# THEORETISCHE UND METHODISCHE ÜBERLEGUNGEN ZUR MESSUNG UND DARSTELLUNG VON EINKOMMENS-VERHÄLTNISSEN

Forschungsbericht Nr. 2

Bergbauerninstitut
des BM.f.L.u.F.
Grinzinger Allee 74
1196 Wien
Tel. (0222) 32 57 42

# THEORETISCHE UND METHODISCHE ÜBERLEGUNGEN ZUR MESSUNG UND DARSTELLUNG VON EINKOMMENSVERHÄLTNISSEN

Forschungsbericht Nr. 2

Rudolf NIESSLER

Wien, November 1980

INHALTSVERZEICHNIS

# I. Bemerkungen zur allgemeinen Theorie der Einkommensverteilung

## 1. Die historische Entwicklung der Theorie der Einkommensverteilung

In der Geschichte der Nationalökonomie hat die Theorie der Einkommensverteilung eine relativ junge Geschichte. Die Notwendigkeit, sich nicht nur mit der Entstehung und Verwendung des Einkommens, sondern auch mit dessen Verteilung auseinanderzusetzen, wurde zunächst von den Klassikern erkannt. Adam Smith widmet sich in seinem ersten Buch "The Wealth of Nations" den "Ursachen der Vervollkommnung der Produktivkräfte der Arbeit und der Ordnung, nach welcher ihr Produkt sich naturgemäß unter die verschiedenen Volksklassen verteilt".[1] Smith zielt also hiermit auf den Vergleich der Einkommenssituation zwischen verschiedenen Gruppen der Gesellschaft, wie Unternehmer und Arbeiter und Bauern ab. In der Tradition der Klassiker vollzieht Marx den Übergang zum Zweiklassenmodell, in dem er die Verteilung des Gesamteinkommens in Lohn und Profit als Bestimmungsgrund der Nachfrage ansieht. Die Frage, welche Einkommen die <u>einzelnen</u> Wirtschaftssubjekte, oder Haushalte beziehen, wird bis Endes des letzen Jahrhunderts nicht untersucht. <u>Ansätze</u> zu einer eigenständigen Theorie der personellen Einkommensverteilung finden sich bei Pareto, der 1895 eine Untersuchung der Steuerstatistiken veröffentlichte.[2] Da die Frage der Bedürfnisgerechtigkeit in der ökonomischen Theorie im Hintergrund blieb gelang  es der

---

1) A. Smith: Eine Untersuchung über Natur und Wesen des Volkswohlstandes, 3.Auflage, S. XVII,Jena 1923.

2) V. Pareto: La Courbe de la Répartition de la Richesse. Lausanne 1896

Theorie der personellen Einkommensverteilung nur äußerst langsam sich in der wissenschaftlich-ökonomischen Theorie zu etablieren. Die Entwicklung des Wirtschaftswachstums und seine Theorie waren sehr eng verknüpft mit der funktionellen Verteilungstheorie. Die Erkenntnis, daß wirtschaftliches Wachstum in den Industrieländern nicht mehr im bisherigen Ausmaß möglich erscheint, hat die Wachstumstheorie, die dadurch selbst in eine Krise geraten war, wieder mehr in den Hintergrund treten lassen. Damit verlor jedoch die makro-ökonomische Theorie der funktionalen Verteilung als wichtige Stütze der Wachstumstheorie an Bedeutung. Gleichzeitig zeichnet sich bereits ab, daß mit nachlassendem Wirtschaftswachstum die Auseinandersetzung um die Verteilungsgerechtigkeit ein zentrales Problem in den Sozialwissenschaften wird.[1] Mit dem Problem der Verteilungsgerechtigkeit kommt aber der Frage nach den Bestimmungsgründen der personellen Einkommensverteilung wieder größere Bedeutung zu. Allerdings besteht allgemein Übereinstimmung, daß der Stand der Theorie der personellen Einkommensverteilung noch sehr unbefriedigend ist.

## 2. Die Bedeutung der Theorie der Einkommensverteilung

Obwohl die Bedeutung der Verteilungstheorie immer wieder betont wurde, gibt es in der Tat einige Gründe, die ihr Vorrücken in den Mittelpunkt wissenschaftlichen Interesses verhindert haben. In der Grenzproduktivitätstheorie wird die Verteilung nur als Teilaspekt der Preistheorie und die funktionelle Einkommensverteilung als leistungsgerecht angesehen. Die Wohlfahrtstheorie hingegen schließt die Behandlung der Einkommensverteilung unter dem Gesichtspunkt der Bedürfnisgerechtigkeit aus. Eine Erhöhung der

---

1) G. Bombach: Neue Dimensionen der Lehre der Einkommensverteilung, Basler Universitätsreden, Heft 66, S. 32 f, Basel 1972

Wohlfahrt würde im Sinne dieser Theorie beispielsweise dann erreicht, wenn bei Gleichstellung aller übrigen mindestens einer besser gestellt wird, egal ob dieser bedürftig ist oder nicht.

Von den liberalen Ökonomen wurde versucht die Einkommensverteilung zu tabuisieren.[1] Da die Einkommensverteilung folglich im Sinne der Neoliberalen nicht Ziel der Wirtschaftspolitik sein darf, ist die Erklärung ihres Zustandekommens nicht dringlich, ja sogar nicht erwünscht, da dadurch ihre Beeinflussung möglich würde. Die Bedeutung der Verteilungstheorie blieb solange im Hintergrund, wie in Anbetracht hoher Wachstumsraten die Einkommenssteigerung für alle wichtiger als die Umverteilung zugunsten weniger erscheint. Durch die Verlangsamung des Wachstumsprozesses erhält aber nun gesamtwirtschaftlich die Verteilungsproblematik einen zentralen Stellenwert.

Wie schon oben erwähnt ist die Verteilung hauptsächlich in zweierlei Hinsicht zu untersuchen. Erstens als Verteilung der Einkommensanteile am Sozialprodukt. Arbeits- und Kapitaleinkommen stellen die beiden Einkommensaggregate dar. Der Einfachheit halber wird der Faktor Boden von den Klassikern vernachlässigt. Zu den Arbeitseinkommen zählen Löhne, Gehälter und der Unternehmerlohn. Kapitaleinkommen ist die Summe aus Pacht und Zinsen. Diese Verteilungstheorie bezeichnet man als Theorie der "funktionalen Verteilung" im weiteren Sinne.[2] Mit dem Durchbruch der Keynesianischen Theorie wurde auf die Bestimmungsgründe der Nachfrage, das heißt Konsum- und Investitionsverhalten besonderer Wert gelegt. Das Verhalten

1) F. Hayek: "Alle Bestrebungen eine 'gerechte' Verteilung sicherzustellen, müssen darum darauf gerichtet sein, die spontane Ordnung des Marktes umzuwandeln .....in eine totalitäre Ordnung". in: Freiburger Studien, S. 119, Tübingen 1969.

2) Blümle: Theorie der Einkommensverteilung S. 11ff

der Wirtschaftssubjekte zur Bestimmung der effektiven Nachfrage war das entscheidende Kriterium für die Sichtweise der Einkommensverteilung. Man bezeichnet dies auch als Verteilung auf sozio-ökonomische Gruppen. Da sich im wesentlichen die Funktion der Einkommen einer bestimmten Empfängergruppe zuordnen läßt, wird der Begriff der funktionalen Verteilung oft unkorrekter Weise synonym für die Gruppenverteilung verwendet.

Während jedoch Untersuchungen der Verteilung auf soziale Gruppen meist auf die Anteile am Sozialprodukt ausgerichtet sind, hat die Theorie der "personellen Einkommensverteilung" die Einkommenshöhe sowie die Streuung der Einkommen nach ihrer Höhe zum Gegenstand.[1] Die Erklärung der Häufigkeitsverteilung der Einkommen unterschiedlicher Höhe ist der ältere Ansatz und wird auch schlechthin als Theorie der personellen Einkommensverteilung aufgefaßt. Die Untersuchung der Verteilung auf Wirtschaftsbereiche, beispielsweise Industrie, Handwerk und Landwirtschaft, wobei also nach Art der Güter oder der Produktion klassifiziert wird, stellt zweifelsohne ein wichtiges Anliegen dar. Werden die Einkommen nach dem Wohnsitz oder dem Beschäftigungsort der Einkommensempfänger aggregiert, so liegt eine Fragestellung im Hinblick auf die regionale Einkommensverteilung vor.

## 3. Kriterien zur Beurteilung von Einkommensverteilungen

Als grundlegendes Kriterium zur Beurteilung von Einkommensverteilungen muß die Frage der Verteilungsgerechtigkeit näher spezifisiert werden. Wenn man nicht von vornherein normative Ökonomie betreiben will, muß man sich im klaren darüber sein, daß die Grenzen zwischen Erklärung und Rechtfertigung bzw. Verurteilung fließend sind. Außerdem

---

1) Blümle: Theorie der Einkommensverteilung, S. 13

ist es nicht ausreichend, Unterschiede am Einkommen durch rein ökonomische Sachverhalte zu erklären (z.B. Grenzproduktivitätstheorie). Bei einer starken Ausweitung des Datenkranzes wird allerdings die Zahl der möglichen Ursachen zu groß, sodaß oft, um die Vielfalt der Ursachen zu komprimieren, kaum mehr zu definierende Begriffe wie die "gesellschaftlichen Verhältnisse" als Erklärung von Einkommensunterschieden dargestellt werden. Dem Wissenschafter bleibt als Ausweg die empirische Forschung, die ihm durch Messen von Zusammenhängen wichtige Hilfestellungen zur Entwicklung einer Theorie der Einkommensverteilung bietet. Es obliegt allerdings den Sozialwissenschaften nicht nur Erklärungen anzubieten, sondern die Theorie bzw. die Verteilungen selbst kritisch an gesellschaftlichen Zielvorstellungen zu prüfen. Unter dem Aspekt der Verteilungsgerechtigkeit können wichtige normative Beurteilungskriterien wie Leistungsgerechtigkeit und Bedürfnisgerechtigkeit nicht umgangen werden.[1)]

Der Vergleich der Einkommen, das heißt von Verfügungsgewalt über ökonomische Mittel zum Zwecke der Bedürfnisbefriedigung, ist deswegen relevant, da kardinale oder auch nur ordinale Nutzenmessungen nicht möglich sind. Daraus folgt, daß die grundsätzliche Fähigkeit des Einkommens, Nutzen zu stiften, bei den zu vergleichenden Individuen gleich ist. Weiters sollte angenommen werden, daß jeder zum Vergleich herangezogene Einkommensbezieher im wesentlichen dem gleichen Güterangebot gegenübersteht und daß die Güterpreise für alle Individuen dieselben sind. Problematisch ist die Tatsache, daß für verschiedene Individuen unterschiedliche Güterpreise bestehen können. Für die Landwirtschaft gelten beispielsweise geringere Preise für selbsterzeugte bzw. unverarbeitete Agrarprodukte und für Wohnraum als für sonstige Berufsgruppen. Allerdings können Industrieprodukte für die Landwirte teurer sein als

---

1)Blümle: Ob die Theorie durch diese Umgehung besser wird, bleibt fraglich, denn ohne eine Präzisierung der Norm Chancengleichheit bleibt eine solche Analyse fragwürdig, kann also auf den normativen Aspekt gar nicht verzichten. S.17

für andere Berufsgruppen. Ein Vergleich von Preisindizes verschiedener Berufsgruppen wurde bislang noch nicht abgeschätzt. Prinzipiell ist aber, da zumindest keiner Berufsgruppe mangelnde Leistungsbereitschaft unterstellt werden kann, die Bedürfnisgerechtigkeit das entscheidende Kriterium zur Beurteilung der Einkommensverteilung. Das heißt in diesem Sinne müßte man eine Gleichverteilung der personellen Einkommen als normativen Idealtypus der Verteilung ansehen. Daß dies im Sinne der Leistungsgerechtigkeit natürlich problematisch ist, ist offensichtlich. Doch ist es in der Einkommenstheorie allgemein üblich, die soziale- bzw. Einkommenssituation je nach Höhe der Abweichung von der Gleichverteilung entsprechend zu beurteilen.

Als Ursache der Einkommensdifferenzierung in der Landwirtschaft werden oft "objektive und subjektive Gründe" genannt. Unter objektiven Ursachen versteht man natürliche Produktionsverhältnisse, Betriebsgröße, Marktlage und Einflüsse der Wirtschaftspolitik. Als subjektive Ursache gilt der Einfluß der Betriebsleiterfähigkeit.

Da die Betriebsleiterfähigkeit eine Folge der Ausbildung und anderer objektiver Einflüsse ist, kann dieses scheinbar subjektive Kriterium auf objektive Ursachen zurückgeführt werden[1]. Von einigen Agrarökonomen[2] wird die Betriebsleiterfähigkeit als die bedeutendste einkommensdifferenzierende Ursache erachtet. Dies ist aber im sozialwissenschaftlichen Sinne nicht haltbar, da die Betriebsleiterfähigkeit weitestgehend von objektiven Faktoren determiniert

1) Einkommensdifferenzen zwischen verschiedenen Gruppen der Landwirte (z.B. Bergbauern - Ackerbauern) durch unterschiedliche Betriebsleiterbefähigungen zu erklären erscheint uns als eine unzulässige Rechtfertigung sozialer Ungleichheit.

2) Siehe v.a. W. Ort: Die Ursachen der Einkommensunterschiede in landwirtschaftlichen Betrieben und ihre Quantifizierung, Volkswirtschaftliche Schriften 164, Dunckler & Humblot, Berlin 1971.

wird. Die Beweiskraft der empirischen Forschung bei der Frage nach den Ursachen von Einkommensungleichheit wird oft überschätzt.

Viele Einflußvariable sind für die Fragestellung ungeeignet definiert, oder oft sogar nicht einmal meßbar. Über die subjektiven Einflußfaktoren auf das Einkommen lassen sich durch die empirische Analyse nur wenig verlässliche Aussagen gewinnen.

# II. Beschreibung und Darstellung von Einkommensverteilungen

## 1. Allgemeine Darstellung von Verteilungen

### a) Tabellen

Die am häufigsten verwendete Darstellung von Verteilungen ist die Tabellenform. Die Merkmalsausprägungen eines stetigen quantitativen Merkmals können nicht unmittelbar eine Klasseneinteilung erzeugen. Vielmehr hat man eine künstliche Klasseneinteilung zu schaffen, indem man auf der Zahlengerade geeignete Intervalle bildet, in die man alle vorhandenen Merkmalsausprägungen einordnet. Das Intervall, indem sich alle Merkmalsausprägungen befinden, wird in möglichst gleich große Teilintervalle untergliedert. Für die Anzahl der Intervalle ist die Überlegung maßgeblich, möglichst wenig Information zu verlieren (durch zu wenige Klassen) und das Bild der Verteilung übersichtlich zu gestalten (d.h. nicht zu viele Intervalle). Wenn z.B., wie dies für die Einkommensverteilungen zutrifft, die Merkmalsausprägungen sehr ungleich dicht liegen, kann es vorteilhaft sein, ungleiche Klassenbreiten zu wählen. Offene Klassen sollten nach Möglichkeit vermieden werden, da weitere Berechnungen dann auf Schwierigkeiten stoßen. Den extremen Bereichen einer Verteilung sollte, da diese vor allem in der Einkommensverteilung aus inhaltlichen Gründen besondere Beachtung verdienen, besondere Sorgfalt in der Darstellung entgegen gebracht werden.

Zur Darstellung der Klassen seien folgende Bezeichnungen eingeführt:

| | |
|---|---|
| $k$... | Anzahl der Klassen für $i=1,\dots,k$ |
| $e_{i-1}$... | Klassenuntergrenze der i-ten Klasse |
| $e_i$... | Klassenobergrenze der i-ten Klasse |
| $d_i = e_i - e_{i-1}$... | Klassenbreite |
| $x_i = \frac{1}{2}(e_i + e_{i-1})$ | Klassenmitte |

$f_i$...absolute Häufigkeit einer Merkmalsausprägung
$p_i$...relative Häufigkeit einer Merkmalsausprägung

b) Graphische Darstellung von Verteilungen

Die gebräuchlichste Art der graphischen Darstellung von Verteilungen ist sowohl bei diskreten als auch bei stetigen Merkmalen das Histogramm. Es zeigt die Häufigkeiten $f_i$ bzw. $p_i$ in einem halboffenen Intervall $[a,b)$ an. Darstellungsmittel ist die Fläche eines Rechteckes, das mit diesem Intervall als Basis gekennzeichnet wird. Bei stetigen Merkmalen werden die durch die Klasseneinteilung geschaffenen Intervalle verwendet.

Eine weitere oft verwendete Art der Darstellung ist das Häufigkeitspolygon. Dieses wird direkt aus dem Histogramm abgeleitet, indem man die Mitten der oberen Rechtecksbegrenzungen miteinander verbindet.

Die folgende Darstellung eines Histogramms und des dazugehörigen Häufigkeitspolygons aus einer Häufigkeitstabelle möge das soeben Erläuterte veranschaulichen.

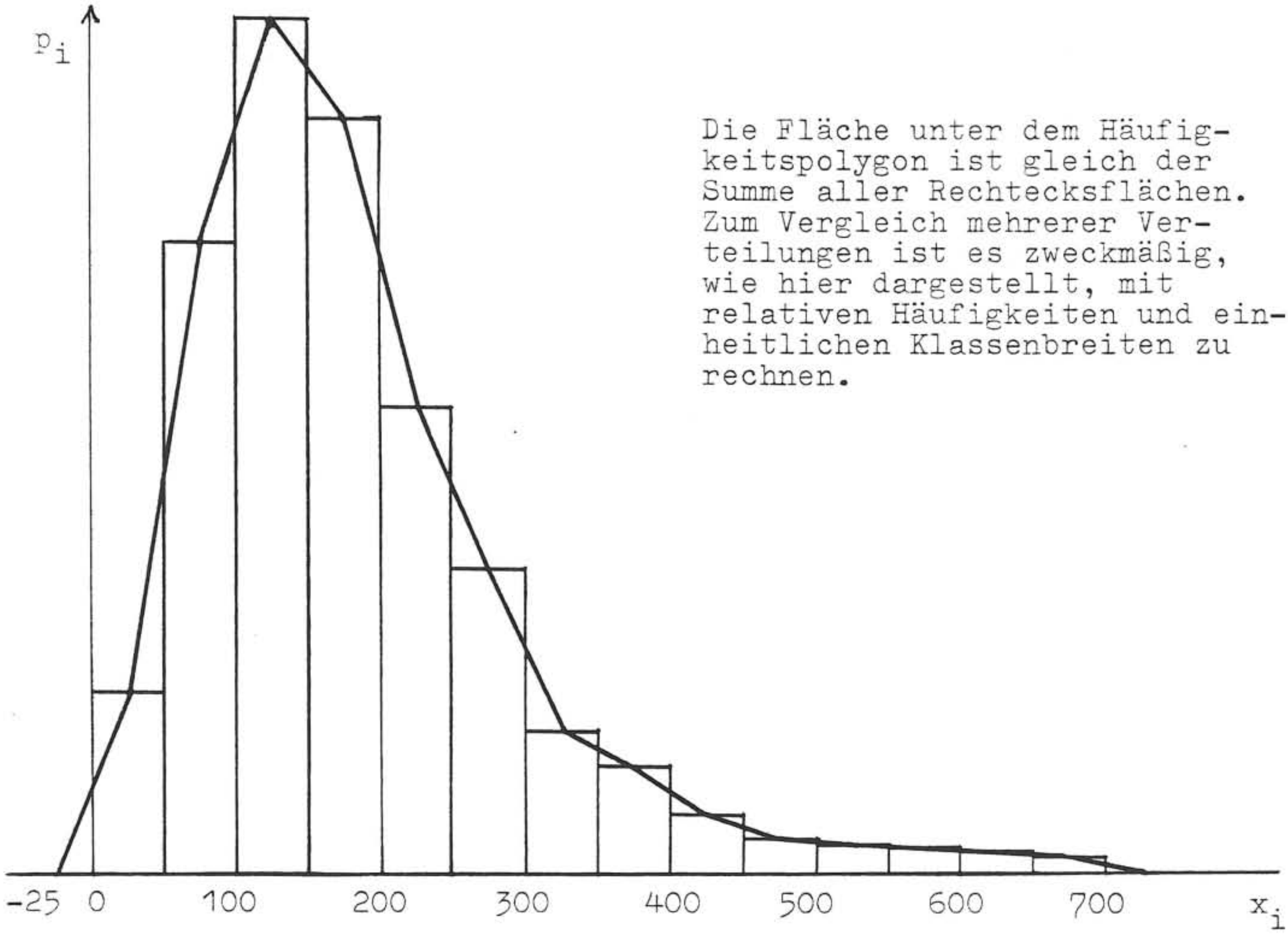
Tabelle 1: <u>Die Streuung der Betriebe nach dem Gesamteinkommen je Betrieb im Bundesmittel 1976</u> (Zum Zwecke der anschaulichen Darstellung vereinfacht)

| Einkommen-Klassengrenzen $e_{i-1}$ | $e_i$ | $d_i$ | $p_i$ | $x_i$ | $F_i$ |
|---|---|---|---|---|---|
| 0*) | 50 | 50*) | 4,89 | 25 | 4,89 |
| 50 | 100 | 50 | 17,38 | 75 | 22,27 |
| 100 | 150 | 50 | 23,64 | 125 | 45,91 |
| 150 | 200 | 50 | 20,90 | 175 | 66,81 |
| 200 | 250 | 50 | 12,88 | 225 | 79,69 |
| 250 | 300 | 50 | 8,33 | 275 | 88,02 |
| 300 | 350 | 50 | 3,92 | 325 | 91,94 |
| 350 | 400 | 50 | 2,91 | 375 | 94,85 |
| 400 | 450 | 50 | 1,65 | 425 | 96,50 |
| 450 | 500 | 50 | 0,96 | 475 | 97,46 |
| 500 | 550 | 50*) | 0,84*) | 525 | 98,30 |
| 550 | 600 | 50*) | 0,70*) | 575 | 99,00 |
| 600 | 650 | 50*) | 0,60*) | 625 | 99,60 |
| 650 | 700 | 50*) | 0,40*) | 675 | 100,00 |

Alle Häufigkeiten in Prozent, Einkommen in tausend

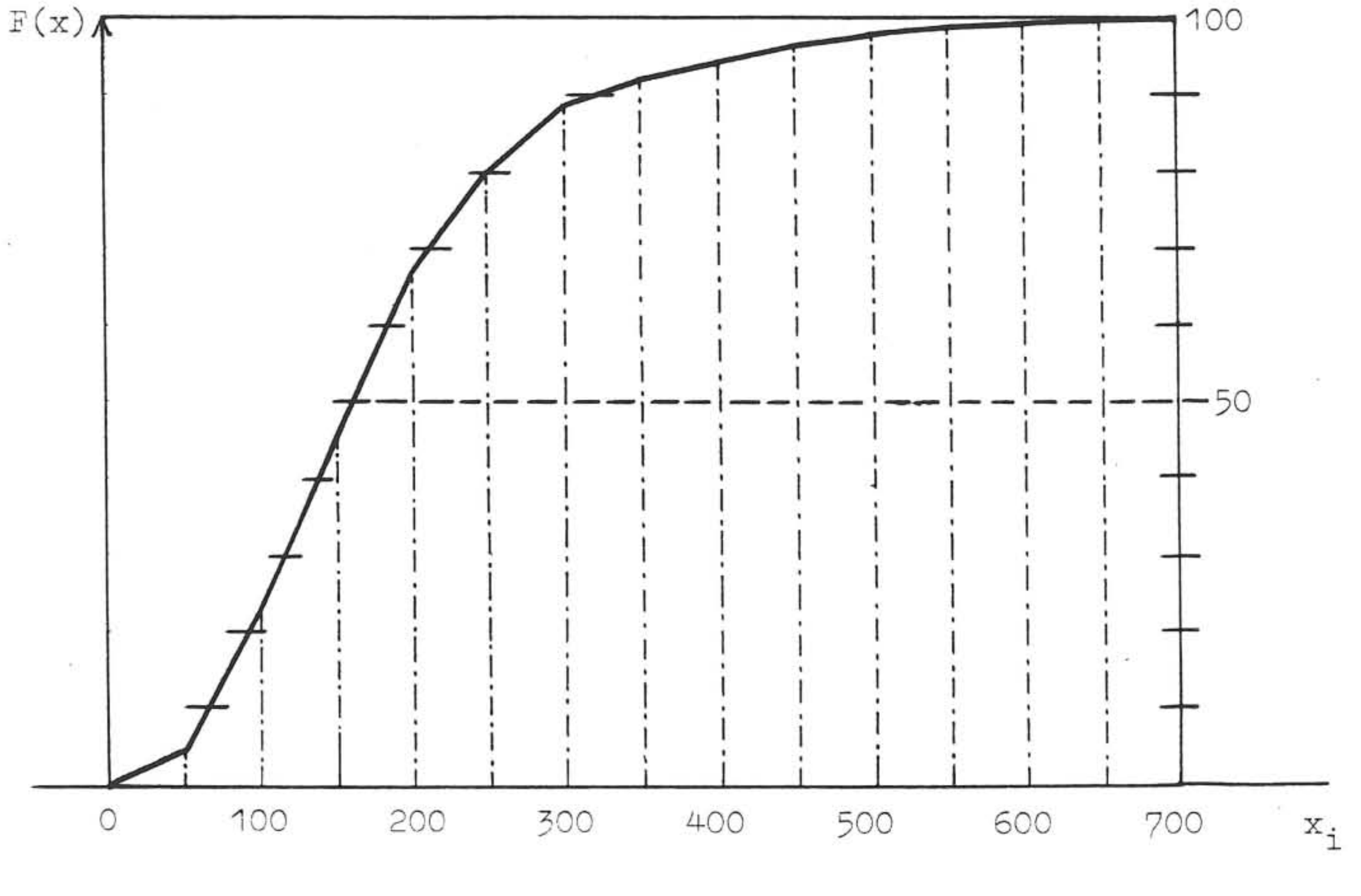
*) Nur der einfachen Darstellung wegen wurde mit gleichen Klassenbreiten gerechnet. Die Häufigkeitsverteilung im Bereich größer 500 wurde angenommen. Bei einer exakten Darstellung ist diese Vorgangsweise <u>nicht zulässig</u>.

Histogramm und Häufigkeitspolygon obiger Tabelle




Die Fläche unter dem Häufigkeitspolygon ist gleich der Summe aller Rechtecksflächen. Zum Vergleich mehrerer Verteilungen ist es zweckmäßig, wie hier dargestellt, mit relativen Häufigkeiten und einheitlichen Klassenbreiten zu rechnen.

Die Verteilungsfunktion F(x)



Eine weitere Art der Darstellung von Verteilungen stellt auf die kumulierten Häufigkeiten ab. Hierzu benötigt man die sogenannte (empirische) Verteilungsfunktion $F(x)$ des Merkmals x. $F(x)$ ist die relative Häufigkeit der Elemente der Grundgesamtheit, deren Merkmalsausprägungen kleiner oder gleich x sind. Die Summenkurve wird näherungsweise durch einen monoton steigenden Polygonzug gebildet. Der Streckenzug ergibt sich in Tabelle 1 aus den Abszissen $e_i$ und den Ordinaten $F_i$. Neben diesen grundsätzlichen Methoden der Darstellung von Verteilungen ist es nun zweckmäßig,Einkommensverteilungen nach verschiedenen Kriterien zu untersuchen. Die Lage der Verteilung sowie deren Streuung, Schiefe, Steilheit, Konzentration etc. seien an dieser Stelle nur kurz erwähnt. All die in der Folge dargestellten Maßzahlen und Graphiken dienen dazu, unsere Information über die Verteilung in verschiedener Hinsicht zu verdichten.

## 2. Verteilungsmaßzahlen

Verteilungsmaßzahlen oder Verteilungsparameter sollen dazu dienen, gewisse Eigenschaften von Verteilungen mit quantitativen Merkmalen zu erfassen. An Präzision vermögen sie das anschauliche Mittel der graphischen Darstellung meist zu übertreffen. Die Statistik braucht sich nicht damit zu begnügen, Mittelwerte aus den verschiedenen empirischen Phänomenen zu berechnen, sie ist durchaus in der Lage, auch andere Fragestellungen quantitativ, d.h. mit Maßzahlen zu erfassen. Alle Maßzahlen sind in ihrer einfachsten Form dargestellt; für die tatsächlichen Berechnungen müssen je nach Datengrundlage entsprechende Varianten verwendet werden.

### a) Lageparameter

Auf diese Klasse der Verteilungsmaßzahlen sei wegen ihrer Einfachheit und allgemeinen Gebräuchlichkeit nur knapp eingegangen. Bei Anwendung der Maßzahlen für gruppierte

Daten wird Gleichverteilung innerhalb der Klassen angenommen.

| | |
|---|---|
| $x_i$ | bezeichne in der Folge Merkmalsausprägungen bei diskreten Merkmalen und Klassenmitten bei gruppierten Daten (stetiges oder diskretes Merkmal) |
| $N$ | Anzahl der Messungen |
| $\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$ | Arithmetisches Mittel einer Meßreihe |

Das arithmetische Mittel wird schlechthin oft als "der Mittelwert" bezeichnet. Für praktische Berechnungen bei der Einkommensanalyse verwendet man die Formel für gruppierte Daten und relative Häufigkeiten.

$$\bar{x} = \sum_{i=1}^{k} p_i x_i$$

$x_i$... Klassenmitte

k... Anzahl der Klassen

Sind offene Klassen vorhanden, so sollten zumindest Angaben über die Merkmalssummen in diesen Klassen vorhanden sein. Willkürliche Annahmen über die Klassenbreite können zu groben Verzerrungen der Maßzahlen führen.
Das geometrische Mittel wird oft in höheren Verteilungsmaßzahlen in Beziehung zum arithmetischen Mittel gesetzt. Es ist folgendermaßen definiert:

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} = \sqrt[n]{\prod_{i=1}^{n} x_i}$$ Geometrisches Mittel

Es gilt generell, daß das geometrische Mittel kleiner ausfällt als das arithmetische Mittel.

Für die Beschreibung von Einkommensverteilungen wird oft noch das quadratische Mittel herangezogen. Es wird dieses aber speziell zur Konstruktion von Streuungsmaßzahlen verwendet.

$$Q = \sqrt{\frac{1}{n}\left(x_1^2 + x_2^2 + \ldots + x_n^2\right)} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2}$$ Quadratisches Mittel

Für die drei beschriebenen Mittelwerte gilt die Größenrelation:

$$Q \gg \overline{X} \gg G$$

Speziell für Zwecke der Beurteilung von Einkommensverteilungen ist der Median eine besonders aussagefähige Größe. Er basiert auf folgendem Konzept: Man ordnet die Merkmalsausprägungen $x_i$ der Größe nach. Innerhalb dieser geordneten Zahlenreihe suchen wir nun den Wert, der in der Mitte dieser Reihe liegt. Oberhalb und unterhalb sollen gleich viele Reihenwerte liegen. Der Median ist definiert:

für n ungerade: $\tilde{X}_{0,5} = X_{\left(\frac{n+1}{2}\right)}$

für n gerade: $\tilde{X}_{0,5} = \frac{1}{2}\left[X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right]$

Für gruppierte Daten errechnet man den Median aus der Verteilungsfunktion $F(x)$. Es sei $F(x)$ stetig und streng monoton steigend, mit Ausnahme der Bereiche in denen $F(x) = 0$ und $F(x) = 1$, dann gilt die Gleichung:[1]

$$F(\tilde{X}_{0,5}) = \frac{1}{2}$$

Als Näherungswert ergibt sich daraus:

$$\tilde{X}_{0,5} = e_{m-1} + \frac{d_m}{f_m}\left[\frac{N}{2} - F_{m-1}\right]$$

wobei m...Nummer der Medianklasse für die gilt:

$$F_{m-1} \leq \frac{N}{2} \leq F_m$$

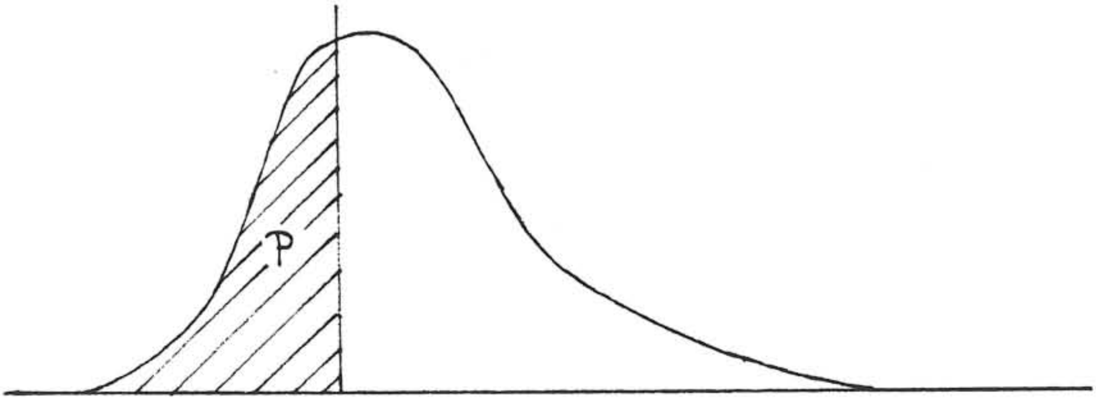
Die Brauchbarkeit des Medians ist vielfältig. Er ist unempfindlich gegen "Ausreißer" und er kann auch für Verteilungen mit offenen Klassen berechnet werden. Im direkten

---

1) M. Fisz: Wahrscheinlichkeitsrechnung und mathematische Statistik, Deutscher Verlag der Wissenschaften, Berlin 1978, S 101ff.

Vergleich zwischen Median und Mittelwert können Aussagen über die Gestalt der Verteilung getroffen werden.

Der Median kann als Spezialfall der p-Quantile interpretiert werden. Ein p-Quantil ist jene Zahl, welche den Bruchteil p einer Verteilung abtrennt. Der Median ist somit das p-Quantil mit p = 0,5. Der Einfachheit halber sei das p-Quantil an einer glatten Kurve dargestellt.




In Analogie zum Median ergibt sich daher für allgemeine p-Quantile:

$$\tilde{x}_p = e_{i(p)-1} + \frac{d_{i(p)}}{f_{i(p)}} \left[ N_p - F_{i(p)-1} \right] \text{ wobei}$$

$i(p)$ ...Nummer der Klasse für die gilt

$$F_{i(p)-1} \leq N_p \leq F_{i(p)}$$

In der Praxis rechnet man meist nicht mit allgemeinen p-Quantilen sondern mit Spezialfällen wie Quartilen und Dezilen. In der Formel sind dann für den p-Wert die entsprechenden Zahlenwerte (z.B. 0,25; 0,10; 0,90) einzusetzen.

Wegen der besonders häufigen Verwendung zur Beschreibung von Einkommensverteilungen sei noch etwas näher auf Quantile eingegangen. Das Kriterium Anschaulichkeit stellt eine wichtige Forderung an die Verteilungsmessung dar, wenn mit zunehmender Bedeutung der Verteilungsgerechtigkeit Zielsetzung und Zielerreichung im Rahmen der Wirtschaftspolitik begründet und kontrolliert werden soll.[1] Quantile entsprechen dieser Anforderung. Allerdings weisen sie den Nachteil auf, daß sie nicht alle Verteilungsänderungen berücksichtigen, wenn eine Beschränkung auf ein Quantil erfolgt und nur ein Teil der Einkommensempfänger erfaßt wird. Eine Charakterisierung der ganzen Verteilung durch einige Quantile bringt aber den Nachteil mit sich, daß Umverteilungen innerhalb der Quantile nicht zum Ausdruck kommen. Soll eine Einkommensverteilung durch eine oder wenige Maßzahlen gekennzeichnet werden, so sind Quantile ungeeignet, da sie die Information zugrundeliegender Statistiken nicht vollständig wiedergeben. Bei fundierten Einkommensanalysen, die nicht nur mit komprimierten Verteilungsmaßzahlen operieren, werden Quantile sehr häufig herangezogen.

Oft werden aus Quantilen weitere Lagemaße konstruiert, so z.B. das p-Quantilsmittel:

$$\frac{1}{2}\left(\tilde{x}_p + \tilde{x}_{1-p}\right)$$

Als Spezialfälle sind das Quartilsmittel (p = 0,25) und das Dezilmittel (p = 0,1) zu nennen.

Der Modalwert, auch Modus genannt, ist der häufigste Wert einer Verteilung. Es ist dies eine Merkmalsausprägung, bei der die Kurve der Häufigkeitsdichte ein lokales Maximum erreicht. Existiert nur ein lokales Maximum einer Dichtefunktion, so spricht man von einer unimodalen Verteilung, hat sie mehrere, von einer multimodalen. Für Einkommensverteilungen ist oft ein Vorhöcker im unteren Einkommensbereich der Dichtefunktion charakteristisch. Es handelt sich hierbei meist um Teilzeitbeschäftigte, und die

1) Blümle: Theorie der Einkommensverteilung, Heidelberger Taschenbuch 173, Berlin 1975, S 35

Verteilungsfunktion besitzt zwei lokale Maxima und ist daher von bimodalen Typus.

Für gruppierte Daten errechnet sich der Modus (h) näherungsweise aus dem Histogramm.

$$h \approx e_{\alpha - 1} + \frac{f_\alpha - f_{\alpha - 1}}{2f_\alpha - f_{\alpha - 1} + f_{\alpha + 1}}$$ wobei $\alpha$ ...Nummer der modalen Klasse

b) Einfache Streuungsmaßzahlen

Der Begriff der Streuung bezieht sich ganz allgemein darauf, wie eng oder wie weit die Merkmalsausprägungen eines quantitativen Merkmals einer Grundgesamtheit "zusammen" liegen. Lagemaße allein charakterisieren eine Verteilung nicht ausreichend. Die Streuung ist ein zentrales Phänomen der Statistik.

Streuungsmaße können in drei Klassen eingeteilt werden. Kriterium der Unterscheidung ist die Abstandsbildung:

i) Abstände zweier geeigneter Ranggrößen

ii) Abstände aller Merkmalsausprägungen voneinander

iii) Abstände der Merkmalsausprägungen von einem Lagemaß

Maße jeder Klasse finden in der Analyse von Einkommensverteilungen Verwendung.

Zu den Maßzahlen der ersten Klasse zählen jene Streuungsmaße, die von Quantilen abhängen.

Die Semi-Interquantilsdistanz: $\Delta_p = \frac{1}{2}(\tilde{x}_{1-p} - \tilde{x}_p)$ für $p < 0{,}5$

mittelt die Distanz der beiden äußeren Quantile.

Das Maß hat den Vorteil, daß es gegenüber Ausreißern unempfindlich ist. Das Gegenteil ist der Fall bei der Spannweite. Sie berücksichtigt nur die Extremwerte. Um sich eine Vorstellung von der Schwankungsbreite der Einkommen zu machen, wird sie häufig ausgewiesen.

$$R \ldots\ x_{(n)} - x_{(1)} \quad \text{Spannweite}$$

Ein für die Analyse von Einkommensverteilungen vielfach vorgeschlagenes Maß[1] der Klasse 2 ist der Gini-Koeffizient.[2] Für den diskreten Fall ist:

$$\Delta_R = \frac{\left(\frac{1}{2n^2}\right) \sum\limits^{i} \sum\limits^{j} |x_i - x_j|}{\frac{1}{n} \sum\limits^{i} x_i}$$

Er stellt wohl das bekannteste und in empirischen Untersuchungen häufigst verwendete Ungleichheitsmaß dar. Eine anschauliche Interpretation dieser Kennzahl zur Messung der relativen Konzentration liefert Pyatt.[3] Er stellt den Gini-Koeffizient als Ergebnis eines Spieles dar, indem sich sonst identische Spieler nur durch die Höhe ihres Einkommens x unterscheiden. Die Spielregeln werden folgendermaßen definiert:

- o Für jeden Spieler wird ein Einkommen $x'$ aus der Gesamtpopulation $x_1$, $x_2 \ldots$, $x_n$ zufällig ausgewählt.
- o Übersteigt das Einkommen $x'$ das Einkommen des Spielers, so erhält er die Differenz als Spielgewinn ausbezahlt.
- o Die Teilnahme an dem Spiel ist kostenlos.

---

1) G. Bruckmann: "Einige Bemerkungen zur statistischen Messung der Konzentration" in Metrika Jg. 14, 1969
2) Meist wird der Gini-Koeffizient als "Konzentrations" bzw. "Ungleichheitsmaß" betrachtet. Im Rahmen der von mir vorgenommenen Klassifikation der Streuungsmaße nach F.Ferschl kann er aber auch als Streuungsmaßzahl angesehen werden.
3) G. Pyatt: On the Interpretation and Disaggregation of Gini-Coefficients. Econ. Journal 86 (1976) S. 243-255

Alle Spieler mit Ausnahme des reichsten, können in diesem Spiel gewinnen - und selbst der Spieler mit dem höchsten Einkommen kann nichts verlieren. Der Gini-Koeffizient ist der Erwartungswert des durchschnittlichen Spielgewinnes bezogen auf das mittlere Einkommen in der Gesamtpopulation.

Da für die Analyse der Einkommensverteilung die Daten meist in gruppierter Form vorliegen, wird zur Berechnung des Gini-Koeffizienten die Lorenzkurve herangezogen. Bei der Lorenzkurve werden wir daher nochmals auf den Gini-Koeffizienten zurückkommen.

In Klasse 3 der Streuungsmaßzahlen fallen die meisten herkömmlichen Streuungsparameter. In Anbetracht der Fülle der konstruierbaren Parameter beschränken wir uns auf die Darstellung der wichtigsten in Einkommensanalysen verwendbaren Maße.

Die <u>Varianz</u> ist einer der wichtigsten Begriffe der Statistik. Sie ist definiert als

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Die Varianz hat die Dimension eines Quadrates und hat daher in der deskriptiven Statistik den Nachteil, wenig anschaulich zu sein. Erst in der induktiven Statistik entfaltet die Varianz ihre volle Bedeutung. In der Deskriptivstatistik wird daher meist die <u>Standardabweichung</u>, die Wurzel aus der Varianz herangezogen.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Sie ist eine eindimensionale Größe und somit in ihrer Interpretation recht anschaulich. Als Lagemaß, von dem die Abstände gemessen werden, wird hierbei das arithmetische Mittel verwendet. Wenn man nun Streuungsmaße in Beziehung zu Lagemaßen setzt, kommt man zur Klasse der Dispersionsmaße. Das wichtigste und in Einkommensanalysen häufig angewendete Dispersionsmaß ist der Variationskoeffizient. Er ist folgendermaßen definiert:

$$V = \frac{\sigma}{\bar{x}} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\bar{x}}$$

Der Variationskoeffizient hat gegenüber der Standardabweichung den Vorteil, daß er normiert ist. Bei einer Verdoppelung aller Einkommen würde beispielsweise die Standardabweichung auch doppelt so hoch wie früher ausfallen. Bei der Messung der Einkommensverteilung hat es sich jedoch eingebürgert, Verteilungen bei welchen die Einkommensverhältnisse übereinstimmen, als im gleichen Maße gleich bzw. ungleich anzusehen, was auch als "Gesetz von Bresciani-Turroni" bezeichnet wird.[1] Aus dieser Forderung folgt, daß die Einkommensniveaus bei der Verteilungsmessung nicht berücksichtigt werden dürfen. Soll ein Verteilungsmaß zur Wohlfahrtsmessung herangezogen werden, ist obige Forderung äußerst problematisch. In wachsenden Volkswirtschaften mit steigenden Pro-Kopf-Einkommen erhöhen sich, bei gleicher Verteilung der relativen Einkommen, die absoluten Einkommensunterschiede zusehends. Neuere Ansätze der Verteilungsmessung setzen sich daher speziell mit der Problematik der absoluten Einkommensunterschiede auseinander. Bei der Besprechung der neueren Ansätze zur Ungleichheitsmessung werden wir nochmals darauf zurückkommen. Für verschiedene Zwecke der Einkommensanalyse, z.B. zum intertemporalen oder interregionalen Vergleich, erscheint allerdings, da auf eine Umrechnung in vergleichbare Absolutwerte verzichtet werden kann, das Konzept

1) Blümle: Theorie der Einkommensverteilung, Heidelberger Taschenbuch 173, Berlin 1975, S. 38

der Messung relativer Einkommensunterschiede durchaus sinnvoll. Der Variationskoeffizient widerspricht aber, da er Einkommenstransfers unabhängig von der Einkommenshöhe in gleicher Weise berücksichtigt, jeglichen nutzentheoretischen Ansätzen einer "gerechten" Einkommensverteilung.

Der Vollständigkeit halber sei an dieser Stelle noch die Formel für die Varianz bei gruppierten Daten und relativen Häufigkeiten angegeben. Standardabweichung und Variationskoeffizient ergeben sich daraus unmittelbar.

$$\sigma_x^2 = \sum_{i=1}^{k} p_i (x_i - x)^2 \quad \text{wobei } x_i \ldots \text{Klassenmitten}$$

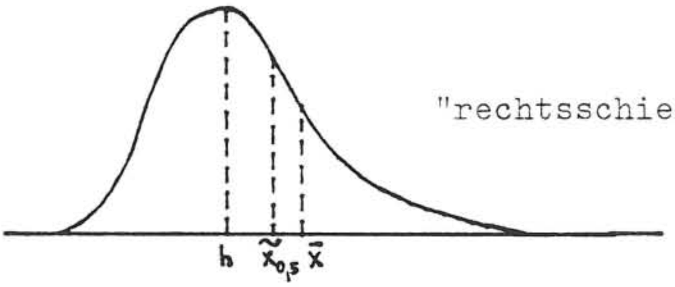
c) Höhere Verteilungsmaßzahlen

Für praktische Berechnungen werden höhere Verteilungsmaßzahlen eher selten verwendet. Der Vollständigkeit halber seien die wichtigsten kurz besprochen. Oft sind Verteilungen durch gleichzeitige Verwendung von Lage- und Streuungsparameter schon ausreichend charakterisiert. Allerdings können Verteilungen mit gleichem Mittelwert und gleicher Varianz sehr unterschiedliche Gestalt aufweisen. Es ist daher für viele Fragestellungen unentbehrlich, höhere Verteilungsmomente heranzuziehen. Für viele Verteilungen, insbesondere Einkommensverteilungen, ist eine starke Asymmetrie charakteristisch. Um diese Besonderheit zahlenmäßig abzubilden, werden Schiefemaße konstruiert. Grundsätzliche Aussagen über die Asymmetrie lassen sich aufgrund eines direkten Vergleiches der Lageparameter Modus, Median und arithmetisches Mittel treffen.

So gilt für rechtsschiefe Verteilungen folgende Relation:

$$h < \tilde{x}_{0,5} < \bar{x}$$

Dies entspricht einer typischen Dichtefunktion der Einkommen.

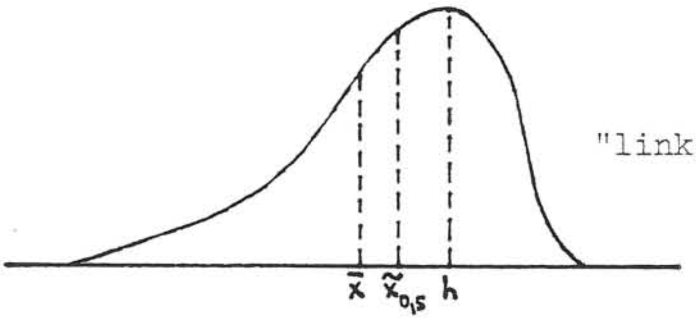



"rechtsschief"

$h$...Modalwert
$\tilde{x}_{0,5}$...Median
$\bar{x}$...arithmetisches Mittel

Für linksschiefe Verteilungen gilt: $h > \tilde{x}_{0,5} > \bar{x}$

Graphisch dargestellt zeigt die Kurve folgende Gestalt:




"linksschief"

Schiefemaße werden nun so konstruiert, daß sie für rechtsschiefe Verteilungen positive und für linksschiefe negative Maßzahlen erhalten.

Die beiden Pearson`schen Schiefekoeffizienten sind folgendermaßen definiert:

$$Sk_1 = \frac{\bar{x} - h}{\sigma} \qquad Sk_2 = \frac{3(\bar{x} - \tilde{x}_{0,5})}{\sigma}$$

$Sk_2$ wird verwendet wenn die Berechnung des Modalwertes Schwierigkeiten bereitet. Die Formel beruht auf der Tatsache, daß mit guter Näherung gilt:

$$\bar{x} - h \approx 3(\bar{x} - \tilde{x}_{0,5})$$

Der wegen des großen Rechenaufwandes meist nicht verwendete Momentenkoeffizient der Schiefe[1] ist definiert als

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \qquad \text{wobei} \qquad \mu_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$$

Um Informationen über die Gestalt der Verteilung zu gewinnen zieht man Kurtosismaße heran (Wölbungs- oder Steilheitsmaße). Man unterscheidet abgeplattete (platykurtische) und spitze (leptokurtische) Verteilungen. Das Kurtosismaß von R.A. Fischer charakterisiert eine Verteilung hinsichtlich ihrer Steilheit.

Momentenkoeffizient der Kurtosis:

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 \qquad \text{wobei} \qquad \mu_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4$$

Man bezeichnet Verteilungen mit:

$\gamma_2 > 0$ als leptokurtisch
$\gamma_2 = 0$ als mesokurtisch
$\gamma_2 < 0$ als platykurtisch

d) Die Messung der Konzentration

Die Streuungsanalyse untersucht, wie die Merkmalsausprägungen der Elemente der Grundgesamtheit zueinander liegen. Die Frage der Konzentration stellt auf die Gesamtsumme der Merkmalsausprägungen ab und fragt, wie sich die Summe auf die einzelnen Elemente aufteilt, ob sie sich bei wenigen Elementen konzentriert oder ob sie gleichmäßig auf alle Elemente aufgeteilt ist. Daß Konzentrationsmaße für die Analyse von Einkommensverteilungen in höchstem Maße relevant sind liegt auf

1) M. Fisz: Wahrscheinlichkeitsrechnung und mathematische Statistik, Deutscher Verlag der Wissenschaften, Berlin 1978, S. 97

der Hand.

Eine äußerst anschauliche Darstellung der Einkommens-konzentration ist die Lorenzkurve. Man ordnet die Einkommensempfänger nach der Größe der Einkommen und stellt die kumulierten Anteile der Einkommensempfänger den kumulierten Anteilen der Einkommen gegenüber. Formal wird die Lorenzkurve folgendermaßen konstruiert:

Es sei $x_{(1)}, x_{(2)} \dots x_{(n)}$ eine geordnete Reihe nicht-negativer Zahlen[1]

Die kumulierten Anteile der Elemente sind dann:

$$u_i = \frac{i}{n} \quad \text{wobei } u_0 = 0$$

Die kumulierten Anteile der Merkmalssummen er-rechnen sich aus:

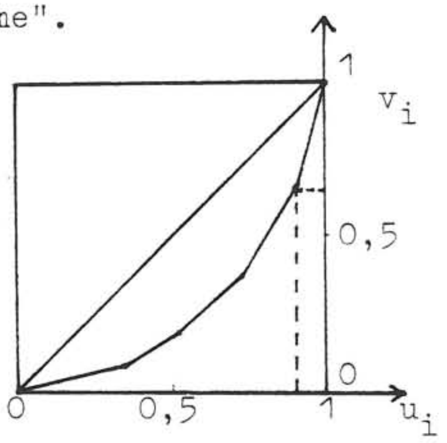
$$v_i = \frac{x_{(1)} + x_{(2)} + \dots + x_{(i-1)} + x_{(i)}}{\sum_{k=1}^{n} x_k} = \frac{\sum_{j=1}^{i} x_{(j)}}{n\bar{x}} \quad \text{für } i = 1, \dots n$$

Die Lorenzkurve bildet nun den Streckenzug, der die Punkte mit den Koordinaten $(u_i, v_i)$ $i = 0,1 \dots, n$ miteinander verbindet.
Für gruppierte Daten gilt:

$$u_i = \sum_{j=1}^{i} p_i \quad \text{und} \quad v_i = \frac{\sum_{j=1}^{i} p_i x_j}{\sum_{i=1}^{n} p_i x_i} \qquad \text{wobei } u_0 = 0,\ v_0 = 0.$$

1) Auf die spezielle Problematik, die durch das Auftreten negativer Einkommen in der Landwirtschaft hervorgerufen wird, wird in Kapitel 4 eingegangen.

Die Lorenzkurve liefert in sehr anschaulicher Weise Aussagen in der Form:
"Die 10 Prozent der obersten Einkommensbezieher vereinigen auf sich 30 Prozent der gesamten Einkommenssumme".




Aus der Lorenzkurve kann nun ein sehr wichtiges Konzentrationsmaß gewonnen werden. Wie schon in Kapitel b) erwähnt, kann der Gini-Koeffizient als das Verhältnis der Fläche zwischen der Lorenzkurve und der Diagonale des Einheitsquadrates zur Gesamtfläche unter der Diagonale gedeutet werden. Zur exakten Darstellung des Gini-Koeffizienten muß daher die nur in einigen Stützstellen bekannte Lorenzkurve interpoliert werden. Eine lineare Interpolation führt zu einer systematischen Unterschätzung des Gini-Koeffizienten. Mathematisch ist das Problem, eine stetig-konvexe Lorenzkurve zu erzeugen gelöst.[1] Im Rahmen unserer Arbeit konnten wir uns mit einer linearen Interpolation zufrieden geben, da eine relativ große Anzahl von Stützstellen vorhanden war, sodaß sich der Fehler in kleinen Grenzen hält. In der Literatur wird oft der Gini-Koeffizient, der sich nach linearer Interpolation der Lorenzkurve ergibt als Lorenz-Münzner'scher Konzentrationskoeffizient bezeichnet.

1) J.L. Gastwirth, M. Glauberman: The Interpolation of the Lorenz Curve and Gini Index from Grouped Data, Econometrica 46 (1976), S. 479ff.

Für gruppierte Daten ist das Konzentrationsmaß von Lorenz-Münzner definiert als

$$\mathcal{X} = 1 - 2 \sum_{i=1}^{k} p_i \bar{v}_i \quad \text{wobei} \quad \bar{v}_i = \frac{1}{2}(v_{i-1} + v_i) \text{ für } i = 1, \ldots k$$

Eine Bewertung der Lorenzkurve bzw. des Gini-Koeffizienten im Hinblick auf seine Sensibilität bezüglich Einkommenstransfers findet sich in Kapitel 3. Besondere Probleme für die Anwendung auf landwirtschaftliche Einkommensdaten werden in Kapitel 4 erörtert. Daß die Analyse der Einkommensverteilung besondere methodische Sorgfalt erfordert, darauf wurde in obiger Darstellung mehrfach hingewiesen. Dies folgt unmittelbar aus der zentralen sozio-ökonomischen Funktion des Einkommens als Verfügungsgewalt über ökonomische Mittel zum Zwecke der Bedürfnisbefriedigung. Der Gedanke der Bedürfnisgerechtigkeit fordert nun aber eine nutzentheoretische Bewertung der Verteilung. Unter diesem Blickwinkel werden vor allem die Extrembereiche der Verteilung interessant. In neuerer Zeit wurden daher spezielle Methoden entwickelt, um diesen Forderungen einer nutzentheoretischen Interpretation zu entsprechen.[1]

## e) Spezielle Maßzahlen zur Beurteilung von Einkommensverteilungen

An erster Stelle sei eine relativ alte Maßzahl erwähnt. Es handelt sich um Paretos $\alpha$. Pareto entdeckte Ende des 19. Jahrhunderts bei der Untersuchung verschiedener Einkommensverteilungen eine Gemeinsamkeit in den Häufigkeitsverteilungen. Er glaubte, daß Einkommensverteilungen "naturgesetzlich" bestimmten Funktionalbeziehungen gehorchen. Es sei N die Zahl derer, die ein Einkommen größer gleich Y beziehen.

---

1) Der Gedanke dieser nutzentheoretischen Interpretation der Einkommen ist schon sehr alt, wurde aber speziell in jüngster Zeit und auch nicht im gleichen Sinne wieder aufgegriffen.

Er fand für seine Verteilungsstatistiken folgenden Zusammenhang:

$$N = cY^{-\alpha}$$

Die Gleichung wird logarithmiert als Gerade geschätzt wobei durch $\alpha$ die Verteilung gemessen werden soll.

$$\log N = \log c - \alpha \log Y$$

Der Ansatz weist einen entscheidenden Mangel auf. Durch das Kumulieren einer und das Logarithmieren beider Variablen spricht dieses zwar sehr empfindlich auf Verteilungsänderungen im Bereich der oberen Einkommen an, bleibt aber zu unempfindlich für Änderungen im unteren Bereich.[1] Obwohl das Zustandekommen einer Paretoverteilung ökonomisch begründet werden kann, ergibt sich kaum eine anschauliche Interpretation des Maßes $\alpha$.

Beim folgenden Maß, der logarithmischen Varianz liegt nun deutlich der Gedanke zugrunde, daß bei der Verteilungsmessung der Nutzen des Einkommens verglichen werden muß. Es herrscht unter Wissenschaftern die Ansicht vor, daß Nutzen interpersonell nicht vergleichbar ist. Tatsache ist aber, daß es in gewissen Grenzen allgemeine Vorstellungen von Nutzenfunktionen gibt und daß die Einkommensverteilung eine entscheidende Äußerung sozialer Gerechtigkeit und entsprechend vom Politiker zu beurteilen ist.
Der Variationskoeffizient impliziert **eine** für **alle** Individuen gleiche Nutzenfunktion. Unabhängig von der Einkommenshöhe werden gleiche Gewichtungen gleicher Umverteilungsbeträge angenommen, d.h. es wird von einem

1) Blümle: Theorie der Einkommensverteilung, Heidelberger Taschenbuch 173, Berlin 1975, S. 34

konstanten Grenznutzen des Einkommens ausgegangen. Im allgemeinen wird jedoch angenommen, daß mit zunehmender Sättigung durch ein Gut - hier das Einkommen - die durch die zusätzliche Einheit des Gutes bedingte Nutzenzunahme abnimmt. Dies wird auch als das Gesetz vom abnehmenden Grenznutzen (1. Gossen`sches Gesetz) bezeichnet. Um das Gesetz vom abnehmenden Grenznutzen für ein Verteilungsmaß verwenden zu können muß es quantifiziert werden. Da nun die Logarithmusfunktion die Eigenschaft besitzt, daß relative Zunahmen einer Variablen zu gleichen absoluten Zunahmen ihrer Logarithmen führen, kann damit der Nutzen des Einkommens bestimmt werden.

$U = a \ln Y$ d.h. der Nutzen U ist direkt proportional dem Logarithmus des Einkommens.

Das Streuungsmaß, das diesen Nutzenvorstellungen entsprechend die Einkommensverteilung als Nutzenstreuung mißt, ist mit der logarithmischen Varianz (Standardabweichung) zugleich einer der Parameter von Gibrats Lognormalverteilung. Diese wird der Schiefe der tatsächlichen Häufigkeitsverteilung gerecht und erscheint somit zu deren Darstellung besser geeignet als die einfache Standardabweichung.
Die logarithmische Varianz ist definiert als:

$$LV = \frac{1}{N} \sum_i (u_i - \bar{u})^2$$

Diese Maßzahl gewichtet Transfers im unteren Einkommensbereich stärker und reagiert daher stärker auf Änderungen im untersten Einkommensbereich. Durch die logarithmische Transformation kann jedoch im obersten Einkommensbereich die Konkavitätsbedingung und damit die Pigou-Dalton-Bedingung verletzt werden.[1] Das Kriterium von Pigou-Dalton

1) Holzmann, Pflug, Vetschera: Untersuchung zur österr. Einkommensverteilung (I), Institut für Statistik der Universität Wien, Ergebnisbericht Nr. 4, 1978

fordert, daß wenn eine Verteilung durch die Änderung zweier Einkommensgrößen derart verändert wird, daß deren Summe gleichbleibt, eine Zu- bzw. Abnahme oder ein Gleichbleiben des Ungleichheitsmaßes je nach der Änderung der absoluten Differenz der beiden Einkommen (Zu- Abnahme, Konstanz).[1] Die Quadration der Logarithmen kann die Erfüllung der Bedingung nicht gewährleisten.
Ein wegen seiner günstigen Disaggregationseigenschaften manchmal verwendetes Ungleichheitsmaß ist der Theil-Koeffizient.[2] Er wird aus dem informationstheoretischen Konzept der Entropie abgeleitet. Die Entropie H ist definiert als:

$$H = \sum_{i=1}^{k} p_i \, lg \frac{1}{p_i}$$

mit lgx als Logarithmus zur Basis 2.

Für H gilt allgemein:

- H wird ein Maximum wenn $p_1 = p_2 = \dots = p_k = \frac{1}{k}$
- Konzentriert sich die gesamte relative Häufigkeit auf eine Merkmalsausprägung, gilt also für ein $i : p_i = 1$, so verschwindet H.
- Bei einer Gleichverteilung aller Elemente auf die K Klassen wächst H mit steigendem K.
  (K... Anzahl der Klassen)

Aus diesem allgemeinen Konzept der Entropie als Streuungsmaß, kann nun für den Fall allgemeiner Anteile $p_i$ ein Konzentrationsmaß gewonnen werden.
Das Theil-Maß ist bestimmt durch:

$$T = \sum_{i=1}^{k} p_i \frac{x_i}{\bar{x}} \ln\left(\frac{x_i}{\bar{x}}\right)$$

---

1) Champernowne D.G.: A Comparison of Measures of Inequality of Income Distribution, Ec. Journal 84, 1974, S 789ff.
2) Peters W.: Interregionale Einkommensunterschiede in der Landwirtschaft in der BRD, Agrarwirtschaft 1975, 7, S.183ff.

Als letztes Maß der Einkommenskonzentration sei der Atkinson-Koeffizient[1] besprochen. Atkinson versucht, Einkommenskonzentration unter Berücksichtigung einer Wohlfahrtsfunktion zu messen. Diese Wohlfahrtsfunktion ist in den individuellen Einkommen symmetrisch und additiv.

Sie sei definiert mit:

$$W \equiv \int_0^{\bar{y}} U(y) f(y)\, dy \qquad \text{für } 0 \leqslant y \leqslant \bar{y}$$

wobei f(y) ... Häufigkeitsverteilung des Attributes y
und u(y) ... Nutzenfunktion

Die Spezifikation der Nutzenfunktion bestimmt somit die Rangordnung von Verteilungen. Wenn für u(y) gilt $u'(y) > 0$ und $u''(y) \leqslant 0$ so kann eine eindeutige Rangordnung der Verteilungen nur dann erzielt werden, wenn sich die Lorenzkurven der Verteilungen nicht schneiden. Andernfalls können immer monoton steigende, konkave Funktionen u(y) gefunden werden, die eine unterschiedliche Rangordnung von Verteilungen erzeugen.[2] Der Atkinson-Koeffizient basiert auf der Beziehung zwischen tatsächlichem Wohlfahrtsniveau und jenem, das bei Gleichverteilung der Einkommen realisiert würde. Das Gleichverteilungsäquivalenzeinkommen $y_e$ errechnet sich aus der Gleichung:

$$U(y_e) \int_0^{\bar{y}} f(y)\, dy = \int_0^{\bar{y}} U(y) f(y)\, dy$$

1) Atkinson A.B.: On the Measurement of Inequality, Journal of economic theory 2, 1970, S. 244 - 263

2) Pichelmann K.: "Einige Bemerkungen zur Messung von Einkommenskonzentration", IHS, August 1979, S. 6ff

Für A folgt daraus:

$$A = 1 - \frac{y_e}{\mu} \quad \text{wobei } \mu = \int_0^{\bar{y}} y f(y) dy$$

Die Nutzenfunktion u(y) wird durch

$$u(y) = \frac{y^{1-e}}{1-e} \quad \text{für } e \neq 1$$

bzw. $u(y) = \ln y$ für e = 1 definiert, wobei stets gelten muß: $e \geqslant 0$ zur Sicherung der Konkavitätsbedingung.

Für diskrete Verteilungen ist das Atkinsonmaß wie folgt definiert:

$$A_e = 1 - \left[ \sum_{i=1}^{N} \left(\frac{y_i}{\mu}\right)^{1-e} f(y) \right]^{\frac{1}{1-e}} \quad \text{für } e \neq 1$$

$$A = 1 - \frac{\hat{x}}{\mu} \quad \text{mit} \quad \hat{x} = \prod_{i=1}^{N} y_i f(y_i) \quad \text{für } e = 1$$

Die Gestalt der Grenznutzenfunktion u'(y) bestimmt die Bewertung von Einkommenstransfers, die der Atkinson-Koeffizient vornimmt. U'(y) weist daher, wie leicht ersichtlich ist, die Konstante Elastizität -e auf. Wenn also ein individuelles Einkommen um 1 % steigt, so fällt der entsprechende Grenznutzen um e Prozent seines vorherigen Wertes. Wenn der "Ungleichheitsaversionsparameter" e steigt, so steigen auch die Gewichte, die den Transfers zu niedrigen Einkommensgruppen zugemessen werden. Für e = 0 ergibt sich eine lineare Nutzenfunktion, die keine sinnvolle Interpretation des dazugehörigen Atkinsonkoeffizienten erlaubt.

Der Vorteil von $A_e$, als einem auf einer sozialen Wohlfahrtsfunktion aufbauenden Maß, im Gegensatz zu den traditionellen Konzentrationsmaßen, ist offensichtlich. Unter der Voraussetzung, daß allgemein anerkannte Vorstellungen über eine soziale Wohlfahrtsfunktion herrschen, können mit Hilfe des Atkinson-Koeffizienten durch Variation des Parameters e normative Postulate der Ungleichheitsaversion explizit in der Bewertung der Einkommensungleichheit berücksichtigt werden.

### 3. Ein Vergleich der Methoden

Eine ausführliche Analyse der gebräuchlichsten Ungleichheitsindizes liefert Champernowne.[1] Sieben Kriterien dienten als Beurteilungsgrundlage der verschiedenen Ungleichheitsmaße.

a) Leichte Anwendbarkeit auf vorhandenes statistisches Datenmaterial und geringer Rechenaufwand.
b) Das Maß hängt allein von der Häufigkeitsverteilung der Einkommen ab.
c) Unabhängigkeit von der Größe der Population.
d) Skaleninvarianz (Unabhängigkeit von der Skalierung der Variablen).
e) Pigou-Dalton-Kriterium (siehe III.1. e: logarithmische Varianz)
f) Normierter Schwankungsbereich des Index.
g) Anwendung für spezielle Fragestellung der Ungleichheitsmessung.

Von den oben dargestellten Indizes wurden folgende in den Vergleich einbezogen:

$I_1$... Variationskoeffizient
$I_2$... logarithmische Varianz bzw. Standardabweichung
$I_3$... Relation des geometrischen Mittels zum arithmetischen Mittel[2]

---

1) Champernowne D.G.: A Comparison of Measures of Inequality of Income Distribution, The econ. Journal 84, 1974, S.787-816

2) $I_3$ ist ein Spezialfall des Atkinson-Index und zwar unter der Annahme, daß der Grenznutzen des Einkommens umgekehrt proportional zum Einkommen ist. Siehe auch III. 2. e) Darstellung für A.

$I_5$... Gini-Koeffizient
$I_6$... die Theil`sche Entropie
$\alpha$ ... Paretos Alpha
$\gamma_1$... Momentenkoeffizient der Schiefe
$\gamma_2$... Momentenkoeffizient der Kurtosis

$I_1$, $I_2$, $I_3$, $I_5$, $I_6$ werden als allgemeine Ungleichheitsmaße betrachtet; $\alpha$, $\gamma_1$, $\gamma_2$ werden zur Analyse spezieller Phänomene herangezogen. Auf eine genaue Darstellung der Methode der Analyse sei im Rahmen dieser Arbeit verzichtet. Die Sensibilität der Indizes wurde an mehreren theoretischen Verteilungen getestet. Hierfür mußten sie, speziell um Bedingung f) zu genügen, entsprechend transformiert werden. Die Ergebnisse der vergleichenden Analyse und somit die speziellen Bedingungen der Anwendbarkeit der Indizes sei nun kurz zusammengefaßt.

$I_2$ erfüllt die Pigou-Dalton Bedingung (e) nicht für jede beliebige Verteilung.

$I_1$ bewertet Verteilungen mit $\alpha \leqslant 2$ immer stets gleich. Die meisten empirischen Einkommensverteilungen die untersucht wurden hatten aber ein $\alpha \leqslant 2$.

Die logarithmische <u>Standardabweichung</u>($I_2$)war das Maß, das am sensibelsten auf Änderungen im unteren Einkommensbereich reagierte. Der Variationskoeffizient($I_1$)und der Theil`sche Ungleichheitskoeffizient ($I_6$) eigneten sich am besten zur Analyse der extrem hohen Einkommensbereiche. Im mittleren Einkommensbereich zeigte die Anwendung der logarithmischen Standardabweichung ($I_2$) und das Maß $I_3$ (Relation der Mittelwerte) sowie der Gini-Koeffizient ($I_5$) die höchste Empfindlichkeit. Generell kann gesagt werden, daß die <u>spezielle Fragestellung</u> das entscheidende Kriterium für die jeweilige Anwendung der Ungleichheitsmaße ist. So sollte man z.B. bei der Analyse der hohen Einkommen einen speziellen Index, wie z.B. Paretos $\alpha$ heranziehen, um etwaige Änderungen in diesem Bereich leicht zu erkennen. Allerdings müßte dafür die Verteilung dem Bildungsgesetz von Pareto gut entsprechen. Zur genauen Analyse der niedrigen Einkommensbereiche wird man $I_2$ heranziehen und sicherlich nicht $I_6$.

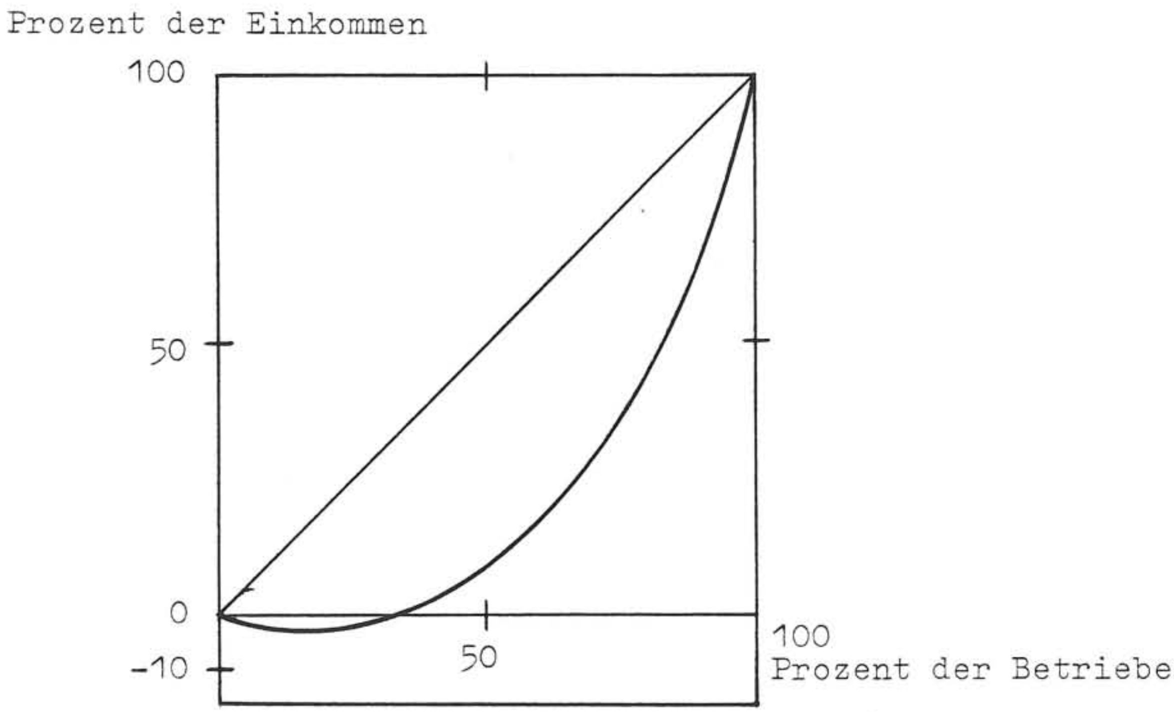
## 4. Die statistische Beurteilung der Datenbasis

### a) Formales Kriterium: Negatives Einkommen

Durch den betriebswirtschaftlichen Einkommensbegriff ergeben sich bei der Analyse landwirtschaftlicher Einkommensdaten nicht nur inhaltliche Probleme, sondern auch Schwierigkeiten im Hinblick auf die Anwendung der Methoden. Die wesentlichsten Schwierigkeiten bereitet die enorme Spannweite der Einkommensdaten und hier vor allem der Tatbestand, daß selbst negative Einkommen auftreten. Zunächst seien die Auswirkungen von Negativeinkommen auf einige Ungleichheitsmaße besprochen. Negative Merkmalsbeiträge sind mit den entwickelten Modellvorstellungen von vollständiger Konzentration nicht vereinbar.[1] Daher wird in der Konzentrationsmessung in der Regel von ihrer Existenz abgesehen und gegebenenfalls auftretende negative Merkmalsausprägungen werden gleich Null gesetzt. Ein derartiges Vorgehen kann aber keineswegs befriedigen. Konzentrationserscheinungen sind sicherlich unterschiedlich einzuschätzen, je nachdem, ob negative Einkommen auftreten oder nicht. So wird beispielsweise für zwei Verteilungen (eine mit negativen Merkmalsausprägungen und eine ohne) dieselbe Konzentration ausgewiesen, wenn negative Merkmalsbeiträge nicht berücksichtigt werden.

1) Nach H.v. Witzke: S. 25ff; Personelle Einkommensverteilung in der Landwirtschaft. Beim Auftreten negativer Merkmalsausprägungen geht die Normierung der Indizes verloren, sodaß eine sinnvolle Interpretation erschwert wird.

Formal lassen sich bei der Berechnung des Gini-Koeffizienten negative Merkmalsbeiträge berücksichtigen. Allerdings muß ein solcher Gini-Koeffizient als auch die dazugehörige Lorenzkurve anders interpretiert werden. Eine derartige Lorenzkurve hat folgende Gestalt:




So lange die Lorenzkurve fällt, sind die Merkmalsbeiträge negativ. Mit dem Steigen der Lorenzkurve machen sich die ersten positiven Beiträge bemerkbar. Die Kurve verläuft aber noch so lange im negativen Bereich, bis die negativen Merkmalsbeiträge durch die kleinsten positiven gerade kompensiert werden. Der Schnittpunkt der Lorenzkurve mit der Abszisse kann dann auch nicht mehr als Anteil der Merkmalsbeiträge, die kein Einkommen erzielen, interpretiert werden. Im Falle der Berücksichtigung negativer Einkommen geht beim Gini-Koeffizient die Normierung auf das Intervall (0, 1) verloren. Der Koeffizient kann dann Werte zwischen 0 und $+\infty$ annehmen. Außerdem vergrößert die Berücksichtigung negativer Merkmalsausprägungen die Probleme der Mehrdeutigkeit des

Gini-Koeffizienten. So kann $\Delta_R$ den Wert eins sowohl dann annehmen, wenn vollständige Konzentration im eingangs definierten Sinn auftritt[1], als auch dann, wenn nicht vollständige Konzentration vorliegt und im entsprechenden Umfang negative Merkmalsbeiträge auftreten. Durch die Berücksichtigung negativer Merkmalsbeiträge geht auch die Eigenschaft der graphisch so anschaulichen Interpretation des Gini-Koeffizienten verloren. Der Gini-Koeffizient unter Berücksichtigung negativer Merkmalsausprägungen sollte aus den genannten Gründen nicht allein verwendet werden. Er kann nur dazu dienen, die Aussage des eigentlichen Gini-Koeffizienten und der Lorenzkurvendarstellung zu ergänzen.

Formal nicht definiert ist die Berechnung der Logarithmen von Null und negativen Zahlen, sodaß die Verwendung der logarithmischen Varianz für negative Einkommen nicht möglich ist. In diesem Fall müßten alle negativen Einkommen größer Null gesetzt werden. Selbiges gilt für den Theil-Koeffizienten. Auch der Atkinson-Koeffizient ist generell nur für Einkommen $0 \leqslant y \leqslant \overline{y}$ definiert. Die Probleme mit dem negativen Einkommen fallen vor allem beim landwirtschaftlichen Einkommen ins Gewicht. Das landwirtschaftliche Einkommen zeigt für manche Hauptproduktionsgebiete relativ große Häufigkeiten im negativen Einkommensbereich.[2] Für die Bereiche können daher das Atkinsonmaß, das Theilmaß sowie alle logarithmischen Maße nicht generell angewendet werden. Für manche Fragestellungen empfiehlt sich aber ein Gebrauch der Maße unter Ausschaltung der negativen Einkommen. Beim Gesamteinkommen stehen der Anwendung der neueren Verteilungsmaßzahlen keine methodischen Hindernisse im Weg. Obgleich auch das Gesamteinkommen negative Werte erreichen kann, so liegt

1) Siehe auch III. 2. d)

2) Siehe z.B. Hochalpengebiet

doch der Klassenmittelwert der untersten Klasse stets im positiven Bereich.[1] Bei der inhaltlichen Interpretation von Ergebnissen muß aber stets auf die Besonderheiten des Einkommensbegriffes Rücksicht genommen werden.

## b) Empirische und theoretische Verteilungen

Die vorliegenden Daten der Einkommensstatistik für die Landwirtschaft sind im wesentlichen zensurierte Verteilungen, das heißt es liegen nur Daten aus einem bestimmten Bereich der Verteilung vor.[2] Die Verteilungen in der Landwirtschaft sind beidseitig offen. Je nach betrachteter Einkommensvariable ergeben sich unterschiedlich große offene Intervalle oben oder unten. Aufgrund der Häufigkeiten ist beim Gesamteinkommen der problematischere Bereich meist der obere, beim landwirtschaftlichen Einkommen sind es meist beide Bereiche. Die Abschätzung der Verteilung im oberen Bereich bringt aufgrund sehr guter modelltheoretischer Anpassung weniger Schwierigkeiten mit sich als der untere Bereich, da auf negative Einkommen die vorhandenen Modellansätze nicht anwendbar sind. In Anbetracht der methodischen Schwierigkeiten im Bereich der negativen Einkommen läßt sich daher die genaue Gestalt der Verteilung nicht spezifizieren. Ein Teil der oben besprochenen Ungleichheitsmaße kann daher nur mit Einschränkungen verwendet werden.

Mehr Schwierigkeiten bietet die Multimodalität vieler in der landwirtschaftlichen Statistik ausgewiesenen Einkommensverteilungen. Wenn keine sinnvolle ökonomische

---

1) Häufigkeitstabellen und eigene Berechnungen
2) Man beachte den Unterschied zu einer abgeschnittenen Verteilung. Eine solche wird nur in einem vorgegebenen Intervall definiert und wenn notwendig so normiert, daß die Gesamtwahrscheinlichkeit 1 beträgt.

Erklärung des Zustandekommens einer derartigen Verteilung existiert, empfiehlt es sich die Verteilung an das theoretische Modell anzupassen. Insbesondere trifft dies zu, wenn die Ursache einer multimodalen Verteilung eine zu geringe Besetzung in der Stichprobe ist. Ob Anpassung oder nicht, ist im konkreten Fall stets nach inhaltlichen und formalen Kriterien zu prüfen.

Um die beiden angesprochenen Probleme der offenen Klassen und der Multimodalität zu veranschaulichen sei folgende Verteilung dargestellt:

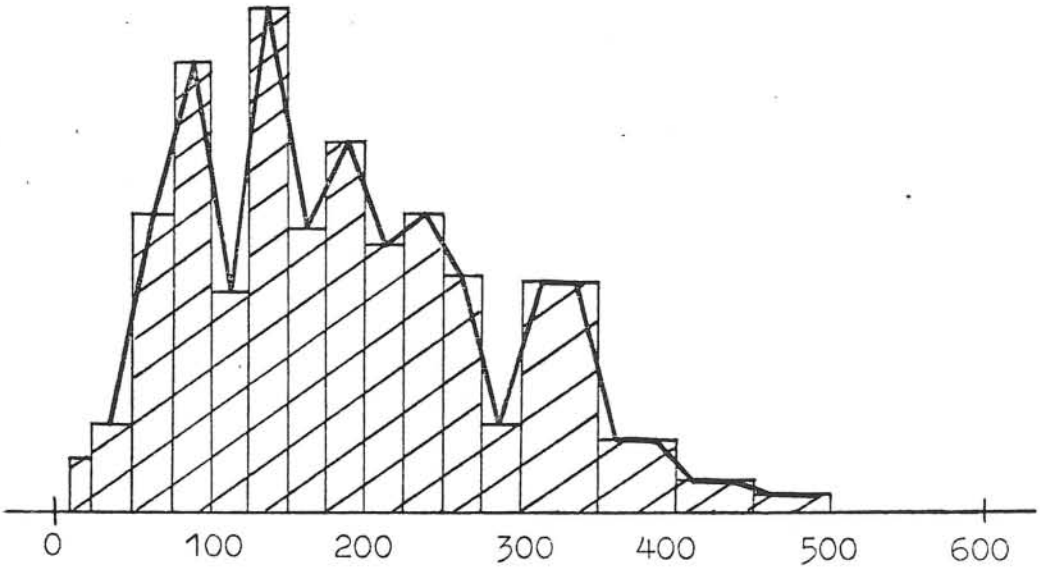
Streuung der Betriebe nach dem Gesamteinkommen je Betrieb 1979 für das Wald- und Mühlviertel:

Tabelle :

| Einkommensklassen in Tausend | Häufigkeiten in % $p_i$ | Klassenbreite $d_i$ | Korrigierte Häufigkeiten (Höhe des Histogramms) $p_i'$ |
|---|---|---|---|
| bis 10 | 0,9 | -- | -- |
| 10 - 25 | 0,8 | 15 | 1,3 |
| 25 - 50 | 2,2 | 25 | 2,2 |
| 50 - 75 | 7,6 | 25 | 7,6 |
| 75 - 100 | 11,4 | 25 | 11,4 |
| 100 - 125 | 5,5 | 25 | 5,5 |
| 125 - 150 | 12,8 | 25 | 12,8 |
| 150 - 175 | 7,2 | 25 | 7,2 |
| 175 - 200 | 9,4 | 25 | 9,4 |
| 200 - 225 | 6,8 | 25 | 6,8 |
| 225 - 250 | 7,3 | 25 | 7,3 |
| 250 - 275 | 6,0 | 25 | 6,0 |
| 275 - 300 | 2,2 | 25 | 2,2 |
| 300 - 350 | 11,8 | 50 | 5,9 |
| 350 - 400 | 3,6 | 50 | 1,8 |
| 400 - 450 | 1,6 | 50 | 0,8 |
| 450 - 500 | 0,9 | 50 | 0,45 |
| 500 und mehr | 2,0 | -- | -- |

$$d = 25,\ p_i' = p_i \cdot \varphi_i,\ \varphi_i = \frac{d}{d_i}$$

Histogramm: <u>Häufigkeitsverteilung der Betriebe nach dem Gesamteinkommen je Betrieb 1979 für das Wald- und Mühlviertel</u>




Die dargestellte Verteilung ist zwar rechtsschief hat aber fünf Modalwerte und insgesamt 2,9 % der Betriebe liegen in offenen Klassen. Die offenen Klassen, obwohl hier sehr klein, sind in der Konzentrationsmessung deshalb besonders wichtig, weil mit Einkommenssummen gerechnet wird. Die obersten 2 % der Betriebe können z.B. ein Vielfaches dieses Prozentsatzes der Einkommenssumme auf sich vereinigen. Ihrer Bedeutung muß daher in der Analyse entsprechende Beachtung geschenkt werden.

Um das offene Intervall im oberen Einkommensbereich mit in die Schätzung von Ungleichheitsmaßen einzubeziehen muß die Verteilung in diesem Bereich näher spezifiziert werden. Das kann geschehen, indem man aus den vorhandenen Daten, mit Ausnahme der offenen Klassen, die <u>Parameter einer theoretischen Verteilung</u> schätzt und daraus die Ungleichheitsmaße berechnet. Ein derartiges Vorgehen ist deshalb möglich, da eindeutige Beziehungen zwischen den Parametern einer theoretischen Verteilung und den gebräuchlichsten Ungleichheitsmaßen bestehen.

Die in der Praxis häufiger angewendete Methode ist die Extrapolation der Häufigkeiten in der offenen Klasse. Man schätzt nur aus dem oberen Einkommensbereich der Daten die Parameter einer theoretischen Verteilung und damit die Besetzung weiterer Einkommensklassen. Die beiden Methoden unterscheiden sich in der Form der Zusammenfassung der in den einzelnen Klassenbesetzungen enthaltenen Information zur Gesamtaussage eines Ungleichheitsmaßes. Im ersten Fall wird versucht, die Abweichungen der Klassenbesetzungen von der theoretischen Besetzung zu minimieren. In weiterer Folge wird dann mit der theoretischen Besetzung weitergerechnet. Im zweiten Fall verwendet man zunächst nur Informationen aus dem oberen Bereich der Einkommensverteilung, um über diese hinaus zu extrapolieren. Bei der Berechnung von Ungleichheitsmaßen wird die gesamte Information der beobachteten und geschätzten Verteilung verwendet. Da die theoretischen Verteilungen aber oftmals erhebliche Abweichungen von den beobachteten aufweisen[1] verwendet man daher meist zur Extrapolation die zweite Methode.

In der ökonomischen Theorie werden vornehmlich die Lognormalverteilung und die Paretoverteilung als Modelle für die Einkommensverteilung verwendet. Da eine Entscheidung, welche der beiden Verteilungen für die Zwecke der Einkommensanalyse in der Landwirtschaft heranzuziehen ist, erst aufgrund der mit beiden Verfahren erzielten Anpassungen getroffen werden kann, seien beide kurz erläutert.

1) Dies gilt insbesondere für die Einkommensverteilungen in der Landwirtschaft. (Multimodalität) Es ist in jedem konkreten Fall zu prüfen, ob die empirische Verteilung sinnvoll ökonomisch begründet werden kann. Von vorneherein mit theoretischen Verteilungen zu rechnen erscheint uns wegen des Informationsverlustes nicht sinnvoll. Siehe auch nächstes Beispiel.

i) Die Lognormalverteilung

Die zweiparametrige Lognormalverteilung[1] ist eine unimodale rechtsschiefe Verteilung mit der Dichte

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \quad \text{für } x > 0$$

wobei: $\mu$.........Logarithmus des Medians
$\sigma^2$.........logarithmische Varianz

Für den Zusammenhang zwischen Normalverteilung und Lognormalverteilung gilt die Beziehung:

$$L(x) = N(\log x) \quad \text{für } x > 0$$

Hieraus und aus dem zentralen Grenzwertsatz läßt sich eine theoretische Begründung für die Verwendung der Lognormalverteilung als Modell der Einkommensverteilung ableiten. Es gilt, da eine additive Verknüpfung mehrerer beliebig verteilter Zufallsvariable eine normalverteilte Zufallsvariable ergibt, daß deren multiplikative Verknüpfung, wie sie bei den das Einkommen determinierenden Faktoren plausibel ist, eine lognormalverteilte Zufallsvariable ergibt.[2]

Eine einfache Methode zur Parameterschätzung der Lognormalverteilung ist die graphische Methode. Hierbei wird die Verteilungsfunktion auf der Einkommensachse

---

1) M. Fisz: Wahrscheinlichkeitsrechnung und mathematische Statistik, Deutscher Verlag der Wissenschaften, Berlin 1978, S. 205

2) Holzmann, Pflug, Vetschera: Untersuchung zur österr. Einkommensverteilung, Institut für Statistik, Wien 1978, S. 12

auf Logarithmen und auf der Wahrscheinlichkeitsachse auf Quantilsgrenzen der standardisierten Normalverteilung transformiert. Die Parameter der Lognormalverteilung sind dann der Anstieg ($\sigma$) und der Abschnitt auf der log x-Achse ($\mu$) der so entstandenen Geraden. $\mu$ und $\sigma$ können mit Hilfe der linearen Regression geschätzt werden. Ein Vorteil dieser Methode liegt darin, durch geeignete Auswahl der Regressionspunkte zu bestimmen, aus welchem Bereich der Verteilung die Parameter geschätzt werden sollen. Die Besetzung der k-ten Einkommensklasse wird nun näherungsweise abgeschätzt. Die Anzahl der Einkommensbezieher mit Einkommen zwischen $u_k$ und $u_{k+1}$ ergibt sich unter Annahme der Lognormalverteilung durch

$$N \int_{u_k}^{u_{k+1}} dL(x|\mu,\sigma^2)$$

wobei N die Gesamtzahl der Einkommensbezieher darstellt.

Mit Hilfe der Trapezregel läßt sich das Integral mit guter Näherung abschätzen. Für die Besetzung der k-ten Einkommensklasse ergibt sich dann

$$f_i \approx N \cdot \frac{1}{2}\left(L(u_k|\mu,\sigma^2) + L(u_{k+1}|\mu,\sigma^2)\right) \cdot (u_{k+1} - u_k)$$

Da die Lognormalverteilung im oberen Bereich stückweise sehr gut durch eine Gerade dargestellt werden kann, ist diese einfache Annäherung ausreichend.

ii) <u>Die Paretoverteilung</u>

Wie oben schon erwähnt hat die Paretoverteilung nur im obersten Einkommensbereich Gültigkeit.[1] Manche Autoren nehmen an, daß dieser Bereich mit den beiden

---

1) Siehe Kapitel 3.2.e) Paretos $\alpha$

obersten Dezilen der Einkommensbezieher begrenzt ist.[1] Aus diesem Bereich müssen nun die Parameter geschätzt werden.[2]

Die Annahme, die der Paretoverteilung zugrunde liegt, ist ein funktionaler Zusammenhang zwischen der Zahl der Einkommensbezieher und der Einkommenshöhe von der Gestalt:

$$Q(x) = ß \cdot x^{-\alpha}$$

Q(x) ist die Zahl der Einkommensbezieher deren Einkommen größer als x ist. Die Funktion ist für alle positiven x definiert aber nur für ein Einkommen größer $x_0$ empirisch nachweisbar. Auf unser Problem wirkt sich diese Einschränkung nicht nachteilig aus. Die sehr gute Anpassung der Paretoverteilung im oberen Bereich der empirischen Einkommensverteilungen läßt sie zur Extrapolation der Einkommensverteilung als durchaus geeignet erscheinen.

Zur Parameterschätzung ist es sinnvoll die Daten auf Anteile der gesamten Einkommensbezieher zu normieren.

$$h(x) = \frac{Q(x)}{N} = ß^* x^{-\alpha}$$

Beidseitige logarithmische Transformation führt auf die Gleichung:

$$\log h(x) = \log ß^* - \alpha \log x$$

1) H. Lydall: The Structure of Earnings, Oxford 1968

2) Die Klasseneinteilung der Einkommensvariable in der Landwirtschaft deckt meist den größten Teil der oberen Dezile ab, sodaß einer Schätzung genügend Stützstellen zur Verfügung stehen.

$\alpha$ und ß* können nun direkt durch lineare Regression geschätzt werden. Es ergibt sich somit für jedes x der relative Anteil der Einkommensbezieher von Einkommen größer gleich x an der Gesamtheit der Einkommensbezieher. Die Besetzung des Intervalls ($u_i$, $u_{i+1}$) ergibt sich durch:

$$f_i = N.(h(u_i) - h(u_{i+1}))$$

Die Paretoverteilung kann in Einkommensanalysen wegen mangelnder Besetzung im Gültigkeitsbereich der Funktion oftmals nicht verwendet werden.[1] Auch die Konsistenz der Parameter im Zeitablauf ist Anlaß berechtigter Zweifel an der Verwendung der Paretoverteilung. Beide Kritikpunkte tangieren die Einkommensstatistiken für die Landwirtschaft nur schwach. Die Klassenbesetzung wird bis in extrem hohe Einkommensbereiche relativ exakt ausgewiesen und eine Analyse der Einkommensungleichheit im Zeitablauf kann aufgrund der erst seit 1976 vorhandenen zuverlässigen Daten nicht durchgeführt werden.[2] Als Vorteil gegenüber der Lognormalverteilung erweist sich sogar die Beschränkung auf den oberen Bereich. Der mittlere Bereich der Verteilung, der oft mehrere Modalwerte zeigt, geht von vornherein nicht in die Schätzung ein. Die Problematik ist also in der Analyse von landwirtschaftlichen Einkommensdaten etwas anders gelagert als z.B. bei Einkommensverteilungen der Unselbständigen.

Neben den empirischen Vorteilen der Paretoverteilung ist der theoretische Einkommensbegriff, der der Anwendung der Paretoverteilung zugrunde liegt, für die

---

1) Siehe Lohnstufenstatistik etc.

2) Um Ernteschwankungen etc. auszuschalten muß mit mehrjährigen (mindestens 3-jährigen) Durchschnitten gerechnet werden. Vier Beobachtungsjahre lassen einen Zeitvergleich mit dreijährigen Durchschnitten nicht zu.

Landwirtschaft eher zutreffend als für die unselbständigen Einkommen. Mandelbrot[1] etwa geht bei der Erklärung der Paretoverteilung davon aus, daß das Gesamteinkommen eine Summe von verschiedenen Einkommenskategorien, wie landwirtschaftliches Einkommen, außerlandwirtschaftliches Erwerbseinkommen und Sozialeinkommen ist. Somit ist das Modell für landwirtschaftliche Einkommensdaten anwendbar.

An zwei Verteilungen seien die beiden Extrapolationsmöglichkeiten dargestellt.

1) B. Mandelbrot: The Pareto-Lévy Law and the distribution of income, Internat. Econ. Review 1, (1960) S. 79ff.

Extrapolation der oben offenen Klasse mit der
Lognormalverteilung:

Tabelle : Gesamteinkommen/Betrieb im Bundesmittel 1979

| | $x$ | $F_i$ | $\hat{F}_i(x)$ | Extrapolation $x$ | $\hat{F}_i(x)$ | $\hat{h}(x)$ |
|---|---|---|---|---|---|---|
| Einkommen in tausend | 200 | 56,1 | 55 | 600 | 98,6 | 1,4 |
| | 250 | 71,0 | 70,8 | 700 | 99 | 1,0 |
| | 300 | 79,9 | 81 | 800 | 99,4 | 0,6 |
| | 400 | 92,1 | 92 | 1000 | 99,7 | 0,3 |
| | 500 | 96,9 | 97 | 1200 | 99,9 | 0,1 |

$\hat{\sigma} = 1,89$
$\hat{\mu} = 5,2$

Tabelle : Gesamteinkommen/Betrieb im Alpenvorland 1979

| | $x$ | $F_i$ | $\hat{F}_i(x)$ | Extrapolation $x$ | $\hat{F}_i(x)$ | $\hat{h}(x)$ |
|---|---|---|---|---|---|---|
| Einkommen in tausend | 250 | 60,5 | 58,5 | 600 | 95,5 | 4,5 |
| | 300 | 70,7 | 71 | 700 | 98,3 | 1,7 |
| | 400 | 85,8 | 86,3 | 800 | 99,1 | 0,9 |
| | 500 | 93,8 | 93,5 | 1000 | 99,7 | 0,3 |
| | | | | 1200 | 99,9 | 0,1 |

$\hat{\sigma} = 1,86$
$\hat{\mu} = 5,4$

Anmerkung: Werte die mit einem $\wedge$ gekennzeichnet sind bezeichnen Schätzwerte. Alle Häufigkeiten in Prozent.

Extrapolation der oben offenen Klasse mit der
Paretoverteilung:

Tabelle : Gesamteinkommen/Betrieb im Bundesmittel 1979

| Einkommens-klasse in tausend | $p_i$ | $h(x)$ | $\hat{h}(x)$ | Extrapolation $x$ | $\hat{h}(x)$ | $\hat{F}_i$ |
|---|---|---|---|---|---|---|
| 300 bis 350 | 8,2 | 20,1 | 20,8 | 600 | 1,7 | 98,3 |
| 350 bis 400 | 4,0 | 11,9 | 11,9 | 700 | 0,97 | 99,03 |
| 400 bis 450 | 3,0 | 7,9 | 7,3 | 800 | 0,60 | 99,4 |
| 450 bis 500 | 1,8 | 4,9 | 4,8 | 1000 | 0,27 | 99,73 |
| 500 und mehr | 3,1 | 3,1 | 3,3 | 1200 | 0,14 | 99,86 |

$\hat{\alpha} = 3{,}617$
$\widehat{\log \beta^*} = 23{,}66$

Tabelle : Gesamteinkommen/Betrieb im Alpenvorland 1979

| Einkommens-klasse in tausend | $p_i$ | $h(x)$ | $\hat{h}(x)$ | Extrapolation $x$ | $\hat{h}(x)$ | $\hat{F}_i$ |
|---|---|---|---|---|---|---|
| 350 bis 400 | 5,3 | 19,5 | 20,4 | 600 | 3,5 | 96,5 |
| 400 bis 450 | 5,2 | 14,2 | 13,2 | 700 | 2,1 | 97,9 |
| 450 bis 500 | 2,8 | 9,0 | 8,9 | 800 | 1,37 | 98,63 |
| 500 und mehr | 6,2 | 6,2 | 6,3 | 1000 | 0,66 | 99,34 |
| | | | | 1200 | 0,37 | 99,63 |

$\hat{\alpha} = 3{,}262$
$\widehat{\log \beta^*} = 22{,}13$

Anmerkung: Werte die mit einem ^ gekennzeichnet sind be-
zeichnen Schätzwerte. Alle Häufigkeiten in Prozent.

Beide Verteilungen zeigen sehr gute Anpassung an die empirischen Werte und durchaus plausible Extrapolationsergebnisse. Bei der Paretoverteilung wurde der Bereich der beiden obersten Dezile zur Schätzung herangezogen. Bei der Lognormalverteilung führte eine Auswahl der Regressionspunkte zu einer wesentlichen Verbesserung der Anpassung. Eine Schätzung über den gesamten Bereich ist daher auch bei der Lognormalverteilung nicht sinnvoll. Generell zeigt sich, daß die Paretoverteilung die Besetzung im extremsten Bereich höher schätzt als die Lognormalverteilung. Unter Zuhilfenahme des empirischen Klassenmittelwertes der offenen Klasse läßt sich nun der Verlauf der Verteilung im obersten Bereich gut wiedergeben.

Für die Interpolation des unten offenen Intervalls sind beide Verteilungen aufgrund ihrer methodischen Eigenschaften nicht verwendbar. Wenn keine weiteren Informationen über die Verteilung in diesem Bereich vorliegen, ist eine lineare Interpolation aus dem untersten Bereich der Verteilungsfunktion eine einfache Möglichkeit der Abschätzung im negativen Einkommensbereich.

LITERATURVERZEICHNIS

AGRARISCHE RUNDSCHAU Nr. 5, Einkommenspolitik für die Land- und Forstwirtschaft 1978.

AITCHISON, J./BROWN, J.A.C., The Lognormal Distribution, Cambridge 1957.

AITCHISON, J./BROWN, J.A.C., The Lognormal Distribution with special reference to its uses in economics, Cambridge 1969.

ATKINSON, A.B., On the Measurement of Inequality, Journal of Economic Theory 2, 1970.

BLÜMLE, G., Theorie der Einkommensverteilung, Berlin 1975.

BRUCKMANN, G., Einige Bemerkungen zur statistischen Messung der Konzentration, Metrika Jhg. 14, 1969.

BUCHEGGER, R., Zur Theorie und Messung der personellen Einkommensverteilung, Österreichische Zeitschrift für Politikwissenschaften Nr. 1, 1975.

CHAMPERNOWNE, D.G., A Comparison of Measures of Inequality of Income Distribution, The Economic Journal 84, 1974.

CHRISTL, J./SUPPANZ, H., Zur Entwicklung der Einkommen und der Einkommensverteilung in Österreich, Wien 1978.

FAHRNBERGER, A./NEUKOMM, W./SCHUH, A., Einkommensdisparität der Landwirtschaft, Schriftenreihe für Agrarpolitik und Agrarsoziologie Bd. 28, Linz 1980.

FERSCHL, F., Statistik - Manuskript, Institut für Statistik, Wien 1973.

FISZ, M., Wahrscheinlichkeitsrechnung und mathematische Statistik (Übersetzung aus dem Polnischen), Berlin 1978.

FIELDS, G.S./FEI, J.C.H., On Inequality Comparisons Econometrica 46, 1978.

GASTWIRTH, J.L., The Estimation of the Lorenz Curve and Gini-Index, Review of Econometrics and Statistics 54, 1972.

GASTWIRTH, J.L./GLAUBERMANN, M., The Interpolation of the Lorenz Curve and Gini-Index from Grouped Data, Econometrica 46, 1976.

HAIMBÖCK, H., Untersuchungen über einen Einkommensvergleich zwischen landwirtschaftlichen und außerlandwirtschaftlichen Bevölkerungsgruppen in Österreich, Bodenkultur Heft 5, Wien 1974.

HOLZMANN, R./PFLUG, G./VETSCHERA, R., Untersuchungen zur österreichischen Einkommensverteilung (I), Ergebnisbericht 4, Institut für Statistik, Wien 1978.

KARG, G., Einkommensverteilung und Einkommenselastizität der Nachfrage nach Nahrungsmitteln, Agrarwirtschaft Nr. 6, 1973.

KÖHNE, M., Die Analyse der intrasektoralen Einkommenslage als Informationsgrundlage der Agrarpolitik, Agrarwirtschaft, 1973.

LANG, G., Die Buchführungsergebnisse land- und forstwirtschaftlicher Betriebe, Agrarische Rundschau Nr. 5, 1979.

LYDALL, H., The structure of earnings, Oxford 1968.

MANDELBROT, B., The Pareto-Lévy Law and the Distribution of Income, International Economic Review 1, 1960.

ORT, W., Die Ursachen der Einkommensunterschiede in landwirtschaftlichen Betrieben und ihre Quantifizierung, Volkswirtschaftliche Schriften Heft 164, Berlin 1971.

ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG, Prognose der Wirtschaftsentwicklung der Achtziger Jahre, Wien 1980.

PETERS, W./SCHMITT, G., Interregionale Einkommensunterschiede in der Landwirtschaft, Agrarwirtschaft Nr. 11, 1973.

PETERS, W., Interregionale Einkommensunterschiede in der Landwirtschaft in der Bundesrepublik Deutschland, Agrarwirtschaft Nr. 7, 1975.

PFÄHLER, F., Vorschläge zur Darstellung der Einkommenslage in der Landwirtschaft und erste Ergebnisse einer Auswertung der Testbetriebe, Agrarwirtschaft Nr. 1, 1974.

PICHELMANN, K./WAGNER, M., Zur Entwicklung der Einkommensverteilung in Österreich, Institut für Höhere Studien, Wien 1978.

PICHELMANN, K., Zur Messung der Einkommenskonzentration, Institut für Höhere Studien, Wien 1979.

PYATT, G., On the Interpretation and Disaggregation of Gini-Coefficients, The Econometric Journal 86, 1976.

RINTELEN, P., Die Einkommensdisparität in der deutschen Landwirtschaft und ihre Ursachen, Agrarwirtschaft, 1968.

ROBINSON, J./EATWELL, J., Einführung in die Volkswirtschaftslehre, München 1974.

SACHS, L., Angewandte Statistik, Berlin 1978.

SCHMIDT, H., Die Erfolgsdisparität innerhalb der Landwirtschaft, Agrarwirtschaft, 1967.

SEN, A., On Economics Inequality, Oxford 1973.

SUPPANZ, H./WAGNER, M., Die Einkommensverteilung in Österreich, Institut für Höhere Studien, Forschungsbericht Nr. 143, Wien 1979.

WITZKE, H.v., Zur Aussagefähigkeit inter- und intrasektoraler Einkommensvergleiche in den Agrarberichten, Agrarwirtschaft Nr. 7, 1975.

WITZKE, H.v., Agrarpreise, EG-Agrarpreispolitik und personelle Einkommensverteilung in der Landwirtschaft, Zeitschrift für Wirtschafts- und Sozialwissenschaften, Jhg. 99, 1979.

WITZKE, H.v., Personelle Einkommensverteilung in der Landwirtschaft und Agrarpreise, Volkswirtschaftliche Schriften, Heft 281, Berlin 1979.

YAMANE, T., Statistik Band 1,2, Frankfurt/Main 1976.

FORSCHUNGSBERICHTE DER BUNDESANSTALT FÜR BERGBAUERNFRAGEN

Nr. 1: Landwirtschaftliche Entwicklungs- und Strukturdaten des Waldviertels
(von Josef Krammer - Mai 1980) vergriffen

Nr. 2: Theoretische und methodische Überlegungen zur Messung und Darstellung von Einkommensverhältnissen
(von Rudolf Niessler - November 1980) Preis*): S 46,-

Nr. 3: Analyse der Buchführungsergebnisse von Betrieben mit negativen landwirtschaftlichen Einkommen
(von Josef Krammer/Rudolf Niessler - November 1980)
Preis*): S 40,-

Nr. 4: Strukturentwicklung und Einkommenssituation der Milchproduktionsbetriebe
(von Josef Krammer - April 1981) Sonderheft: Der Förderungsdienst Nr.1/81

Nr. 5: Der Einkommensbegriff in der Landwirtschaft
(von Rudolf Niessler - Mai 1981) Preis*): S 66,-

Nr. 6: Die Entwicklung der Bergbauerneinkommen
(von Rudolf Niessler - September 1981) Preis*): S 75,-

Nr. 7: Die Einkommensverteilung in der österreichischen Landwirtschaft
(von Rudolf Niessler/Josef Krammer - Juni 1982)
Preis*): S 96,-

Nr. 8: Der Maschinen- und Betriebshilfering aus der Sicht der Mitglieder - 2 Fallstudien
(von Ignaz Knöbl - Dezember 1981) Preis*): S 116,-

Nr. 9: Die Einkommensentwicklung in der österreichischen Landwirtschaft 1976 bis 1983 (Trendanalyse) 2. aktualisierte Auflage
(von Maria Asamer/Rudolf Niessler - 1984) Preis:*) S 51,-

Nr. 10: Bergbauernförderung in Österreich: Direktzahlungen von Bund und Ländern - 2. aktualisierte Auflage
(von Ignaz Knöbl - November 1983) Preis:*) S 40,-

Nr. 11: Struktur- und Einkommensentwicklung in der Schweinehaltung
(von Robert Schnattinger - September 1983) Preis:*) S 80,-

Nr. 12: Agrarpolitik in Norwegen
(von Josef Krammer - Dezember 1983) Preis*): S 40,-

Nr. 13: Einkommenspolitische Strategien
(von Rudolf Niessler - 1984) Preis*): S 50,-

Nr. 14: Produktionskosten der Milch nach Bestandesgröße und Bewirtschaftungserschwernis
(von Maria Asamer - 1984) noch nicht erhältlich

Nr. 15: Faserflachsanbau in Österreich (Betriebs- und volkswirtschaftliche Analyse)
(von Robert Schnattinger - 1985) Preis*): S 75,-

(*) zuzüglich Versandspesen bei Postversand

## BERGBAUERNKARTEN DER BUNDESANSTALT FÜR BERGBAUERNFRAGEN

Karte 1: "Bergbauerngebiet"
Abgrenzung des Bergbauerngebietes gemäß Verordnung des Bundesministeriums für Land- und Forstwirtschaft vom 14. Dezember 1979, mit der die Bergbauernbetriebe in den einzelnen Bundesländern bestimmt werden (BGBL. vom 31. Dezember 1979, 188. Stk.)

Karte 2: "Bewirtschaftungserschwernis nach Gemeinden"
erstellt aus der Bergbauernzonierung, Stand 1980