=1}^{\infty}$ be a bounded, real valued utility stream - on $(\Omega, \mathcal{F}_t, \mathbb{P} = \mathbb{P}^* + \mathbb{P}_\infty)$.

- **Insensitivity to the future**: $\mathbb{P}_\infty(W(\mathbf{u})) = 0$ if $\mathbb{P}_\infty$ is a *purely finitely additive* measure.
- **Insensitivity to the present**: $E_{\mathbb{P}^*}[W(\mathbf{u})] = 0$ for all countably additive measures $\mathbb{P}^*$

Sustainability Characterized

Letting $\mathbf{u} = \{u_t\}_{t=1}^{\infty}$ be a bounded, real valued utility stream - on $(\Omega, \mathcal{F}_t, \mathbb{P} = \mathbb{P}^* + \mathbb{P}_\infty)$.

- **Insensitivity to the future**: $\mathbb{P}_\infty(W(\mathbf{u})) = 0$ if $\mathbb{P}_\infty$ is a *purely finitely additive* measure.
- **Insensitivity to the present**: $E_{\mathbb{P}^*}[W(\mathbf{u})] = 0$ for all countably additive measures $\mathbb{P}^*$

Chichilnisky's axiomatization

Chichilnisky's axiomatization [1996,1997]:

$$W(\mathbf{u}) = \alpha \int_{R^+} u(c_t) d\mathbb{P}^*(t) + (1-\alpha)\mathbb{P}_\infty(\mathbf{u})$$

where $\mathbb{P}_\infty$ is measure zero on finite sets. This characterization ensures equity to present and unforseen generations.

Chichilnisky's axiomatization

The combination of measures which are singular w.r.t each other disallows ordinary optimization procedures.

$$\mathbb{P}^* \perp \mathbb{P}_\infty$$

Chichilnisky's axiomatization

However:
Both $d\mathbb{P}^*$ and $d\mathbb{P}_\infty$ are absolutely continuous with respect to $d\mathbb{P}$

Statistical Estimation

The goal here is to introduce statistical estimators for a sustainable development path - or utility stream - via a representation of Kullback-Leibler divergence.

K-L Divergence

Recall:

$$KL(d\mathbb{P}^k, d\mathbb{P}) = E_{d\mathbb{P}}[log(\frac{d\mathbb{P}^k}{d\mathbb{P}})]$$

K-L Divergence

Let $d\mathbb{P}^k$ be the probability density induced by the filtration $\mathcal{F}$ at the $k-th$ cutoff of the utility stream.
Then

$$KL(d\mathbb{P}^k, d\mathbb{P}) = \sum^{k} d\mathbb{P} log(\frac{d\mathbb{P}_k}{d\mathbb{P}})$$

K-L Divergence

But $d\mathbb{P} = d\mathbb{P}^* + d\mathbb{P}_\infty$. So

$$KL(d\mathbb{P}^k, d\mathbb{P}) =$$

$$= \sum^k d\mathbb{P}^* log(\frac{d\mathbb{P}^k}{d\mathbb{P}^* + d\mathbb{P}_\infty}) + \sum^k d\mathbb{P}_\infty log(\frac{d\mathbb{P}^k}{d\mathbb{P}^* + d\mathbb{P}_\infty})$$

K-L Divergence

$$= \sum^{k}(d\mathbb{P}^* + d\mathbb{P}_\infty)log(d\mathbb{P}_k) - \sum^{k}(d\mathbb{P}^* + d\mathbb{P}_\infty)log(d\mathbb{P}^* + d\mathbb{P}_\infty)$$

The second term is just the entropy of the full measure. Looking at the first term...

K-L Divergence

...and taking the conditional expectation

$$= E(\sum^{k}(d\mathbb{P}^{*} + d\mathbb{P}_{\infty})log(d\mathbb{P}_{k})|\mathcal{F}_{k})$$

$$= E(\sum^{k} d\mathbb{P}^{*}log(d\mathbb{P}_{k})|\mathcal{F}_{k}) + E(\sum^{k} d\mathbb{P}_{\infty}log(d\mathbb{P}_{k})|\mathcal{F}_{k})$$

K-L Divergence

yields

$$= \sum^{k} log(d\mathbb{P}_k)E(d\mathbb{P}^*) + \sum^{k} log(d\mathbb{P}_k)E(d\mathbb{P}_\infty)$$

This is $:= data \cdot parameters + data \cdot parameters$, as well as a convex sum of singular measures meeting Chichilnisky's criteria.

K-L Divergence

Minimizing

$$= \sum^{k} log(d\mathbb{P}_k)E(d\mathbb{P}^*) + \sum^{k} log(d\mathbb{P}_k)E(d\mathbb{P}_\infty)$$

with respect to the parameters of the measures (which can include mixing parameter α) yields estimating equations which yield inference on utility/developments (i.e. consumption) paths

Next steps

- Derive examples for singular measure pairs
- Investigate distribution of K-L sum, estimating equations, possible CUSUM test.
- Thank you.

Next steps

- Derive examples for singular measure pairs
- Investigate distribution of K-L sum, estimating equations, possible CUSUM test.
- Thank you.

Next steps

- Derive examples for singular measure pairs
- Investigate distribution of K-L sum, estimating equations, possible CUSUM test.
- Thank you.