Zeitschrift: Mitteilungen / Schweizerische Vereinigung der Versicherungsmathematiker = Bulletin / Association Suisse des Actuaires = Bulletin / Swiss Association of Actuaries

Herausgeber: Schweizerische Vereinigung der Versicherungsmathematiker

Band: - (1990)

Heft: 1

Artikel: Gemeinsames bei der Messung von Ungleichheit, Streuung, Risiko und Information

Autor: Eichhorn, Wolfgang / Vogt, Arthur

DOI: https://doi.org/10.5169/seals-967229

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 21.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

WOLFGANG EICHHORN, Karlsruhe und ARTHUR VOGT, Bern

Gemeinsames bei der Messung von Ungleichheit, Streuung, Risiko und Information*

1. Einleitung

In der Wissenschaft können zwei Richtungen unterschieden werden, einerseits das Immer-weiter-Treiben der Spezialisierung in immer feinere Verästelungen, andererseits das Entdecken von Zusammenhängen zwischen verschiedenen Gebieten. *Koestler* (1967) nennt die erste "arboration" (Verzweigung) und die zweite "reticulation" (Vernetzung).
Beim Entdecken von Zusammenhängen sind wiederum zwei Spielarten denkbar. Einerseits können Methoden eines Gebietes im "Neuland" eines anderen Gebietes verwendet werden, womit eine Doppelarbeit vermieden wird. Als kleines Beispiel sei die Übertragung der Methoden der Preismessung auf die Morbiditätsmessung genannt (*Vogt,* 1986). Anderseits können aber auch beide Gebiete bereits unabhängig voneinander untersucht worden sein, und nachträglich kann eine formale Analogie der beiden verwendeten Methoden festgesetellt werden. Als prominentes Beispiel hierzu nennt unser Jubilar in seiner Antrittsrede (*Bühlmann,* 1964) die mathematische Statistik, die sich aus der Verbindung der (deduktiven) Wahrscheinlichkeitsrechnung mit der (induktiven) Statistik ergibt:

> Als mathematische Statistik bezeichnen wir heute jene Wissenschaft, die zwischen Beobachtungsgrössen einerseits und erklärenden gedanklichen Modellen anderseits die Brücke objektiven Schliessens schlägt .... Wann ist aber diese Querverbindung zum ersten Mal von einem Mathematiker erkannt worden? Diese entscheidende erkenntnistheoretische Tat darf sicher als die Geburtsstunde der mathematischen Statistik gewürdigt werden ... Der Mathematiker, der den genialen Brückenschlag vollzog: *Jakob Bernoulli;* das Buch, welches die Querverbindung zwischen Wahrscheinlichkeitsrechnung und Statistik herstellt: "Ars conjectandi".

* Ein besonderer Dank gebührt den Herren Peter Bader vom Bundesamt für Privatversicherungswesen, Bern, und Hugo Huber von der Krankenkasse KKB, Bern, für die Berechnung der im folgenden behandelten Masszahlen der Lebens- bzw. Krankenversicherung. Wir danken auch dem Bundesamt für Sozialversicherung, Bern, für die Bereitstellung von Daten zur AHV-Einkommensstatistik.

Die vorliegende Arbeit ist vorwiegend der letztgenannten Spielart zuzuordnen. Sie befasst sich mit den vier Begriffen Ungleichheit, Streuung, Risiko und Information und deren Vernetzung, und zwar von einem messtheoretischen Standpunkt aus. Zur Frage des Messens innerhalb der vier betrachteten Problemkreise gibt es eine Fülle von Literatur. Dabei geht es fast immer nur um genau einen der vier Begriffe. Eine Zusammenschau folgt in Abschnitt 2, dem theoretischen Teil. Unser Ziel ist es zu zeigen, dass die Probleme und Methoden des Vergleichens und Messens, wie sie im Kontext der Begriffe

(1) *Ungleichheit* (Beispiel: Vergleich von Einkommensverteilungen in der Wirtschaftswissenschaft),
(2) *Streuung* (Beispiel: Vergleich von Verteilungen in der Statistik),
(3) *Risiko* (Beispiel: Vergleich von Lotterien) und
(4) *Information* (Beispiel: Vergleich von Informationsstrukturen in der Informationstheorie).

auftreten, eine nahezu identische Struktur besitzen. Was das Vergleichen anbelangt, stützen wir uns auf eine unveröffentlichte Vortragsunterlage (*Nermuth,* 1988).
Elementare Vergleichskriterien führen in den betrachteten Mengen zu keiner vollständigen Ordnung. Mit anderen Worten: Nicht für jedes Paar von Elementen aus einer solchen Menge M ist aufgrund dieser Kriterien entscheidbar, welches der beiden Elemente vor dem anderen anzuordnen ist.
Dieses Dilemma kann behoben werden, und zwar durch die Einführung sogenannter Masse

$$F : M \longrightarrow R \quad (R = \text{Menge der reellen Zahlen}),$$

die neben einer Reihe zweckmässiger Eigenschaften vor allem die folgende Eigenschaft haben: Dominiert von zwei Elementen x und y aus M nach den oben angesprochenen Vergleichskriterien das Element y das Element x streng (oder schwach), so gilt

$$F(x) \underset{(=)}{<} F(y).$$

Als Beispiel aus einem anderen Zusammenhang seien dazu Preisindices genannt. Wenn alle Preise in der Situation 0 kleiner sind als in der Situation 1,

so sind die Lebenshaltungskosten in der Situation 0 offensichtlich geringer als in der Situation 1. Wenn die Preise nur eines Teiles der Güter in der Situation 0 und die der übrigen Güter in der Situation 1 niedriger sind, ist ein Vergleich nicht ohne weiteres möglich. Dann werden Preisindices eingeführt; vergleiche etwa *Eichhorn* (1976), *Eichhorn/Voeller* (1976), *Vogt* (1979). Ein Preisindex ist ein Mass, das eine Reihe naheliegender Bedingungen erfüllt und insbesondere für je zwei Preissituationen zu entscheiden erlaubt, welche die "teurere" bzw. die "billigere" ist.
Bei der Beschäftigung mit Eigenschaften derartiger Masse ist eine "ars metiendi" gefragt, oder – modern ausgedrückt – eine axiomatische Fundierung. Der Abschnitt 3 ist praxisbezogen. Unter Ökonomen ist die Messung der Ungleichheit, etwa einer Einkommensverteilung, wohlbekannt. Dabei werden Ungleichheiten z. B. mit der Lorenzkurve dargestellt und mit dem Gini-Ungleichheitsmass gemessen. Mathematiker sprechen dagegen eher von "Streuung", die in der Form der relativen Streuung z. B. mit dem Variationskoeffizienten gemessen wird.

2. Ars metiendi

2.1 Ungleichheit

Es sei R_+ die Menge der nichtnegativen reellen Zahlen und R_+^n das n-fache karthesische Produkt solcher Mengen. Gegeben seien zwei Einkommensverteilungen

$$x = (x_1, \dots, x_n) \in R_+^n \quad \text{und} \quad y = (y_1, \dots, y_n) \in R_+^n$$

mit

$$x_1 + \cdots + x_n = y_1 + \cdots + y_n.$$

Die Verteilung x heisst *weniger ungleich* als die Verteilung y, wenn x durch eine endliche Folge von Transfers der folgenden Art aus y kompensiert werden kann: Eines der Einkommen in y, sagen wir y_i, wird auf y_i^* reduziert, $y_i^* = y_i - z_i$ $(y_i \geq z_i \geq 0)$. Gleichzeitig wird ein anderes Einkommen um den Betrag z_i erhöht. Das so erhöhte Einkommen darf dabei nicht grösser als y_i sein.

Diese Definition ist gleichbedeutend – vgl. etwa *Marshall/Olkin* (1979) – mit: Es existiert eine bistochastische Matrix B, das heisst eine Matrix mit nichtnegativen Elementen und Summe der Elemente jeder Zeile gleich 1 sowie Summe der Elemente jeder Spalte gleich 1, derart dass

$$x = By. \tag{2.1.1}$$

Anmerkung: Wenn hier und im folgenden ein Vektor rechts neben einer Matrix oder einem Vektor steht, ist er als ein Spaltenvektor aufzufassen.
Die hier durch "weniger ungleich" eingeführte Vergleichsmöglichkeit zwischen zwei Einkommensverteilungen hat mehrere Mängel, die für die praktischen Anwendungen gravierend sind:

(1) Es gibt Einkommensverteilungen $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ mit $x_1 + \cdots + x_n = y_1 + \cdots + y_n$, die bezüglich "weniger ungleich" nicht miteinander verglichen werden können. Selbst für solche Vektoren liegt also keine vollständige, sondern nur eine partielle Ordnung vor (Beispiel für $n = 3$: *Weder* (2,4,8) ist weniger ungleich als (1,6,7) noch gilt das Umgekehrte).

(2) Zwei Einkommensverteilungen x und y mit $x_1 + \cdots + x_n \neq y_1 + \cdots + y_n$ sind bezüglich "weniger ungleich" nicht vergleichbar.

(3) Ist $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_m)$ mit $n \neq m$, so gibt es bezüglich "weniger ungleich" keine Vergleichsmöglichkeit für die beiden Einkommensverteilungen x und y.

Um diese Mängel zu beheben, führt man sogenannte Ungleichheitsmasse ein. Ein Ungleichheitsmass ist eine Funktionenfolge

$$F_n : R_+^n \longrightarrow R \tag{2.1.2}$$

mit den folgenden Eigenschaften (2.1.3), (2.1.4):

$$F_n(By) \leq F_n(y) \tag{2.1.3}$$

für alle natürlichen $n \geq 2$ und alle $y \in R_+^n$ und alle bistochastischen Matrizen B,

$$F_n(By) < F_n(y) \tag{2.1.4}$$

genau dann, wenn die Komponenten des Vektors By keine Permutation der Komponenten des Vektors y sind.

Eine Funktion (2.1.2) mit (2.1.3) heisst Schur-konvex, eine mit (2.1.3) und (2.1.4) heisst streng Schur-konvex. Offenbar hat ein solches Ungleichheitsmass die folgenden Eigenschaften:

(a) Es hat keine der Mängel (1), (2), (3).

(b) Ist x weniger ungleich als y im Sinne der Definition (2.1.1), dann wegen (2.1.3), (2.1.4) auch bezüglich des Ungleichheitsmasses F_n. (Wegen der Eigenschaften der Matrizen B sind die extremen Komponenten des Vektors $x = By$ näher oder gleich weit voneinander entfernt als die des Vektors y.)

(c) Es ist symmetrisch; das folgt aus der Tatsache, dass die Permutationsmatrizen eine Teilmenge der bistochastischen Matrizen sind.

(d) Es hat die Mittelwerteigenschaft

$$F_n(a, \dots, a) \leq F_n(y_1, \dots, y_n) \leq F_n(0, \dots 0, y_1 + \cdots + y_n)$$

für $a = (y_1 + \cdots + y_n)/n$ bzw.

$$F_n(a, \dots, a) < F_n(y_1, \dots, y_n) < F_n(0, \dots, 0, y_1 + \cdots + y_n)$$

für $(y_1, \dots, y_n) \neq (a, \dots, a)$
und $(y_1, \dots, y_n) \neq$ Permutation von $(0, \dots, 0, y_1 + \cdots + y_n)$

Oft werden die Forderungen (2.1.3), (2.1.4) an ein Ungleichheitsmass noch ergänzt durch

– Normierung

$$F_n(b, \dots, b) = 0 \tag{2.1.5}$$

für alle $n \geq 2$ und alle $b \in R_+$

und

– Extensibilität:

$$F_n(y_1, \dots, y_n) \leq F_{n+1}(y_1, \dots, y_n, 0) \tag{2.1.6}$$

für alle $n \geq 2$ und alle $(y_1, \dots, y_n) \in R_+^n$;

vgl. etwa *Eichhorn/Gehrig* (1982). Da aber in der einschlägigen Literatur spezielle Ungleichheitsmasse vorgeschlagen werden, die (2.1.5) und (2.1.6) nicht erfüllen, nehmen wir diese Forderungen nicht in die Definition eines Ungleichheitsmasses auf.
Welches spezielle Ungleichheitsmass man für praktische Anwendungen verwendet, hängt vom Zweck ab, den man mit der Messung verbindet. Dieser Zweck drückt sich in weiteren Forderungen (neben (2.1.3), (2.1.4)) an das Mass F_n aus mit dem Ziel, durch (2.1.3) und (2.1.4) und die weiteren Forderungen genau ein spezielles Ungleichheitsmass zu charakterisieren.

Charakterisierungen von Massen wie des

– Gini-Ungleichheitsmasses

$$G_n(y_1, \dots, y_n) = \frac{1}{2n^2 a} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j| \tag{2.1.7}$$

und der

– Normierten Varianz

$$V_n(y_1, \dots, y_n) = \frac{1}{na^2} \sum_{i=1}^{n} (y_i - a)^2 \tag{2.1.8}$$

und anderer spezieller Ungleichheitsmasse finden sich in der einschlägigen Literatur.

Wegen $a = (y_1 + \cdots + y_n)/n$ gelten

$$G_n(\mu y_1, \dots, \mu y_n) = G_n(y_1, \dots, y_n)$$

und

$$V_n(\mu y_1, \dots, \mu y_n) = V_n(y_1, \dots, y_n)$$

für alle $\mu > 0$ und alle $(y_1, \dots, y_n) \in R_+^n$.

Ungleichheitsmasse F_n mit

$$F_n(\mu y) = F_n(y) \tag{2.1.9}$$

für alle $\mu > 0$ und alle $y \in R_+^n$

heissen relative oder (vom Grade 0) homogene Ungleichheitsmasse.

2.2 *Streuung*

Sei X ein Banachraum. Wir schreiben

$$u = (u_1, \dots, u_n) \in R_+^n, \quad u_1 + \cdots + u_n = 1, \quad x = (x_1, \dots, x_n), \quad x_i \in X \tag{2.2.1}$$

und bezeichnen mit (u, x) die Verteilung, die dem Punkt $x_i \in X$ die Wahrscheinlichkeit $u_i \geq 0$ zuordnet $(i = 1, \dots, n)$. Analog sei (v, y) eine Verteilung, die dem Punkt y_j die Wahrscheinlichkeit v_j zuordnet $(j = 1, \dots, m)$. Es gelte

$$x_1 + \cdots + x_n = y_1 + \cdots + y_m. \tag{2.2.2}$$

Eine Verteilung (u, x) *streut weniger* als eine Verteilung (v, y), wenn

$$\sum_{i=1}^{n} u_i f(x_i) \leq \sum_{j=1}^{m} v_j f(y_j) \tag{2.2.3}$$

für alle konvexen Funktionen $f : X \longrightarrow R$.

Diese Definition ist äquivalent – siehe etwa *Bourgin* (1983) – zu der folgenden:

(u, x) *streut weniger* als (v, y), wenn eine zeilenstochastische Matrix B, das heisst eine Matrix mit nichtnegativen Elementen und Summe der Elemente jeder Zeile gleich 1, existiert derart, dass

$$v = uB \qquad \text{und} \qquad x = By. \tag{2.2.4}$$

Die beiden betrachteten Verteilungen haben wegen $ux = uBy = vy$ den gleichen Mittelwert. Der Spezialfall

$$n = m, \quad u = v = \left(\frac{1}{n}, \dots, \frac{1}{n}\right) \quad (B \text{ also bistochastisch}),\ X = R_+$$

stellt gerade die in Abschnitt 2.1 behandelte Situation dar.
Es liegt nahe, analog wie in 2.1 unter einem Streuungsmass eine Funktionenfolge

$$F_{2n} : R_+^n \times X^n \longrightarrow R_+ \tag{2.2.5}$$

mit den folgenden Eigenschaften (2.2.6) und (2.2.7) zu verstehen:

$$F_{2n}(u, By) \leq F_{2m}(uB, y) \tag{2.2.6}$$

für alle $n \geq 2$, $m \geq 2$ sowie für alle $u \in R_+^n$, $y \in X^m$ und alle zeilenstochastischen nxm-Matrizen B,

$$F_{2n}(\mu u, x) = F_{2n}(u, x) \tag{2.2.7}$$

für alle $\mu > 0$ und alle $u \in R_+^n, x \in X^n$.
Eigenschaft (2.2.7) ist wegen der Interpretation der Komponenten $u_1, \dots, u_n$ von $u \in R_+^n$ als Wahrscheinlichkeiten erforderlich:

$$F_{2n}\left(\frac{u}{u_1 + \cdots + u_n}, x\right) = F_{2n}(u, x).$$

Bei den Streuungsmassen handelt es sich offenbar um eine spezielle Klasse bezüglich $y \in X^n$ Schur-konvexer und bezüglich $u \in R_+^n$ Schur-konkaver Funktionen. Dies geht aus (2.2.6) im Falle $m = n$ hervor, wobei zu beachten ist, dass die Menge der zeilenstochastischen $n \times n$-Matrizen die der bistochastischen enthält.

2.3 *Risiko*

Wir interpretieren die Komponenten des Vektors

$$y = (y_1, \dots, y_n) \in R_+^n$$

als die Preise einer Lotterie. Diese Lotterie ist durch die Eintrittswahrscheinlichkeiten

$$(u_1, \dots, u_n) = u \in R_+^n, \quad u_1 + \cdots + u_n = 1,$$

der Preise charakterisiert. Zwei beliebige Lotterien mit beliebigen Preisvektoren können durch Zulassen von Nullwahrscheinlichkeiten stets zu zwei Lotterien $u \in R_+^n$ und $v \in R_+^n$ mit gleichem Preisvektor $y \in R_+^n$ gemacht werden.
Die Lotterie (u, y) heisst *weniger riskant* als die Lotterie (v, y) wenn

$$\sum_{i=1}^{n} u_i U(y_i) \geq \sum_{i=1}^{n} v_i U(y_i) \tag{2.3.1}$$

für alle konkaven Funktionen $U : R_+ \longrightarrow R$

gilt. Mit (2.3.1) werden die Erwartungswerte der (Nutzen)funktion U bezüglich der Verteilungen u und v verglichen.
Diese Definition ist äquivalent – siehe etwa *Mosler* (1982) – zu:
Die Lotterie (u, y) heisst *weniger riskant* als die Lotterie (v, y), wenn eine zeilenstochastische Matrix B existiert derart, dass

$$v = uB \qquad \text{und} \qquad y = By.$$

Diese Lotterien haben den gleichen Erwartungswert, denn

$$vy = uBy = uy.$$

Analog wie in 2.1 und 2.2 definieren wir als Risikomass einer Lotterie (u, y) eine Funktionenfolge

$$F_{2n} : R_+^{2n} \longrightarrow R \tag{2.3.2}$$

mit den Eigenschaften (2.3.3) und (2.3.4):

$$F_{2n}(u, By) \leq F_{2n}(uB, y) \tag{2.3.3}$$

für alle $n \geq 2$ und alle $u \in R_+^n$ sowie für alle y und alle zeilenstoachastischen Matrizen B mit $By = y$,

$$F_{2n}(\mu u, y) = F_{2n}(u, y) \tag{2.3.4}$$

für alle $n \geq 2$ sowie für alle $\mu > 0$ und alle $u \in R^n_+$, $y \in R^n_+$.

2.4 *Information*

Jetzt sei XCR^q die Menge aller Wahrscheinlichkeitsverteilungen auf einem endlichen Zustandsraum Ω. Die Informationsstruktur (u, x) gibt mit Wahrscheinlichkeit u_i das Signal i, und in diesem Fall ist dann

$$x_i = (x_{i1}, \ldots, x_{iq})$$

die a-posteriori Wahrscheinlichkeitsverteilung auf Ω. Mit anderen Worten ist u_{ik} die a-posteriori-Wahrscheinlichkeit, dass der Zustand k eintritt, vorausgesetzt, das Signal i wurde beobachtet.
Ferner sei A eine endliche Menge möglicher Aktionen und $g : A \times \Omega \longrightarrow R$ eine von Neumann-Morgenstern-Nutzenfunktion. Gegeben ein Signal i, wählt der Akteur eine Aktion a, die seinen erwarteten Nutzen maximiert. Wir schreiben

$$h(x_i) := \max_{a \in A} \sum_{k=1}^{q} x_{ik} g(a, k). \tag{2.4.1}$$

Die Funktion $h : X \longrightarrow R$ ist konvex für jedes g.
Der *Wert der Informationsstruktur* (u, x) bei der Nutzenfunktion g ist definiert durch

$$W(u, x, g) := \sum_{i=1}^{n} u_i h(x_i) \tag{2.4.2}$$

Eine Informationsstruktur (v, y) heisst *informativer* als eine Informationsstruktur (u, x), wenn gilt:

$$W(u, x, g) \leq W(v, y, g) \quad \text{für alle Nutzenfunktionen } g. \tag{2.4.3}$$

Satz (Blackwell): Die folgenden Bedingungen sind äquivalent:

(v, y) ist informativer als (u, x), (2.4.4)

(u, x) streut im Sinne von (2.2.4) weniger als (v, y). (2.4.5)

Man beachte, dass die stärker streuende Struktur informativer ist. Wo gar keine Streuung ist, ist überhaupt keine Information (maximale Entropie, Wärmetod).

Als einschlägige Literatur zu diesem Thema seien die Arbeiten von *Blackwell* (1953), *Blackwell/Girshick* (1954) und *Nermuth* (1982) genannt.

Informationsmasse sind mathematisch-statistisch gesehen offenbar nichts anderes als Streuungsmasse, wie sie in Abschnitt 2.2 behandelt wurden. Zur Charakterisierung von speziellen Informationsmassen in einem etwas weniger allgemeinen Zusammenhang siehe *Aczél/Daróczy* (1975).

3. Anwendungen in der Versicherungsmathematik

3.1 Das Gini-Ungleichheitmass

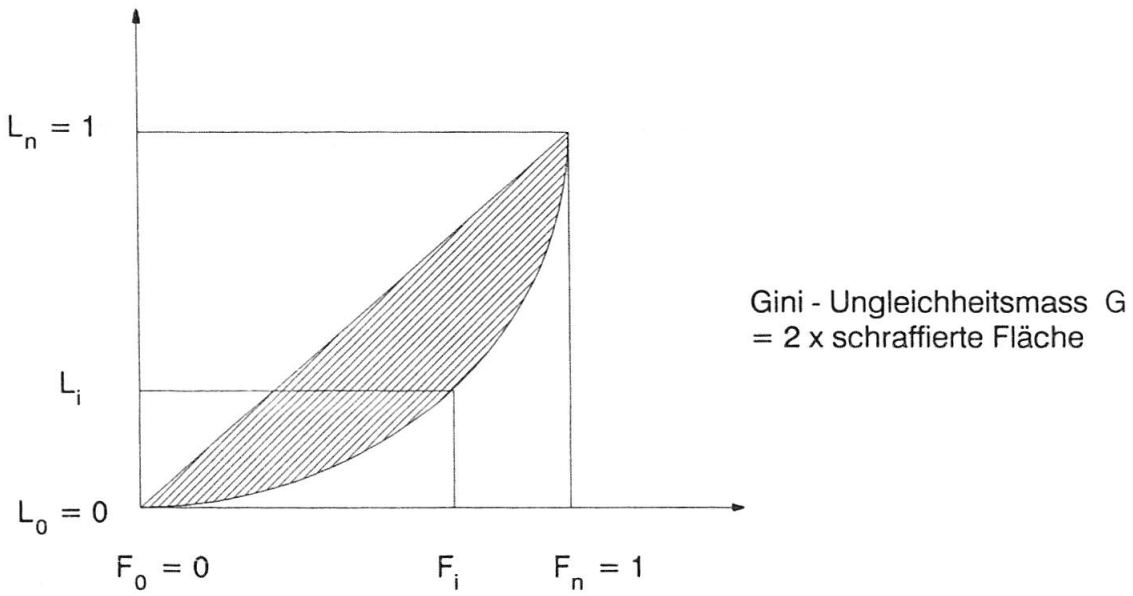
Das bekannteste Mass, um die "Ungleichheit einer Verteilung" zu messen, beruht auf der *Lorenzkurve*. Lorenz hat 1905 diese Kurve eingeführt, um zu beurteilen, in welchem von zwei Ländern die Einkommen gleichmässiger verteilt sind oder ob in einem Land mit der Zeit die Verteilung gleichmässiger wird oder nicht (*Dagum,* 1983b, 1987). Bei dieser Kurve gibt die Abszisse den Anteil der ärmsten Personen an der Gesamtbevölkerung an, deren Anteil am gesamten Einkommen der Ordinate entspricht. Die Lorenzkurve ist vor allem beliebt, um Einkommens- und Vermögenskonzentrationen zu beurteilen. Wir glauben jedoch, dass sie allgemein bei extensiven Zufallsgrössen, die man addieren kann, sinnvoll ist (z. B. nicht bei der Temperatur, die eine intensive Grösse ist). Wir wollen die Lorenzkurve und das daraus abgeleitete Gini-Ungleichheitsmass im folgenden in drei Gebieten anwenden. Wir wollen untersuchen, wie ungleich

- die Lebensdauer der Individuen gemäss verschiedenen Sterbetafeln verteilt sind,
- die Krankheitskosten auf die verschiedenen sie verursachenden Personen verteilt sind und
- die Einkommen gemäss AHV-Lohnstatistik verteilt sind.

Im Jahre 1914 hat Gini das nach ihm benannte Ungleichheitsmass (Gini-Index, Gini-Koeffizient, *Gini-Ungleichheitsmass)*

$$G = 1 - 2 \int_0^1 L(F)\, dF \tag{3.1.1}$$

eingeführt (*Dagum* 1983b), das dem doppelten Flächeninhalt zwischen der 45°-Geraden $L = F$ und der Lorenzkurve $L(F)$ entspricht. G ist 0 bei absolut gleichmässiger Verteilung und 1 wenn z. B. das gesamte Vermögen einem einzelnen gehört. Es lässt sich übrigens beweisen, dass die Formeln (3.1.1) und (2.1.7) äquivalent sind.



Bei sich nicht schneidenden Lorenzkurven kann man unmittelbar eine Aussage machen, wie sie nach Ungleichheit anzuordnen sind. Wenn sie sich jedoch schneiden, ist man auf ein Mass angewiesen, beispielsweise auf das obige Gini-Ungleichheitsmass.

Formel (3.1.1) gilt sowohl für stetige wie auch für diskrete Verteilungen. In der Praxis hat man meist geschichtete Daten, d. h. zum Beispiel Einkommensklassen von 0 bis 999 Franken, von 1000 bis 1999 Franken usw. Mit der Trapezregel erhält man aus (3.1.1) die entsprechende Formel für derart geschichtete Daten (mit $F_0 = L_0 = 0$, $F_n = L_n = 1$):

$$G = 1 - 2 \sum_{i=1}^{n} (F_i - F_{i-1}) \frac{L_i + L_{i-1}}{2}. \tag{3.1.2}$$

Wir wollen (3.1.2) zunächst in der *Lebensversicherung* anwenden. Für die Bezeichnungen siehe man etwa *Gerber* (1986, S. 114 ff.). Die F_i geben da den Anteil der Personen an der Gesamtbevölkerung an, die vor dem Alter i sterben (in diesem Fall $n = \omega + 1$):

$$F_i = \frac{\sum_{x=0}^{i-1} d_x}{\sum_{x=0}^{\omega} d_x} = \frac{\sum_{x=0}^{i-1}(l_x - l_{x+1})}{\sum_{x=0}^{\omega}(l_x - l_{x+1})} = \frac{l_0 - l_i}{l_0} \tag{3.1.3}$$

und die L_i geben den Anteil dieses Bevölkerungsanteils an der gesamten Lebenszeit der Gesamtbevölkerung an, also

$$L_i = \frac{\sum_{x=0}^{i-1}(x + 1/2)d_x}{\sum_{x=0}^{\omega}(x + 1/2)d_x} = \frac{\sum_{x=0}^{i-1}(x + 1/2)(l_x - l_{x+1})}{\sum_{x=0}^{\omega}(x + 1/2)(l_x - l_{x+1})}. \tag{3.1.4}$$

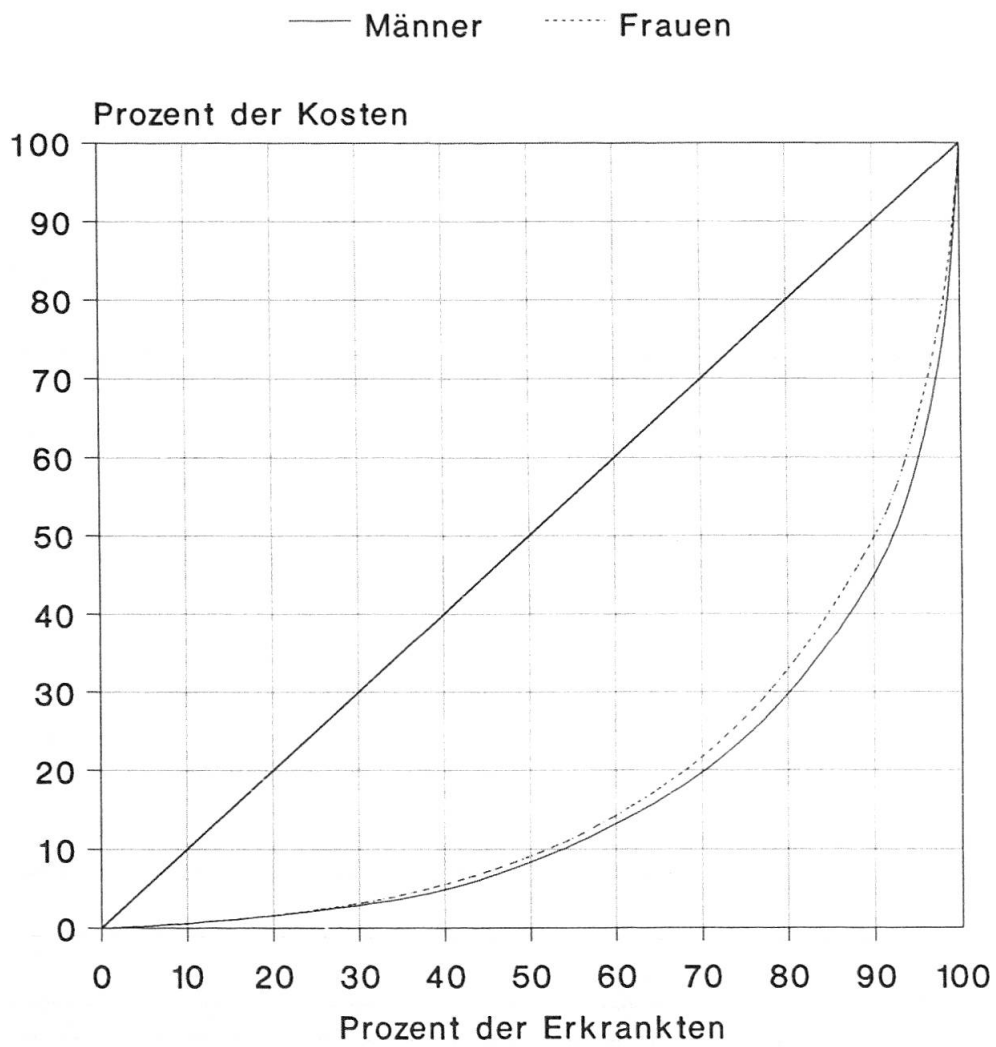
Setzen wir (3.1.3) und (3.1.4) in (3.1.2) ein, so ergibt sich das Gini-Ungleichheitsmass

$$G = 1 - 2\sum_{i=1}^{\omega+1}\left(\frac{d_{i-1}}{l_0}\frac{\sum_{x=0}^{i-2}(x + 1/2)d_x + \frac{(i-1/2)d_i}{2}}{\sum_{x=0}^{\omega}(x + 1/2)d_x}\right) \tag{3.1.5}$$

$$= 1 - \frac{2}{l_0\sum_{x=0}^{\omega}(x + 1/2)d_x}\sum_{i=0}^{\omega} d_i\left(\sum_{x=0}^{i-1}(x + 1/2)d_x + \frac{(i + 1/2)d_i}{2}\right) \tag{3.1.6}$$

In der Tabelle im Anhang finden sich die numerischen Werte von G für verschiedene Sterbetafeln. Es zeigt sich unter anderem, dass die Lebensdauern von Frauen gleichmässiger verteilt sind als die von Männern und dass bei beiden Geschlechtern die Lebensdauern sich mit der Zeit gleichmässiger verteilen.

Bei der Anwendung in der Krankenversicherung gehen wir direkt von Formel (3.1.2) aus. Die F_i geben dabei den Anteil der erkrankten Personen an, die Kosten der Klasse i oder weniger verursachen und die L_i den Anteil an den Gesamtkosten, den dieser Personenanteil verursacht. Wir haben zur graphischen Darstellung der Lorenzkurve auf der nächsten Seite das Beispiel der Krankenversicherung gewählt, da hier – in den Worten eines Journalisten, allerdings im Zusammenhang mit der Vermögensverteilung in der Bundesrepublik Deutschland – "der weite Bogen um die Gerechtigkeit"(vgl. "Die Zeit", 3.11.1978) am längsten ist. Es sind die Lorenzkurven für die Heilungskosten (ohne Zusatzversicherung) pro Versicherten der Krankenkasse



Quelle: KKB, Rechnungsjahr 1988
16.05.89 / hb

KKB für das Rechnungsjahr 1988 aufgezeichnet. In der Tabelle im Anhang sind die Werte des Gini-Ungleichheitsmasses angegeben, je für Männer und Frauen separat. Es zeigt sich, dass die Krankheitskosten der Frauen leicht gleichmässiger verteilt sind als die der Männer. Da diese Kurven in früheren Jahren praktisch gleich verlaufen, haben wir sie nicht angegeben. Die Versicherten, die gar keine Kosten verursachen, sind nicht einbezogen. Falls man das auch tut, verläuft die Lorenzkurve zuerst horizontal bis zur Quote der Versicherten mit Kosten Null.

In *Kunz* (1988) sind die Lorenzkurven der *Einkommen gemäss AHV-Lohnstatistik* für die Jahre 1953, 1968 und 1985 angegeben. Es zeigt sich, dass diese Verteilungen mit der Zeit gleichmässiger wurden. Die entsprechenden Werte des Gini-Ungleichheitsmasses finden sich in der Anhangtabelle. – Insgesamt zeigt sich: Am gleichmässigsten sind die Lebensdauern verteilt, dann kommen die Einkommen und am ungleichmässigsten sind die Krankheitskosten verteilt.

In diesem Zusammenhang ist auf das Phänomen eines Trends aufmerksam zu machen, den man unter Umständen eliminieren möchte. *Lüthi* (1981) weist im Kapitel "Dynamische Ungleichheitsmessung" darauf hin, dass wegen der Änderung des Einkommens oder des Vermögens der Individuen innerhalb des Lebenszyklus die statische Ungleichheitsmessung nicht nur durch die Differenzen zwischen den Individuen, sondern auch durch die Altersverteilung bestimmt wird. Möchte man nur die Differenzen zwischen den Individuen erfassen, so muss der Einfluss der Altersstruktur der Bevölkerung eliminiert werden. Als Lösungsmöglichkeit dazu nennt *Lüthi* (1981) die Neutralisierung der Altersstruktur dadurch, dass aus nach dem Alter geschichteten Daten mit Hilfe des Gini-Ungleichheitsmasses die Ungleichheit infolge unterschiedlichen Alters geschätzt wird. Die so entstandene Lorenzkurve infolge unterschiedlichen Alters wird nun als neue Referenzkurve für Gleichheit benutzt (anstelle der 45-Grad-Linie): Gleichheit besteht dann, wenn alle Individuen über den Lebenszyklus hinweg dieselben Einkommens- oder Vermögensprofile aufweisen. Damit ergibt sich aus der Differenz zwischen dem gebräuchlichen Gini-Mass und dem Gini-Mass infolge unterschiedlichen Alters ein neues Ungleichheitsmass. – Für den Fall der Stromgrösse "Einkommen", könnte man auch den Endwert der Einkommen über den Lebenszyklus mit dem gebräuchlichen Gini-Ungleichheitsmass untersuchen. Dann braucht man zur "Gleichheit" nicht gleiche Einkommensprofile; wer in jungen Jahren viel verdient und später

wenig, wäre dann "gleich gut gestellt" wie jemand, dem es umgekehrt ergeht (was wohl nicht den psychologischen Sachverhalt wiedergibt). – Die *Eliminierung eines Trends* wäre bei den Krankheitskosten übrigens noch viel wichtiger, da die Krankheitskosten mit dem Alter leider noch mehr zunehmen als die Einkommen. Die Autoren sind überzeugt, dass hier noch interessante Resultate zu finden sind.

3.2 *Die Messung der relativen Streuung*

Es liege ein System verschiedener, aber vergleichbarer Grössen vor. Ist dieses für Ökonomen von Relevanz und wollen sie die Unterschiede messen, so sprechen sie von "Ungleichheits-Messung". Im Unterschied dazu sprechen mathematische Statistiker und Physiker eher von "Streuung". Beidemale ist dasselbe gemeint. Oder was ist der Unterschied zwischen den beiden Aussagen:

Die Einkommen sind stark ungleich verteilt

und

Die Einkommen streuen stark?

Neben dem Gini-Ungleichheitsmass (3.1.1) könnte man ein relatives, von der Geldeinheit unabhängiges Streuungsmass, z. B. den Variationskoeffizienten

$$VK = \frac{\text{Standardabweichung}}{\text{Erwartungswert}}, \tag{3.2.1}$$

verwenden. Mit der Lebenserwartung eines Neugeborenen, nämlich

$$e_0 = \frac{\sum_{x=0}^{\omega} d_x(x + 1/2)}{l_0}, \tag{3.2.2}$$

lautet das in der Sprache der Sterbetafeln

$$VK = \frac{1}{e_0}\sqrt{\sum_{x=0}^{\omega}\left(x + \frac{1}{2} - e_0\right)^2 \frac{d_x}{l_0}}. \tag{3.2.3}$$

Bei e_0 handelt es sich übrigens um ein sogenanntes Lagemass der Sterbetafel. Mit dem Gini-Ungleichheitsmass (3.1.1) und dem Variationskoeffizienten

haben wir nun auch Dispersionsmasse dafür. Beim Vergleich der Resultate des Variationskoeffizienten mit denjenigen des Gini-Ungleichheitsmasses, das bei extremer Ungleichverteilung den Wert 1 annimmt, ist zu berücksichtigen, dass (3.2.3) für extreme Streuung den Wert $\sqrt{n}$ annimmt. Dies ist aber auch die einzige der sieben Eigenschaften, die *Dagum* (1983a) bei Ungleichheitsmassen betrachtet, in denen sich diese beiden Masse unterscheiden. *Piesch* (1975) schlägt vor, Ungleichheitsmasse, die nicht "normalisiert" sind, d. h. nicht den Wert 0 für Gleichverteilung und den Wert 1 für extreme Ungleichverteilung haben, durch eine lineare Transformation zu *"normalisieren"*. Normalisiert man den Variationskoeffizienten in diesem Sinn, so erhält man die Wurzel der normierten Varianz gemäss (2.1.8).
Es ergibt sich aus der Tabelle, dass die *Reihenfolge* "im Grossen" bei den drei Anwendungen Lebensdauern, Einkommen und Krankheitskosten bezüglich beider Masse gleich ist. Das ist aber nicht notwendig. Verschiedene Masse haben neben gemeinsamen auch abweichende Eigenschaften. Insbesondere das Gini-Ungleichheitsmass und der Variationskoeffizient weisen neben den obgenannten sechs gemeinsamen Eigenschaften auch Unterschiede auf. Bei den Datenkonstellationen innerhalb der drei betrachteten Anwendungsgebiete führt dies zu Inversionen der Reihenfolge nach dem Gini-Ungleichheitsmass und dem Variationskoeffizienten.

Wolfgang Eichhorn
Institut für Wirtschaftstheorie
und Operations Research
Universität Karlsruhe
D – 7500 Karlsruhe 1

Arthur Vogt
Bundesamt für
Privatversicherungswesen
3003 Bern

Literatur

Aczél, J./Daróczy, Z. (1975): On Measures of Information and Their Characterizations. Academic Press, New York.

Blackwell, D. (1953): Equivalent Comparisons of Experiments. Ann. Math. Stat., 24.

Blackwell, D./Girshick, M.A. (1954): Theory of Games and Statistical Decisions. Wiley, New York.

Bourgin, R.D. (1983): Geometric Aspects of Convex Sets with the Radon-Nikodym Property. Springer Lecture Notes in Mathematics, Band 993, Springer-Verlag, Berlin.

Bühlmann, H. (1964): Die "Geburtsstunde" der mathematischen Statistik, Antrittsrede, gehalten am 11. Juli 1964 an der Universität Zürich, Vierteljahresschrift der Naturforschenden Gesellschaft Zürich.

Dagum, C. (1983a): Income Inequality Measures. In: Encyclopaedia of Statistical Sciences.

Dagum, C. (1983b): Lorenz Curve. In: Encyclopaedia of the Statistical Sciences.

Dagum, C. (1987): Gini Ratio. In: The new Palgrave.

Eichhorn, W. (1976): Fisher's Tests Revisited, Econometrica *44*, 247–256.

Eichhorn, W./Gehrig, W. (1982): Measurement of Inequality in Economics. In Modern Applied Mathematics – Optimization and Operation Research. Herausgegeben von Korte B. North-Holland Publishing Company, Amsterdam.

Eichhorn, W./Voeller, J. (1976): Theory of the Price Index. Lecture Notes in Economics and Mathematical Systems, Band 140, Springer Verlag, Berlin.

Encyclopaedia of Statistical Sciences (1983): John Wiley and Sons, New York.

Gerber, H.U. (1986): Lebensversicherungsmathematik. Springer-Verlag, Berlin.

Koestler, A. (1967): The Ghost in the Machine. Picador edition published 1975 by Pan Books Ltd, London.

Kunz, P. (1988): Das AHV-pflichtige Einkommen aus statistischer Sicht. Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker, 149–162.

Lüthi, A. (1981): Messung wirtschaftlicher Ungleichheit. Lecture Notes in Economics and Mathematical Systems, Band 189, Springer-Verlag, Berlin.

Marshall, A.W./Olkin, I. (1979): Inequalities: Theory of Majorization and Its Applications. Academic Press, New York.

Mosler, K.C. (1982): Entscheidungsregeln bei Risiko: multivariate stochastische Dominanz. Springer Lecture Notes in Economics and Mathematical Systems, Band 204, Springer-Verlag, Berlin.

Nermuth, M. (1982): Information Structures in Economics, Springer Lecture Notes in Economics and Mathematical Systems, Band 196, Springer Verlag, Berlin.

Nermuth, M. (1988): Verschiedene ökonomische Theorien mit gleicher formaler Struktur: Risikomessung, Einkommensungleichheit, Informationsstrukturen usw. Vortragsunterlage.

Piesch, W. (1975): Statistische Konzentrationsmasse. J.C.B. Mohr, Tübingen.

The New Palgrave (1987): Dictionary of Economics. Macmillan Press, London; Stockton Press, New York.

Vogt, A. (1979): Das statistische Indexproblem im Zwei-Situationen- Fall. Dissertation Nr. 6448, erstellt an der ETH Zürich unter der Leitung von Prof. Bühlmann (Referent) und Prof. Eichhorn (Korreferent). Kann beim zweiten Autor bezogen werden.

Vogt, A. (1986): Eine Formel zur Ursachenanalyse der Kostensteigerung im Gesundheitswesen. Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker, 95–105.

Anhang: *Numerische Illustrationen*

a) Überlebensordnungen

Rentnertafeln Einzelversicherung

	1990	1980	1970	1953
Männer				
Gini-Mass G	0,1458	0,1476	0,1556	0,1621
Variationskœffizient VK	0,1874	0,1890	0,2072	0,2188
Lebenserwartung e_0	81,281	78,554	76,130	72,952
Frauen				
Gini-Mass G	0,1361	0,1418	0,1503	0,1461
Variationskœffizient VK	0,1502	0,1643	0,1887	0,1784
Lebenserwartung e_0	86,608	82,907	80,044	77,321

Gruppenversicherungstafeln

	GKM/F80	GRM/F80	GKM/F70	GRM/F70
Männer				
Gini-Mass G	0,1627	0,1582	0,1663	0,1619
Variationskoeffizient VK	0,2218	0,2143	0,2293	0,2187
Lebenserwartung e_0	71,611	76,053	70,614	74,046
Frauen				
Gini-Mass G	0,1573	0,1481	0,1622	0,1536
Variationskoeffizient VK	0,2163	0,1865	0,2250	0,1985
Lebenserwartung e_0	78,141	81,078	75,260	77,911

b) Krankenversicherung: Pflegekosten pro Erkrankten

	1988	1978
Männer		
Gini-Mass *G*	0,679	0,656
Variationskoeffizient *VK*	2,27	2,28
Anzahl Erkrankte	110 291	90 084
Frauen		
Gini-Mass *G*	0,649	0,634
Variationskoeffizient *VK*	1,84	1,83
Anzahl Erkrankte	147 353	121 184
Insgesamt		
Gini-Mass *G*	0,663	0,646
Variationskoeffizient *VK*	2,01	2,00
Anzahl Erkrankte	257 644	211 268

c) AHV-Einkommen nach Kunz (1988)

	1985	1968	1953
Gini-Mass *G*	0,38	0,45	0,46
Variationskoeffizient *VK*	0,99	1,36	1,08
Anzahl Beitragspflichtige in Mio	3,62	3,12	2,43

Zusammenfassung

Die Arbeit befasst sich zunächst theoretisch mit dem Vergleichen und Messen im Kontext der Begriffe Ungleichheit, Streuung, Risiko und Information. Es zeigt sich, dass die dabei auftretenden Strukturen nahezu identisch sind. Anschliessend werden das Gini-Ungleichheitsmass sowie der Variationskoeffizient in den drei Gebieten Sterbetafeln, AHV-Einkommensverteilung und Kosten pro Erkrankten in der Krankenversicherung praktisch angewandt. Es zeigt sich, dass die Ungleichheiten in dieser Reihenfolge zunehmen. Das heisst, dass die Lebensdauern der Individuen gleichmässiger als die Einkommen und diese gleichmässiger als die Krankheitskosten verteilt sind.

Résumé

Le travail aborde en premier lieu et du point de vue théorique les opérations de mesure et de comparaison appliquées aux notions d'inégalité, de dispersion, de risque et d'information. Il s'avère que les structures en question sont pratiquement identiques. Les auteurs appliquent par la suite la mesure d'inégalité de Gini ainsi que le coefficient de variation aux trois domaines suivants relevant de la pratique: les tables de mortalité, la répartition des revenus dans l'AVS et les frais de guérison en assurance-maladie. Le résultat est que l'inégalité augmente en considérant les trois domaines dans l'ordre indiqué, soit que la distribution des durées de vie des individus est moins prononcée que celle des revenus et que cette dernière est également moins prononcée que celle des coûts des maladies.

Summary

First the notions of inequality, dispersion, risk and information are compared with one another and their qualitative assessment is discussed. This reveals that the inherent structures are almost identical. Then the so-called Gini-measure of inequality and the coefficient of variation are applied to mortality tables, income distributions from the Swiss Federal Social Security Scheme and medical costs per sick person in health insurance. It turns out that inequality increases in this order, i.e., that individuals' lengths of life are more regularly distributed than incomes and these in turn more regularly than medical costs.