

Zur mathematischen Theorie der Einkommensverteilung in Abhängigkeit von Alter und Zeit

Autor(en): **Leuenberger, Franz**

Objektyp: **Article**

Zeitschrift: **Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries**

Band (Jahr): **55 (1955)**

PDF erstellt am: **27.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-551789>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Zur mathematischen Theorie der Einkommensverteilung in Abhängigkeit von Alter und Zeit

Von *Franz Leuenberger*, Bern

Einleitung

Durch die Entwicklung der Sozialpolitik in den letzten 50 Jahren ist unter anderem die Forderung nach einem möglichst genauen Wissen um die Verteilung einer bestimmten Personengesamtheit nach dem Einkommen entstanden. So hat sich auch die mathematische Statistik dieser Frage angenommen. Ihre Literatur ist reich an Arbeiten über diesen Gegenstand. Die meisten Autoren haben sich vor allem darum bemüht, auf empirischem oder deduktivem Wege Einkommensverteilungsfunktionen aufzustellen. Eine Zusammenstellung der bekanntesten dieser Gesetze, welche samt und sonders Funktionen einer einzigen Variablen, eben des Einkommens sind, findet sich bei M. Krämer [1]¹⁾.

In neuester Zeit macht sich eine wesentliche Erweiterung der Problemstellung geltend, in der Schweiz vor allem durch die Einführung der eidgenössischen Alters- und Hinterlassenenversicherung (AHV). Da die zum Teil nach Alter abgestuften AHV-Renten zum geleisteten durchschnittlichen Jahresbeitrag und demzufolge zum durchschnittlichen Erwerbseinkommen nicht proportional sind (polygonale Natur des Rentensystems), ist die Kenntnis der Verteilung der gleichaltrigen Beitragspflichtigen auf die verschiedenen Einkommen erforderlich. Das Alter tritt also als neue stochastische Variable auf. Es versteht sich von selbst, dass auch die Veränderung des Einkommens in der Zeit ins Auge gefasst werden muss, wenn das finanzielle Gleichgewicht auf die

¹⁾ Die Zahlen in eckiger Klammer beziehen sich auf den Literaturnachweis am Schlusse der Arbeit.

Dauer gewahrt werden soll. Wir sehen uns auf diese Weise einer Verteilung der Bevölkerung nach den drei Merkmalen Einkommen, Alter und Kalenderzeit gegenüber, während man bis in neuester Zeit unter einer Einkommensverteilung stets die univariable Verteilung einer Gesamtheit nach dem Einkommen, unabhängig vom Alter der einzelnen Individuen und ohne Rücksicht auf die zeitlichen Veränderungen, verstand.

Die genannte dreidimensionale Verteilung hat E. Kaiser [2] als erster einem umfassenden Studium unterzogen. Seine Arbeit ist für die vorliegende Studie richtungweisend.

In einem ersten Kapitel entwickeln wir nach einigen allgemeinen Betrachtungen kurz die Grundzüge der Theorie von E. Kaiser, damit die in den weiteren Kapiteln folgenden Ausführungen auch ohne vorherige Lektüre der Abhandlung von E. Kaiser verständlich sind. Im zweiten und dritten Kapitel tragen wir Erweiterungen der erwähnten Theorie vor, und im vierten Kapitel greifen wir einige Einzelprobleme heraus, welche uns im Hinblick auf Theorie und Praxis einer eingehenden Untersuchung würdig scheinen.

I. Kapitel

Zum gleichen Typ gehörige Schar von Einkommensverteilungen $f(u/x)$

§ 1. Vorbetrachtung

1. Wir haben in der Einleitung erwähnt, dass wir uns für eine Personengesamtheit interessieren wollen, welche nach den drei Merkmalen *Zeit* t , *Einkommen* u und *Alter* x verteilt ist. Legen wir unsern Betrachtungen eine endliche Zeitspanne zugrunde, so können wir formal sofort die dreidimensionale Häufigkeitsfunktion $k(t, u, x)$ anschreiben, welche definiert sei über einem Bereich B , dort nicht-negativ und stetig ist. Ferner sei

$$\int \int \int_B k(t, u, x) dx du dt = 1,$$

das ist die übliche Forderung an eine Häufigkeitsfunktion.

Es ist offensichtlich unmöglich, $k(t,u,x)$ auf empirischem Wege durch Auswertung statistischer Ergebnisse zu gewinnen, falls B so geartet ist, dass die in Betracht fallende Zeit t nicht nur die Vergangenheit betrifft, sondern über die Gegenwart hinausreicht. Nun sind aber die Probleme, welche von der Praxis her an uns herantreten, stets so gestaltet, dass sie sich auf die Zukunft erstrecken. Zur Bestimmung von $k(t,u,x)$ müssen wir daher notgedrungen in die Zukunft schliessen, d. h. eine Voraussage bevölkerungs- und wirtschaftstheoretischer Art treffen. Erleichtert werden dem Mathematiker Vermutungen über den voraussichtlichen Verlauf einer Funktion in der Zukunft dann, wenn ihre Entwicklung in der Vergangenheit verfolgt werden kann. In einem solchen Falle ist es oft möglich, gewisse mehr oder weniger verborgene Gesetzmässigkeiten in der zeitlichen Entwicklung aufzudecken. Wollen wir diese Methode wenigstens zum Teil auf das vorliegende Problem, $k(t,u,x)$ zu bestimmen, anwenden, so müssen wir die biometrischen und ökonomischen Gegebenheiten nach Möglichkeit getrennt in den Aufbau der Funktion eingehen lassen. Dies deshalb, weil wir über die Veränderung der Einkommensverteilung in der Vergangenheit in der Literatur fast keine Angaben vorfinden, da dieser Frage erst seit kurzer Zeit Aktualität zukommt. Die wirtschaftliche Komponente, welche unsere Funktion $k(t,u,x)$ mitbestimmen hilft, würde somit den Schluss von der Vergangenheit in die Zukunft gänzlich verhindern, wenn sie nicht losgelöst von der biometrischen Komponente in die Darstellung eingehen könnte. Diese Trennung ist aber auf einfache Art zu bewerkstelligen und ermöglicht uns, die bevölkerungstheoretischen Untersuchungen über die zukünftige Entwicklung einer Personengesamtheit, worüber eine reichhaltige Literatur besteht, zu nutzen. Hier wurden Gesetzmässigkeiten aufgedeckt, welche man mathematisch erfasst und zur Vorausberechnung eines künftigen Bevölkerungsverlaufs verwendet hat.

Die angedeutete Auseinanderhaltung von biometrischen und ökonomischen Faktoren wird erreicht, indem man zur Bestimmung von $k(t,u,x)$ diejenigen ihrer Rand- oder Marginal- und Parallelverteilungen heranzieht, welche sich zu diesem Zwecke ganz besonders eignen. Auf diese Weise hat E. Kaiser eine in sich geschlossene Theorie, welche wir kurz resümieren wollen, entwickelt.

Vorgängig sei noch bemerkt, dass man bei H. Cramér [3] eine erschöpfende Behandlung der mehrdimensionalen Verteilungen findet. Wir halten für das bessere Verständnis des Folgenden nur fest, dass

die k -dimensionale Randverteilung aus der n -dimensionalen Verteilung hervorgeht, indem man über die $n - k$ verbleibenden Variablen der Häufigkeitsfunktion integriert.

§ 2. Die Funktionenschar $\varphi_t(u, x)$ und die drei Fundamental-
funktionen $\lambda(x)$, $g(u)$ und $s(x)$

2. Wir betrachten eine *Personengesamtheit*, welche nach zwei Merkmalen verteilt ist, dem *Einkommen u und dem Alter x* , dies zum *Zeitpunkt t_0* . Die Verteilung soll durch die Häufigkeitsfunktion $\varphi_{t_0}(u, x)$ gekennzeichnet sein. $\varphi_{t_0}(u, x)$ sei definiert in $0 \leq u < \infty$, $x_0 \leq x \leq x_1$; dort nicht-negativ, stetig und differenzierbar in bezug auf die beiden Variablen u und x . Als Häufigkeitsfunktion ist $\varphi_{t_0}(u, x)$ in ihrem Definitionsbereich normiert, d. h. es gilt:

$$\int_0^\infty \int_{x_0}^{x_1} \varphi_{t_0}(u, x) dx du = 1. \quad (1)$$

Wir machen darauf aufmerksam, dass durch Vorgabe von $\varphi_t(u, x)$ für alle in Betracht fallenden t auch $k(t, u, x)$ bestimmt ist, falls wir noch die Kenntnis der Verteilung unserer Gesamtheit in bezug auf die Kalenderzeit t voraussetzen. Die Multiplikation der betreffenden Häufigkeitsfunktion mit $\varphi_t(u, x)$ liefert $k(t, u, x)$. Die $\varphi_t(u, x)$ stellen nämlich eine Schar zweidimensionaler Parallelverteilungen von $k(t, u, x)$ dar. In der Folge werden wir uns nicht mehr mit $k(t, u, x)$ selbst, sondern allein mit der Darstellung der Funktionen $\varphi_t(u, x)$ beschäftigen.

3. Zur Bestimmung von $\varphi_{t_0}(u, x)$ – den Parameter t_0 lassen wir der Einfachheit halber vorläufig weg und schreiben $\varphi(u, x)$ – verwendet E. Kaiser die folgenden drei *Fundamentalfunktionen*, welche er als gegeben betrachtet:

1° $\lambda(x)$, die *Altersstruktur*, das ist die Häufigkeitsfunktion, welche die Verteilung unserer Personengesamtheit auf die verschiedenen Alter wiedergibt; sie ist mit der Zeit veränderlich zu denken. Wir können sie als Randverteilung von $\varphi(u, x)$ auffassen:

$$\lambda(x) = \int_0^\infty \varphi(u, x) du. \quad (2)$$

Aus (1) geht hervor, dass $\lambda(x)$ normiert ist im Intervall $x_0 \leq x \leq x_1$.

2° $f(u/x_0) = g(u)$, die Einkommensverteilung der x_0 -Jährigen, eine Funktion, welche mit $\varphi(u, x)$ auf folgende Weise zusammenhängt: Wollen wir aus $\varphi(u, x)$ die Einkommensverteilung für ein bestimmtes Alter x gewinnen, so erhält man die gesuchte Häufigkeitsfunktion $f(u/x)$ durch die Quotientenbildung

$$f(u/x) = \frac{\varphi(u, x)}{\lambda(x)}. \quad (3)$$

Aus (2) folgt tatsächlich

$$\int_0^{\infty} f(u/x) du = 1.$$

Lassen wir x die Werte zwischen den Grenzen x_0 und x_1 annehmen, so erhalten wir eine *Schar eindimensionaler Parallelverteilungen*, von denen jede die Einkommensverteilung bezüglich eines bestimmten Alters x wiedergibt. In Anlehnung an Cramér [3] verwenden wir die Schreibweise $f(u/x)$ und nicht einfach $f(u, x)$, um darzutun, dass es sich nicht etwa um eine zweidimensionale, sondern um eine Schar von eindimensionalen Häufigkeitsfunktionen handelt, wobei x lediglich die Qualität eines Parameters hat. Im übrigen halten wir uns an die Bezeichnungsweise von E. Kaiser. Die Bedeutung von $f(u/x_0) = g(u)$ ist nun ohne weiteres klar.

3° $s(x)$, die *Skala s*. Der Name ist eine Verdeutschung der schon von G. King [4] verwendeten «salary scale», und $s(x)$ ist definiert durch den Quotienten der Mittel, das sind die Momente 1. Ordnung in bezug auf den Ursprung:

$$s(x) = \frac{\int_0^{\infty} u f(u/x) du}{\int_0^{\infty} u f(u/x_0) du} = \frac{M(x)}{M(x_0)}. \quad (4)$$

$s(x)$ gibt also den Verlauf (die «Steigerung») des mittleren Einkommens mit dem Alter wieder.

Unmittelbar aus der Definition geht hervor, dass $s(x_0) = 1$ ist. Die Fundamentalfunktion I^0 gibt demographisches Geschehen wieder, während 2° und 3° wirtschaftliche Einflüsse erfassen.

§ 3. Die Hypothesen A und B

4. $\varphi(u, x)$ bestimmt sich nach E. Kaiser durch Vorgabe unserer drei Funktionen auf Grund der *Hypothese A*:

$$f(u/x) = \frac{1}{c(x)} g \left[\frac{u}{c(x)} \right]. \quad (5)$$

Häufigkeitsfunktionen, welche sich solchergestalt nur durch die Präzision, d. h. den Maßstab, unterscheiden, heissen zum gleichen Typ gehörig (Lévy [5], Haller [6]). Jede Funktion $f(u/x)$ kann als Repräsentant des Typs angesehen werden, insbesondere $f(u/x_0) = g(u)$. Es ist leicht zu verifizieren, dass

$$c(x) = \frac{M(x)}{M(x_0)}, \quad \text{d. h. aber} \quad c(x) = s(x),$$

so dass wir die Hypothese A in der Form schreiben können

$$f(u/x) = \frac{1}{s(x)} g \left[\frac{u}{s(x)} \right]. \quad (5')$$

Aus (3) und (5') folgt sofort

$$\varphi(u, x) = \frac{\lambda(x)}{s(x)} g \left[\frac{u}{s(x)} \right]. \quad (6)$$

Links steht die gesuchte Funktion, der Ausdruck rechts ist gegeben durch die drei Fundamentalfunktionen.

5. Wir erinnern uns wieder daran, dass sich die besprochenen Vorgänge in der Ebene $(u, x)_{t_0}$ abspielen. Den Übergang von t_0 zu einem beliebigen Zeitpunkt t wird nach E. Kaiser vollzogen durch die *Hypothese B*:

$$f_t(u/x) = f_{t_0}(u/x) = f(u/x) \quad (7)$$

und somit

$$\varphi_t(u, x) = \lambda_t(x) f_t(u/x) = \frac{\lambda_t(x)}{s(x)} g \left[\frac{u}{s(x)} \right]. \quad (8)$$

Die drei Fundamentalfunktionen genügen so zur Darstellung der Verhältnisse in jeder $(u, x)_t$ -Ebene, wobei nur zu beachten ist, dass die Altersstruktur für jeden Zeitpunkt t bekannt sein muss, während die beiden andern Fundamentalfunktionen in bezug auf die Kalenderzeit invariant sind.

§ 4. Die Globalverteilung

6. Wenn wir die eine der Randverteilungen von $\varphi_t(u, x)$, nämlich $\lambda_t(x) = \int_0^\infty \varphi_t(u, x) du$, schon eingeführt und interpretiert haben, wenden wir uns jetzt der Berechnung und Bedeutung der andern zu:

$$\bar{f}_t(u) = \int_{x_0}^{x_1} \varphi_t(u, x) dx = \int_{x_0}^{x_1} \frac{\lambda_t(x)}{s(x)} g\left[\frac{u}{s(x)}\right] dx. \quad (9)$$

Da $\varphi_t(u, x)$ normiert ist, so gilt dies ebenfalls für $\bar{f}_t(u)$. $\bar{f}_t(u)$ ist die sogenannte *globale Einkommensverteilung zum Zeitpunkt t*, das ist die Häufigkeitsfunktion, welche die Verteilung der Personengesamtheit zur Zeit t in bezug auf das Einkommen allein, unabhängig vom Alter x , darstellt.

7. Die Berechnung von $\bar{f}_t(u)$ – den Parameter t lassen wir in der Folge der Einfachheit halber weg, da sich die Einkommensverteilungen in den beiden Ebenen $(u, x)_t$ und $(u, x)_{t'}$, ja nur durch die biometrischen Funktionen $\lambda_t(x)$ und $\lambda_{t'}(x)$ voneinander unterscheiden – gestaltet sich nicht mehr so einfach, wenn man die Voraussetzungen, welche zur Beziehung (1) führten, modifiziert. Ursprünglich haben wir nämlich $f(u/x_0) = g(u)$ den Definitionsbereich $0 \leq u < \infty$ zugrunde gelegt. Unter den für Einkommensverteilungen gebräuchlichen Häufigkeitsfunktionen gibt es aber solche, welche nur für $a \leq u \leq b$ erklärt sind, wo $a, b > 0$. Die gesamte statistische Masse verteilt sich alsdann auf dieses Intervall. Die Begrenzung nach unten kann z. B. durch das Vorhandensein eines *Existenzminimums* erklärt werden. Unsere Rechnung gestaltet sich für eine derart begrenzte Verteilung wie folgt:

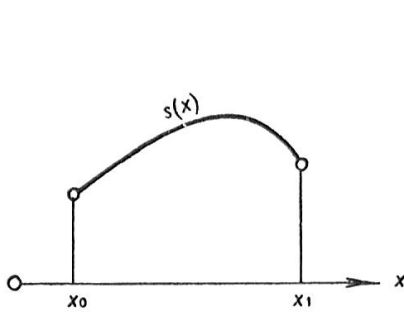
$$f(u/x_0) = g(u) \text{ sei in } a_0 \leq u \leq b_0 \text{ normiert, d. h. } \int_{a_0}^{b_0} g(u) du = 1.$$

Die Anwendung der Hypothese A hat zur Folge, dass

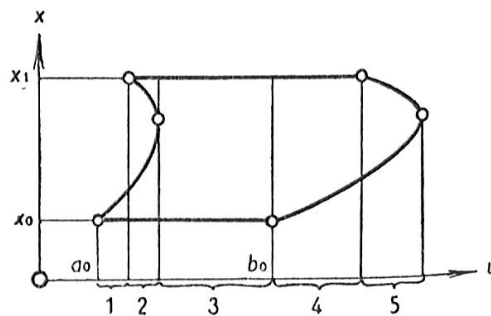
$$f(u/x) = \frac{1}{s(x)} g\left[\frac{u}{s(x)}\right]$$

erklärt und normiert ist in $a_0 s(x) \leq u \leq b_0 s(x)$, d. h. aber: Falls $s(x) \neq 1$, so ist $\varphi(u/x) = \lambda(x) f(u/x)$ nicht mehr über einem Rechteck definiert. Die unangenehmste Folge davon ist, dass die Globalverteilung

$\bar{f}(u)$ aus *mehreren Ästen* zusammengesetzt werden muss, deren Anzahl sich aus dem Verlauf von $s(x)$ bestimmt und welche getrennt berechnet werden müssen. Häufig treffen wir den Fall einer nach oben konvexen Funktion $s(x)$ an, wie sie in Figur 1 dargestellt ist. Eine erste Auswertung der individuellen Beitragskonten der AHV hält sich auch in diesem Rahmen. Ein solcher Verlauf von $s(x)$ liefert eine *fünfstufige globale Einkommensverteilung*, da der Definitionsbereich von $\varphi(u, x)$ dann nämlich etwa so aussieht, wie Figur 2 es zeigt.



Figur 1



Figur 2

Sofort anschreiben können wir den Ausdruck für den Ast 3. Dort gilt offenbar

$$\bar{f}_3(u) = \int_{x_0}^{x_1} \varphi(u, x) dx.$$

Zur Bestimmung der von u abhängigen Integrationsgrenzen in der Darstellung der vier andern Äste müssen die Begrenzungskurven $a_0 s(x)$ und $b_0 s(x)$ als Funktionen von u dargestellt werden. Die Zweideutigkeit der Umkehrfunktion von $s(x)$ in unserm Beispiel führt dazu, dass \bar{f}_1 und \bar{f}_2 bzw. \bar{f}_4 und \bar{f}_5 getrennt berechnet werden müssen. Bei einer monoton verlaufenden, z. B. einer linearen Skala s wird die Globalverteilung offenbar dreiastig, da die Intervalle 1 und 2 sowie 4 und 5 zusammengefasst werden können.

Da die Anschrift der Integraldarstellung für $\bar{f}_1, \bar{f}_2, \bar{f}_4$ und \bar{f}_5 einerseits keine besonderen Schwierigkeiten mehr bietet, andererseits etwas weitläufig ist, verweisen wir den Interessenten auf die Arbeit von E. Kaiser. Dort finden sich auch die Bedingungen für den Kontakt erster Ordnung in den Zusammensetzstellen, wie sie sich aus der Gleichsetzung der ersten Ableitung in den Kontaktpunkten ergeben. Überdies sind in der erwähnten Studie die Summen- und Komplementärfunktionen

für die Parallelverteilungen und für die Globalverteilung dargestellt und werden ausführlich besprochen.

Zwei Modelle, welche zur Veranschaulichung der in diesem Kapitel eingeführten Funktionen dienen können, findet man in einer Publikation von W. Wegmüller [7] (Seiten 311 und 315).

II. Kapitel

Zur gleichen Klasse gehörige Schar von Einkommensverteilungen $f(u/x)$

§ 1. Zielsetzung

8. Wenn man von einer Ästhetik in mathematischen Dingen sprechen kann, so ist die von E. Kaiser geschaffene Theorie der Einkommensverteilung als ausgesprochen schön zu taxieren, solange die eine der drei Fundamentalfunktionen, wir meinen $g(u)$, im Intervall $0 \leq u < \infty$ definiert ist. Eine kleine Einbusse erleidet diese Schönheit aber dann, wenn sich die entsprechende statistische Masse nicht mehr auf die ganze positive Halbachse verteilt. Wir denken an die mehrstige Globalverteilung, welche sich in diesem Falle einstellt. Diese Erscheinung hat ihre Ursache in der Hypothese A , welche besagt, dass sämtliche Parallelverteilungen $f(u/x)$ zum selben Typ gehören sollen. An der Plausibilität dieser Hypothese zweifeln wir indessen nicht; sie ist offensichtlich. Es wäre ein sinnloses Unterfangen, andere Hypothesen aufstellen zu wollen, welche den erwähnten Nachteil nicht in sich bergen, dafür aber nicht mehr plausibel sind. Dagegen legen wir uns die Frage vor, ob es nicht möglich wäre, *beide* Gesichtspunkte miteinander zu verbinden, d. h. an Stelle der Hypothese A eine andere zu setzen, welche zugleich vernünftig ist und eine *einastige globale Einkommensverteilung* liefert. In diesem und im nächsten Kapitel kommen Versuche, welche nach dieser Richtung hin tendieren, zur Sprache. Dabei begnügen wir uns mit der Betrachtung des Falles, wo sich $g(u)$ asymptotisch der u -Achse nähert. Wohl kann man einwenden, das Einkommen könne eine bestimmte Grenze nicht überschreiten; dem ist aber entgegenzuhalten, dass bei ähnlichem Sachverhalt auf andern Gebieten der mathematischen Statistik und Versicherungsmathematik dieser Tatbestand ebenfalls übergangen wird. Man denke etwa an die analytische Dar-

stellung einer Überlebensordnung nach Makeham oder an die Anwendung des Gaußschen Fehlergesetzes in Fällen, wo der Fehler eine bestimmte Grenze nicht überschreiten kann usw. Zudem haftet der Festsetzung einer obern Einkommensgrenze meist etwelche Willkür an. Die Forderung nach der Existenz des Integrals über $g(u)$ sorgt von selbst dafür, dass durch die Nichtbegrenzung nach oben bei der Anwendung der Theorie auf Probleme der Sozialversicherung nicht etwa paradoxe Resultate zu erwarten sind. Zudem ist ja auch die Existenz des Mittels gewährleistet (Skala s).

§ 2. Die Zugehörigkeit der Schar $f(u/x)$ zur gleichen Klasse im allgemeinen Fall

9. Nach B. Haller [6] und anderen heissen Häufigkeitsfunktionen, welche sich nur durch die Wahl des Maßstabes und die Lage des Nullpunktes voneinander unterscheiden, zur gleichen Klasse gehörig. P. Lévy [5] schlägt die Bezeichnung Typ im weitern Sinne («type au sens large» oder «type généralisé») vor, um das Wort Klasse für eine allgemeinere Verknüpfung zu reservieren. Unsere Benennungsweise schliesst sich jener von B. Haller an und hat zur Folge, dass die in dieser Arbeit nachstehend verwendeten Begriffe «Klasse» und «Parameterklasse» gebührend auseinandergehalten werden müssen. Zwei Häufigkeitsfunktionen $f(x)$ und $f_1(x)$ gehören dann zur selben Klasse, wenn die Relation gilt:

$$f_1(x) = \frac{1}{q} f\left(\frac{x-p}{q}\right).$$

Wir wollen nun fordern, dass sämtliche unserer Parallelverteilungen $f(u/x)$ (zum Zeitpunkt t_0) derselben Klasse angehören sollen und drücken dieses *Postulat* aus in der Beziehung

$$f(u/x) = \frac{1}{b(x)} g\left[\frac{u-a(x)}{b(x)}\right]; \quad \begin{matrix} a(x_0) = 0 \\ b(x_0) = 1 \end{matrix} \quad (10)$$

mit der Nebenbedingung

$$\frac{M(x)}{M(x_0)} = s(x). \quad (10')$$

Diese Nebenbedingung hat zur Folge, dass wir nur einen der beiden Parameter $a(x)$ und $b(x)$ frei wählen können. Als erstes wollen wir denn

die Beziehung zwischen a , b und s aufzeigen. Dabei erinnern wir uns daran, dass $g(u)$ nach links begrenzt ist: $0 < u_0(x_0) \leq u < \infty$. Gehen wir über zu $f(u/x)$, so stellt sich die untere Grenze als Funktion des Parameters x dar, und zwar durch

$$u_0(x) = u_0(x_0) b(x) + a(x). \quad (11)$$

Es ist dann

$$M(x) = \frac{1}{b(x)} \int_{u_0(x)}^{\infty} u g\left[\frac{u-a(x)}{b(x)}\right] du.$$

Setze

$$\frac{u-a(x)}{b(x)} = v; \quad u = v b(x) + a(x),$$

daraus folgt

$$du = b(x) dv,$$

so dass

$$M(x) = \int_{u_0(x_0)}^{\infty} \{v b(x) + a(x)\} g(v) dv = b(x) \int_{u_0(x_0)}^{\infty} v g(v) dv + a(x) \int_{u_0(x_0)}^{\infty} g(v) dv.$$

Es ist aber

$$\int_{u_0(x_0)}^{\infty} v g(v) dv = M(x_0) \quad \text{und} \quad \int_{u_0(x_0)}^{\infty} g(v) dv = 1,$$

das heisst

$$M(x) = b(x) M(x_0) + a(x).$$

Die Division durch $M(x_0)$, welches durch Vorgabe von $g(u)$ festgesetzt worden ist, ergibt die gesuchte Beziehung zwischen $s(x)$, $a(x)$ und $b(x)$:

$$s(x) = b(x) + \frac{a(x)}{M(x_0)}. \quad (12)$$

Das gewonnene Resultat wollen wir am Spezialfall $a(x) \equiv 0$ überprüfen. Die Beziehung (10) muss dann mit der Hypothese A übereinstimmen, d. h. $b(x)$ muss mit $s(x)$ identisch sein, was sich aus (12) tatsächlich herausliest. In Ergänzung von (12) können wir setzen

$$a(x) = M(x_0) \{s(x) - b(x)\} \quad (12')$$

und

$$b(x) = s(x) - \frac{a(x)}{M(x_0)}. \quad (12'')$$

§ 3. Spezialfall: Die Hypothese A_1

10. Jede bestimmte Wahl, die wir in bezug auf $a(x)$ bzw. $b(x)$ treffen, kann als Aufstellung einer Hypothese gewertet werden, welche formal an die Stelle der Hypothese A tritt. Das Postulat (10) können wir ansehen als Repräsentant der Gesamtheit dieser Hypothesen. Sichten wir diese Gesamtheit im Hinblick auf unser Ziel, solche Hypothesen aufzustellen, welche *die globale Einkommensverteilung durch einen einzigen Integralausdruck darzustellen* gestatten, so finden wir genau eine, die diese Forderung erfüllt. Dieses Resultat gewinnen wir durch die einfache Überlegung, dass sämtliche Parallelverteilungen $f(u/x)$ für ein- und dasselbe u -Intervall definiert sein müssen, soll die Globalverteilung einastig werden. $\varphi(u, x)$ ist dann über einem Streifen dargestellt, welcher nach links durch eine Gerade parallel zur x -Achse begrenzt ist. Die Invarianz der untern Intervallgrenze gegenüber der Transformation (11) schreibt sich

$$u_0(x) = u_0(x_0) = u_0(x_0) b(x) + a(x).$$

Setzen wir der Einfachheit halber $u_0(x_0) = u_0$, so folgt

$$a(x) = u_0(1 - b(x)). \quad (13)$$

Da bei der aktiven Bevölkerung $b(x) \geq 1$, gilt in der Regel $a(x) \leq 0$. Jetzt können wir in (10) $a(x)$ durch die rechte Seite von (13) ersetzen, so dass die Parallelverteilungen gegeben sind durch die *Hypothese A_1*

$$f(u/x) = \frac{1}{b(x)} g \left[\frac{u - u_0(1 - b(x))}{b(x)} \right], \quad (14)$$

wobei $b(x)$ durch Vorgabe der Fundamentalfunktion $s(x)$ eindeutig bestimmt ist und mit dieser gemäss (12) und (13) in folgender Beziehung steht:

$$b(x) = \frac{s(x) M(x_0) - u}{M(x_0) - u_0} = \frac{M(x) - u_0}{M(x_0) - u_0}. \quad (14')$$

Es zeigt sich hier, dass es vorteilhaft ist, bei der Verwendung der Hypothese A_1 neben $\lambda(x)$ und $g(u)$ nicht mehr $s(x)$, sondern $b(x)$ als dritte Fundamentalfunktion zu betrachten. $b(x)$ kann als «gekürzte» Skala s in bezug auf den Ursprung u_0 interpretiert werden.

11. Sind wir auf dem Umweg über das Postulat (10) zu unserer neuen Hypothese A_1 gelangt, so wollen wir doch noch zeigen, wie man durch eine andere Überlegung auf sie stossen kann: Nachdem wir gesehen haben, dass bei der Verwendung der Hypothese A die Globalverteilung dann einastig wird, wenn $g(u)$ im Intervall $0 \leq u < \infty$ erklärt ist, scheint es naheliegend, in unserm jetzigen Falle, wo wir $u_0 \leq u < \infty$ haben, die Verschiebung

$$u' = u - u_0 \tag{15}$$

vorzunehmen. Das hat nämlich zur Folge, dass wir an Stelle von $g(u)$ eine Häufigkeitsfunktion $\overset{*}{g}(u')$, normiert in $0 \leq u' < \infty$, als Ausgangsfunktion in den Kreis unserer Betrachtungen ziehen können. Dabei gibt $\overset{*}{g}(u')$ die Verteilung der um u_0 reduzierten Einkommen für die x_0 -Jährigen wieder. Wenden wir die Hypothese A auf $\overset{*}{g}(u')$ sinngemäss an, so ergibt sich die Beziehung

$$\overset{*}{f}(u'/x) = \frac{1}{\overset{*}{s}(x)} g \left[\frac{u'}{\overset{*}{s}(x)} \right]; \quad 0 \leq u' < \infty. \tag{16}$$

Jedem Alter x ist dadurch die Verteilung bezüglich der um u_0 verminderten Einkommen u zugewiesen. Da

$$\overset{*}{s}(x) = \frac{\overset{*}{M}(x)}{\overset{*}{M}(x_0)}$$

nicht unmittelbar gegeben ist, wollen wir abklären, wie es sich aus $s(x)$ bestimmen lässt. Es ist nicht schwer einzusehen, dass das Mittel $\overset{*}{M}(x)$ der reduzierten Einkommen u' gegenüber dem Mittel $M(x)$ der wirklichen Einkommen ebenfalls die Einbusse um den Betrag u_0 erleidet:

$$\overset{*}{M}(x) = M(x) - u_0.$$

Somit erhalten wir aber für $\overset{*}{s}(x)$ gerade den Ausdruck, den wir in (14') mit $b(x)$ bezeichnet haben. Machen wir jetzt noch die Verschiebung (15) rückgängig, so treten an Stelle der Schar (16) die Parallelverteilungen

$$f(u/x) = \frac{1}{b(x)} g \left[\frac{u - u_0}{b(x)} + u_0 \right]; \quad u_0 \leq u < \infty. \tag{17}$$

Ein Vergleich mit (14) zeigt, dass (17) mit der Hypothese A_1 identisch ist.

12. Wir sind nun in der Lage, die globale Einkommensverteilung $\bar{f}(u)$ in einen einzigen Integralausdruck zu fassen. Die zweidimensionale Häufigkeitsfunktion $\varphi(u, x) = \lambda(x) f(u/x)$ ist nämlich über einem Streifen definiert, welcher nach links durch die Gerade $u = u_0$ begrenzt ist; es gilt

$$\int_{u_0}^{\infty} \int_{x_0}^{x_1} \varphi(u, x) dx du = 1,$$

und für $\bar{f}(u)$ ergibt sich

$$\bar{f}(u) = \int_{x_0}^{x_1} \varphi(u, x) dx = \int_{x_0}^{x_1} \frac{\lambda(x)}{b(x)} g\left[\frac{u - u_0(1 - b(x))}{b(x)}\right] dx. \quad (18)$$

Die Globalverteilung ist einastig und Kontaktprobleme, wie sie sich bei Verwendung von Hypothese A stellen, existieren nicht.

13. Es bleibt festzustellen, inwieweit die Hypothese A_1 Anspruch auf Plausibilität erheben kann: Wir erinnern uns daran, dass die Annahme, $g(u) = f(u/x_0)$ sei in einem Intervall definiert, welches nach links durch $u_0 > 0$ begrenzt ist, durch das Vorhandensein eines absoluten Existenzminimums erklärt werden kann. Die Hypothese A_1 besagt dann im Gegensatz zur Hypothese A , dass dieses Mindesteinkommen, welches eine Person haben muss, um überhaupt leben zu können, ganz unabhängig vom Alter stets im Betrag u_0 besteht. Es will uns scheinen, diese Invarianz bezüglich des Alters sei durchaus zu vertreten. Dass dagegen das mittlere Einkommen vom Alter x abhängig ist, steht ausser Zweifel, und wir haben diesem Umstand gebührend Rechnung getragen durch die Beibehaltung der Skala s , resp. $b(x)$.

Betrachten wir endlich die Hypothese A_1 im Verein mit der Hypothese B , so ist leicht zu erkennen, dass die Gesamtheit der zum Zeitpunkt t_0 x_0 -Jährigen im Laufe der Zeit einem Einkommenssteigerungsgesetz unterworfen ist, welches mit $s(x)$ übereinstimmt. Dabei ist allerdings zu beachten, dass $s(x)$ nicht mehr individuell interpretiert werden kann, wie dies bei Anwendung der Hypothese A der Fall ist. Vom theoretischen Standpunkt aus gesehen stellt diese Tatsache unseres Erachtens keinen grossen Nachteil dar, da wir es mit einem Problem stochastischer Natur zu tun haben und derartige Probleme ja geradezu dadurch gekennzeichnet sind, dass das Geschehen nur kollektiv, nicht aber im einzelnen, bestimmt ist. Immerhin kommt der Dehnungs-

charakter, welcher der Hypothese A eigen ist, auch bei der Hypothese A_1 noch zum Ausdruck.

In den beiden folgenden Paragraphen zeigen wir, wie sich die Hypothese A_1 auswirkt, wenn für $g(u)$ die Häufigkeitsfunktionen der hyperbolischen Verteilung von Pareto und der semi-normalen Verteilung eingesetzt werden. Wir wählen zur Veranschaulichung gerade diese beiden Gesetze, weil einerseits der Paretoverteilung in bezug auf die Verteilung der Gesamteinkommen ganz besondere historische Bedeutung zukommt, andererseits die semi-normale Verteilung sich heute bei der Behandlung von Lohnschichtungsfragen als geeignet erweist (siehe z. B. E. Kaiser [8]).

§ 4. Anwendung der Hypothese A_1 auf die Paretoverteilung

14. $g(u)$ sei gegeben durch die *Häufigkeitsfunktion der klassischen Paretoverteilung*, d. h.

$$f(u/x_0) = g(u) = \alpha u_0^\alpha u^{-\alpha-1}; \quad \alpha > 1. \quad (19)$$

$g(u)$ ist normiert im Intervall $0 < u_0 \leq u < \infty$. Die Nebenbedingung $\alpha > 1$ rührt daher, dass wir die Existenz des Mittels $M(x_0) = \int_{u_0}^{\infty} u g(u) du$ fordern. Aus (19) ergibt sich nämlich sofort

$$M(x_0) = \frac{\alpha}{\alpha-1} u_0. \quad (20)$$

Gehen wir über zur Darstellung der Parallelverteilungen $f(u/x)$, so finden wir unter Verwendung von (14) und durch die Setzungen

$$\left. \begin{aligned} u_0 b(x) &= \beta(x) \\ u_0 [b(x) - 1] &= \bar{\beta}(x) \end{aligned} \right\} \quad (21)$$

die Beziehung $f(u/x) = \alpha \beta^\alpha(x) [u + \bar{\beta}(x)]^{-\alpha-1}. \quad (22)$

Dass jede Funktion der Schar $f(u/x)$ im gleichen Intervall $u_0 \leq u < \infty$ normiert ist, lässt sich leicht verifizieren:

$$J = \int_{u_0}^{\infty} \alpha \beta^\alpha(x) [u + \bar{\beta}(x)]^{-\alpha-1} du = - \left[\frac{\beta(x)}{u + \bar{\beta}(x)} \right]^\alpha \Big|_{u_0}^{\infty} = \left[\frac{\beta(x)}{u_0 + \bar{\beta}(x)} \right]^\alpha.$$

Ersetzen wir $\beta(x)$ und $\bar{\beta}(x)$ gemäss (21), so erhalten wir wirklich das erwartete Resultat $J = 1$.

15. Begrenzen wir das Integral J nach oben durch die Variable u , so erhalten wir die den Häufigkeitsfunktionen $f(u/x)$ zugeordneten Verteilungsfunktionen $F(u/x)$ in der Form

$$F(u/x) = 1 - \left[\frac{\beta(x)}{u + \bar{\beta}(x)} \right]^\alpha. \quad (23)$$

Bezeichnen wir die Komplementärfunktionen mit $H(u/x)$, d. h.

$$H(u/x) = \int_u^\infty f(v/x) dv = 1 - F(u/x),$$

so lesen wir aus (23), dass

$$H(u/x) = \left[\frac{\beta(x)}{u + \bar{\beta}(x)} \right]^\alpha. \quad (23')$$

Durch Logarithmieren dieser Gleichung können wir die zugehörige Schar der sogenannten Paretogeraden erhalten.

16. Die *globale Einkommensverteilung* endlich, gekennzeichnet durch die Häufigkeitsfunktion $\bar{f}(u)$, drückt sich folgendermassen aus:

$$\bar{f}(u) = \alpha \int_{x_0}^{x_1} \lambda(x) \beta^\alpha(x) [u + \bar{\beta}(x)]^{-\alpha-1} dx. \quad (24)$$

Ob dieses Integral geschlossen darstellbar ist, auf bekannte tabellierte Funktionen zurückgeführt werden kann oder numerisch ausgewertet werden muss, hängt von den Annahmen ab, die wir in bezug auf $\lambda(x)$, $b(x)$ und α treffen. Ist insbesondere α ganzzahlig, so genügt es, dass $\lambda(x)$ und $b(x)$ Polynome sind, damit $\bar{f}(u)$ in geschlossener Form angegeben werden kann. Die Ordnung der Polynome spielt dabei keine Rolle (vgl. H. v. Mangoldt [9], S. 69 ff.). Wir machen bei dieser Gelegenheit darauf aufmerksam, dass $b(x)$, $\beta(x)$ und $\bar{\beta}(x)$ sämtlich Polynome n -ter Ordnung sind, falls auch $s(x)$ ein solches ist. Die Wahl etwa einer linearen Skala s hat zur Folge, dass b , β und $\bar{\beta}$ ebenfalls lineare Funktionen des Alters x sind.

Es wird immer wieder und mit Recht darauf hingewiesen, dass sich das Gesetz von Pareto in der Form (19) nur zur Darstellung der mittleren und höheren Einkommen eignet, so dass wir hier u_0 nicht mehr als Existenzminimum interpretieren können. Liegt ein Problem

vor – sei es volkswirtschaftlicher, sozialpolitischer oder irgendwelcher Art, – das die Darstellung nur eben dieser mittleren und höheren Einkommen verlangt, so stellt u_0 ganz einfach die untere Grenze der zu erfassenden Einkommen dar. Die Eigenschaft der Hypothese A_1 , für jedes Alter x die Verteilung der Einkommen im gleichen u -Intervall $u_0 \leq u < \infty$ wiederzugeben, macht sich in einem solchen Fall besonders angenehm bemerkbar.

§ 5. Anwendung der Hypothese A_1 auf die semi-normale Verteilung

17. Unsere Fundamentalfunktion $g(u)$ sei diesmal gegeben durch eine *Pearson-Kurve vom Typus III*

$$g(u) = \frac{\gamma^\varepsilon}{\Gamma(\varepsilon)} (u - u_0)^{\varepsilon-1} e^{-\gamma(u-u_0)}; \quad \begin{array}{l} \varepsilon > 0 \\ \gamma > 0. \end{array} \quad (25)$$

Die Benennung «semi-normale Verteilung» geht auf J. F. Steffensen [10] zurück. Er gibt sie allerdings in einer etwas andern Form und zeigt, dass $g(u)$ für $\varepsilon \rightarrow \infty$ gegen eine Normalverteilung strebt. Ferner ist bekannt, dass (25) als Spezialfall die sogenannte χ^2 -Verteilung enthält. Auch können die semi-normalen Häufigkeitsfunktionen im Falle $\varepsilon \geq 1$ als eine Art Interpolation der Gesetze, welche zur Poissonschen Klasse gehören, aufgefasst werden, da für ganzzahliges ε : $\Gamma(\varepsilon) = (\varepsilon - 1)!$ Für weitere Besonderheiten verweisen wir auf die zitierte Arbeit von Steffensen. Wir begnügen uns damit, die Hypothese A_1 auf (25) anzuwenden. Die Parallelverteilungen $f(u/x)$ zeigen dann folgende Gestalt:

$$f(u/x) = \frac{\left(\frac{\gamma}{b(x)}\right)^\varepsilon}{\Gamma(\varepsilon)} (u - u_0)^{\varepsilon-1} e^{-\frac{\gamma}{b(x)}(u-u_0)}. \quad (26)$$

Wir bezeichnen noch $\frac{\gamma}{b(x)}$ mit $\gamma(x)$ und stellen fest, dass die Funktionen der angeschriebenen Schar nur insoweit vom Parameter x abhängig sind, als $\gamma(x)$ es ist. Dass u_0 festbleiben würde, war zu erwarten. Dagegen bleibt hervorzuheben, dass auch ε gegenüber dem Alter invariant ist. Letzteres gilt übrigens auch bei der Anwendung der Hypothese A , wie E. Kaiser gezeigt hat.

Jeder Häufigkeitsfunktion $f(u/x)$ ist eine Verteilungsfunktion $F(u/x)$ zugeordnet. Gemäss (26) präsentiert sich die Schar dieser Verteilungsfunktionen durch

$$F(u/x) = \frac{P[\gamma(x)(u-u_0), \varepsilon]}{\Gamma(\varepsilon)}. \quad (27)$$

Im Zähler steht die oben unvollständige Gammafunktion, wobei der erste Term in der eckigen Klammer die obere Integrationsgrenze bezeichnet. Das ausgezeichnete Tabellenwerk von K. Pearson [11] gestaltet die Berechnung von (27) einfach.

18. Gewichten wir die Parallelverteilungen (26) mit $\lambda(x)$, so liefert uns das die zweidimensionale Häufigkeitsfunktion $\varphi(u, x)$, normiert in $u_0 \leq u < \infty, x_0 \leq x \leq x_1$; hieraus gewinnen wir schliesslich die *globale Einkommensverteilung*

$$\bar{f}(u) = \frac{(u-u_0)^{\varepsilon-1}}{\Gamma(\varepsilon)} \int_{x_0}^{x_1} \lambda(x) \gamma^\varepsilon(x) e^{-\gamma(x)(u-u_0)} dx. \quad (28)$$

Erwähnenswert ist, dass bei der Durchführung der Rechnung auf der Basis der Hypothese A die Integralausdrücke für jeden einzelnen Ast von wesentlich komplizierterer Gestalt sind als bei unserer einastigen Verteilung. Dort ist der am einfachsten zu berechnende «mittlere» Ast \bar{f}_m (vgl. erstes Kapitel, Figur 2, Ast 3) nämlich gegeben durch

$$\bar{f}_m(u) = \frac{1}{\Gamma(\varepsilon)} \int_{x_0}^{x_1} \lambda(x) \bar{\gamma}^\varepsilon(x) (u-u_0(x))^{\varepsilon-1} e^{-\bar{\gamma}(x)(u-u_0(x))} dx,$$

wo

$$u_0(x) = u_0 s(x) \quad \text{und} \quad \bar{\gamma}(x) = \frac{\gamma}{s(x)}.$$

E. Kaiser weist denn auch darauf hin, dass dieses und die übrigen Integrale, welche einen Beitrag an die Globalverteilung liefern, selbst im Fall, wo $\lambda(x)$ und $s(x)$ linear sind, nicht mehr auf die Form klassischer Funktionen gebracht werden können, sich insbesondere nicht durch unvollständige Gammafunktionen ausdrücken lassen. Wird dagegen $\bar{f}(u)$ auf Grund der Hypothese A_1 bestimmt, d. h. durch die Beziehung (28) dargestellt, so besteht bei besonders günstigen Annahmen bezüglich $\lambda(x)$ und $s(x)$, bzw. $b(x)$, die Möglichkeit des Zurückführens auf die genannten Funktionen.

19. Diese Zurückführung gelingt stets dann, wenn für die Skala s und die Altersstruktur λ in erster Näherung *lineare Funktionen eingesetzt werden*, was wir hier nachweisen wollen. Es sei

$$s(x) = 1 + a(x - x_0), \quad \text{so dass} \quad s(x_0) = 1, \quad (29)$$

eine Annahme, wie sie auch bei den ersten Studien des finanziellen Gleichgewichts der AHV getroffen wurde; vgl. E. Kaiser [8]. Laut (14') ist dann

$$b(x) = 1 + A(x - x_0), \quad \text{wobei} \quad A = \frac{a M(x_0)}{M(x_0) - u_0}. \quad (29')$$

$M(x_0)$ kann natürlich durch ε , γ und u_0 ausgedrückt werden, und zwar in der folgenden Weise:

$$M(x_0) = \frac{\varepsilon}{\gamma} + u_0,$$

so dass der Zusammenhang zwischen A und a auch dargestellt werden kann durch

$$A = a \left(1 + \frac{\gamma}{\varepsilon} u_0 \right).$$

Ferner sei

$$\lambda(x) = \frac{cx + d}{\int_{x_0}^{x_1} (cx + d) dx} = Cx + D \quad (30)$$

(C wird in der Regel negativ sein).

Lassen wir die getroffenen Annahmen unter Berücksichtigung von

$$\gamma(x) = \frac{\gamma}{b(x)} = \frac{\gamma}{Ax + E}, \quad \text{wo } E = 1 - Ax_0$$

gemäss (29'), in (28) eingehen, so bekommt $\bar{f}(u)$ die Gestalt

$$\bar{f}(u) = \frac{(u - u_0)^{\varepsilon-1}}{\Gamma(\varepsilon)} \int_{x_0}^{x_1} (Cx + D) \left(\frac{\gamma}{Ax + E} \right)^{\varepsilon} e^{-\frac{\gamma(u-u_0)}{Ax+E}} dx. \quad (31)$$

Die Substitution

$$\frac{\gamma(u - u_0)}{Ax + E} = y$$

scheint hier angebracht zu sein.

Dadurch folgt $Cx + D = \frac{1}{A} \left\{ \frac{C \gamma(u - u_0)}{y} + AD - CE \right\},$

$$\left(\frac{\gamma}{Ax + E} \right) = \left(\frac{y}{u - u_0} \right)^\varepsilon$$

und $dx = -\frac{\gamma(u - u_0)}{A} \cdot \frac{dy}{y^2}.$

Nach möglichster Vereinfachung ist dann

$$\bar{f}(u) = \frac{\gamma}{A^2 \Gamma(\varepsilon)} \int_{y_0}^{y_1} \{CE - AD - C \gamma(u - u_0) y^{-1}\} y^{\varepsilon-2} e^{-y} dy$$

und hieraus durch Aufteilen

$$\bar{f}(u) = \frac{\gamma}{A^2 \Gamma(\varepsilon)} \left\{ (CE - AD) \int_{y_0}^{y_1} y^{\varepsilon-2} e^{-y} dy - C \gamma(u - u_0) \int_{y_0}^{y_1} y^{\varepsilon-3} e^{-y} dy \right\}, \quad (32)$$

wobei $y_0 = \gamma(u - u_0); \quad y_1 = \gamma(x_1) (u - u_0).$

Ist $\varepsilon > 2$, so lassen sich die beiden Integrale je durch die Differenz zweier oben unvollständiger Gammafunktionen ausdrücken. Wir erhalten

$$\begin{aligned} \bar{f}(u) = \frac{\gamma}{A^2 \Gamma(\varepsilon)} \{ & (CE - AD) \{ P[\gamma(x_1)(u - u_0), \varepsilon - 1] - P[\gamma(u - u_0), \varepsilon - 1] \} \\ & - C \gamma(u - u_0) \{ P[\gamma(x_1)(u - u_0), \varepsilon - 2] - P[\gamma(u - u_0), \varepsilon - 2] \} \}. \end{aligned} \quad (33)$$

Die Diskussion von (25) lehrt, dass die Bedingung $\varepsilon > 2$ gleichwertig ist der Forderung, $g(u)$ solle in $u = u_0$ eine horizontale Tangente haben. Nur in einem solchen Fall ist (33) verwendbar.

Streben wir eine Darstellung an, welche auch für das ε -Intervall $0 < \varepsilon \leq 2$ Gültigkeit hat, so müssen wir uns der unten unvollständigen Gammafunktion

$$Q[z, \alpha] = \int_z^\infty \zeta^{\alpha-1} e^{-\zeta} d\zeta \quad \begin{array}{l} z > 0 \\ -\infty < \alpha < \infty \end{array}$$

bedienen. Die Beziehung (32) nimmt dann folgende Gestalt an:

$$\begin{aligned} \bar{f}(u) = \frac{\gamma}{A^2 \Gamma(\varepsilon)} \{ & (CE - AD) \{ Q[\gamma(u - u_0), \varepsilon - 1] - Q[\gamma(x_1)(u - u_0), \varepsilon - 1] \} \\ & - C \gamma(u - u_0) \{ Q[\gamma(u - u_0), \varepsilon - 2] - Q[\gamma(x_1)(u - u_0), \varepsilon - 2] \} \}. \end{aligned} \quad (34)$$

Die Bestimmung des Mittels

$$\bar{M} = \int_{x_0}^{x_1} \lambda(x) M(x) dx = M(x_0) \int_{x_0}^{x_1} \lambda(x) s(x) dx$$

läuft hier auf die Integration einer quadratischen Parabel hinaus.

Damit haben wir unsern Spezialfall in seiner vollen Allgemeinheit gelöst und wollen das Resultat der Rechnung in der Aussage festhalten: Ist $f(u/x_0) = g(u)$ durch eine semi-normale Verteilung gegeben und werden die Parallelverteilungen $f(u/x)$ auf Grund der Hypothese A_1 bestimmt, so kann die Globalverteilung $\bar{f}(u)$ jedenfalls dann durch unvollständige Gammafunktionen ausgedrückt werden, wenn die beiden andern Fundamentalfunktionen $\lambda(x)$ und $s(x)$ linear sind. Unter diesen Voraussetzungen und für $\varepsilon > 2$ kann $\bar{f}(u)$ sowohl durch (33) als auch durch (34) dargestellt werden, wogegen für $\varepsilon \leq 2$ nur mehr durch (34).

Zur numerischen Bestimmung der globalen Häufigkeiten kann neben der bereits erwähnten Tafel von K. Pearson die Thalmantabelle der Prymschen Funktion [12] herangezogen werden. Gegebenenfalls wird uns die (evtl. iterierte) Anwendung der Rekursionsformeln

$$\left. \begin{aligned} P[z, \alpha + 1] &= \alpha P[z, \alpha] - z^\alpha e^{-z} \\ Q[z, \alpha + 1] &= \alpha Q[z, \alpha] + z^\alpha e^{-z} \end{aligned} \right\} \quad (35)$$

gestatten, in das Tabellierungsintervall zu gelangen. Besonders zu beachten ist allerdings, dass für $\varepsilon = n$ ($n = 1, 2$) die Q -Rekursion in der gewünschten Richtung – vom kleineren zum grösseren Exponenten – nur bis zu

$$Q[y_m, 0] = \int_{y_m}^{\infty} \frac{e^{-y}}{y} dy; \quad m = 0, 1$$

wirksam ist. Für die zahlenmässige Ermittlung von $\bar{f}(u)$ haben wir dann im wesentlichen dieses sogenannte Exponentialintegral auszuwerten, eine transzendente Funktion, welche wir z. B. bei Jahnke und Emde [13] (S. 6 ff.) tabelliert finden. Interessanterweise zeichnen sich auch diese beiden Fälle schon bei der Diskussion der Fundamentalfunktion $g(u)$, gegeben durch (25), aus. Nur gerade für $\varepsilon = 1, 2$ nämlich hat die Tangente in $u = u_0$ ein von Null verschiedenes endliches Steigungsmass.

III. Kapitel

Zur gleichen Parameterklasse gehörige Schar von Einkommensverteilungen $f(u/x)$

§ 1. Die Zugehörigkeit der Schar $f(u/x)$ zur gleichen Parameterklasse im allgemeinen Fall

20. Wir werden in diesem Kapitel noch eine Modifikation der Hypothese A kennenlernen, welche als Verallgemeinerung der *Hypothese* A_1 angesehen werden kann. Dabei ist es vielleicht angebracht, einmal grundsätzlich festzuhalten, von welchen Gedanken wir uns eigentlich leiten lassen müssen, wenn wir mit Hilfe der beiden Fundamentalfunktionen $g(u)$ und $s(x)$ durch Einschalten einer plausiblen Hypothese die gesamte Schar der $f(u/x)$ bestimmen sollen. Dass die Einführung einer Hypothese eine unumgängliche Notwendigkeit darstellt, ist an sich klar, sind doch die Funktionen der gesuchten Schar durch die Vorgabe der genannten Grundfunktionen in keiner Weise bestimmt. Einem festen Alter $x \neq x_0$ könnte man nämlich vorerst eine ganz beliebige Häufigkeitsfunktion zuordnen, wobei einzig darauf zu achten wäre, dass die Bedingung $M(x) = s(x) M(x_0)$ realisiert sein muss, was ganz einfach durch die Anpassung der Kurvenparameter bewerkstelligt werden könnte. Die Einschränkung der Wahlfreiheit bestünde nur eben darin. Diese übergrösse Freiheit aber müssen wir vermeiden, denn schon bei oberflächlicher Betrachtung des praktischen Problems ist es evident, dass alle Funktionen der Schar $f(u/x)$ untereinander eine gewisse «Ähnlichkeit» (im landläufigen Sinn) aufweisen müssen. Und diese Tatsache ist es letztlich, welche uns die aufzustellende Hypothese suggerieren soll. Wie kann nun die geforderte «Ähnlichkeit» mathematisch ausgedrückt werden? Wohl sicher einmal durch die Postulierungen, dass sämtliche Scharcurven zum gleichen Typ (Hypothese A), oder zur gleichen Klasse (Hypothese A_1), gehören sollen.

21. Wir können jedoch eine noch allgemeinere Annahme zur Sprache bringen: Sämtliche Funktionen $f(u/x)$ sind Repräsentanten ein und derselben Parameterklasse. Das ist so zu verstehen, dass wir bei der vorgegebenen Häufigkeitsfunktion $f(u/x_0)$ irgendeinen Parameter —

nicht mehr speziell die Präzision – auszeichnen und dies dokumentieren durch die Schreibweise

$$f(u/x_0) = g(u, c),$$

wobei wie bis anhin gelten soll

$$\int_{u_0}^{\infty} g(u, c) du = 1.$$

Der Übergang zur Häufigkeitsverteilung für das Alter x besteht jetzt einfach darin, dass an die Stelle von c ein Parameter $c(x)$ so gesetzt wird, dass die Einhaltung der vorgegebenen Skala s gewährleistet ist.

§ 2. Spezialfall: Die Hypothese A_2

22. Bis hierher wurde noch nicht gefordert, dass $u_0(x) = u_0$ sei. Diese Bedingung führen wir jetzt ein. Dann verteilt sich die statistische Masse stets auf das gesamte Intervall $u_0 \leq u < \infty$. Wir haben gesehen, dass sich die letztgenannte Forderung auf die Bestimmung der globalen Einkommensverteilung günstig auswirkt, wollen aber vielleicht gerade diesen Vorteil nicht zum Leitmotiv für die Invarianzforderung – u_0 von x unabhängig – wählen. Es ist der Problemlage eher angemessen, wenn wir von den beiden Interpretationsmöglichkeiten für u_0 , welche wir kennengelernt haben, ausgehen und ganz einfach feststellen, dass der Spezialfall der Annahme einer *konstanten untern Grenze natürlich ist*. Dass sich dann überdies die Darstellung der Globalverteilung einfacher gestaltet, sehen wir als willkommene Beigabe an.

Aus dem Gesagten ist jetzt ersichtlich, wie sich die *Hypothese A_2* präsentieren wird:

$$f(u/x) = g[u, c(x)]; \quad c(x_0) = c \quad (36)$$

mit den beiden Nebenbedingungen

$$\int_{u_0}^{\infty} g[u, c(x)] du = 1 \quad (36')$$

und

$$M(x) = \int_{u_0}^{\infty} u g[u, c(x)] du = s(x) \int_{u_0}^{\infty} u g(u, c) du = s(x) M(x_0). \quad (36'')$$

Unmittelbar folgt: $\varphi(u, x) = \lambda(x) g[u, c(x)]$ (37)
 mit $\int_{u_0}^{\infty} \int_{x_0}^{x_1} \lambda(x) g[u, c(x)] dx du = 1$ }

und $\bar{f}(u) = \int_{x_0}^{x_1} \lambda(x) g[u, c(x)] dx$. (38)

Es ist leicht einzusehen, dass $\int_{u_0}^{\infty} \bar{f}(u) du = 1$ ist. Da der Zusammenhang zwischen $s(x)$ und $c(x)$ nicht in expliziter Form angegeben werden kann, wollen wir die Sachlage an zwei Beispielen erörtern. Wir wählen dazu wieder die Pareto- und die semi-normale Verteilung, wodurch uns Vergleichsmöglichkeiten mit dem Vorangegangenen erwachsen werden.

§ 3. Anwendung der Hypothese A_2 auf die Paretoverteilung

23. Wir nehmen an, die Einkommensverteilung für die Gesamtheit der x_0 -Jährigen sei gegeben durch die Beziehung (19). Bei der Auszeichnung eines Parameters bleibt uns keine Wahl; es kommt nur α in Frage, da u_0 vom Alter x unabhängig sein soll. Die Schar der Parallelverteilungen hat dann gemäss der Hypothese A_2 die Form

$$f(u/x) = g[u, \alpha(x)] = \alpha(x) u_0^{\alpha(x)} u^{-\alpha(x)-1}; \alpha(x) > 1, \quad (39)$$

wobei $\alpha(x)$ eindeutig bestimmt ist durch die Nebenbedingung (36''). Zu seiner Berechnung haben wir die Gleichung

$$M(x) = \frac{\alpha(x)}{\alpha(x) - 1} u_0$$

nach $\alpha(x)$ aufzulösen, was

$$\alpha(x) = \frac{M(x)}{M(x) - u_0} = \frac{s(x) M(x_0)}{s(x) M(x_0) - u_0} = \frac{\alpha s(x)}{\alpha s(x) - \alpha + 1} \quad (39')$$

ergibt.

Es besteht natürlich die Möglichkeit, nicht mehr $s(x)$, sondern $\alpha(x)$ als Fundamentalfunktion aufzufassen. Wir stellen uns das etwa so vor, dass man ein vorliegendes statistisches Material für einige Alter x_i

nach Pareto ausgleicht und durch die auf solche Weise bestimmten $\alpha(x_i)$ eine glatte Kurve legt. Die Skala s kann man erhalten durch Auflösen von (39').

Die Anschrift der anderen zur Diskussion stehenden Funktionen in der (u, x) -Ebene erübrigt sich, da die formale Darstellung keine Schwierigkeiten bereitet. Die Globalverteilung ist natürlich einastig, wird aber im allgemeinen nicht in geschlossener Form darstellbar sein; dagegen ist leicht nachzuweisen, dass $\bar{f}(u)$ stets durch unvollständige Gammafunktionen ausgedrückt werden kann, falls $\lambda(x)$ und $\alpha(x)$ linear sind.

24. Dass wir hier einer ganz andern Situation gegenüberstehen als im Falle, wo die Hypothese A_1 zur Anwendung kommt, wollen wir an einem einfachen Zahlenbeispiel illustrieren:

Als Einheit betrachten wir die untere Grenze der zu erfassenden Einkommen, d. h. $u_0 = 1$. Das Mittel für die x_0 -Jährigen sei doppelt so gross: $M(x_0) = 2$. Berechnen wir $f(u/x)$ für ein bestimmtes Alter x' , dem $s(x') = 1,5$ zugeordnet sei, so ergibt sich unter Verwendung von

$$a) \text{ Hypothese } A_1: f_1(u/x') = 8(u + 1)^{-3},$$

$$b) \text{ Hypothese } A_2: f_2(u/x') = \frac{3}{2} u^{-\frac{5}{2}}.$$

Beide Funktionen, f_1 und f_2 , sind hyperbolischer Natur, im Intervall $1 \leq u < \infty$ normiert und weisen das gleiche Mittel $M(x') = 3$ auf.

§ 4. Anwendung der Hypothese A_2 auf die semi-normale Verteilung

25. Stellen wir die Fundamentalfunktion $f(u/x_0)$ durch eine semi-normale Verteilung dar, wie sie durch (25) gegeben ist, so können wir entweder γ oder ε als veränderlichen Parameter betrachten. Wählen wir vorerst einmal γ , so bestimmt sich $\gamma(x)$ auf Grund von

$$M(x) = \frac{\varepsilon}{\gamma(x)} + u_0$$

zu

$$\gamma(x) = \frac{\varepsilon}{M(x) - u_0}.$$

Da aber gilt

$$\varepsilon = \gamma(M(x_0) - u_0),$$

so kann $\gamma(x)$ geschrieben werden als

$$\gamma(x) = \gamma \frac{M(x_0) - u_0}{M(x) - u_0},$$

und unter Berücksichtigung von (14') erhalten wir schliesslich

$$\gamma(x) = \frac{\gamma}{b(x)}. \quad (40)$$

Ein Vergleich mit (26) klärt uns sofort darüber auf, dass die Hypothese A_2 in diesem Falle zum genau gleichen Resultat führt wie die Hypothese A_1 , was uns nicht verwundert, vielmehr durch eine einfache Eindeutigkeitsbetrachtung vorauszusehen war.

26. Wesentlich anders verhält es sich, wenn wir ε auszeichnen. Die Schar der Parallelverteilungen wird dann dargestellt durch

$$f(u/x) = \frac{\gamma^{\varepsilon(x)}}{\Gamma[\varepsilon(x)]} (u - u_0)^{\varepsilon(x)-1} e^{-\gamma(u-u_0)}; \quad \begin{array}{l} \varepsilon(x) > 0 \\ \gamma > 0 \end{array} \quad (41)$$

wo

$$\varepsilon(x) = \gamma(M(x) - u_0) = \varepsilon \frac{M(x) - u_0}{M(x_0) - u_0} = \varepsilon b(x). \quad (41')$$

Die eigentlichen Verteilungsfunktionen können durch die Beziehung

$$F(u/x) = \frac{P[\gamma(u - u_0), \varepsilon(x)]}{\Gamma[\varepsilon(x)]} \quad (42)$$

sehr einfach ausgedrückt werden.

Was endlich die globalen Häufigkeiten anbetrifft, so sind sie gegeben durch die Funktion

$$\bar{f}(u) = e^{-\gamma(u-u_0)} \int_{x_0}^{x_1} \lambda(x) \frac{\gamma^{\varepsilon b(x)}}{\Gamma[\varepsilon b(x)]} (u - u_0)^{\varepsilon b(x)-1} dx. \quad (43)$$

Trotzdem dieser Ausdruck recht kompliziert gebaut ist, gestaltet sich die numerische Auswertung nicht allzu schwierig. Sie lässt sich relativ leicht durchführen durch Gewichten der Parallelverteilungen (41) mit $\lambda(x)$.

Ein besonderes Augenmerk ist darauf zu richten, dass unter Umständen – je nach der Wahl von ε und der Beschaffenheit der Skala s – die geforderte «Ähnlichkeit» der Scharkurven nicht mehr gewährleistet ist (vgl. etwa die Ausführungen über die Abhängigkeit des Verlaufs einer semi-normalen Häufigkeitsfunktion von ε bei E. Kaiser [2], S. 296). Allen Eventualitäten können wir jedenfalls dann begegnen, wenn auf Grund des vorliegenden statistischen Materials $\varepsilon(x) > 2$ gesetzt werden darf. Sonst aber müssen wir die Verhältnisse von Fall zu Fall prüfen.

§ 5. Streuungsvergleiche

27. In Theorie und Praxis der mathematischen Statistik spielt die sogenannte mittlere quadratische Abweichung, das ist das Moment zweiter Ordnung bezogen auf das Mittel M , eine bedeutende Rolle. Wir bezeichnen sie mit \mathfrak{M}_2 . Im Gegensatz dazu drücken wir das zweite Moment in bezug auf den Ursprung durch M_2 aus. Die positive Wurzel aus \mathfrak{M}_2 wird Streuung σ (standard deviation) genannt. (A. Linder [14] verwendet die Benennung «Streuung» sowohl für \mathfrak{M}_2 als auch für $\sqrt{\mathfrak{M}_2}$). Durch die Vorgabe der Fundamentalfunktion $f(u/x_0)$ darf $\mathfrak{M}_2(x_0)$ als bekannt vorausgesetzt werden. Wir wollen aufzeigen, wie sich bei Anwendung der Hypothesen A , A_1 und A_2 jeweils $\mathfrak{M}_2(x)$ als Funktion von $\mathfrak{M}_2(x_0)$ und $s(x)$, resp. $b(x)$, berechnen lässt und so die Auswirkung der Hypothesen in dieser Hinsicht vergleichen.

Für die *Hypothese A* hat E. Kaiser nachgewiesen, dass allgemein gilt:

$$M_k(x) = s^k(x) M_k(x_0),$$

also insbesondere

$$M_2(x) = s^2(x) M_2(x_0).$$

Hieraus aber folgt unter Verwendung der wohlbekannteren und leicht herzuleitenden Relation $\mathfrak{M}_2 = M_2 - M^2$:

$$\mathfrak{M}_2(x) = s^2(x) [M_2(x_0) - M^2(x_0)] = s^2(x) \mathfrak{M}_2(x_0). \quad (44)$$

Wird dagegen $f(u/x)$ auf Grund der *Hypothese A₁* bestimmt, so ist $\mathfrak{M}_2(x)$ durch

$$\mathfrak{M}_2(x) = \int_{u_0}^{\infty} \frac{[u - M(x)]^2}{b(x)} g \left[\frac{u - u_0(1 - b(x))}{b(x)} \right] du$$

dargestellt.

Substituieren wir $\frac{u - u_0(1 - b(x))}{b(x)} = v; \quad du = b(x) dv,$

so ergibt sich vorerst

$$\mathfrak{M}_2(x) = \int_{v_0=u_0}^{\infty} \{v b(x) + u_0[1 - b(x)] - M(x)\}^2 g(v) dv.$$

Lösen wir (14') nach $M(x)$ auf, so stoßen wir auf die Beziehung

$$M(x) = b(x) M(x_0) + u_0[1 - b(x)].$$

Dadurch lässt sich der Ausdruck in der geschweiften Klammer wesentlich vereinfachen. Es ist

$$\{ \}^2 = \{b(x)[v - M(x_0)]\}^2 = b^2(x)[v - M(x_0)]^2,$$

so dass sich $\mathfrak{M}_2(x)$ berechnet zu

$$\mathfrak{M}_2(x) = b^2(x) \int_{u_0}^{\infty} [v - M(x_0)]^2 g(v) dv = b^2(x) \mathfrak{M}_2(x_0). \quad (45)$$

Die *Hypothese* A_2 endlich lässt eine explizite Darstellung für $\mathfrak{M}_2(x)$ nicht zu.

Die gewonnenen Resultate stellen wir unter Einschluss der Anwendung auf die Pareto- und die semi-normale Verteilung in der folgenden Tabelle zusammen:

$\sigma(x) = \sqrt{\mathfrak{M}_2(x)}$			
	Hypothese A	Hypothese A_1	Hypothese A_2
	$s(x) \sigma(x_0)$	$b(x) \sigma(x_0)$	$\sqrt{\int_{u_0}^{\infty} u^2 g[u, c(x)] du - s^2(x) M^2(x_0)}$
Pareto ($\alpha > 2$)	$\frac{u_0 s(x)}{\alpha - 1} \sqrt{\frac{\alpha}{\alpha - 2}}$	$\frac{u_0 b(x)}{\alpha - 1} \sqrt{\frac{\alpha}{\alpha - 2}}$	$\frac{u_0 s(x)}{\alpha - 1} \sqrt{\frac{\alpha^3 [s(x) - 1]^2 + 2\alpha^2 [s(x) - 1] + \alpha}{s(x) \{2\alpha - \alpha s(x) - 2\}}}$
Pearson III	$\frac{s(x) \sqrt{\varepsilon}}{\gamma}$	$\frac{b(x) \sqrt{\varepsilon}}{\gamma}$	$\frac{b(x) \sqrt{\varepsilon}}{\gamma}$ (Anpassungsparameter γ) $\frac{\sqrt{\varepsilon b(x)}}{\gamma}$ (Anpassungsparameter ε)

28. Wie wir leicht zeigen können, führt die «Typenhypothese» A bei der aktiven Bevölkerung, für welche $M(x) > M(x_0)$, zu einer kleineren Streuung als die «Klassenhypothese» A_1 , sofern $u_0 > 0$. Aus der Quotientenbildung

$$\frac{b(x)}{s(x)} = \frac{M(x) M(x_0) - u_0 M(x_0)}{M(x) M(x_0) - u_0 M(x)}$$

lässt sich nämlich herauslesen, dass für $u_0 > 0$

$$\frac{b(x)}{s(x)} \underset{(\geq)}{\underset{(\leq)}{>}} 1 \quad \text{für} \quad M(x_0) \underset{(\leq)}{\underset{(\geq)}{<}} M(x),$$

woraus aber sofort folgt

$$b(x) \sigma(x_0) \underset{(\geq)}{\underset{(\leq)}{>}} s(x) \sigma(x_0) \quad \text{für} \quad M(x_0) \underset{(\leq)}{\underset{(\geq)}{<}} M(x). \quad (46)$$

Ist $u_0 = 0$, so gilt stets das Gleichheitszeichen, was nicht verwunderlich ist, sind doch in diesem Grenzfall die Hypothesen A und A_1 identisch. Dagegen besteht für $u_0 > 0$ die Gleichheit nur gerade im trivialen Fall $M(x) = M(x_0)$.

Ein analoger allgemeiner Vergleich mit der Hypothese A_2 ist naturgemäss nicht möglich. Die Situation für die beiden speziellen Anwendungsbeispiele ist aus der vorstehenden Tabelle ersichtlich.

In diesem Zusammenhang wollen wir noch auf folgenden Umstand hinweisen: Bei der Besprechung der Hypothese A haben wir $s(x)$ als Präzision bezeichnet und meinten damit den Maßstab relativ zu demjenigen, welcher $f(u/x_0) = g(u)$ zugrunde liegt. Aus $\sigma(x) = s(x) \sigma(x_0)$ ist ersichtlich, dass $s(x)$ ebensogut als Streuung betrachtet werden kann, falls wir $\sigma(x_0)$ als deren Einheit wählen. Diese enge Beziehung zwischen Maßstab und mittlerer quadratischer Abweichung ist nicht unbekannt und das Gesagte kann sinngemäss auch auf Hypothese A_1 und $b(x)$ übertragen werden.

IV. Kapitel

Erhaltungsfragen

§ 1. Formulierung des Hauptproblems

29. Wir wollen annehmen, die Parallelverteilungen würden auf Grund der Hypothese A_2 bestimmt, gemäss der die Häufigkeitsfunktionen $f(u/x)$ zur gleichen Parameterklasse gehören. Es drängt sich die Frage auf, ob dann $\bar{f}(u)$ – durch (38) dargestellt – ebenfalls ein Vertreter derselben Parameterklasse sei. Ein Rückblick auf die im vorigen Kapitel durchgerechneten Beispiele klärt uns sofort darüber auf, dass dies im allgemeinen nicht zutrifft. Es bleibt die Frage offen, ob die «Erhaltung der Parameterklasse» wenigstens in speziellen Fällen gewährleistet ist, d. h. dann, wenn unsere Parallelverteilungen zu einer ganz bestimmten, noch zu charakterisierenden Parameterklasse gehören. Die Abklärung dieser Frage soll uns im folgenden beschäftigen. Natürlich treten bei Zugrundelegung der Hypothesen A und A_1 entsprechende Probleme auf, welche aber ohne nennenswerte Schwierigkeiten gelöst werden können, wenn die eben gestellte Frage einmal beantwortet ist. Deshalb dürfen wir diese mit gutem Recht als Hauptproblem bezeichnen.

30. Die *exakte Formulierung des Hauptproblems* lautet: Welchen Bedingungen muss $g(u, c)$ genügen, damit bei vorgegebenem $s(x)$ die Integralrelation

$$\bar{f}(u) = \int_{x_0}^{x_1} \lambda(x) g[u, c(x)] dx = g(u, m); m = m(\lambda) \quad (47)$$

simultan für «alle» λ zu Recht besteht, d. h. für die der nachstehenden Voraussetzung (III) genügenden λ ? Es ist zu beachten, dass die Bedeutung der auftretenden Funktionen und die im ersten Kapitel gestellten allgemeinen Anforderungen an $\varphi(u, x)$ folgende *Voraussetzungen* einschliessen:

(I) $g(u, c)$ im Gültigkeitsbereich $u_0 \leq u < \infty$ stetig, daselbst nicht-

negativ; $\int_{u_0}^{\infty} g[u, c(x)] du = 1,$

(II) $c(x)$ stetig,

(III) $\lambda(x) \geq 0$, stetig; $\int_{x_0}^{x_1} \lambda(x) dx = 1.$

§ 2. Lösung des Hauptproblems

31. Fall 1: Es sei $m(\lambda) \equiv m_0$. Unter Beachtung von (III) lässt sich die Funktionalgleichung (47) auf die Form

$$(a) \quad \int_{x_0}^{x_1} \lambda(x) \{g[u, c(x)] - g(u, m_0)\} dx = 0$$

bringen. Diese Gleichung muss für alle zulässigen λ identisch erfüllt sein. Dabei ist wesentlich, dass die Wahl von λ im betrachteten Fall keinen Einfluss hat auf den Ausdruck in der geschweiften Klammer. Dieser kann als *Funktion von x allein* aufgefasst werden, wobei u die Rolle eines Parameters spielt. Indem wir uns auf eine in der Variationsrechnung oft geübte Schlussweise berufen, folgern wir:

$$(b) \quad g[u, c(x)] - g(u, m_0) \equiv 0 \quad \text{für alle } x.$$

Für Leser, denen die erwähnte Schlussweise nicht geläufig sein sollte, skizzieren wir den *Beweis für (b)* grob: Wir formulieren die Gegenannahme so, dass $\gamma(u, x) = g[u, c(x)] - g(u, m_0)$ etwa an der Stelle $x = x_p$ positiv sei. Da $\gamma(u, x)$ nach Voraussetzung stetig ist, lässt sich eine ganze Umgebung von x_p angeben, in welcher $\gamma(u, x)$ auch noch positiv ist. Bei geeigneter Wahl von λ (nadelförmig) liefert dann nur gerade diese Umgebung einen «wesentlichen» Beitrag an den Integralwert, d. h. die übrigen Anteile können beliebig klein gemacht werden. Durch eine solche Konstruktion wird aber das Integral (a) positiv, womit die Gegenannahme auf einen Widerspruch führt. $\gamma(u, x)$ muss tatsächlich identisch verschwinden.

Aus (b) folgt
$$c(x) \equiv c(x_0) = m_0$$

oder in Worten: Soll (47), mit der Nebenbedingung $m(\lambda)$ konstant, für «alle» λ Gültigkeit besitzen, so muss die Skala s notwendig konstant $s(x) \equiv s(x_0) = 1$ sein. Dass diese Bedingung auch hinreichend ist, ist trivial.

32. Fall 2: Es sei $m(\lambda)$ nicht konstant. Wir wählen $\lambda_1(x)$ und $\lambda_2(x)$ so, dass

$$g(u, m_1) = \int_{x_0}^{x_1} \lambda_1(x) g[u, c(x)] dx,$$

$$g(u, m_2) = \int_{x_0}^{x_1} \lambda_2(x) g[u, c(x)] dx,$$

wo $m_1 \neq m_2$.

Aus der Bildung

$$\frac{\alpha g(u, m_1) + \beta g(u, m_2)}{\alpha + \beta} = \int_{x_0}^{x_1} \left\{ \frac{\alpha \lambda_1(x) + \beta \lambda_2(x)}{\alpha + \beta} \right\} g[u, c(x)] dx$$

mit $\alpha, \beta > 0$ folgt unter Berücksichtigung von (47) und der Voraussetzung (III) unmittelbar die Beziehung

$$(c) \quad \frac{\alpha g(u, m_1) + \beta g(u, m_2)}{\alpha + \beta} = g[u, m(\alpha, \beta)].$$

D. h. aber: g hat die Eigenschaft, dass mit $g(u, m_1)$ und $g(u, m_2)$ zugleich jede normierte Linearkombination zur gleichen Parameterklasse gehört.

Setzen wir etwa

$$\alpha + \beta = p + q = 1,$$

so geht (c) über in

$$p g(u, m_1) + q g(u, m_2) = g[u, m(p)].$$

Gemäss unserer Setzung schreibt sich diese Gleichung auch als

$$(d) \quad p \{g(u, m_1) - g(u, m_2)\} + g(u, m_2) = g[u, m(p)].$$

Die richtige Interpretation von (d) gestattet uns, zu folgern, dass die gesuchte Funktion g notwendig von der Form

$$g(u, c) = g_1(c) g_2(u) + g_3(u) \tag{48}$$

sein muss.

Es ist zu diskutieren, inwieweit (48) Lösung der Funktionalgleichung (47) ist. Wohl verifiziert man durch Einsetzen, dass die notwendige Bedingung (48) auch hinreichend ist zur Erfüllung besagter Funktionalgleichung. Die Voraussetzung (I) verlangt aber überdies, dass für alle x gelte:

$$\int_{u_0}^{\infty} g[u, c(x)] du = g_1[c(x)] \int_{u_0}^{\infty} g_2(u) du + \int_{u_0}^{\infty} g_3(u) du \equiv 1.$$

Diese Forderung ist aber nur bei Vorgabe einer trivialen Skala s erfüllt, d. h. wenn $s(x) \equiv 1$. Mit dieser letzten Feststellung ist auch der Fall 2 abgeklärt, und wir können zusammenfassend festhalten:

Werden die Häufigkeitsfunktionen $f(u|x)$ auf Grund der Hypothese A_2 bestimmt, so gehört die Globalverteilung dann und nur dann für «alle» λ zur gleichen Parameterklasse wie die Parallelverteilungen, falls gilt: $s(x) \equiv s(x_0) = 1$.

Wir werden aber sehen, dass (48) bei der Behandlung eines Spezialproblems auch im Falle einer nichttrivialen Skala s sinnvoll ist und positiv verwertet werden kann.

§ 3. Die Erhaltungsfrage bei Verwendung der Hypothese A_1

33. Die Schar $f(u/x)$ bestimme sich aus den Fundamentalfunktionen $f(u/x_0)$ und $b(x)$ gemäss (14) und (14'). Es stellt sich die Frage: Lässt sich ein $g(u) = f(u/x_0)$ so angeben, dass die globale Häufigkeitsfunktion $\bar{f}(u)$, welche gemäss (18) berechnet wird, im allgemeinen, d. h. für «alle» λ , zur gleichen Klasse gehört wie $g(u)$?

Da die Hypothese A_1 offenbar als Spezialfall der Hypothese A_2 angesehen werden kann, muss die Antwort wie folgt ausfallen: Liegt eine triviale Skala s vor, d. h. ist $b(x) \equiv 1$, so gehört $\bar{f}(u)$ stets zur gleichen Klasse wie $f(u/x_0)$; es gilt insbesondere $f(u/x_0) = f(u/x) = \bar{f}(u)$. Andernfalls aber existiert kein $g(u)$ so, dass die in der gestellten Frage enthaltene Forderung erfüllt wäre.

§ 4. Erhaltungsfragen für die Hypothese A

Wir treffen folgende Unterscheidung:

1. $s(x) \equiv 1$.

Es gilt trivialerweise stets $f(u/x_0) = f(u/x) = \bar{f}(u)$.

2. $s(x) \not\equiv 1$.

a) Ist $u_0 = 0$ und $f(u/x_0)$ nach rechts nicht begrenzt, so geht aus der Lösung des Hauptproblems unmittelbar hervor, dass es keine Funktion $g(u) = f(u/x_0)$ so gibt, dass die durch (9) gegebene globale Häufigkeitsfunktion unabhängig von $\lambda(x)$ zum gleichen Typ gehört wie g .

b) Liegt eine «abgeschnittene» Fundamentalfunktion $f(u/x_0)$ vor, so dass die Globalverteilung mehrastig wird, so kann die Hypothese A nicht mehr als Spezialfall der Hypothese A_2 betrachtet werden, und es lässt sich dann folgende Erhaltungsfrage stellen: Wie muss $g(u) = f(u/x_0)$ beschaffen sein, damit der «mittlere» Ast der Globalverteilung (vgl. erstes Kapitel, Figur 2, Ast 3)

Teilstück einer Häufigkeitsfunktion ist, welche demselben Typ wie $g(u)$ angehört, d. h. der Bedingungsgleichung

$$\bar{f}_m(u) = \int_{x_0}^{x_1} \frac{\lambda(x)}{s(x)} g\left[\frac{u}{s(x)}\right] dx = \frac{1}{m} g\left(\frac{u}{m}\right); \quad m = m(\lambda) \quad (49)$$

für «alle» λ genügt?

Zur näheren Charakterisierung von g stützen wir uns auf die Beziehung (48). Setzen wir $\frac{1}{s(x)} = r(x)$, so erscheint sie hier, wo die additive Zusatzfunktion $g_3(u)$ offenbar in Wegfall kommt, in der speziellen Gestalt

$$g[ur(x)] = g_1(u) g_2[r(x)] \quad (50)$$

und ist so zu interpretieren, dass (50) bei vorgegebenem r für alle x ($x_0 \leq x \leq x_1$) notwendig gelten muss, falls (49) erfüllt sein soll. $r(x)$ ist stetig und wegen $s(x) \neq 1$ nicht konstant.

Es steht somit die *Funktionalgleichung*

$$g_1(u) g_2(r) = g(ur), \quad (50')$$

wo sowohl u als auch r in einem nicht ausgearteten Intervall definiert sind, zur Diskussion.

H. W. Pexider [15] hat gezeigt, dass stetige Funktionen, welche (50') erfüllen, von *folgender Form* sein müssen:

$$\left. \begin{aligned} g_1(u) &= B_1 u^A \\ g_2(r) &= B_2 r^A. \end{aligned} \right\} \quad (51)$$

Es gilt somit die notwendige Bedingung

$$f(u/x_0) = g(u) = B u^A. \quad (52)$$

Über B und A kann frei verfügt werden, d. h. wir dürfen diese beiden Parameter tatsächlich so wählen, dass $B u^A$ im Definitionsintervall von $f(u/x_0)$ normiert ist. Je nach der Beschaffenheit dieses Intervalls resultiert eine *hyperbolische* (Pareto) oder *parabolische* (E. Kaiser [2]) Verteilung. Bei einer beidseitig «abgeschnittenen» Fundamentalfunktion $f(u/x_0)$ stehen beide Möglichkeiten offen.

Endlich verifizieren wir durch Einsetzen von (52) in (49) leicht, dass unsere notwendige Bedingung auch hinreichend ist. Hierzu sei vermerkt, dass schon E. Kaiser die Bedingungsgleichung (52) als hinreichendes Kriterium angegeben, die Frage aber, ob (52) zur Erfüllung von (49) zugleich auch notwendig sei, offengelassen hat. Durch den Rückgriff auf das Hauptproblem und seine Lösung liess sich der vorliegende Beweis der Notwendigkeit und damit die vollständige Charakterisierung ohne viel Umtriebe durchführen. Konnte die Bedingung (48) dort wegen der Normierungsvoraussetzung (I) zu keinen verwendbaren Resultaten führen, so hat sie hier ihren guten Sinn, da diese Vorschrift in Wegfall kommt. Der Ast $\bar{f}_m(u)$ und diejenigen Teile der Funktionen $f(u/x)$, welche einen Beitrag an seine Berechnung liefern, sind ja nicht normiert.

§ 5. Das Erhaltungsproblem für die zeitlich veränderliche Globalverteilung

35. Bringen wir die in Beziehung (7) ausgedrückte *Hypothese B* zum Einsatz, so wird die *globale Einkommensverteilung zur Kalenderzeit t* durch

$$\bar{f}_t(u) = \int_{\bar{G}} \lambda_t(x) f(u/x) dx \quad (53)$$

dargestellt. In dieser allgemeinen Gestalt hat die Formel für jede der drei Hypothesen A , A_1 und A_2 Gültigkeit. Ist $\bar{f}_t(u)$ für die beiden erstgenannten Hypothesen durch die Vorgabe der drei Fundamentalfunktionen g , s und λ_t für jeden Zeitpunkt eindeutig bestimmt, so muss bei Verwendung der Hypothese A_2 noch der auszuzeichnende Parameter angegeben werden. Ferner hat man in (53) mit der einzigen Ausnahme, wo die Typenhypothese auf ein «abgeschnittenes» g angewendet wird, wo die Typenhypothese auf ein «abgeschnittenes» g angewendet wird, x_0 und x_1 als Integrationsgrenzen einzusetzen.

36. Damit wir das Problem, welches hier aufgeworfen und gelöst werden soll, konkret formulieren können, wollen wir vorerst einmal annehmen, $f(u/x)$ würde auf Grund von (36), (36') und (36''), das ist gemäss der *Hypothese A₂*, bestimmt. Ferner sei etwa $\bar{f}_{t_0}(u) = h(u, m)$, wodurch sich (53) zur Zeit t_0 schreibt als

$$\bar{f}_{t_0}(u) = h(u, m) = \int_{x_0}^{x_1} \lambda_{t_0}(x) g[u, c(x)] dx.$$

Mit dieser Vorbereitung haben wir schon angedeutet, dass wir uns die Frage vorlegen wollen, ob bei geeigneter Wahl von g stets, d. h. zu jedem Zeitpunkt t , gelte

$$\bar{f}_t(u) = \int_{x_0}^{x_1} \lambda_t(x) g[u, c(x)] dx = h(u, m_t); \quad m_{t_0} = m. \quad (54)$$

Diese Gleichung ist in trivialer Weise für jede Funktion g erfüllt, wenn sich die betrachtete Personengesamtheit im *Beharrungszustand* befindet: $\lambda_t(x) = \lambda_{t_0}(x) = \lambda(x)$ und selbstverständlich auch dann, wenn $s(x) \equiv 1$, d. h. $c(x) = c$ ist. Andernfalls können wir das Problem so angehen, dass wir g aus der Funktionalgleichung

$$\int_{x_0}^{x_1} \lambda(x) g[u, c(x)] dx = h[u, m(\lambda)] \quad (54')$$

für «alle» λ zu bestimmen versuchen. (54') erinnert stark an (47) und damit an das Hauptproblem. Auch hier soll der Integralausdruck Gesetze darstellen, welche unter sich zur gleichen Parameterklasse gehören. Der einzige Unterschied besteht darin, dass dort $g = h$ gilt, hier aber nicht. (54) und (54') sagen nur aus, dass die Globalverteilung im Laufe der Zeit ihre Parameterklasse nicht ändern solle, aber keineswegs etwa, dass diese Parameterklasse mit derjenigen, welcher die Parallelverteilungen $f(u/x)$ angehören, übereinstimmen müsse. Bei der Lösung des Hauptproblems haben wir aber gerade diese Bedingung $g = h$ nirgends entscheidend verwendet. Somit wissen wir, dass sich abgesehen von den beiden besprochenen Trivialfällen keine Fundamentalfunktion g so finden lässt, dass die globale Einkommensverteilung zur Zeit t im allgemeinen derselben Parameterklasse angehört, wie diejenige zur Zeit t_0 .

Auf Grund dieser Ausführungen ist jetzt auch ersichtlich, wie die entsprechenden Fragen zu beantworten sind, welche sich bei Berechnung von $f(u/x)$ gemäss der Typen- oder Klassenhypothese stellen. Offenbar können die Resultate von IV. § 3 und § 4 einfach übernommen und sinngemäss interpretiert werden.

Schlussbemerkungen

37. Im zweiten und dritten Kapitel haben wir der Hypothese *A* von E. Kaiser zwei formal gleichberechtigte Hypothesen A_1 und A_2 zur Seite gestellt. Es ist nicht unsere Meinung, dass einer der drei Hypothesen eine absolute Vorzugsstellung eingeräumt werden soll. Vielmehr stellen wir uns vor, dass bei der Wahl einer dieser Hypothesen in einem praktischen Fall primär das vorliegende statistische Material ausschlaggebend sein soll, sekundär etwa die Rücksicht auf Einsparung an Rechnungsarbeit. Dadurch, dass zwischen mehreren Hypothesen gewählt werden kann, ist die gesamte Theorie elastischer, anpassungsfähiger geworden. So verstanden hat die durchgeführte Erweiterung wohl ihre Berechtigung.

38. In diesem Zusammenhang möchten wir noch einige Bemerkungen zur *Hypothese B*, welche sich mit der Veränderung der Einkommensverteilung im Laufe der Zeit beschäftigt, anbringen. E. Kaiser hat ihr eine absolute Geldeinheit (konstante Kaufkraft) zugrunde gelegt und angenommen, die Verteilung sei unter diesem Gesichtspunkt für eine bestimmte Altersklasse stets dieselbe. Hält man an dieser Stabilität fest, berücksichtigt dagegen eine rein nominelle Veränderung der Einkommen, so zeigt sich die Hypothese *B* in der allgemeineren Gestalt (E. Kaiser [16]):

$$f_t(u/x) = \frac{1}{\sigma(t)} f_{t_0}\left(\frac{u}{\sigma(t)} / x\right).$$

$\sigma(t)$ stellt dabei das mittlere Einkommen als Funktion der Zeit t dar, sofern das Einkommensmittel zur Zeit t_0 als Einheit betrachtet wird.

Wollte man das heute vieldiskutierte «Prinzip der gleitenden Lohnskala» (Indexlohn) in eine mathematische Form kleiden, so müsste es durch die obige Gleichung geschehen. Die Erfahrung lehrt aber, dass die natürliche Entwicklung – ohne staatlichen Eingriff – diesem Prinzip nur bedingt nachkommt. Wenn wir in der vorliegenden Arbeit trotzdem und absichtlich darauf verzichtet haben, eine Modifikation der Hypothese *B* ins Auge zu fassen, so wurden wir dabei durch folgende Überlegungen geleitet:

Erstens einmal ist die Extrapolation der Einkommensverteilung in die Zukunft hinaus naturgemäss eine sehr problematische Angelegenheit. Es sind dabei neben rein wirtschaftlichen Aspekten u. a. auch Verschiebungen im Verhältnis zwischen Leistungs- und Soziallohn zu berücksichtigen, Faktoren, deren Vorausbestimmung einen Akt von nicht zu unterschätzender Willkür darstellt. Durch die Annahme der Stabilität zieht man sich unter solchen Umständen wohl am besten aus der Affäre. Wir gehen mit E. Kaiser einig, wenn er sagt: «Il serait bien difficile de remplacer l'hypothèse B par une autre, plus plausible encore». Der zweite Grund, der uns von einer Modifikation absehen liess, besteht in folgendem: Wollte man die Hypothese B durch eine andere ersetzen, so würde man etwa annehmen, $f_t(u/x)$ gehöre zur gleichen Klasse oder Parameterklasse wie $f_{t_0}(u/x)$, damit andeutend, dass die beiden Verteilungen in ihrem Verlauf nicht allzu stark voneinander abweichen. Derart könnte eine *Hypothese* B_1 , resp. B_2 aufgestellt werden, wobei in der $(u,t)_x$ -Ebene genau die gleichen Rechnungen durchzuführen wären, welche wir bei der Einführung der Hypothesen A_1 und A_2 in der $(u,x)_t$ -Ebene geleistet haben. Somit ist aber die theoretische Grundlage im Prinzip schon geschaffen.

39. Zum Schluss möchten wir noch kurz Stellung nehmen zu den Ergebnissen, welche die Untersuchungen im vierten Kapitel gezeitigt haben. Eine Gesamtschau der Resultate legt nahe, die kritische Würdigung in globo und etwa folgendermassen vorzunehmen: Die *Fragen nach Erhaltung* müssen – vielleicht entgegen der primären Erwartung – im grossen und ganzen *verneint* werden. Zweifellos hat sich die durchgeführte Untersuchung aufgedrängt, und sie hat uns eine volle Abklärung der Sachlage gebracht.

Literaturnachweis

- [1] *M. Krämer*: «Das Solidaritätsprinzip in der sozialen Altersversicherung.» (Diss. Bern 1947.)
- [2] *E. Kaiser*: «La distribution des revenus dans la technique mathématique de la sécurité sociale.» (Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker, 50. Band, Heft 2, Bern 1950.)
- [3] *H. Cramér*: «Mathematical methods of statistics.» (Uppsala 1945.)
- [4] *G. King*: «On Staff Pension Funds.» (Journal of the Institute of Actuaries, Vol. XXXIX, London 1905.)
- [5] *P. Lévy*: «Théorie de l'addition des variables aléatoires.» (Paris 1937.)
- [6] *B. Haller*: «Verteilungsfunktionen und ihre Auszeichnung durch Funktionalgleichungen.» (Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker, 45. Band, Heft 1, Bern 1945.)
- [7] *W. Wegmüller*: «Mathematik und Statistik im Dienste der Eidgenössischen Alters- und Hinterlassenenversicherung.» (Schweizerische Zeitschrift für Volkswirtschaft und Statistik, 86. Jahrgang, Heft 4, 1950.)
- [8] *E. Kaiser*: «Demographische und wirtschaftliche Rechnungsgrundlagen zur Eidgenössischen Alters- und Hinterlassenenversicherung.» (Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker, 47. Band, Heft 1, Bern 1947.)
- [9] *H. v. Mangoldt*: «Einführung in die höhere Mathematik.» (3. Band, Leipzig 1914.)
- [10] *J.F. Steffensen*: «On the semi-normal distribution.» (Uppsala 1937.)
- [11] *K. Pearson*: «Tables of the incomplete F -Function.» (Published by the Office of Biometrika, London, Re-issue 1946.)
- [12] *W. Thalmann*: «Zahlenwerte der Prymschen Funktion zur Berechnung von Rentenbarwerten.» (Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker, Heft 26, Bern 1931.)
- [13] *Jahnke und Emde*: «Tables of Functions.» (Fourth edition, New York 1945.)
- [14] *A. Linder*: «Statistische Methoden für Naturwissenschaftler, Mediziner und Ingenieure.» (Basel 1945.)
- [15] *H. W. Pexider*: Monatshefte für Mathematik und Physik, Heft 14, 1902.
- [16] *E. Kaiser*: «Système de répartition et distributions de revenus variant avec le temps.» (Bericht des XIII. Internationalen Kongresses der Versicherungsmathematiker in Scheveningen. I, 1951.)
- L. Féraud*: «Les instruments mathématiques de la statistique.» (Lausanne, Paris 1946.)
- M. Fréchet*: «Sur les formules de répartition des revenus.» (Revue de l'Institut international de statistique, La Haye 1939.)
- P. Lévy*: «Calcul des probabilités.» (Paris 1925.)

