Mesurer les inégalités de revenu

Matti Langel

Université de Neuchâtel

Workshop sur la mesure des inégalités de revenu

Neuchâtel, 15 juin 2012

Plan

Plan

Définir les inégalités

- Terme difficile à définir.
- Toute déviation de l'allocation parfaitement égale d'une ressource entre les éléments d'une population.
- Critère objectif : l'inégalité n'a pas besoin d'être associé à la notion d'injustice ($\neq$équité).
- La mesure des inégalités repose sur des choix normatifs.
- Peut s'appliquer à l'allocation de toutes ressources entre tous types d'unités statistiques.

Définir le revenu

- Aspect multidimensionnel et complexe.
- Un flux (fortune/patrimoine = un stock).
- Combinaison de plusieurs sources (salaire, allocations, rente, etc.).
  - Est-il purement monétaire ?
  - Doit-il être calculé au niveau individuel, au niveau du ménage ?
  - Doit-il être ventilé en fonction de la structure du ménage ?
  - Faut-il prendre en compte les revenus nuls ou négatifs ?
  - Comment comparer les revenus entre des pays aux prestations sociales différentes ?
- Le revenu n'est qu'une dimension du "bien-être" au sens économique (welfare).

Mesurer les inégalités : Pourquoi?

- Analyser la structure de la répartition du revenu dans une population.
- Rendre compte du niveau, ou du degré d'inégalité.
- Comparer cette structure avec d'autres populations, d'autres périodes.
- Indicateur important lorsque l'on traite de développement, de bien-être, de conditions de vie.
- Permet d'étudier l'impact de politiques sociales et économiques.

Mesurer les inégalités : Comment?

1. En étudiant l'ensemble la distribution du revenu dans la population.
   - Fonction de répartition,
   - Courbe de Lorenz.
2. En construisant un indicateur synthétique qui résume la dispersion de la répartition des revenus au sein d'une population.
   - mesures d'inégalité relatives (indices de Gini, de Theil, d'Atkinson, QSR).
   - mesures d'inégalité absolues (variance, indices de Kolm).

Quelques articles/ouvrages fondateurs

- Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Assiociation*, 9, 209-219.
- Gini, C. (1912). *Variabilità e mutabilità.* Tipografia di Paolo Cuppin.
- Pigou, A. C. (1912). *Wealth and Welfare*, Mc Millan.
- Dalton, H. (1920). The measurement of the inequality of incomes, *The Economic Journal*, 30, 348-361.

- La mesure des inégalités est un sujet très vaste.
- Fait l'objet de beaucoup de recherche, principalement depuis les années 1970.

Quelques ouvrages de référence

- Sen, A. (1997). *On Economic Inequality (2nd edition)*, Clarendon Press. (1st edition: 1973)
- Cowell, F. A. (2011). *Measuring Inequality (3rd edition)*, Oxford University Press. (1st edition: 1977)
- Silber, J. (ed.) (1999). *Handbook of Income Inequality Measurement*. Kluwer Academic Publishers.
- Atkinson, A. B. and Bourguignon, F. (eds.) (2000). *Handbook of Income Distribution, vol. 1*. Elsevier.

Plan

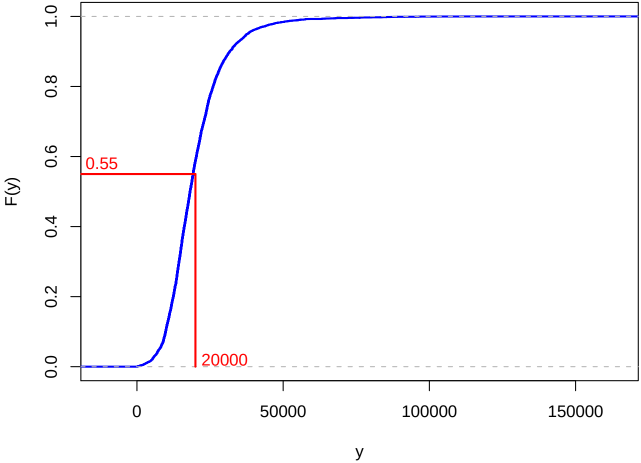
Fonction de répartition

- On note Y une variable aléatoire continue représentant le revenu.
- Pour simplifier, considérons Y strictement positive (pas de revenus nuls ou négatifs).
- $f(y)$ est sa densité de probabilité,
- $F(y)$ est sa fonction de répartition :

$$F(y) = \int_0^y f(u)du = \Pr(Y \leq y).$$

- $F^{-1}(.)$, est la fonction inverse de $F(.)$ ou fonction quantile,
- $\mu = \int_0^\infty yf(y)dy$ est l'espérance de Y (sa moyenne).

Fonction de répartition (empirique)



- Exemple d'interprétation: dans cette population, 55% des individus ont un revenu inférieur ou égal à 20'000.

Un outil central: la courbe de Lorenz

- Plusieurs mesures sont basées sur la courbe de Lorenz.
- La courbe de Lorenz $L(\alpha)$ représente la part du revenu total détenue par la proportion $\alpha \in [0, 1]$ d'individus les plus pauvres :

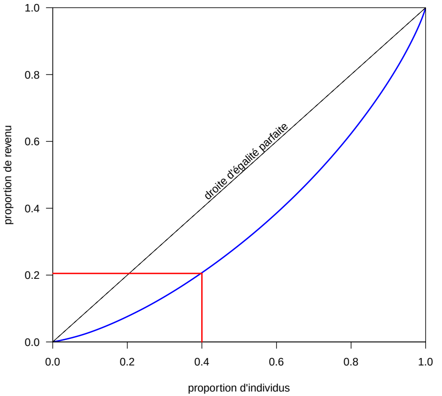
$$L(\alpha) = \frac{\text{revenu total des } \alpha \text{ plus pauvres}}{\text{revenu total}}.$$

Courbe de Lorenz (cas continue)

$$L(\alpha) = \frac{\int_0^{F^{-1}(\alpha)} yf(y)dy}{\int_0^{\infty} yf(y)dy} = \frac{1}{\mu} \int_0^{\alpha} F^{-1}(u)du.$$

- Si Y est continue et strictement positive et $\mu < \infty$ alors $L(\alpha)$ est croissante et convexe sur l'intervalle (0,1).

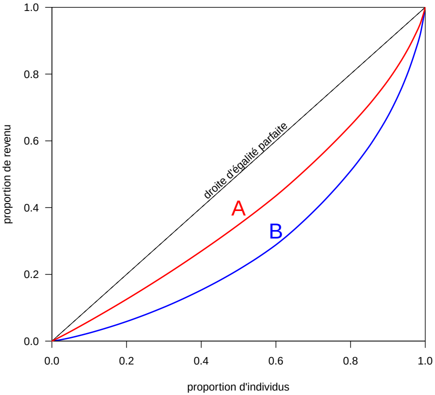
Courbe de Lorenz



- Exemple d'interprétation: dans cette population, les 40% les plus pauvres possèdent ensemble environ 20% du revenu total.
- Plus la courbe de Lorenz s'éloigne de la droite d'égalité parfaite, plus on dira que le niveau d'inégalité est élevé.

Dominance de Lorenz

- On peut comparer deux distributions en étudiant leur courbe de Lorenz.



- Dominance relative de Lorenz : A est plus égalitaire que B.
- Deux courbes de Lorenz peuvent se croiser.

Mesures (ou indices) d'inégalité

- Valeur numérique (scalaire) résumant la dispersion de la répartition des revenus au sein d'une population.
- Il existe de nombreuses mesures d'inégalité différentes.
- Types de mesure :
  - mesures relatives.
  - mesures absolues.
- Différentes approches :
  - axiomatique : satisfaire à des propriétés spécifiques.
  - interprétation : orienté vers la facilité de compréhension.

Plan

Population finie : notation

- On note $U = \{1, ...k, ...N\}$ une population finie de taille N.
- On note y_k le revenu de l'individu k.
- On note $y_{(k)}$ la statistique d'ordre (le k-ième plus petit revenu).
- On note $\mathbf{y} = (y_1, ..., y_k, ..., y_N)$ le vecteur de revenus.
- Revenu total : $Y = \sum_{k \in U} y_k$,
- Revenu moyen : $\overline{Y} = \frac{Y}{N}$,
- Variance des revenus : $\sigma_y^2 = \frac{1}{N} \sum_{k \in U} (y_k - \overline{Y})^2$.

Propriétés

- Approche axiomatique : Dalton (1920), Kolm (1976), Shorrocks (1980), Cowell and Kuga (1981), Chakravarty (1999).
- Notons I une mesure d'inégalité et $I(\mathbf{y})$ la valeur prise par I sur le vecteur $\mathbf{y}$.
- On mesure le niveau d'inégalité : plus I est grand, plus le niveau d'inégalité est élevé.
- On s'intéresse à un certain nombre de propriétés souhaitables pour I.

Propriété 1 : Normalisation

Normalisation

Soit $\bar{\mathbf{y}} = \overline{Y}\mathbf{1}$ avec $\mathbf{1}$ un vecteur de 1 de dimension N. La mesure I satisfait au principe de normalisation si

$$I(\bar{\mathbf{y}}) = 0.$$

- En situation d'égalité parfaite (tous les revenus sont égaux), la mesure d'inégalité prend la valeur 0.

Propriété 2 : Symétrie

Symétrie

Soit $\mathbf{y}_\lambda$ une replication d'ordre λ de $\mathbf{y}$ avec $\lambda \in \mathbb{N}^*$. La mesure I satisfait au principe de symétrie si

$$I(\mathbf{y}) = I(\mathbf{y}_\lambda).$$

- L'agrégation de deux ou plusieurs populations parfaitement identiques, conserve le niveau d'inégalité inchangé.

Propriété 3 : Principe de transferts

Principe de transferts (Pigou, 1912; Dalton, 1920)

Soit

$$\tilde{\mathbf{y}} = (y_1, ..., y_k + \delta, ..., y_\ell - \delta, ..., y_N),$$

avec $y_k \leq y_\ell$ et $0 < \delta < (y_\ell - y_k)/2$.
La mesure I satisfait au principe de transferts si

$$I(\mathbf{y}) \geq I(\tilde{\mathbf{y}}).$$

- Un transfert de revenu d'un individu plus riche vers un individu plus pauvre (sans que ce transfert ne rende le receveur plus riche que le donneur) se traduit sur l'indice I par une baisse du niveau d'inégalité.

Propriété 4 : Invariance à l'échelle

Invariance à l'échelle

Soit $\lambda \in \mathbb{R}_{+}^{*}$. La mesure I satisfait au principe d'invariance à l'échelle de mesure si
$$I(\lambda \mathbf{y}) = I(\mathbf{y}).$$

- La multiplication de tous les revenus par une constante positive n'a pas d'effet sur le niveau d'inégalité.
- La mesure I qui satisfait à cette propriété est sans unité. Elle est indépendante de la devise utilisée pour les revenus.
- On dira également que I est homogène de degré 0.
- Les mesures invariantes à l'échelle sont dites relatives.

Propriété 5 : Invariance par translation

Invariance par translation

Soit $c \in \mathbb{R}_+^*$. La mesure I satisfait au principe d'invariance par translation si

$$I(\mathbf{y} + c\mathbf{1}) = I(\mathbf{y}).$$

- L'addition d'une constante positive à tous les revenus n'a pas d'effet sur le niveau d'inégalité.
- La mesure I qui satisfait à cette propriété dépend de l'unité des observations.
- Les mesures invariantes par translation sont dites absolues.

Mesures absolues vs mesures relatives

- Une mesure d'inégalité ne peut pas être simultanément relative (invariante à l'échelle) et absolue (invariante par translation).
- Deux conceptions différentes des inégalités.
- Choix normatif :
  - Tous les revenus sont augmentés de 10%.
    mesure relative $\rightarrow$ niveau d'inégalité inchangé.
    mesure absolue $\rightarrow$ augmentation des inégalités.
  - Tous les revenus sont augmentés de CHF 1000.
    mesure relative $\rightarrow$ diminution des inégalités.
    mesure absolue $\rightarrow$ niveau d'inégalité inchangé.

Mesures relatives d'inégalité

- On ne se concentre ici que sur les mesures relatives.
- De part leur indépendance à l'échelle de mesure, les mesures relatives sont plus répandues dans la pratique.
- Quelques mesures connues : Gini, Atkinson, Theil, Zenga, QSR.

Plan

Indice de Gini

Indice de Gini (cas continu)

$$G = 1 - 2\int_0^1 L(\alpha)d\alpha.$$

- La mesure d'inégalité la plus célèbre et la plus utilisée (Gini, 1912).
- Recommandé par l'Union Européenne et Eurostat (indicateur secondaire).
- Prend une valeur entre 0 (égalité parfaite) et 1 (inégalité parfaite).
- Graphiquement: deux fois la surface entre la droite d'égalité parfaite et la courbe de Lorenz.

Indice de Gini

Indice de Gini (population finie)

$$G = \frac{2}{NY}\sum_{k\in U} k y_{(k)} - \frac{N+1}{N} = \frac{N^{-2}\sum_{k\in U}\sum_{\ell\in U}|y_k - y_\ell|}{2\overline{Y}}.$$

- au numérateur : moyenne des écarts absolus entre les revenus pris deux à deux (absolute mean difference).

  $\rightarrow$ l'indice de Gini est une mesure relative.

  $\rightarrow$ son numérateur est une mesure absolue d'inégalité.

Indice de Gini pour quelques pays d'Europe

Pays	Indice de Gini estimé
Norvège	0.236
Hongrie	0.241
Suède	0.241
Pays-Bas	0.255
Belgique	0.266
Allemagne	0.293
Suisse	0.295
France	0.299
Pologne	0.311
Italie	0.312
Royaume-Uni	0.330
Bulgarie	0.332
Roumanie	0.333
Espagne	0.339

(Eurostat, 2010)

Indices d'entropie généralisée

Indices d'entropie généralisée

$$\mathrm{GE}_\varphi = \begin{cases} \dfrac{1}{\varphi(\varphi-1)}\left[\dfrac{1}{N}\sum_{k\in U}\left(\dfrac{y_k}{\overline{Y}}\right)^\varphi - 1\right], & \text{if} \quad \varphi \neq 0, \varphi \neq 1, \\ \dfrac{1}{N}\sum_{k\in U}\log\dfrac{Y}{Ny_k}, & \text{if} \quad \varphi = 0. \\ \sum_{k\in U}\dfrac{y_k}{Y}\log\dfrac{Ny_k}{Y}, & \text{if} \quad \varphi = 1. \end{cases}$$

- Le paramètre φ permet de modifier la sensibilité de l'indice à l'une ou l'autre des extrémités de la distribution.
- Plus φ est grand, plus l'indice sera sensible aux changements en haut de la distribution.

Indice de Theil

Indice de Theil

$$\mathrm{GE}_1 = \mathrm{T} = \sum_{k \in U} \frac{y_k}{Y} \log \frac{N y_k}{Y}.$$

- L'indice de Theil (1967) fait partie de la famille des indices d'entropie généralisée (cas $\varphi = 1$).
- Utilisé pour sa faculté de décomposition additive en sous-groupes.

Indice d'Atkinson

Indice d'Atkinson (1970)

$$A_\epsilon = \begin{cases} 1 - \left[\dfrac{1}{N}\displaystyle\sum_{k\in U}\left(\dfrac{y_k}{\overline{\overline{Y}}}\right)^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}, & \text{if} \quad \epsilon \geq 0, \epsilon \neq 1, \\ 1 - \displaystyle\prod_{k\in U}\left(\dfrac{y_k}{\overline{\overline{Y}}}\right)^{\frac{1}{N}}, & \text{if} \quad \epsilon = 1. \end{cases}$$

- Construit à partir de la notion de fonction d'utilité sociale.
- Le paramètre ϵ se définit comme un paramètre d'aversion aux inégalités.
- Plus la valeur choisie pour ϵ est grande, plus l'indice est sensible aux inégalités.
- Dans la pratique, on utilise des valeurs de ϵ entre 0.5 et 1.5.

Indice de Zenga

Indice de Zenga (cas continu)

$$Z = 1 - \int_0^1 \frac{L(\alpha)}{\alpha} \cdot \frac{1-\alpha}{1-L(\alpha)} d\alpha.$$

- Nouvelle mesure d'inégalité proposée par Zenga (2007).
- Comme l'indice de Gini, l'indice de Zenga prend une valeur entre 0 et 1.
- Lié à la courbe de Lorenz.

Coefficient de variation

Coefficient de variation

$$\mathrm{CV} = \frac{\sigma_y}{\overline{Y}}.$$

- Un indice d'inégalité est une mesure de dispersion.
- La variance σ_y^2 ou l'écart-type σ_y peuvent être vus comme des mesures absolues d'inégalité.
- Le coefficient de variation peut être vu comme une mesure relative d'inégalité.

Rapports interquantiles

- Les quantiles permettent de construire des mesures de dispersion ou d'inégalité.
- On note Q_α, le quantile d'ordre α avec $\alpha \in [0; 1]$.

Rapports interquantiles

$$\mathrm{QR} = \frac{Q_{1-\alpha}}{Q_\alpha},$$

avec $\alpha \leq 0.5$.

- Les rapports interquantiles prennent une valeur 1 en cas d'égalité parfaite (violation de la propriété de normalisation).

Rapports interquantiles

- Les rapports de quantiles ont deux avantages :
  - facilité d'interprétation et de compréhension.
  - robustesse (insensibilité aux valeurs extrêmes).
- Dans la pratique les rapports suivants sont souvent rencontrés :

Rapport interdéciles et Rapport interquartiles

$$\mathrm{IDR} = \frac{Q_{0.9}}{Q_{0.1}}, \qquad \mathrm{IQR} = \frac{Q_{0.75}}{Q_{0.25}}.$$

- Exemple : En France en 2007, le rapport interdéciles est estimé à 3.39. *(Insee, 2008)*

Quintile Share Ratio (QSR ou S80/S20)

Quintile Share Ratio (QSR)

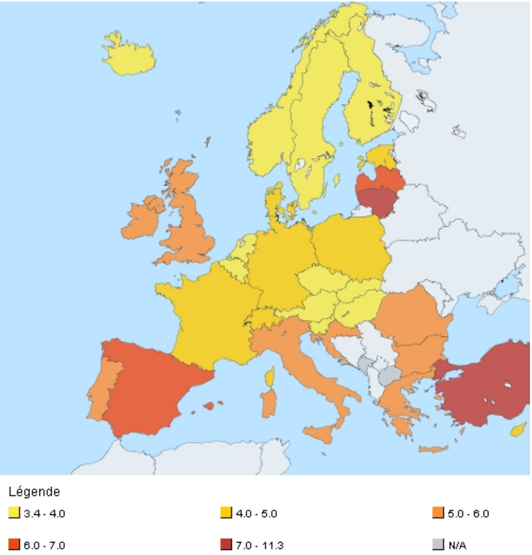
$$\text{QSR} = \frac{1 - L(0.8)}{L(0.2)}.$$

- Recommandé par l'Union Européenne et Eurostat (indicateur primaire).
- Facile à interpréter et à comprendre:

$$\text{QSR} = \frac{\text{revenu moyen des 20\% les plus riches}}{\text{revenu moyen des 20\% les plus pauvres}}.$$

- Prend la valeur minimale 1 si égalité parfaite.
- Peut se généraliser à n'importe quels quantiles (déciles, quartiles).

Le QSR en Europe



(Eurostat, 2010)

Les mesures relatives et leurs propriétés

	Gini	Theil	Atkin.	Zenga	CV	IDR	QSR
normalisation	✓	✓	✓	✓	✓	X	X
symétrie	✓	✓	✓	✓	✓	X	X
transferts	✓	✓	✓	✓	✓	X	X
invariance à l'échelle	✓	✓	✓	✓	✓	✓	✓
invariance par translat.	X	X	X	X	X	X	X

$\rightarrow$ les mesures faisant intervenir les quantiles ne sont pas conforme à l'approche axiomatique.

Décomposition

- Une autre propriété appréciée est la capacité d'une mesure à se décomposer.
- Deux types de décomposition :
  - En sous-groupes : partition de la population (age, région, sexe).
  - Par source de revenus : partition du revenu (salaire, rente).
- Permet d'avoir une vue plus détaillée de l'inégalité.
- Seuls les indices d'entropie généralisée admettent à la fois les propriétés usuelles de l'approche axiomatique et la décomposition additive en sous-groupes (Bourguignon, 1979; Shorrocks, 1980).
- Beaucoup de recherches sur la décomposition des autres indices.

Plan

Estimation

- Dans la pratique les mesures d'inégalité sont estimées dans la population à l'aide d'un échantillon.
- L'estimation des indices d'inégalité implique souvent l'estimation de la fonction de répartition F ou de la courbe de Lorenz.
- Plusieurs approches:
  - Non-paramétrique,
  - Paramétrique,
  - Semi-paramétrique.

Estimation : approche non-paramétrique

- Approche basée sur les seules données observées, sans référence à un modèle.
- Aucune contrainte à priori sur la forme de la distribution de revenus.
- Avantage : ne repose pas sur les hypothèses d'un modèle.
- Estimation directe des indices d'inégalité, de la courbe de Lorenz, des quantiles.

Exemple : un estimateur de l'indice de Gini

$$\widehat{G} = \frac{2\sum_{i=1}^{n} i y_{(i)}}{n\sum_{i=1}^{n} y_{(i)}} - \frac{n+1}{n}.$$

(cas iid, échantillon de taille n).

Estimation : approche paramétrique

- Approche basée sur la modélisation de la distribution de revenu à l'aide d'une loi de probabilité (Lognormale, Dagum, Generalized Beta).
- Les indices d'inégalité s'expriment comme des fonctions des paramètres de la distribution.
- L'estimation consiste ensuite à estimer ces paramètres.

Exemple : loi Lognormale et indice de Gini

La variable aléatoire Y suit une Loi Lognormale de paramètres μ et σ^2.
L'indice de Gini s'écrit alors :

$$G(\sigma) = 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$$

avec $\Phi(.)$, la fonction de répartition d'une Normale standardisée.

Estimation : approche semi-paramétrique

- Approche basée sur la modélisation des extrémités de la distribution de revenu à l'aide d'une loi de probabilité.
- La loi de Pareto est fréquemment utilisée pour modéliser le haut de la distribution.
- Le recours à un modèle pour étudier notamment les très hauts revenus permet de résoudre partiellement les problèmes de robustesse.

Estimation : les enjeux (I)

1. La fonction à estimer est non-linéaire.
   - Méthodes spécifiques pour l'estimation de variance et la construction d'intervalle de confiance.
   - Deux familles de méthodes :
     $\rightarrow$ linéarisation (méthodes asymptotiques),
     $\rightarrow$ replication (bootstrap, jackknife).
2. L'échantillon peut être issu d'un plan de sondage complexe.
   - La méthode d'estimation doit prendre en compte le plan de sondage.

Estimation : les enjeux (II)

3. Les données peuvent être groupées.
   - Données de revenu disponibles sous forme de classes (intervalles).
   - Prendre en compte la dispersion à l'intérieur des classes.
4. Le revenu peut être nul ou négatif.
   - Certaines mesures ne tolèrent que des revenus positifs.
5. Le revenu d'un individu est une information sensible.
   - Problèmes de non-réponse non-ignorable.
   - Erreurs de mesure.

Estimation : les enjeux (III)

6. Les mesures d'inégalité sont sensibles aux valeurs extrêmes.
   - Distributions fortement asymétriques, valeurs extrêmes (très haut revenus).
   - La fonction d'influence permet d'étudier la robustesse des mesures d'inégalité.
   - Seuls les rapports interquantiles sont robustes, d'où leur intérêt.
7. Mesure de l'évolution du niveau d'inégalité.
   - Enquêtes longitudinales.
   - Comment mesurer l'évolution et établir si cette évolution est significative.

Plan

Conclusion

- Choix de la mesure :
  - Objectif, public cible, propriétés.
- Mesures relatives ou absolues ?
  - Comment doit-on percevoir les inégalités ?
  - Les individus semblent souvent percevoir les inégalités comme absolues (Amiel and Cowell, 1992, 1997).
- Quelle approche choisir pour l'estimation ?
  - Possibilité d'éviter la dépendance à un modèle.
  - Nécessité de modéliser pour des questions de robustesse ?
- Sujets connexes :
  - mesures de pauvreté (par ex. taux de risque de pauvreté).
  - indicateurs composites (par ex. indice de Sen, IDH).