METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL REVIEW OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT

Vol. V - N. 3. **I - XII - 1925.**

SOMMARIO — SOMMAIRE — CONTENTS — INHALT

PADOVA
AMMINISTRAZIONE DEL "METRON"
R. UNIVERSITÀ - ISTITUTO DI STATISTICA

La Rivista internazionale di Statistica METRON esce in quattro numeri all'anno, che costituiscono complessivamente un volume di 700-800 pagine.

METRON accoglie articoli originali di metodologia statistica e di applicazioni statistiche alle varie discipline, e rassegne o discussioni di risultati raggiunti col metodo statistico in diversi campi della scienza o tali da poter interessare il cultore della statistica. Pubblica altresì una bibliografia di tutte le opere e riviste ricevute in omaggio od in cambio.

Articoli e rassegne potranno essere scritti in italiano, francese, inglese o tedesco. I manoscritti in lingua francese, inglese o tedesca dovranno essere dattilografati.

La collaborazione non è retribuita. Gli autori riceveranno gratuitamente 25 estratti dei lavori pubblicati.

I manoscritti per la pubblicazione dovranno essere indirizzati al *Prof. Corrado Gini, R. Università di Roma - Istituto di Statistica e Politica Economica*, oppure al membro del Comitato direttivo che rappresenta lo Stato a cui l'autore appartiene. Gli autori sono pregati di conservare copia del manoscritto inviato, poichè, nel caso che questo non venga pubblicato, la Direzione non ne garantisce la restituzione.

Al Prof. Corrado Gini dovranno pure essere indirizzate le richieste di cambi da parte di riviste o di altri periodici e ogni pubblicazione inviata in cambio od in omaggio.

Le richieste di abbonamenti, del pari che i versamenti, dovranno invece essere indirizzati alla *Amministrazione del Metron*, presso l'*Istituto di Statistica. R. Università di Padova.*

Il prezzo di abbonamento per ciascun Volume è di **20 scellini** in Europa e di **5 dollari** fuori di Europa, porto compreso, il prezzo di un fascicolo è rispettivamente di **6 scellini** e di **1 $^1/_2$ dollari** porto compreso. Per l'Italia e i paesi a cambio più sfavorevole, il prezzo del volume è di **100 lire italiane** e quello del fascicolo di **30 lire italiane**, porto compreso.

La Revue Internationale de Statistique METRON paraît en quatre fascicules, par an formant en tout un volume de 700-800 pages.

METRON publie des articles originaux de méthodologie statistique et d'applications statistiques aux différentes disciplines, ainsi que des revues ou des discussions des résultats obtenus par la méthode statistique dans toutes les sciences ou bien intéressant les savants qui s'occupent de statistique.

METRON publie aussi une bibliographie de tous les ouvrages et Revues reçues en hommage ou en échange.

Les articles et les revues pourront être écrites en français, en italien, en anglais ou en allemand. Les manuscrits en français, en anglais ou en allemand doivent être envoyés dactylographiés.

On enverra gratis aux auteurs 25 copies tirées à part de leurs travaux après publication.

On adressera les manuscrits pour la publication à *M. le Prof. Corrado Gini, Istituto di Statistica e Politica Economica. R. Università di Roma* (*Italie*), ou bien au membre du comité de direction représentant le pays de l'auteur. On prie les auteurs de garder une copie du manuscrit qu'ils adressent à la Revue, car, en cas de non publication, la rédaction ne garantit pas de pouvoir le renvoyer.

Les demandes d'echange de la part des Revues et des autres périodiques ainsi que toutes les publications envoyées en échange ou en hommage doivent aussi être adressées au Prof. Corrado Gini.

Les demandes des nouveaux abonnements, ainsi que tout payement, devront être adressés à l'*Administration du Metron*, auprès de l'*Institut de Statistique de l'Université Royale de Padoue - Italie.*

Le prix d'abonnement par volume est fixé à **20 sh.** (chèque) dans les pays européens et à **5 dollars** (chèque) dans les pays extra-européens, frais d'envoi compris. Le prix par fascicule est respectivement de **6 sh.** et de **1 $^1/_2$ dollars**, frais d'envoi compris. Pour l'Italie et les pays ayant un change plus défavorable, le prix du Volume est de **100 lires it.** et le prix par fascicule est de **30 lires it.** frais d'envoi compris.

METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL REVIEW OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT

Vol. V - N. 3. **I - XII - 1925.**

SOMMARIO — SOMMAIRE — CONTENTS — INHALT

PADOVA
AMMINISTRAZIONE DEL "METRON"
R. UNIVERSITÀ - ISTITUTO DI STATISTICA

ARTICOLI GIUNTI ALLA RIVISTA CHE VERRANNO PUBBLICATI NEI PROSSIMI NUMERI.
(*Secondo l'ordine d'arrivo*).

ARTIKEL DIE AN DIE ZEITSCRIFT ANGELANGT SIND UND WELCHE IN DEN NACHFOLGENDEN NUMMERN ERSCHEINEN WERDEN.
(*Nach der Reihenfolge des Eingangs*).

ARTICLES REÇUS PAR LA REVUE ET À PARAITRE PROCHAINEMENT.
(*D'après la date de reception*).

ARTICLES RECEIVED BY THE REVIEW WHICH WILL BE PUBLISHED IN FUTURE ISSUES.
(*According to date of receipt*).

C. Gini. *Sulle leggi della frequenza e delle combinazioni sessuali dei parti plurimi.*

C. Gini e M. Boldrini. *Il centro della popolazione italiana.*

K. Popoff. *La prédominance des naissances masculines (D'après les données de la Statistique du Royaume de Bulgarie).*

C. Gini. *La richesse et les revenus nationaux des Indes Britanniques.*

G. Findlay Shirras. *Production in India before and after the War.*

A. R. Crathorne. *A weighted rank correlation problem.*

V. Romanowsky. *On the moments of standard deviation and of correlation coefficient in samples from normal.*

J. Vandellós. *La richesse et le revenu de la péninsule Iberique.*

J. Lestsckinsky. *Probleme der Bevölkerungs Bewegung bei den Juden.*

Città di Castello (Umbria) - Società Anonima Tipografica « Leonardo da Vinci »

EUGEN SLUTSKY.

Ueber stochastische Asymptoten und Grenzwerte.

Einleitung.

Den Ausgangspunkt dieser Studie bildeten einige terminologische Ueberlegungen inbetreff einer in der Wahrscheinlichkeitsrechnung sich so oft darbietenden Wortverbindung: « Die Versuchszahl (n) kann so gross gemacht werden, dass die Wahrscheinlichkeit für die Grösse x zwischen den Grenzen $c - \varepsilon$ und $c + \varepsilon$, wo ε eine beliebig kleine positive Grösse ist, enthalten zu sein, dem Werth 1, somit der Gewissheit beliebig nahe kommt » (1). Es kann schwerlich irgend eine Meinungsverschiedenheit darüber bestehen, dass die Notwendigkeit, eine solche oder ihr ähnliche Wendung bei jeder Gelegenheit zu wiederholen in demselben Masse unrationel ist, als wenn man statt von « Limes », « Ableitung », « Integral » u. s. w. zu reden, bzw., ihre symbolische Bezeichnungen zu gebrauchen sich ihrer vollen Definitionen zu bedienen hätte (2).

Die Bemühungen um eine schärfere Fassung des Begriffs des Grenzwertes im Sinne der Wahrscheinlichkeitstheorie, oder, wie man kürzer sagen kann, des stochastischen (3) Grenzwertsbegriffs haben

---

(1) Fast wörtlich nach L. v. BORTKIEWICZ, *Die Iterationen*, Berlin, 1917, p. 42.

(2) Diese Ueberlegungen wurden durch Prof. W. ROMANOWSKY's Bezeichnungsweise: « $x = c$, modo Bernoulliano » angeregt, der ich aber nicht beistimmen konnte, da es hier in erster Linie nicht die Analogie mit dem Gleichheits-, « sondern mit dem Limesverhältnisse durch die logische Struktur entsprechender Sachverhalte nahegelegt wird. (W. ROMANOWSKY, *Über lineäre Korrelation zweier Grössen, Westnik Statistiki,* Bd. XII, Moskau, 1923, p. 25, passim, russisch). Siehe übrigens S. D. POISSON, *Recherches sur la probabilité des jugements*, Paris, 1837, p. 139, 143.

(3) Diesen Terminus gebrauche ich in dem Sinne von « Wahrscheinlichkeitstheoretisch ». Vergl. J. BERNOULLI, *Ars conjectandi*, Basileae, 1713, p. 213; L. v. BORTKIEWICZ, *Die Iterationen*, Berlin, 1917, p. 3; AL. A. TSCHUPROW, *On the mathe-*

mich dann auf eine Begriffsanalyse der betreffenden Sachverhalte geführt, bis ich einsah, dass der geläufige obwohl bis dahin namenlose Begriff einer Erweiterung und Verallgemeinerung nach verschiedenen Richtungen bedürftig und fähig ist. So entstand der Begriff der stochastischen Asymptote und es eröffnete sich eine Reihe von Fragen über die Beschaffenheiten dieser stochastischen Gebilde. Ausser dem Bedürfniss, verschiedenen bekannten Sätzen der Wahrscheinlichkeitsrechnung eine allgemeinere Formulierung zu geben, entstanden dabei ganz neue Aufgaben. So wurde durch die Analogie mit dem gewöhnlichen Grenzwertsbegriff die Frage nahegelegt, ob die Summe (geometrische, bzw., algebraische) zweier oder mehrerer stochastischen Grenzwerte (bzw. Asymptoten) wieder ein stochastischer Grenzwert (oder Asymptote) einer Summe von entsprechenden zufälligen Grössen sein muss; dann eine ähnliche Frage über Funktionen der stochastischen Grenzwerte überhaupt, u. s. f. Die Natur der Grössen, bzw., der Funktionen, die zu einer gegebenen zufälligen Variablen in einem Grenzwert-, bzw., Asymptotenverhältniss im stochastischen Sinne stehen können, bildete ein anderes Thema, das zur Zeit, m. W., nicht einmal gestellt wurde. Dazu gesellte sich weiter die Frage über die hinreichenden und notwendigen Bedingungen des Bestandes stochastischer Asymptoten und Grenzwerte, und es gelang mir auch in diesem mehr untersuchten Problem einige weitere Schritte zu machen. So kam diese Studie als eine vorläufige Skitze eines auch für die theoretische Statistik wichtigen Kapitels der Wahrscheinlichkeitsrechnung zu Stande, das bis jetzt, m. E., noch niemals als ein besonderer Zweig dieser Disziplin betrachtet und ausgelegt wurde.

Diese Arbeit war schon in ihren Grundzügen recht vorgeschritten (jetzige Kapiteln I-III und die Sätze 14, 16 *a*, 16 *d* des Kap. IV), als ich einige Schriften von F. P. CANTELLI kennen lernte, die sich auf dasselbe Thema des stochastischen Grenzwerts beziehen (1). Obgleich

---

matical expectation of the moments of frequency distributions in the case of correlated observations, «Metron», vol. II, n. 3, 4, p. 461; AL. A. TSCHUPROW, *Ziele und Wege der stochastischen Grundlegung der statistischen Theorie*, «Nordisk Statistisk Tidskrift», Bd. III, S. 435. Auf die wichtige aber ziemlich subtile Frage über verschiedene Nuancierungen, die mit diesem Terminus in der neueren Literatur sich verbinden, kann ich hier nicht eingehen. Siehe übrigens noch unten § 5.

(1) Vergl. F. P. CANTELLI: (1) *La tendenza ad un limite nel senso del calcolo delle probabilità* «Rendiconti del circolo matem. di Palermo», t. 41, 1916; (2) *Sulla legge dei grandi numeri* «Memorie della Reale Academia dei Lincei», Serie 5, vol. 11, 1916; (3) *Su due applicazioni di un teorema di G. Boole alla statistica matematica* «Atti d. R. Ac. d. Lincei», Ser. 5, Rendiconti, Classe di scienze fisiche,

die Priorität in Betreff der Idee des stochastischen Grenzwerts unzweifelhaft Herrn F. P. CANTELLI zukommt, darf ich doch glauben, dass meine Resultate keinesfalls vollständig durch diejenigen seiner Untersuchungen gedeckt werden.

Im ersten Kapitel werden die Begriffe der stochastischen Asymptote und des stochastischen Grenzwerts eingeführt und der Versuch gemacht ihre Beziehungen zu dem sogenannten « Gesetz der grossen Zahlen » einigermassen klar zu legen. Das zweite Kapitel (über mathematische Erwartungen) wird dem mathematischen Hilfsapparat gewidmet. Das dritte bis fünfte Kapitel enthalten eine Reihe von Sätzen über stochastische Asymptoten und Grenzwerte.

Die Darstellung wurde ganz elementar gehalten, und schon der Umstand, dass es mit so einfachen Hilfsmitteln möglich war, einige zum Teil ganz neue Ergebnisse zu erhalten, scheint mir der schlagendste Beweis zu sein, dass wir uns hier wirklich auf einem noch wenig bebauten Gebiete befinden. Vielleicht wird auch mancher Leser mit dem Verfasser dieser Zeilen den Gedanken teilen, dass dadurch auch eine allgemeinere Einsicht bekräftigt wird, nämlich diejenige, dass der *Name* kein blosser Schall ist und dass ein namenloses Wesen sich keiner normalen Existenz erfreuen kann (1).

Erstes Kapitel.

Die stochastische Asymptote und « das Gesetz der grossen Zahlen ».

1. — Fangen wir mit der Einführung einiger Bezeichnungen an. Es sei die Wahrscheinlichkeit der Erfüllung einer Bedingung A mit $P\left\{A\right\}$ bezeichnet (2). Dann setze man:

---

mat. e natur., vol, XXVI 1⁰ sem., 1917; (4) *Sulla probabilità come limite della frequenza* (ibid.); (5) *Sulla oscillazione delle frequenze intorno alla probabilità*, « Metron », vol. III, n. 2, 1923. Von diesen Schriften, auf die ich durch freundliche Mitteilung von Prof. AL. A. TSCHUPROW aufmerksam gemacht wurde, konnte ich bis jetzt in die erste noch nicht Einblick bekommen.

(1) Ich benutze diese Gelegenheit um meinen herzlichsten Dank Herrn Prof. Al. A. TSCHUPROW auszusprechen, der die Freundlichkeit hatte, im Sommer 1923 die erste Skitze dieser Arbeit zu lesen und nachher in einem privaten Briefwechsel eine Reihe kritischer Bemerkungen zu machen, die sich für mich vom grössten Werte erwiesen haben. Al. A. TSCHUPROW's Berufung in einigen seiner letzten Schriften auf meine ungedruckte Arbeit: « Ueber die notwendigen und hinreichenden Bedingungen der Existenz der BERNOULLI'schen Asymptote, bzw., des BERNOULLI'schen Grenzwertes » bezieht sich auf die obengenannte erste Redaktion dieser Arbeit.

(2) W. ROMANOWSKY, *op. cit.*, p. 26.

$$(1) \quad . \quad . \left\{ \begin{array}{l} P_a^b(x) = P\{a < x < b\} \\ P_{(a)}^{(b)}(x) = P\{a \leqq x \leqq b\} \\ P_a^{(b)}(x) = P\{a < x \leqq b\} \\ P_{(a)}^b(x) = P\{a \leqq x < b\} \end{array} \right\} \quad \text{und} \quad \left\{ \begin{array}{l} P_a^{\infty}(x) = P\{a < x\} \\ P_{(a)}^{\infty}(x) = P\{a \leqq x\} \\ P_{-\infty}^{(b)}(x) = P\{x \leqq b\} \\ P_{-\infty}^b(x) = P\{x < b\} \end{array} \right\}.$$

Genügt eine reele Grösse (x) irgend einer von diesen acht Bedingungen: $a < x < b$, $a \leqq x \leqq b$, u. s. f., so sagen wir, dass ihr Wert dem entsprechenden Intervall gehört. Die allgemeine Bezeichnung eines Intervalls sei J_x. Die Wahrscheinlichkeit dafür, dass x einen zu J_x gehörenden Wert annehme, sei die *Wahrscheinlichkeit des Intervalls* J_x genannt und mit $P\{J_x\}$ bezeichnet. Einzelne Werte lassen sich auch unter den Intervallsbegriff subsumieren, da die Bedingung: $x = a$ der Bedingung: $a \leqq x \leqq a$ gleichkommt.

Wird jedem Intervall einer Grösse (x) die ihm zukommende Wahrscheinlichkeit zugeordnet, so nenne ich die in solcher Weise geordnete Menge der Wahrscheinlichkeitswerte die *Wahrscheinlichkeitsverteilung* von x, kurz, die *W.-V.* von x, oder elliptisch die Verteilung schlechthin und bezeichne sie mit $\overline{P}_x$ (1). Eine Grösse, die eine bestimmte *W.-V.* hat, nenne ich die *zufällige Variable* (2). Alle Werte, die sie annehmen kann, sind ihre *möglichen Werte.* Die Wahrscheinlichkeit jedes unmöglichen Wertes, Bzw., Intervalls ist gleich Null (3).

---

(1) Ein Seitenstück zu der *W.-V.* bildet die Verteilung von empirischen Häufigkeitswerten (kurz *H.-V.*). Unter der Verteilung shlechthin (nicht in elliptischer Redeweise) kann man also lediglich das mathematisch formalisierte Gebilde verstehen, das durch seine rein formellen der *W.-V.* ebenso wie der *H.-V.* gemeinsamen Eigenschaften definiert wird. Vergl. GEORG POLYA, *Ueber den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem.* « Math. Zeitschr. », Bd. 8, Heft 3-4, 1920, p. 173.

(2) Dieser Terminus findet sich zuerst, m. W., bei P. A. NEKRASSOW, *Neue Grundlagen der Lehre von Wahrscheinlichkeiten der Summen und der Durchschnittswerte* « Matematitschesky Zbornik », Tom XXI-XXII, Moskau, 1901, p. 9 (russisch). Aber zum Grundpfeiler aller Begriffskonstruktionen der theoretischen Statistik wurde dieser Begriff erst von Prof. A. A. TSCHUPROW erhoben.

(3) Bekanntlich gilt das Entgegengesetzte nicht: ein möglicher Wert kann die Wahrscheinlichkeit 0 haben. Das gleiche gilt auch von der Wahrscheinlichkeit 1. Auf diese und ähnliche subtile Fragen kann ich hier nicht eingehen. Siehe S. N. BERNSTEIN, *Versuch einer axiomatischen Begründung der Wahrscheinlichkeitstheorie,*

Zerlegt man ein Intervall in Teilintervalle, von denen jede zwei keinen Wert gemein haben, so ist nach einem Grundsatze der Wahrscheinlichkeitsrechnung die Summe der Wahrscheinlichkeiten aller Teilintervalle gleich der Wahrscheinlichkeit des ganzen Intervalls. Darum wird es stets möglich sein ein solches Teilsystem von Intervallen zu bilden, dass, erstens, die darin eingehenden Intervalle den Werten der gegebenen zufälligen Variablen eineindeutig zugeordnet, und, zweitens, die Wahrscheinlichkeiten übriger Intervalle durch die Wahrscheinlichkeiten der Intervalle des Systems in eindeutiger Weise bestimmt werden. Die Wahrscheinlichkeit eines zum System gehörigen Intervalls wird dabei eine eindeutige Funktion des Wertes der zufälligen Variablen sein. Ich nenne sie die *Verteilungsfunktion der Wahrscheinlichkeiten*, oder elliptisch die Verteilungsfunktion schlechthin. Jede zufällige Variable hat eine und zwar lediglich eine bestimmte Verteilung; doch was die Verteilungsfunktionen anbelangt, die zu ihr gehören, so gibt es deren eine unendliche Menge; jede von ihnen ist allen anderen äquivalent und kann, den Besonderheiten eines Falls entsprechend, als Untersuchungsgrundlage gewählt werden.

Es gibt verschiede Typen der Verteilungsfunktionen. Der erste kann durch die Funktion

$$P_{-\infty}^{(x)}(x) = W(x)$$

repräsentiert werden. Sie ist monoton, und zwar nicht abnehmend, von rechts stetig und im ganzen Gebiet der reelen Werte von $-\infty$ bis $+\infty$ definiert; es gilt für sie: $\lim\limits_{x \to -\infty} W(x) = o$ und $\lim\limits_{x \to +\infty} W(x) = 1$. Jedwede Verteilung: stetige, ebenso wie unstetige wird durch sie eindeutig bestimmt. Ähnliches gilt für die Funktion $P_{(x)}^{\infty}(x)$ und noch für einige andere Modalitäten desselben Typus. Zur Darstellung der Verteilungen, bei denen die Menge möglicher Werte endlich, oder unendlich abzählbar ist, kann noch die Verteilungsfunktion des zweiten Typus dienen:

$$p_x = w(x),$$

wo mit p_x die Wahrscheinlichkeit $P\{x = a\}$ bezeichnet wird. Für

---

Mitteilungen der «Math. Gesellschaft zu Charkow», Charkow, 1917, (russisch), p. 239 ff. CH. LAGRANGE, *Etude du principe de la limite*, « Bull. de l'Acad. roy. de Belgique », Classe des sciences, n.os 9-10, Bruxelles, 1901.

unmögliche Werte der zufälligen Variablen ist sie gleich O, für die möglichen Werte: $x_1, x_2, \ldots$ (bzw. $, \ldots\ldots x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots\ldots$) gibt sie deren Wahrscheinlichkeiten: $p_1, p_2, \ldots$ (bzw. $, \ldots\ldots p_{-2}, p_{-1}, p_0, p_1, p_2, \ldots\ldots$). Sie kann auch in der Form eines Schemas, z. B.:

$$\left\{\begin{matrix} x_1, & x_2, & \ldots & x_i, & \ldots \\ p_1, & p_2, & \ldots & p_i, & \ldots \end{matrix}\right\}$$

geschrieben werden (1).

Eine zufällige Variable (x) und eine unabhängige Variable (φ) können sich manchmal in einem besonderen Verhältnisse befinden) das darin besteht, dass jedem Werte der letzteren eine bestimmte Wahrscheinlichkeitsverteilung der ersteren eindeutig zugeordnet wird. So entspricht jeder Versuchsnummer in einer Versuchsserie eine bestimmte *W.-V.*, die je nach der Anordnung entweder bei allen Versuchen eine und dieselbe oder, von Versuch zu Versuch, auch eine andere sein kann. Dieser Sachverhalt lässt sich auch so ausdrücken, dass die *W.-V.* einer zufälligen Variable (x) eine Funktion einer unabhängigen Variablen (φ) ist, oder noch kürzer, dass *x mit φ stochastisch verbunden ist.* Wird dabei die *W.-V.* von x bei allen Werten von φ eine und dieselbe sein, so sagen wir, dass x von φ *stochastisch unabhängig ist*, im entgegengesetzten Falle aber, dass x von *φ stochastisch abhängt* (2).

Was die Natur der unabhängigen Variable betrifft, so ist es für obige Feststellungen gänzlich gleichgültig, ob sie eine zufällige, oder eine nichtzufällige ist. Nur so viel sei gesagt, dass, wenn es eine

---

(1) In obigen Definitionen dieses Paragraphen wird der Versuch gemacht einige Konstruktionen von R. v. Mises mit den Begriffsbestimmungen von Al. A. Tschuprow in organische Verbindung zu bringen. Von einer Auseinandersetzung über verschiedene Divergenzpunkte muss ich hier aus Raumrücksichten absehen. Einiges betreffend meiner Ablehnung des R. v. Mises'schen Wahrscheinlichkeitsbegriffs siehe unten § 4. Vergl. Al. A. Tschuprow: *Zur Theorie der Stabilität statistischer Reihen* («Skandinavisk Aktuarietidskrift», 1918, p. 202); *Ueber die Korrelationsfläche der arithmetischen Durchschnitte* («Metron», Vol. 1, n. 4, 1921, p. 41); *Grundbegriffe und Grundprobleme der Korrelationstheorie* («Nordisk statistisk Tidskrift», Bd. 3, H. 2-3, 1924, p. 183 ff.). R. v. Mises: *Fundamentalsätze der Wahrscheinlichkeitsrechnung* («Math. Zeitschr.», Bd. 4, H. 1-2., 1919, p. 70 ff.; *Grundlagen der Wahrscheinlichkeitsrechnung*, («Math. Zeitschr.», Bd. 5, H. 1-2, 1919, p. 52 ff., 66 ff., passim).

(2) Vergl. Al. A. Tschuprow, *Grundbegriffe und Grundprobleme der Korrelationstheorie* (loc. cit., p. 188 ff., 198).

zufällige Variable ist, die die Rolle einer unabhängigen spielt, so bleibt ihr Charakter als einer zufälligen bei der Beurteilung obiger Verhältnisse in erster Linie ganz ohne Belang, da ihre Werte dabei lediglich als bestimmt vorgegebene, nicht aber als wahrscheinliche, in Betracht kommen. Für die Zwecke dieser Arbeit ist es nicht nötig in weitere Problematik stochastischer Zusammenhangs-bzw. Abhängigkeitsverhältnisse einzugehen.

2. — Gehen wir nun zur Definition der Begriffe der stochastischen Asymptote und des stochastischen Grenzwerts über.

Es sei x eine zufällige Variable, die mit einer unabhängigen Variablen φ stochastisch verbunden ist, und es sei $v = f\ (\varphi)$ eine eindeutige Funktion derselben. Wir wollen bei verschiedenen Werten von φ die Wahrscheinlichkeit betrachten, dass x die Werte annehme, deren Abweichungen von v, ihren absoluten Grössen nach, eine beliebige positive Grösse ε nicht übertreffen. Es sei: $\varphi_1, \varphi_2, \ldots \varphi_i, \ldots$ eine unbegrenzte Folge von φ - Werten und

$$\overset{(\varepsilon)}{\underset{(o)}{P}} \mid x - v_1 \mid , \overset{(\varepsilon)}{\underset{(o)}{P}} \mid x - v_2 \mid , \ldots \overset{(\varepsilon)}{\underset{(o)}{P}} \mid x - v_i \mid , \ldots$$

die ihr entsprechende unbegrenzte Folge von Wahrscheinlichkeiten. Wenn bei beliebig kleinen Werten ε und η es ein φ_i gibt, so dass bei allen weiteren φ_k $(k > i)$ die Ungleichung

$$1 - \overset{(\varepsilon)}{\underset{(o)}{P}} \mid x - v \mid < \eta$$

gilt, so lässt sich sagen, dass, bei jedem gegebenen beliebig kleinen positiven ε,

$$(2) \ldots\ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots} \overset{(\varepsilon)}{\underset{(o)}{P}} \mid x - v \mid = 1$$

ist. Die Grösse $v = f(\varphi)$, die diesen Bedingungen genügt, nenne ich die *stochastische* oder die *Bernoulli'sche Asymptote* der zufälligen Variable x für die gegebene Folge: $\varphi_1, \varphi_2, \ldots$ und schreibe

$$(3) \ldots\ldots \qquad as_B\,(x) = v\,, (\varphi_1, \varphi_2, \ldots),$$

wo das symbol as_B eine Verkürzung von « asymptota Bernoulliana » ist, die Bezeichnung aber der unbegrenzten Folge: $\varphi_1, \varphi_2, \ldots$ in entsprechenden Fällen auch durch Symbole $\varphi \to \infty$, bzw., $\varphi \to \varphi_0$ ersetzt, oder, wenn kein Missverständniss zu befürchten ist, auch gänzlich weggelassen werden kann.

Ist $v = f(\varphi) = c$ (*onst.*), so nenne ich die entsprechende stochastische Asymptote den ***stochastischen*** oder den ***Bernoullischen Grenzwert*** und schreibe

$$(4) \ldots\ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots \; B} (x) = c\,,$$

wo das Symbol $\lim_{B}$ die Abkürzung von « limes Bernoullianus » ist (1).

Denselben Sachverhalt kann man auch so ausdrücken, dass eine zufällige Variable x gegen $f(\varphi)$, bzw., c ***stochastisch*** (oder « *modo Bernoulliano* ») ***konvergiere***, was sich symbolisch mit

$$(5) \ldots\ldots \qquad x \underset{(B)}{\rightarrow} f(\varphi)\,, \quad (\varphi_1, \varphi_2, \ldots),$$

bzw., mit

$$(5)' \ldots\ldots \qquad x \underset{(B)}{\rightarrow} c\,, \quad (\varphi_1, \varphi_2, \ldots)$$

schreiben lässt. Von der zufälligen Variablen, die eine stochastische Asymptote, bzw. einen stochastischen Grenzwert hat, kann gesagt werden, dass sie ***stochastisch konvergent*** ist.

Es wäre ein Leichtes alle diese Begriffe auch auf mehrdimensionelle zufällige Grössen auszudehnen, doch sei dies bis zu einer weiteren Stelle aufgeschoben (siehe unten Kap. IV, § 18).

In unseren Definitionen erkennt man leicht die formelle Erweiterung eines seit J. BERNOULLI'S Zeiten wohlbekannten, obschon namenlos gebliebenen, Begriffs. Dass das Verdienst des genialen Forschers nicht lediglich darin bestand, dass er für seinen berühmten Satz einen strengen Beweis erbrachte, sondern, nicht im geringeren Masse, auch in der richtigen Auffassung der ganzen Sachlage und im Auffinden der ganzen stochastischen Struktur des entsprechenden Verhältnisses, — das kann, *m. E.*, nicht bestritten werden; man lese nur die bekannten Ausführungen nach, mit denen der Verfasser den Beweis seines Satzes eingeleitet hat (2). So ist es, glaube ich, eine Sache der geschichtlichen Gerechtigkeit J. BERNOULLI'S Urheberrechte

---

(1) Dem Vorschlage F. P. CANTELLI'S den gewöhnlichen Grenzwert mit dem Symbol « lim », den stochastischen dagegen mit dem Symbol « Lim » zu bezeichnen kann ich nicht beistimmen, da der unterschiedslose Gebrauch beider Bezeichnungen zum festgewurzelten Brauch geworden ist, so dass ihre Differenzierung wohl kaum die Chancen hat, sich zu behaupten, umso weniger da ihre graphische Eigenart in keiner Beziehung zu den zu unterscheidenden Begriffen steht.

(2) JACOBI BERNOULLI, *Ars conjectandi*, Basileae, 1713, p. 225-226 (Die Uebersetzung siehe in OSTWALD'S « Klassiker der exakten Wissenschaften », N° 108, Lpz., 1899, p. 90-92).

nicht lediglich auf das seinen Namen tragende Theorem, sondern auch auf das ihr zu Grunde liegende stochastische Verhältniss anzuerkennen und durch die Gestaltung der entsprechenden Symbole festzuhalten (1).

3. — Bei unserer Formulierung des Begriffs der stochastischen Asymptote ist eine Reihe von Momenten des ursprünglichen Begriffs durch den Erweiterungs-und Formalisierungsprozess betroffen worden.

(1). Anstatt von der Häufigkeit, wie bei J. BERNOULLI selbst, oder vom arithmetischen Durchschnitt, wie bei meisten Verallgemeinerungen seines Satzes, soll man von einer zufälligen Variablen überhaupt sprechen. Das ist aber eigentlich keine Neuerung, da in der betreffenden Literatur bekanntlich schon manchmal die Frage von dem stochastischen Grenzwert verschiedener Funktionen zufälliger Variablen erörtert worden ist, die selbst zufällige Variablen, doch keine arithmetischen Durchschnittswerte sind. So ist der empirische Korrelationskoeffizient, der empirischen Mittelfehler, der Divergenzkoeffizient und mehrere andere Grössen keine Durchschnittswerte, doch haben sie meistens ihre stochastischen Grenzwerte, gegen welche sie mit unbegrenzt wachsender Versuchszahl « modo Bernoulliano » convergieren.

(2). Dass in die allgemeine Definition des stochastischen Grenzwerts der Begriff der mathematischen Erwartung nicht eingeschlossen

---

(1) Eine von der oben besprochenen verschiedene stochastische Relation findet sich bei G. BOHLMANN, *Formulierung und Begründung zweier Hilfsätze der mathematischen Statistik*, « Math. Annalen », Bd. 74, 1913, p. 345, 386, *passim*. Definiert man, wie gewöhnlich, die Aequivalenz zweier Grössen mittelst der Gleichung

$$\lim \left\{ \frac{x}{f(\varphi)} \right\} = 1 \,, \; (\varphi_1 \,, \varphi_2 \,, \ldots) \,,$$

so kann die von G. BOHLMANN angewandte Relation durch die Gleichung

$$\lim_{B} \left\{ \frac{x}{f(\varphi)} \right\} = 1 \,, \; (\varphi_1 \,, \varphi_2 \,, \ldots)$$

definiert und die *stochastische Aequivalenz* genannt werden. Nach Analogie mit der gewöhnlichen Bezeichnung lässt sie sich durch das Symbol

$$x \underset{(B)}{\backsim} f(\varphi) \,, \; (\varphi_1 \,, \varphi_2 \,, \ldots)$$

wiedergeben. Die Untersuchung dieses Verhältnisses schliesse ich aus weiteren Darlegungen aus; nur so viel sei gesagt, dass es dem Umfange nach mit dem stochastischen Asymptotenverhältniss keinesfalls zusammenfällt. So ist, z. B., bei Voraussetzungen des Bernoulli'schen Satzes die Ereigniszahl (m) ihrer mathematischen Erwartung (np) stochastisch äquivalent, doch stellt np bekanntlich keinen stochastischen Grenzwert von m dar.

werden darf, versteht sich von selbst. Denn nicht zum begriffsmässigen Gefüge dieser Konstruktionen, sondern zum theoretischen Inhalt der Lehre von stochastischen Grenzwerten gehört das Entscheiden der Frage, welche Grössen es sind, die in dieser Rolle fungieren müssen, bzw., ob im Falle des Bestandes eines stochastischen Grenzwertes dieser letztere mit der betreffenden mathematischen Erwartung wirklich stets zusammenfallen müsse. Wir werden sehen, dass dies keineswegs immer der Fall ist.

(3). Im Begriff des stochastischen Grenzwerts, bzw., der stochastischen Asymptote soll seine Beziehung zu einer unabhängigen Variablen überhaupt, nicht aber lediglich zu einer Versuchszahl ausgedrückt werden. Denn es ist dieselbe formelle Relation, die eine zufällige Variable sowohl mit einer Versuchszahl, wie mit mancher anderen Grösse, so mit einer Entfernung im Raum oder Zeit, mit einer Wahrscheinlichkeit u. s. f. verbinden kann. Es lassen sich, z. B., stochastisch-physikalische Voraussetzungen aufstellen, bei welchen die mathematische Erwartung der Dichtigkeit der Materie eine stetig abnehmende Funktion der Entfernung von einem Punkte im Raum wäre. Die Dichtigkeit selbst würde dann eine zufällige Variable sein, die mit unbegrenzt wachsender Entfernung sich lediglich im stochastischen Sinne dem Grenzwert 0 nähern müsste. Betrachten wir noch ein Beispiel. Nimmt die Wahrscheinlichkeit des Eintreffens eines Ereignisses bei jedem einzelnen Versuch unbegrenzt zu, so wird bei jeder konstanten Versuchszahl (n) diese letztere ein stochastischer Grenzwert der Ereigniszahl (m) sein, oder in Symbolen: $\lim\limits_{p \to 1}{}_B (m) = n.$ Erreicht p niemals seinen Grenzwert, so darf diese Gleichung nie durch eine schlichte: $m = n$ ersetzt werden.

(4). Dieses Beispiel illustriert noch einen weiteren Umstand, den nämlich, dass eine unabhängige Variable nicht nur ins Unendliche wachsen, sondern auch in einer anderen Weise eine unbegrenzte Folge von Werten durchlaufen kann.

(5). Nun der letzte Punkt. Dass wir uns mit dem Begriff eines Grenzwerts nicht begnügen können, sondern die Annäherung an eine variable Grösse als einen allgemeineren Fall betrachten müssen, liegt ebenfalls in der Sache selbst wesenhaft begründet. Um auch diese Idee an einem Beispiel klar zu machen, wollen wir annehmen, dass an einer unbegrenzten Reihe von stochastisch unabhängigen Variablen: $x_1, x_2, \ldots x_n, \ldots$ eine unbegrenzte Reihe von Versuchen gemacht wird: der erste an der ersten, der zweite an der zweiten Variablen u. s. f., die uns der Reihe nach empirische Werte: x_1',

x_2', ... x_n', ... geben. Bezeichnet man die mathematische Erwartung mit dem Symbol E, und setzt man

$$(6)\ldots\quad x_{(n)} = \frac{1}{n}\sum_{i=1}^{n} x'_i\ ,\ \mu_2^{(i)} = E(x_i - E\,x_i)^2,$$

so hat man bekanntlich:

$$(7)\ldots\quad E\,x_{(n)} = \frac{1}{n}\sum_{i=1}^{n} E\,x_i\ ,$$

$$(8)\ldots\quad \mu_{2\,(n)} = E\,(x_{(n)} - E\,x_{(n)})^2 = \frac{1}{n}\left\{\frac{1}{n}\sum_{i=1}^{n}\mu_2^{(i)}\right\}.$$

Haben nun alle Streuungen ($\mu_2^{(1)}, \mu_2^{(2)}, \ldots \mu_2^{(n)}, \ldots$) eine feste obere Grenze, so geht die Streuung des arithmetischen Durchschnitts ($\mu_{2(n)}$) mit unbegrenzt wachsender Versuchszahl gegen den Grenzwert O, woher bekanntlich der Schluss folgt, dass bei jedem beliebig kleinen positiven ε

$$\lim_{n\to\infty} \overset{(\varepsilon)}{\underset{(o)}{P}} \mid x_{(n)} - E\,x_{(n)} \mid = 1$$

ist (1). Man erkennt leicht, dass dieses Resultat in unserer Schreibweise auch mit

$$(9)\ldots\ldots\quad \lim_{n\to\infty\ B} (x_{(n)} - E\,x_{(n)}) = O$$

bezeichnet werden kann, sowie auch, dass der Satz ganz unabhängig davon gilt, ob die Grösse

$$E\,x_{(n)} = \frac{1}{n}\sum_{i=1}^{n} E\,x_i$$

bei unbegrenzt wachsendem n einen Grenzwert hat oder nicht. Nimmt man, z. B., aus der grossen « Urne der Natur » (2) stets grössere dafür aber meistens heterogenere Stichproben, so wird man oft auf derartige Erfahrungen stossen. Will man also die Relation von $x_{(n)}$ zu E $x_{(n)}$ direkt ausdrücken, so ist man unvermeidlich auf den stochastischen Asymptotenbegriff angewiesen. Die Formel (9) gibt dann ganz allgemein

$$(10)\ldots\quad \underset{n\to\infty}{as_B}\,(x_{(n)}) = E\,x_{(n)}\ .$$

---

(1) A. A. Markow, *Wahrscheinlichkeitsrechnung*, 4 Aufl., 1924, p. 116-118 (russisch).

(2) Nach dem trefflichen Ausdruck von C. Stumpf, *Ueber den Begriff der mathematischen Wahrscheinlichkeit, Sitzungsberichte der philos. philol. u. hist. Classe der k. b.* « Akademie der Wissenschaften zu München », 1892, Heft I, p. 111.

Hat nun E $x_{(n)}$ einen Grenzwert, so ist dann und lediglich dann

$$\lim_{n \to \infty} {}_B (x_{(n)}) = \lim_{n \to \infty} E\, x_{(n)}$$

und nur in dem Fall, wo $E\, x_{(n)}$ eine Konstante ist, kann schlicht

$$\lim_{n \to \infty} {}_B (x_{(n)}) = E\, x_{(n)}$$

geschrieben werden.

Diese Ergebnisse sind auch auf die Häufigkeit des Eintreffens irgend eines Ereignisses A in einer Reihe unabhängiger Versuche anzuwenden. Man verstehe nur unter $x_{(i)}$ die Ereigniszahl bei dem i-ten Versuch, mit möglichen Werten 1, wenn A eintrifft, O, wenn es ausbleibt, und entsprechenden Wahrscheinlichkeiten $p_{(i)}$ und p_i. Dann geht der arithmetische Durchschnitt $x_{(n)}$ in die Häufigkeit m/n und seine mathematische Erwartung in die Durchschnittswahrscheinlichkeit $p_{(n)} = \frac{1}{n}(p_1 + p_2 + \ldots + p_n)$ über. Da die Streuungen von x_1, $x_2, \ldots$ $(p_1 q_1,\ p_2 q_2, \ldots)$ eine feste obere Grenze haben (1/4) die zufälligen Variablen selbst aber voraussetzungsgemäss unabhängig sind, so hat man der Formel (9) zufolge:

$$(11) \ldots \qquad \lim_{n \to \infty} {}_B \left(\frac{m}{n} - p_{(n)} \right) = O$$

und

$$(12) \ldots \qquad as_B\,_{n \to \infty} \left(\frac{m}{n} \right) = p_{(n)} \text{ (1)}.$$

Man erkennt leicht in (9) und (10), bzw. (11) und (12) die bekannten Sätze von POISSON über arithmetische Durchschnitte, bzw. Durchschnittswahrscheinlichkeiten bei unabhängigen Versuchen, die an beliebig wechselnden zufälligen Variablen, bzw., Wahrscheinlichkeiten angestellt werden, nur dass sie von ihm in ganz anderer Weise bewiesen wurden (2). Bleiben wir der Kürze halber beim Satz (12), so finden wir, dass man den Sinn und die Tragweite dieses höchst allgemeinen Satzes, der nur von MARKOFF's Untersuchungen

---

(1) A. A. MARKOFF, *op. cit.*, p. 98-101.

(2) S. D. POISSON, *Recherches sur la probabilité des jugements*, Paris, 1837, p. 246-254, 254-272. « Considérons une série de μ ou $m + n$ épreuves successives, pendant laquelle les chances des deux événements contraires E et F varient d'une manière quelconque » (p. 246). « Soit maintenant A une chose quelconque, susceptible de plusieurs valeurs positives ou négatives . . . Non seulement à chaque épreuve que

über stochastisch zusammenhängende Grössen (1) überholt wurde, verkennt, wenn man die Leistung POISSON's als eine den Fall einer konstantzusammengesetzten Wahrscheinlichkeit betreffende auslegt (2). Will man für diesen Fall einen besonderen Namen haben, so kann es der Fall der *beliebig zusammengesetzten Wahrscheinlichkeit* heissen. Wir haben gesehen, dass der zur angemessenen Formulierung der hier

---

l'on fera pour determiner A, toutes les valeurs possibles seront inégalement probables, mais on supposera, pour plus de généralité, que la chance d'une même valeur varie d'une épreuve à une autre» (p. 254). Diese Festsetzungen sind im weiteren Verlaufe der Beweisgänge durch keine restriktive Voraussetzungen rückgängig gemacht worden, und so haben die bewiesenen Sätze (p. 254, p. 271-272) genau den Sinn, der durch unsere Formeln (11) und (12), bzw., (9) und (10) ausgedrückt wird.

(1) In einer Reihe von Abhandlungen, deren erste: A. A. MARKOFF, *Erweiterung des Gesetzes der grossen Zahlen auf von einander abhängige Grössen* in «Mitt. d. Phys. Math. Ges.» zu Kasan (1907, Ser. 2, XV, N. 4, russisch) erschien. Siehe auch seine Wahrscheinlichkeitsrechnung, 4-te posthume, vom Verfasser umgearbeitete und erweiterte Ausgabe 1924 (russisch), p. 121 ff. Vergl. AL. A. TSCHUPROW Zur Theorie der Stabilität statistischer Reihen, Erste Abhandlung, Ueber den mittleren Fehler des Durchschnittes von gegenseitig nicht unabhängigen «Grössen. Skandinavisk Aktuarietidskrift», 1919. AL. A. TSCHUPROW, *On the mathematical expectation of the moments of frequency distributions in the case of correlated observations.* «Metron», Vol. II, N. 3-4, 1923, p. 19. M. WATANABE, *Extensions of theorems of Bernoulli, Poisson and Tchebycheff*, «The Tôhoku Mathematical Journal» Vol. 12, Nos. 1-2, 1917.

(2) L. v. BORTKIEWICZ, *Kritische Betrachtungen zur theoretischen Statistik*, Erster Artikel, «Jahrb. f. Nat. Oekon. u. Stat. 3 Folge», 8 Bd. 1894, p. 652; Die Iterationen, p. 53 ff. In demselben Sinne interpretiert hier Prof. L. v. BORTKIEWICZ auch die MARKOFF'sche Umformung des POISSON'schen Satzes, wo wegen der durchsichtigen Gedrängtheit der Ableitung das Missverständniss besonders klar auftritt: «Was MARKOFF anlangt, so stellt er, wie die anderen, den Sachverhalt fälschlich so dar, als ob POISSON mit dem Ausdruck» Gesetz der grossen Zahlen den Hilfsatz, betreffend die konstant zusammengesetzte Wahrscheinlichkeit, gemeint hätte» (p. 53). Uns interessiert hier lediglich, dass in obigen Worten unzweifelhaft die Ableitung des Satzes:

$$\lim_B \left(\frac{m}{n} - \frac{p_1 + p_2 + \cdots + p_n}{n} \right) = 0$$

gemeint wird, die im § 15 der 3 Auflage der Wahrscheinlichkeitsrechnung» von A. A. MARKOFF enthalten ist. Die klaren Darlegungen MARKOFF's (p. 71-72) lassen aber keinen Zweifel darüber bestehen, dass hier in die Reihe der Wahrscheinlichkeiten: $p_1, p_2, \ldots$ ganz beliebige Werte eingehen, keinesfalls, dass es bestimmte Werte in *feststehenden* Proportionen sind (verg. Die Iterationen pag. 55). Die Auffassung des MARKOFF'schen Satzes, als eines, der «die konstantzusammengesetzte Wahrscheinlichkeit» betrifft, ist darum unzulässig. Dieser Satz fällt aber mit dem entsprechenden Satze POISSON's (*Recherches*, p. 246-254) zusammen.

obwaltenden Sachverhaltes nötige Begriff der stochastischen Asymptote schon von dem berühmten Verfasser von « Recherches », obschon noch implicite und ohne bewusste Hervorhebung, gebraucht wurde. Dass es dabei nicht zu seiner sprachlichen Formulierung gekommen ist, ändert nichts an der Sache. Desto notwendiger ist es jetzt das Versäumte nachzuholen.

4. — Es klafft zwischen dem gewöhnlichen und dem stochastischen Grenzwertbegriffen eine wahre logiche Kluft, die ohne den Wahrscheinlichkeitsbegriff nicht überbrückbar ist. Die unglückliche Idee, die Wahrscheinlichkeit selbst mittelst des gewöhnlichen Grenzwertsbegriffs aus den Häufigkeiten abzuleiten, ist darum vom Grund aus verkehrt und undurchführbar. Um dies am deutlichsten einzusehen, wollen wir von der scharfsinnigen Frage ausgehen, mit welcher F. P. Cantelli eine seiner interessanten Studien eingeleitet hat (1), und zwar: *Ist die Idee der Wahrscheinlichkeit als eines Grenzwerts der Häufigkeit mit den Grundsätzen der Wahrscheinlichkeitsrechnung vereinbar ?*

Es sei eine unbegrenzt fortsetzbare Reihe unabhängiger Versuche gegeben, und es sei p die konstante Wahrscheinlichkeit eines Ereinisses bei jedem einzelnen Versuch. Die Ereigniszahl in n Versuchen sei mit m und die entsprechende Häufigkeit mit m/n bezeichnet. Diese Voraussetzungen sind die des Bernoulli'schen Theorems, das nun in unserer Schreibweise durch die Gleichung

(13) . . . .
$$\lim_{n \to \infty}{}_B \left(\frac{m}{n}\right) = p$$

ausgedrückt wird. Es fragt sich, ob die Relation

(14) . . . .
$$\lim_{n \to \infty} \left(\frac{m}{n}\right) = p$$

mit der ersteren vereinbar ist. Was bedeutet aber die Idee des Grenzwerts ? Es ist klar: wenn die zweite Gleichung mit Recht bestünde, so könnte man bei jedem vorgegebenen beliebig kleinen ε eine Versuchszahl n_0 auswählen, so gross, dass bei jedem $n > n_0$ die Differenz

$$\left|\frac{m}{n} - p\right| \leq \varepsilon$$

sein müsste. Und dies ist es eben, was mit den Voraussetzungen

(1) F. P. Cantelli, *Sulla probabilità come limite della frequenza.* loc. cit., p. 40.

der Aufgabe unvereinbar ist. In der Tat, es würde bedeuten, dass bei $n > n_0$ die Wahrscheinlichkeit

$$P\left\{\left|\frac{m}{n}-p\right|\leq \varepsilon\right\}=1$$

und die Wahrscheinlichkeit

$$P\left\{\left|\frac{m}{n}-p\right|> \varepsilon\right\}=o$$

sein müsste, was bekanntlich bei keinem endlichen n der Fall ist. Sind also die einzelnen Versuche von einander unabhängig, so ist die Gleichung (14) ein Widersinn; besteht die Gleichung (14) mit Recht, so können die Versuche nicht unabhängig sein. Will man also die Wahrscheinlichkeit als einen Grenzwert der Häufigkeit darstellen, so soll man zuerst die Unmöglichkeit der Unabhängigkeit der Versuche postulieren oder beweisen, und zwar als eine reine à priori'sche Unmöglichkeit, was natürlich wiederum ein Nonsens ist (1).

Aehnliche Gedanken treten bisweilen in einem Gewand auf, in dem man sie nicht so leicht erkennen kann. Es würde sich deshalb lohnen die Auffassung des Gesetzes der grossen Zahlen zu analisieren, die von Prof. AL. A. TSCHUPROW aufgestellt und auf A. COURNOT'S Ansichten zurückgeführt worden ist. (2) Das wesentliche hier gipfelt in der Behauptung, dass die Ableitung des Gesetzes der grossen Zahlen nicht lediglich auf den bekannten Sätzen der Wahrscheinlichkeitsrechnung (von BERNOULLI, POISSON u. s. f.), sondern noch auf einem besonderen Lemma sich begründet, kraft welchem es eigentlich erst möglich wird « aus der Welt der Wahrscheinlichkeiten ob gross, oder klein, sich in die Welt der Häufigkeiten hinüberzutragen » (3). Diese Voraussetzung stellt « die Tatsache des zwischen kleiner Wahrscheinlichkeit und Seltenheit obwaltenden Zusammenhanges » fest, indem sie behauptet, dass « die Ereignisse, deren Wahrscheinlichkeiten sehr klein sind, sich nicht oft wiederholen werden (4).

Will man diesen Satz als einen nomologischen betrachten, so kommt man wieder zum Widerspruch mit der Wahrscheinlichkeits-

---

(1) J. v. KRIES, *Die Principien der Wahrscheinlichkeitsrechnung*, 1886, p. 18-23. L. v. BORTKIEWICZ, *Wahrscheinlichkeitstheorie und Erfahrung*, « Zeitschr. f. Philos. u. philos. Kritik », Bd. 121, 1903. p. 82.

(2) AL. A. TSCHUPROW, *Abhandlungen aus der Theorie der Statistik*, 2 Aufl. 1910, (russisch), p. 227 ff.

(3) *Ibid.*, p. 230.

(4) *Ibid.*, p. 230, 227.

lehre. Es sei die Wahrscheinlichkeit eines Ereignisses so klein, wie man will, in einer Reihe *unabhängiger* Versuche *kann* es dennoch beliebig viele Mal hintereinander auftreten. So ist, z. B., die Wahrscheinlichkeit, dass im Roulette zehntausend millionen Mal hintereinander Rot fällt, nicht die Unmöglichkeit, sondern eine zwar äusserst geringe doch von Null verschiedene Wahrscheinlichkeit, und bei einer hinreichend grossen Anzahl von Reihen zu je 10^{10} Würfen darf das relativ so und so häufige Auftreten solcher Reihen, in welchen alle 10^{10} Male Rot fällt, mit grösster Sicherheit erwartet werden. « Auch die minimalste Wahrscheinlichkeit ist von der Unmöglichkeit noch fundamental unterschieden; und diese Kluft können wir nicht überbrücken, mögen wir die Zahlen auch noch so sehr anwachsen lassen » (1). Nur so viel ist wahr, dass die Wahrscheinlichkeit dafür, dass ein sehr wenig wahrscheinliches Ereigniss *häufig* auftreten wird, eine sehr kleine Grösse weit höherer Ordnung ist, als die seines einmaligen Eintretens. Dass eine solche Auffassung keine Aussage über Häufigkeiten selbst enthält, dürfte, m. E., gerade von dem Standpunkt von A. COURNOT aus nicht behauptet werden, denn seine Ansichten gehen gerade darauf aus, dass jede Behauptung über Wahrscheinlichkeiten von Häufigkeiten eine Behauptung über diese letzteren selbst ist (2). Um das einzusehen braucht man sich in keine physikalischen, bzw., ontologischen Spekulationen einzulassen, sondern nur den schlichten Sinn der entsprechenden Sätze der Wahrscheinlichkeitstheorie sich zur Klarheit zu bringen. Dann sieht man, dass bei betreffenden Voraussetzungen (BERNOULLI'schen, POISSON'schen, MARKOFF'schen) eine *fast volle Gewisheit* besteht, dass « die Ereignisse, deren Wahrscheinlichkeiten sehr klein sind, sich nicht oft wiederholen werden ». Streicht man hier das Wörtchen « fast », so erhält man das COURNOT'sche Lemma, deren Unterschied von dem ersten Satze *nicht also in dem Inhalt, sondern lediglich in der Modalität der*

---

(1) J. v. KRIES, *op. cit.*, p. 21.

(2) « La probabilité mathematique devient alors la mesure de la *possibilité physique* . . . L'avantage de celle-ci c'est d'indiquer nettement l'existance d'un rapport, . . . qui subsiste entre les choses même: rapport, que la nature maintient et que l'observation manifeste lorsque les épreuves se répètent assez ». A. COURNOT, *Essai sur les fondements des nos connaissances,* Nouvelle édition, Paris, 1912, p. 45 (Sperrdruck des. Verfassers). Ueber den für sein ganzes System charakteristischen, doch nichts weniger als bis zum Ende geklärten Begriff « d'impossibilité physique » (oder « d'impossibilité de fait ») siehe seine *Exposition de la théorie des chances et des probabilités* Paris, 1843, p. 79-80, 437-438. Vergl. J. v. KRIES, *Ueber den Begriff der objektiven Möglichkeit,* « Vierteljahrsschrift f. wiss. Philos. », 12 Jahrg., 1888.

Behauptung besteht: was dieser nur mit *fast* voller Gewissheit behauptet, das will das Lemma für eine absolut sichere Erkenntniss ausgeben. Und das ist bei unseren Voraussetzungen sicher falsch.

Nun zu einer anderen Interpretationsmöglichkeit. Obiges Lemma kann noch als ein idiographischer Satz angesehen werden, als eine Behauptung über die tatsächliche Struktur der Welt, bzw., ihres uns umgebenden Teiles (1). Es würde dann besagen, dass obgleich es zwischen allen möglichen Konstellationen der Weltelemente auch solche gibt, die uns allerseltsamste Ereignisse vorspielen müssten, so dass, z. B., alle Zufallsspiele wie von einer dämonischen Kraft verfälscht, warme Körper durch kalte erhitzt, Menschengeschicke durch Sterne geleitet erscheinen würden u. s. f., — doch unsere Welt ist keine von diesen Ausnahmewelten, sondern eine so zu sagen *ordinäre Welt* (2). Man kann diese Voraussetzung für eine vielleicht wohlbegründete halten, doch so viel ist wahr, dass von den Tatsachen der Vergangenheit des chaotischen Geschehens auf seine Zukunft kein Licht fällt. Als ein nomologischer war der Satz falsch, als ein idiographischer ist er nutzlos. Ueber die zukünftigen Geschicke der Welt belehrt er uns nicht, gegen die Möglichkeit eines Sprunges in das Wunderreich der stochastischen Ausnahmezustände gewährt er uns keine Bürgschaft. Wenn wir doch davor keine grosse Angst fühlen, so ist der Grund einfach der, dass wir dieser Möglichkeit nur eine unermässlich kleine Wahrscheinlichkeit anzuerkennen geneigt sind. Nicht die Hypothese der stochastischen Ordinarität der Welt begründet also das Gesetz der grossen Zahlen, sondern letzteres schafft erst die logische Möglichkeit, der ersteren für die Zukunft Glauben zu schenken, da auf der Grundlage alles unseres stochastisch-nomologischen Wissens ihr eine Wahrscheinlichkeit zukommt, die der absoluten Gewissheit praktisch äquivalent ist.

5. — Durch obige Ausführungen scheint der Gedanke nahegelegt zu werden, dass der Begriff der stochastischen Asymptote den logischen Kernpunkt des Gesetzes der grossen Zahlen bildet, sofern diesem letzteren eine entsprechend weite Ausdehnung und Formulierung gegeben ist.

---

(1) Vielleicht kann AL. A. TSCHUPROW's Standpunkt in diesem Sinne verstanden werden: *Abhandlungen aus der Theorie der Statistik*, 2 Aufl., S. 231, (russisch).

(2) Vergl. ZILSEL, *Versuch einer neuen Grundlegung der statistischen Mechanik* « Monatshefte für Math. und Physik » Wien, 1921, Bd. XXXI, p. 153-154. Seine « verallgemeinerte Allagodenhypothese » ist der Hypothese der Ordinarität der Welt äquivalent. Der Verfasser irrt sich aber, sofern er glaubt, dass seine Konstruktionen einen stochastischen Standpunkt entbehrlich machen.

Jede Verallgemeinerung des Bernoullischen Satzes, wo die Versuchszahl die Rolle einer unabhängigen Variable spielt, ist von der Form:

$$(15) \ldots \qquad \underset{n \to \infty}{as_B}(x) = f(n).$$

Jst hier x eine bestimmte zufällige Variable und $f(n)$ eine bestimmte Funktion von n, so ist (15) ein Satz, der entweder wahr oder falsch ist. Lässt man aber x und $f(n)$ unbestimmt bleiben, so wird (15) eine pure logische Form sein, die alle gültigen Sätze, denen sie zukommt, zu einer logisch einheitlichen Gruppe zusammenschliesst. Und so kann, glaube ich, « *das Gesetz der grosen Zahlen* » *als ein allgemeiner Name für jeden gültigen allgemeinen Satz aufgefasst werden, der darin besteht, dass eine so und so beschaffene zufällige Variable bei unbegrenzt wachsender Versuchszahl* (*bzw. — zahlen*) *eine stochastische Asymptote hat.* Alle Verallgemeinerungen auf mehrdimensionelle Grössen und auf andere zusammengesetzten Gebilde seien dabei auch hierin eingeschlossen (siehe unten § 18).

In ganz ähnlichem Sinne gebraucht man schon das Wort « Gesetz » in den Wendungen, wie: « Naturgesetz », « Kausalgesetz », « Wesensgesetz », « Strukturgesetz » u. s. w., durch welche kein bestimmtes Gesetz bezeichnet, sondern ganze Kategorien von solchen umspannt werden. Im Sinne unserer Definition kann man also von verschiedenen Gesetzen der grossen Zahlen reden, so von dem Bernoulli'schen, von den Poisson'schen, von den Markoff' schen u. s. f. (1). Gibt es aber für eine zufällige Variable weder einen stochastischen Grenzwert, noch eine stochastische Asymptote, so darf man elliptisch sagen, dass sie dem Gesetz der grossen Zahlen nicht unterworfen ist (2).

---

(1) « Am besten wäre es vielleicht, der längst eingewurzelten Gewohnheit gemäss, die zu verändern jetzt fast unmöglich wäre, den Namen des Gesetzes der grossen Zahlen als eine allgemeine Benennung für mehrere Theoreme der Wahrscheinlichkeitsrechnung, bei denen eine grosse Zahl gewisser Bedingungen oder Versuche eine wesentliche Rolle spielt, bestehen zu lassen ». W. ROMANOWSKY, *Die Wahrscheinlichkeit und die Statistik* « *Westnik Statistiki* », Bd. XVII, Moskau, 1924, p. 15, Fussn. 1.

(2) Unsere Auffassung des Gesetzes der grossen Zahlen kann man als eine partielle Legitimierung der heutigen ziemlich laxen Verwendung dieses Terminus ansehen, nachdem alle Versuche ihm einen bestimmten einheitlichen Inhalt zu verleihen nicht zu allgemeiner Anerkennung gelangt sind. Am besten orientiert man sich in dieser Frage nach AL. A. TSCHUPROW, *Abhandlungen aus der Theorie der Statistik*, 2 Aufl., 1910 (russisch) und J. M. KEYNES, *A Treatise on Probability* 1921. Siehe ferner L. v. BORTKIEWICZ, *Die Iterationen*, 1917, p. 42 ff. E. SLUTSKY *Zur Frage von dem Gesetze der grossen Zahlen*, (russisch), « Westik Stastiki », Bd. XX, 1925,

Der rechte Sinn dieser Auffassung des Gesetzes der grossen Zahlen würde ganz missverstanden werden, sollte man seinen Inhalt als einen lediglich rein mathematischen auffassen (1). Reine Mathematik weiss beileibe nichts von Ereignissen, Versuchen, vom Zufall, von der Wahrscheinlichkeit. Diese Begriffe in ihrem sachhaltigen, materiellen Inhalt sind der wesenhaft auf Formelles eingestellten reinen Mathematik ebenso fremd, wie Elektronen, Lichtwellen, Himmelskörper, Raum und Zeit. Gruppen, Mengen, Substitutionen, Zahlen, das Formal-Logische überhaupt — *das* ist reine Mathematik. Nach allem Bemühen der letzten Dezennien um die axiomatischen Grundlagen der Mathematik sollte dies schon vollständig klar sein (2). Was mit der Geometrie geschehen ist, ist besonders lehrreich. Die « Geometrie » der HILBERT'schen « Grundlagen» (3) ist schon keine Geometrie mehr, keine Lehre vom Räumlichen; ihre Punkte, Geraden, Flächen sind « Punkte », « Geraden », « Flächen » in Aufführungszeichen — reine Symbole für unbestimmt gedachte Gegenstände im Sinne eines logischen « Etwas ». Nun wäre es ja richtiger eine solche Disziplin nicht mehr Geometrie zu nennen, sondern Mannigfaltigkeitslehre, um so mehr als der richtige Name schon da ist. Die Geometrie, als die Lehre vom Raum ist eine Naturlehre geworden; das ist eine der grössten Leistungen der modernen Physik (4).

---

p. 1-55. Was die heutige Verwendung des Terminus « Gesetz der grossen Zahlen » betrifft, so siehe, z. B. A. MARKOFF, *Wahrscheinlichkeitsrechnung,* 4 Aufl., 1924, p. 116, 117, 121, 134, 173, 174; R. v. MISES, *Grundlagen der Wahrscheinlichkeitsrechnung,* « Math. Zeitschr. » Bd. 5, 1919, p. 91 ff. (Vergl. hier die von unserem Standpunkt bezeichnende Ueberschrift des § 6: *Die Gesetze der grossen Zahlen*); F. P. CANTELLI, *Sulla legge dei grandi numeri,* loc. cit. (Siehe oben p. 4, Fussn. 4). Sogar bei AL. A. TSCHUPROW findet sich gelegentlich eine mit unserer Auffassung ganz vereinbare Ausdrucksweise, so z. B. *Zur Theorie der Stabilität statistischer Reihen,* Erste Abh., loc. cit., p. 208.

(1) « Es ist hieraus ersichtlich, dass die Bezeichnung « Gesetz der grossen Zahlen » eine sehr wenig zutreffende ist. Denn es liegt im Grunde nichts Anderes vor als ein mathematischer Satz » : J. v. KRIES, *Die Principien der Wahrscheinlichkeitsrechnung,* 1886, p. 91, Fussn. 1. Siehe dazu noch eine zustimmende Bemerkung von L. v. BORTKIEWICZ, *Kritische Betrachtungen, loc. cit.,* p. 665, Fussn. 2. Vergl. AL. A. TSCHUPROW, *Abhandlungen aus der Theorie der Statistik,* 2 Aufl., 1910, p. 229: « Sie (Sätze von BERNOULLI, POISSON u. s. f.) sind lediglich Sätze der Kombinatorik ».

(2) E. HUSSERL, *Logische Untersuchungen,* I Bd., 2 Aufl., 1913, p. 247 ff.; DERSELBE, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie,* « Jahrbuch f. Philos. u. phänomenologische Forschung », I Bd., Teil 1, 1913 (auch eine Sonderausgabe, 1922) p. 20-23.

(3) D. HILBERT, *Grundlagen der Geometrie,* Lpz. 2 Aufl. 1913.

(4) Die Geometrie im Sinne der Lehre von dem Raume der reinen Anschauung ist wohl kaum im Entstehen begriffen.

Wollte man also an die Anwendungsprobleme der « Wahrscheinlichkeitsrechnung » mit sachgemässer Gründlichkeit herantreten, so wäre es an erster Stelle unbedingt notwendig sie nach dem Vorbilde der Hilbert'schen Leistung einer vollständigen Formalisierung zu unterwerfen. Keine « Versuche », keine « Ereignisse », keine « Wahrscheinlichkeit », wenn nicht in Aufführungszeichen. Die *Alternative* im Sinne eines Sachverhaltes, dessen Glieder leere « Etwas » sind, die in der Relation der vollen *Disjunktion* stehen (« es besteht entweder A, oder B, . . . oder C ») und denen bestimmte Zahlen (*Valenzen*) zugeordnet sind, so kann man sich das Grundgebilde dieser Lehre vorstellen. Eine Axiomenreihe über Operationen an mannigfaltig sich kreuzenden und ineinanderschachtelnden Alternativen (1), eine andere Axiomenreihe über Relationen von Valenzen verschiedenartig zusammenhängender Alternativen, — das sind die Grundlagen der formalisierten Wahrscheinlichkeitsrechnung, als einer rein mathematischen Disziplin (2). Aller hergebrachte Inhalt der traditionellen « Wahrscheinlichkeitslehre », — sofern er ein rein mathematischer ist, — wird hier seinen Platz und seine Begründung finden. Doch wird es keine Wahrscheinlichkeitslehre im eigentlichen Sinne sein, da sie keine Lehren über *Wahrscheinlichkeit*, obgleich eine Reihe von solchen enthalten wird, die auf Wahrscheinlichkeiten angewandt werden können, in demselben Sinne, wie Arithmetik auf Gegenstände der Statistik und Mannigfaltigkeitslehre auf Raumgegenstände angewandt werden. Ich gebe dieser künftigen Disziplin, einen vorläufigen Namen: die *Disjunktionsrechnung* (3).

Die Wahrscheinlichkeitstheorie, als Lehre vom zufälligen Geschehen, ist die *Stochastik*, — wenn man diesen von Prof. L. v. BORTKIEWICZ neuerdings zum neuen Leben erweckten Bernoulli'schen Terminus in diesem Sinne gebrauchen darf (siehe oben p. 3, Fussn. 3). Es lässt sich glauben, dass die richtige Ahnung des tiefen Unterschieds zwischen der Disjunktionsrechnung und der Stochastik schon

---

(1) Den verdienstlichen Vorstoss in dieser Richtung bildet die R. v. MISES' Lehre von Grundoperationen (*Grundlagen der Wahrscheinlichkeitsrechnung*, loc. cit., p. 73 ff.).

(2) Siehe S. N. BERNSTEIN, *Versuch einer axiomatischen Begründung der Wahrscheinlichkeitstheorie*, « Mitt. der Math. Ges. zu Charkow », 1917 (russisch), p. 209-274. Diese Arbeit verdient recht gründlich studiert zu werden.

(3) E. SLUTSKY, *Zur Frage von den logischen Grundlagen der Wahrscheinlichkeitsrechnung*, (russisch), « Westnik Statistiki », Bd. XII, Moskau, 1922, p. 13-21. Die Name der Disjunktionsrechnung wurde mir dann in einem privaten Briefwechsel von Prof. L. v. BORTKIEWICZ freundlich vorgeschlagen.

von jeher in der Doppelheit der Benennungen « Wahrscheinlichkeitsrechnung » und « Wahrscheinlichkeitstheorie » sich einen Ausdruck fand. Was den Standpunkt von J. BERNOULLI selbst betrifft, so war er durch und durch stochastisch, — in unserem Sinne. « Ars conjectandi sive stochastice » war ihm « die Kunst, so genau als möglich die Wahrscheinlichkeiten der Dinge zu messen » (1). Diese Definition muss aber, in der geschichtlichen Perspektive beurteilt, der Tatsache weichen, dass das ganze Werk BERNOULLI's mit der « Kunst » nur in Anwendungen etwas zu tun hat und dass es seiner ganzen Anlage nach eine systematisch geordnete Darstellung der Lehre von Wahrscheinlichkeiten von Ereignissen, bzw., von Urteilen ist (2). Nicht lediglich rein mathematische Sachverhalte, sondern « infinitam varietatem, quae cum in naturae operibus, tum in actionibus mortalium elucet, quaeque praecipuam hujus Universi pulchritudinem constituit » (3), hatte er in Auge als er sein Werk schrieb. Wenn er in « Proœmium » zum zweiten Teile seines Buches sagt, dass er « totam Doctrinam... ex primis fundamentis eruere » (4) bestrebt war, so ist es für die Eigenart seines ganzen Werkes, charakteristisch. Sein berühmtes Theorem war ihm darum keinesfalls lediglich ein Satz aus der Permutationslehre. Dort, wo er wirklich die mathematischen Grundlagen dieses Satzes in formeller Reinheit ausarbeiten will, kennzeichnet er selbst diesen logischen Schritt in unzweideutigen Worten: « conabor omnia reducere ad abstractam Mathesin » (5). Wenn wir also Ars conjectandi und somit auch die Stochastik der Wahrscheinlichkeitstheorie im oben erläuterten Sinne gleichsetzen, so ist es zum Teil eine treue Wiedergabe, zum Teil eine natürliche Fortentwicklung der Auffassung von J. BERNOULLI selbst.

Diese notwendigerweise zu kurz gefassten Darlegungen werden doch, möchten wir hoffen, den Grundgedanken unserer Auffassung des Gesetzes der grossen Zahlen verständlich machen. Spricht man von *Gesetzen*, so kann man gewiss vernünftiger Weise nicht von rein hypothetisch-formellen Sätzen der Disjunktionsrechnung reden, sondern nur von sachhaltigen stochastischen Sätzen, über deren Begründung man noch so viel streiten möge (ob empirisch, oder apriorisch u. s. w.),

---

(1) *Ars conjectandi*, p. 213.

(2) Vergleiche noch die Bezeichnung « *ars* combinatoria », was unzweifelhaft nichts anderes als « Permutations*lehre* » bedeuten soll. *Op. cit.*, p. 73.

(3) *Ibid.*, p. 72.

(4) *Ibid.*, p. 73.

(5) *Ibid.*, p. 228.

die aber jedenfalls als Gesetze des zufälligen Geschehens gedacht werden müssen. Die Meinung (1), dass es keine Gesetze in diesem Sinne sein können, da sie nicht von Häufigkeiten, sondern lediglich von Wahrscheinlichkeiten etwas aussagen, haben wir schon abgelehnt. « Das vollkommenste Kennzeichen der Richtigkeit (des gefällten Urteils), sagt EDMUND HUSSERL (2), ist die Evidenz, es gilt uns als unmittelbares Innewerden der Wahrheit selbst. In der unvergleichlichen Mehrheit der Fälle entbehren wir dieser absoluten Erkenntnis der Wahrheit, statt ihrer dient uns... die Evidenz für die mehr oder minder hohe Wahrscheinlichkeit des Sachverhalts, an welche sich bei entsprechend « erheblichen » Wahrscheinlichkeitsgraden das fest entschiedene Urteil anzuschliessen pflegt. Alles empirische Wissen ist ein Wissen lediglich in diesem weiteren Sinne: direkt über Wahrcheinlichkeiten, indirekt über entsprechende Sachverhalte (3).

6. — Der neuerdings von Prof. L. v. BORTKIEWICZ vertretenen Auffassung des Gesetzes der grossen Zahlen, als einer *Tatsache* der mehr oder weniger grossen Stabilität der statistischen Häufigkeiten, welche unter der Bedingung der Unveränderlichkeit, bzw., der schwachen Veränderlichkeit allgemeinerer Bedingungen statt hat « sofern ihnen hinreichend grosse Ereigniszahlen zugrunde liegen » (4), kann ich nicht beistimmen, und zwar in erster Linie wegen der, m. E., logisch unzulässigen Gleichsetzung von « Tatsache » und « Gesetz ». Es besteht auch kein terminologisches Bedürfniss danach, da die genannte Tatsache in dem schlichten Ausdruck « die *statistische Stabilität* » eine ganz passende Benennung hat. Wenn L. v. BORTKIEWICZ dabei

---

(1) Siehe oben. 17, Fussn. 2, 3.

(2) *Logische Untersuchungen,* I Bd., 2 Aufl., 1913, p. 13.

(3) Siehe noch *ibid.* p. 14, 254-257. Ich will hier noch bemerken, dass die HUSSERL'sche Idee « der reinen Wahrscheinlichkeitslehre als reiner Theorie der Erfahrungserkenntnis » (*loc. cit.,* Ueberschrift des § 72) mit unserer Auffassung des Gesetzes der grossen Zahlen nicht unvereinbar ist, doch kann ich hier auf diese Frage nicht eingehen. Nur andeutungsweise sei auf die Analogie mit den obersten logischen Gesetzen hingewiesen, denen schon Aristoteles neben einer logischen (im engeren Sinne der apophantischen Logik) noch eine ontologische Formulierung zu geben wusste (Vergl., z. B., ALOIS HÖFLER, *Logik,* 2 Aufl., 1922, p. 546). Jedes ontologische Wesen aber ist « eine *blosse Wesensform,* die zwar ein Wesen, aber ein völlig « *leeres* » ist, ein Wesen, ... das in seiner formalen Allgemeinheit alle, auch die höchsten materialen Allgemeinheiten unter sich hat und ihnen durch die ihr zugehörigen formalen Wahrheiten *Gesetze* vorschreibt » (EDMUND HUSSERL, *Ibidem...* p. 21, Sperrdruck des Verfassers).

(4) L. v. BORTKIEWICZ, *Die Iterationen,* p. 56-57.

sich auf POISSON beruft, bei dem « das Gesetz der grossen Zahlen *seinem Wortlaut nach* nichts anderes als die Stabilität der betreffenden statistischen Zahlenwerte und zwar ohne nähere Angabe des Grades dieser Stabilität » (1) bedeuten solle, so kann ich mich dieser Ansicht nicht anschliessen. POISSON'S Ausdrucksweise ist in ziemlich fahrlässigem, bei weitem nicht strengem Stil gehalten, so dass man seine einzelnen Aussprüche eben nicht ihrem Wortlaut nach, sondern stets nur im vollen Kontext mit anderen zu demselben Thema bezüglichen Stellen beurteilen darf. Den besten Beweis dafür liefert uns gerade die Definition des Gesetzes der grossen Zahlen. Liest man die Erklärung:

« Maintenant, la loi des grands nombres réside dans ces deux équations

$$(16)\ldots\quad \frac{m}{\mu} = \frac{m'}{\mu'}\,,\quad \frac{s}{\mu} = \frac{s'}{\mu'}\,,$$

applicables à tous les cas d'éventualité des choses physiques et des choses morales » (2), (wo mit μ, bzw., μ' die Versuchszahlen in zwei Versuchsserien, mit m/μ und m'/μ' entsprechende Häufigkeiten und mit s/μ, bzw., s'/μ' die Durchschnittswerte irgend einer zufälligen Variable bezeichnet sind), so verfällt man vielleicht der Versuchung diese Formulierungen für gewöhnliche approximative Gleichungen zu halten. Doch nichts dergleichen! Vier Seiten früher liest man, dass für sehr grosse Versuchszahlen μ und μ': " on aura, *à très peu près et très probablement*,

$$\frac{m}{\mu} = \frac{m'}{\mu'}\,, \text{ " (3)}$$

woher es folgt, dass der Verfasser damit nicht eine gewöhnliche, sondern eine *stochastisch* approximative Gleichung gemeint hat. Nimmt man noch andere dazu bezüglichen Stellen in Betracht, so wird es ganz klar, dass diese Formulierungen des Gesetzes der grossen Zahlen nichts anderes als die stochastischen Grenzwertgleichungen

$$(17)\ldots\quad \lim_{\substack{\mu\to\infty\\ \mu'\to\infty}}{}_B \left(\frac{m}{\mu} - \frac{m'}{\mu'}\right) = 0\quad,\quad \lim_{\substack{\mu\to\infty\\ \mu'\to\infty}}{}_B \left(\frac{s}{\mu} - \frac{s'}{\mu'}\right) = 0$$

bedeuten können.

---

(1) *Ibid.*, p. 58.

(2) *Recherches*, p. 143.

(3) *Ibid.*, p. 139 (Von mir gesperrt).

Umsomehr vorsichtiger muss man natürlich bei der Interpretation der vorläufigen Erklärungen POISSON's in der Einleitung (*Préambule*) sein. Ich erlaube mir eine solche Stelle in extenso anzuführen (1):

Les choses de toutes natures sont soumises à une loi universelle qu'on peut appeler *la loi des grands nombres*. Elle consiste en ce que, si l'on observe des nombres très considérables d'événements d'une même nature, dépendants de causes constantes et de causes qui varient irrégulièrement, tantôt dans un sens, tantôt dans l'autre, c'est-à-dire sans que leur variation soit progressive dans aucun sens déterminé, on trouvera, entre ces nombres, des rapports à très peu près constants. Pour chaque nature de choses, ces rapports auront *une valeur spéciale dont il s'écarteront de moins en moins, à mesure que la série des événements observés augmentera davantage*, et qu'ils atteindraient rigoureusement s'il était possible de prolonger cette série à l'infini. Urteilt man wieder «dem Wortlaut nach», so kann man glauben, dass in den gesperrten Worten ein Satz von der Form

$$\lim_{n \to \infty} (x) = c$$

anonciert wird, so, dass das Gesetz der grossen Zahlen durch den gewöhnlichen Grenzwertbegriff ausgedrückt erscheint. Liest man aber weiter die Erläuterungen über die Rolle der Wahrscheinlichkeitsrechnung bei der Interpretation der Versuchsergebnisse (p. 8), und noch weiter das Versprechen des Verfassers das Gesetz der grossen Zahlen in weiteren Teilen seines Werkes zu beweisen (p. 12), so erkennt man, dass hier wieder eine ungenaue Formulierung vorliegt. Was gemeint ist, ist wieder ein stochastisches Grenzverhältniss, das sich in unseren Symbolen als die Gleichung

$$\lim_{n \to \infty} {}_B (x) = c$$

schreiben lässt.

Das Gesetz der grossen Zahlen ist also für POISSON ein Naturgesetz, das für alles zufällige Geschehen (p. 8), sowohl physikalischer, wie auch moralischer Natur gelten muss (p. 7, 12, passim). Es ist aus den Grundsätzen der Wahrscheinlichkeitslehre *à priori* ableitbar (p. 143) und, obschon in der Erfahrung bestätigt (p. 12), doch durch die Erfahrung nicht widerlegbar, da alle Unstimmigkeiten in empirischen Data auf etwaige Veränderungen in den einer unmittelbaren

---

1) *Ibid.*, p. 7 (Die erste Stelle vom Verfasser, die zweite von mir gesperrt).

Beobachtung unzugänglichen « Chancen » zurückgeführt werden können (p. 8, passim).

Wenn diese Grundintentionen des genialen Mathematikers im speziellen Teile seines Werkes keine ganz adäquate Erfüllung gefunden haben, so trägt daran sein allzugrosses Vertrauen in die Allgemeingültigkeit seines Schemas der « Durchschnittswahrscheinlichkeit im eigentlichen Sinne » die Schuld. Nicht diese, die schwächste, Seite seiner Lehren, sondern seine grossen Intentionen sollen für uns massgebend sein. Ich glaube, dass die obige Auffassung des Gesetzes der grossen Zahlen mindestens dem wesenhaften Kerne seiner Idee treu bleibt.

Zweites Kapitel.

Ueber mathematische Erwartungen.

7. — In diesem Kapitel haben wir uns mit einigen Hilfsmitteln zu befassen, deren wir uns weiter bedienen müssen. Ich fange mit der mathematischen Erwartung an, deren Grundeigenschaften als wohlbekannt vorausgesetzt werden können, deren Begriff aber hier in erweiterter an beliebige Verteilungen angepasster Form entwickelt werden soll (1).

Kann eine zufällige Variable x eine endliche Anzahl diskreter Werte $(x_1, x_2, \ldots x_k)$ mit Wahrscheinlichkeiten $(p_1, p_2, \ldots p_k)$ annehmen, so definiert man bekanntlich die mathematische Erwartung einer reelen Funktion $f(x)$ als

$$(1) \ldots \quad E f(x) = \sum_{i=1}^{k} p_i f(x_i) \ .$$

Sind mögliche Werte von x diskret und abzählbar $(\ldots . \ x_{-h}, \ldots x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots x_k, \ldots .)$ ebenso wie ihnen zugeordnete Wahrscheinlichkeiten, so wird die mathematische Erwartung von $f(x)$ durch

$$(2) \ldots \quad E f(x) = lim \sum_{i=-h}^{k} p_i f(x_i) \ , \ (-h \rightarrow -\infty \, ; k \rightarrow \infty)$$

definiert, vorausgesetzt, dass die entsprechende Summe auch wirklich einen Grenzwert ($+\infty$ und $-\infty$ eventuell eingerechnet) hat, der von der Ordnung des Summierens unabhängig ist (2).

---

(1) Vergl. R. v. Mises, *Grundlagen der Wahrscheinlichkeitsrechnung*, « Math. Zeitschr. » 5 Bd, 1-2-H., 1919, § 3, p. 66 ff.

(2) Durch eine Modifizierung des Petersburger Problems erhält man eine zu-

Der Definition der mathematischen Erwartung im allgemeineren Falle kann die Verteilungsfunktion

$$(3)\dots\qquad \underset{-\infty}{P^{(x)}}(x) = W(x)$$

zugrunde gelegt werden (ihre Eigenschaften siehe oben § 1). Die Wahrscheinlichkeit jedes von oben geschlossenen Intervalls ergibt sich daraus von selbst. So setze man

$$(4)\dots\qquad P_{x_1}^{(x_1+\Delta x)}(x) = \Delta W(x_1+\Delta x) = W(x_1+\Delta x) - W(x_1)\,.$$

Beginnen wir mit dem Fall einer zufälligen Variable, deren mögliche Werte in einem endlichen Intervall $(a < x \leqq b)$ enthalten sind. Teilt man die Länge des ganzen Intervalls $(b - a)$ in n gleiche Teile und setzt man $(b - a) : n = \Delta x$, so entsteht eine Reihe von oben geschlossener Intervalle:

$$(a, a+\Delta x >, (a+\Delta x, a+2\Delta x >, \dots (a+\overline{n-1}\,\Delta x, a+n\Delta x >,$$

denen die Wahrscheinlichkeiten

$$\Delta W(a+\Delta x),\ \Delta W(a+2\Delta x), \dots\ \Delta W(a+n\Delta x)$$

zugeordnet werden. Hat nun die Summe

$$(5)\dots\qquad \sum_{i=1}^{n} f(a+i\,\Delta x)\,\Delta W(a+i\,\Delta x)$$

oder in einer kürzeren Schreibweise

$$(5)'\dots\qquad \sum_{a}^{(b)} f(x)\,\Delta W(x)$$

bei unbegrenzt wachsendem n und gegen Null konvergierendem Δx einen Grenzwert, so wird die mathematische Erwartung von $f(x)$ definiert, indem man

$$(6)\dots\qquad E f(x) = \lim_{\Delta x \to 0} \sum_{a}^{(b)} f(x)\,\Delta W(x)$$

---

fällige Variable mit der folgenden Verteilung:

$$\left\{\begin{array}{l} x: \dots -2h, \dots -2^4, -2^3, -2^2, 2^2, 2^3, 2^4, \dots 2k, \dots \\ p: \dots \frac{1}{2h}, \dots \frac{1}{2^4}, \frac{1}{2^3}, \frac{1}{2^2}, \frac{1}{2^2}, \frac{1}{2^3}, \frac{1}{2^4}, \dots \frac{1}{2k}, \dots \end{array}\right\}$$

Man überzeugt sich leicht, dass es hier keine mathematische Erwartung Ex gibt, da dem Symbol

$$\lim_{\substack{h \to \infty \\ k \to \infty}} \left\{ -2h \cdot \frac{1}{2h} - \dots - 2^3 \cdot \frac{1}{2^3} - 2^2 \cdot \frac{1}{2^2} + 2^2 \cdot \frac{1}{2^2} + 2^3 \cdot \frac{1}{2^3} + \dots + 2k \cdot \frac{1}{2k} \right\}$$

keine bestimmte Bedeutung beigelegt werden kann. Die mathematische Erwartung von x^2 aber hat einen bestimmten Sinn und zwar $+\infty$.

setzt. Sind weiter mögliche Werte von x in keinem endlichen Intervall enthalten, so bestimme man eine ähnliche Grösse zuerst für einen beliebigen Intervall und suche dann ihren Grenzwert bei unbegrenzter Erweiterung der Intervallsgrenzen. Erweist sich dieser Grenzwert von der Summierungsordnung unabhängig, so setze man

$$(7)\ldots\quad E f(x) = \lim_{\substack{b \to \infty \\ a \to -\infty}} \left\{ \lim_{\Delta x \to 0} \sum_{a}^{(b)} f(x)\, \Delta\, W(x) \right\} \text{(1)}.$$

Man beachte dabei, dass die Definitionen (6) und (7) die einfacheren Fälle (1) und (2) auch mitumfassen.

Die mathematische Erwartung einer Funktion mehrerer zufälligen Variablen kann in ganz ähnlicher Weise definiert werden. Es seien $x_1, x_2, \ldots x_n$ zufällige Variablen, deren mögliche Werte im Intervall $a_i < x_i \leqq b_i$, $(i = 1, 2, \ldots n)$ enthalten sind. Bezeichnet man $(b_i - a_i)$: $s_i = \Delta x_i$, so hat man $s_1\ s_2 \ldots s_n$ Teilintervalle, von denen jeder durch n Bedingungen

$$(8)\ldots\quad \begin{cases} a_1 + (k_1 - 1)\,\Delta x_1 < x_1 \leqq a_1 + k_1\, \Delta x_1 \\ a_2 + (k_2 - 1)\,\Delta x_2 < x_2 \leqq a_2 + k_2\, \Delta x_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_n + (k_n - 1)\,\Delta x_n < x_n \leqq a_n + k_n\, \Delta x_n \end{cases}$$

gegeben wird, wo jedes k_i von s_i Werten: $1, 2, \ldots s_i$ jeden beliebigen annehmen kann. Bezeichnet man weiter die Wahrscheinlichkeit der Erfüllung von (8) mit $\Delta\, W(x_1, x_2, \ldots x_n)$, so wird die mathematische Erwartung einer Funktion $f\,(x_1, x_2, \ldots x_n)$ durch die Gleichung

$$(9)\ldots\quad E f(x_1, x_2 \ldots x_n) =$$

$$= \lim \sum_{k_1=1}^{s_1} \sum_{k_2=1}^{s_2} \ldots \sum_{k_n=1}^{s_n} f(a_1 + k_1\,\Delta x_1,\ a_2 + k_2\,\Delta x_2, \ldots$$

$$a_n + k_n\,\Delta x_n)\,\Delta\, W(x_1, x_2, \ldots x_n),\ \Delta x_1 \to 0,\ \Delta x_2 = 0, \ldots \Delta x_n \to 0$$

oder in einer kürzeren Schreibweise durch die Gleichung

$$(10)\ldots\quad E f(x_1, x_2, \ldots x_n) =$$

$$= \lim \sum_{ai}^{(b_i)} f(x_1, x_2, \ldots x_n)\,\Delta\, W(x_1, x_2, \ldots x_n),\ (i = 1, 2, \ldots n),\ \Delta x_i \to 0$$

---

(1) Die Ausdrücke rechts in (6) und (7) sind sog. STIELTJES'sche Integrale

$$\int_a^{(b)} f(x)\, d\, W(x)\text{, bzw.,} \int_{-\infty}^{\infty} f(x)\, d\, W(x)\,.$$

Wie weitere Ausführungen zeigen sollen, kann man die dadurch definierte verallgemeinerte Auffassung der mathematischen Erwartung in ganz elementare Beweisgänge einführen, ohne dass man diese Symbole zu benützen braucht.

definiert. Bei diskreten Verteilungen geht (10) wieder in gewöhnliche (n-fache) Summen diskreter Werte, bei stetigen in gewöhnliche (n-fache) Integrale über. Die Voraussetzung der Endlichkeit des Wertbereiches kann in obiger Weise rückgängig gemacht werden.

8. — Jede mathematische Erwartung ist von der Ordnung des Summierens unabhängig. Bei endlichen Anzahlen möglicher Werte ist diese Eigenschaft trivial, in anderen Fällen beruht sie einfach auf dem Umstande, dass bei Abhängigkeit entsprechender Grenzwerte von der Ordnung des Summierens von einer mathematischen Erwartung definitionsgemäss keine Rede sein kann.

Auf diesem Grunde sind die Vorschriften des vorigen Paragraphs über Einteilung des Bereiches möglicher Werte, Summenbildung und Grenzübergänge lediglich als exemplarische zu betrachten; jede Aufgabe kann darüber ihre eigene Forderungen aufstellen. Diese Betrachtungen will ich auf einem Beispiel verdeutlichen, indem ich eine für späteren Gebrauch notwendige Ungleichung ableite.

Es seien $G_1(z)$, $G_2(z)$, ... eindeutige, stetige, nicht abnehmende, gerade Funktionen, die für alle reellen Werte von z definiert sind. Wir betrachten zuerst zwei zufällige Variablen x und y mit ihren Funktionen $G_1(x-c_1)$ und $G_2(y-c_2)$, wo c_1 und c_2 beliebige Konstanten sind.

Es sei $\xi = x - c_1$, $\eta = y - c_2$. Dann setze man:

$$\Delta\xi = \Delta\eta > 0\,,\ \xi_1 = \Delta\xi\,,\ \xi_2 = 2\Delta\xi\,,\ \xi_3 = 3\Delta\xi\,,\ \ldots;$$
$$\eta_1 = \Delta\eta\,,\ \eta_2 = 2\Delta\eta\,,\ \eta_3 = 3\Delta\eta\,,\ \ldots.$$

Die ganze Koordinatenebene (ξ, η) teilen wir in quadratische Streifen ein, deren erster mit dem inneren Quadrat zwischen $(\pm\xi_1, \pm\eta_1)$ zusammenfällt, der i-*te* aber zwischen dem Quadrat $(\pm\xi_{i-1}, \pm\eta_{i-1})$ und dem Quadrat $(\pm\xi_i\,,\ \pm\eta_i)$ liegt. Es sei mit $\Delta W(\xi_i\,,\ \eta_i)$ die Wahrscheinlichkeit sowohl im inneren Raume des i — *ten* Streifens, wie auch auf dem äusseren Rande desselben bezeichnet. Jeder Streifen sei weiter noch nach einer sachgemässen Vorschrift in s_i elementare Quadrate eingeteilt, von denen jedem einer seiner Eckpunckte (ξ_{ik}, η_{ik}) zugeordnet wird, und es sei $\Delta W(\xi_{ik}, \eta_{ik})$ die Wahrscheinlichkeit sowohl im inneren Raume des k — *ten* Quadrats des i — *ten* Streifens, wie auch an zwei in dem genannten Eckpunkte sich kreuzenden Seiten. Man hat dabei

$$(11)\ \ldots\quad \Delta W(\xi_i\,, \eta_i) = \sum_{k=1}^{s_i} \Delta W(\xi_{i,k}\,, \eta_{i,k})\,.$$

Weiter bezeichnen wir die Wahrscheinlichkeit beider unendlichen

Streifen, die durch Ungleichung

$$\xi_{i-1} < |\xi| \leqq \xi_i$$

definiert werden, mit $\Delta W_1(\xi_i)$; eine analoge Wahrscheinlichkeit für die zweite Variable sei $\Delta W_2(\eta_i)$. Unter $\Delta W_1(\xi_1)$ verstehe man dabei die Wahrscheinlichkeit in dem Streifen

$$-\xi_1 \leqq \xi \leqq \xi_1$$

und ähnlich für $\Delta W_2(\eta_1)$. Es ergibt sich dann eine evidente Beziehung

(12) . . . $$\Delta W(\xi_i, \eta_i) \leqq \Delta W_1(\xi_i) + \Delta W_2(\eta_i) \;.$$

Nun bestimmt sich die mathematische Erwartung von dem Produkt $G_1(x-c_1)\, G_2(y-c_2)$, wie folgt:

(13) . . . $$E\, G_1(x-c_1)\, G_2(y-c_2) =$$
$$= lim \sum_{i=1}^{\infty} \sum_{k=1}^{8i} G_1(\xi_{i,k})\, G_2(\eta_{i,k})\, \Delta W(\xi_{i,k}, \eta_{i,k}) \,,\; \Delta\xi = \Delta\eta \to o,$$

wo das rechte Summenzeichen sich auf kleinere Quadrate des i — *ten* Streifens, das linke aber sich auf die unendliche Folge aller quadratischen Streifen: $i = 1, 2, \ldots$ bezieht und das Limeszeichen das Grenzverfahren bedeutet, bei welchem $\Delta\xi$ und $\Delta\eta$ gegen Null konvergieren.

Setzt man nun in alle Summanden, die zum i — *ten* Streifen gehören, statt $\xi_{i,k}$, $\eta_{i,k}$ die Grössen ξ_i, η_i, als die grössten zwischen ihnen, ein, und erinnert man sich, dass G_1 und G_2 nicht abnehmende gerade Funktionen sind, so hat man, nach Massgabe der Formel (11), die Ungleichung

(14) . . . $$\sum_{k=1}^{8i} G_1(\xi_{i,k})\, G_2(\eta_{i,k})\, \Delta W(\xi_{i,k}, \eta_{i,k}) \leqq$$
$$\leqq G_1(\xi_i)\, G_2(\eta_i)\, \Delta W(\xi_i, \eta_i) \,,$$

woher unter der Berücksichtigung der Gleichung $\xi_i = \eta_i$ und der Formel (12) zufolge sich die Beziehung

(15) . . . . $$\sum_{k=1}^{8i} G_1(\xi_{i,k})\, G_2(\eta_{i,k})\, \Delta W(\xi_{i,k}, \eta_{i,k}) \leqq$$
$$\leqq G_1(\xi_i)\, G_2(\xi_i)\, \Delta W_1(\xi_i) + G_1(\eta_i)\, G_2(\eta_i)\, \Delta W_2(\eta_i)$$

ergibt. Setzt man den Ausdruck rechts in die Formel (13) ein, so erhält man gleich die gesuchte Ungleichung:

(16) . . . $$E\, G_1(x-c_1)\, G_2(y-c_2) \leqq E\, G_1(x-c_1)\, G_2(x-c_1) +$$
$$+ E\, G_1(y-c_2)\, G_2(y-c_2).$$

Für einen praktisch wichtigsten Fall hat man daraus:

$$(17)\ldots\quad E\,|\,x-c_1\,|^r\,|\,y-c_2\,|^s \leqq E\,|\,x-c_1\,|^{r+s} + E\,|\,y-c_2)^{r+s}\,.$$

Es leuchtet ein, dass diese Ungleichungen in derselben Weise auch auf beliebig viele Variablen verallgemeinert werden können. Sodann hat man:

$$(18)\ldots\quad E\,|\,x_1-c_1\,|^{r_1}\,|\,x_2-c_2\,|^{r_2}\ldots|\,x_n-c_n\,|^{r_n} \leqq \leqq E\,|\,x_1-c_1\,|^R + E\,|\,x_2-c_2\,|^R + \ldots + E\,|\,x_n-c_n\,|^R\,,$$

wo $R = r_1 + r_2 + \ldots + r_n$ ist, und à fortiori:

$$(19)\ldots\quad E\,(x_1-c_1)^{r_1}\,(x_2-c_2)^{r_2}\ldots(x_n-c_n)^{r_n} \leqq \leqq E\,|\,x_1-c_1\,|^R + E\,|\,x_2-c_2\,|^R + \ldots E\,|\,x_n-c_n\,|^R\,.$$

9. — Ich will jetzt den Begriff der *partiellen mathematischen Erwartung* einführen. Ich definiere sie als eine Summe oder den Grenzwert einer Summe, die nach oben angegebenen Vorschriften gebildet, doch lediglich über ein Teilintervall möglicher Werte einer zufälligen Variable ausgedehnt wird. Die partielle mathematische Erwartung von $f\,(x)$ im Intervall

$$a < x \leqq b$$

sei mit

$$\overset{(b)}{\underset{a}{E}}\, f\,(x)$$

bezeichnet und ähnlich für Intervalle anderer Typen (siehe § 1). Teilt man den ganzen Wertbereich einer zufälligen Variable (x) in Teilintervalle ein, die weder Teile, noch Punkte gemein haben, so wird auch die volle mathematische Erwartung $E\,f\,(x)$ in eine Summe von partiellen zerlegt, z. B.,

$$(19)'\ldots\quad E\,f(x) = \overset{(-\varepsilon)}{\underset{-\infty}{E}}\,f(x) + \overset{(\varepsilon)}{\underset{-\varepsilon}{E}}\,f(x) + \overset{\infty}{\underset{\varepsilon}{E}}\,f(x)\,.$$

Es sei nun $G\,(x)$ eine gerade, eindeutige, stetige, zunehmende Funktion, die für alle reellen Werte ihres Arguments definiert ist. Dann ist sie in jedem Intervall $a < x \leqq b$ grösser als $G\,(a)$ und nicht grösser als $G\,(b)$. Kann $\overset{(b)}{\underset{a}{E}}\;G\,(x)$ als eine Summe von einer endlichen Anzahl von Summanden definiert werden, so hat man gleich nach dem Einsetzen der Intervallsgrenzen statt aller einzelnen Werte von x die Ungleichungen:

$$(20)\ldots \qquad \overset{(b)}{\underset{a}{E}}\, G(x) > G(a)\, \overset{(b)}{\underset{a}{P}}(x)$$

und

$$(21)\ldots \qquad \overset{(b)}{\underset{a}{E}}\, G(x) \leqq G(b)\, \overset{(b)}{\underset{a}{P}}(x) \quad .$$

Kann aber $E\, G(x)$ nur als der Grenzwert einer Summe definiert werden, so beachte man, dass bei jedem Werte von Δx (siehe die Formel 6) die entsprechenden Werte der Summe

$$\sum_{a}^{(b)} G(x)\, \Delta\, W(x)$$

denselben Beziehungen wie (20) und (21) unterworfen werden; darum müssen diese letzteren auch für den Grenzwert (bei $\Delta x \to 0$) gelten Damit sind also die Ungleichungen (20) und (21) auch im allgemeineren Falle bewiesen.

Dazu noch eine Bemerkung. In weiteren Darlegungen werden wir uns meistens mit mathematischen Erwartungen befassen, als deren Argumente die Grössen vom Typus $(x - c)$, bzw, $|\, x - c\, |$ betrachtet werden können. Dabei werden wir den ganzen Wertbereich in Teilintervalle nach dem Schema

$$(22)\ldots \qquad \left\{ \begin{array}{rl} & |\, x - c\, | \leqq \varepsilon \\ \varepsilon < & |\, x - c\, | \leqq H \\ H < & |\, x - c\, | \leqq \infty \end{array} \right.$$

einteilen. *Es sei darum im voraus festgestellt, dass die Grenzen bei dem Symbol einer partiellen mathematischen Erwartung von dem Typus* $Ef\ (x - c)$ *nach Massgabe des Schema* (22) *geschrieben werden, wenn über sie nicht in einer anderen Weise ausdrücklich verfügt wird.* So bedeutet uns also

$$(23)\ldots \qquad \left\{ \begin{array}{l} \overset{(\varepsilon)}{\underset{(o)}{E}} f(x - c) = \overset{x - c \leqq \varepsilon}{\underset{x - c \geqq -\varepsilon}{E}} f(x - c) \\ \overset{(H)}{\underset{\varepsilon}{E}} f(x - c) = \overset{x - c < -\varepsilon}{\underset{x - c \geqq -H}{E}} f(x - c) + \overset{x - c \leqq H}{\underset{x - c > \varepsilon}{E}} f(x - c) \\ \overset{\infty}{\underset{H}{E}} f(x - c) = \overset{x - c < -H}{\underset{x - c \geqq -\infty}{E}} f(x - c + \overset{x - c \leqq \infty}{\underset{x - c > H}{E}} f(x - c)\ , \end{array} \right.$$

wo die erläuternden Bezeichnungen rechts, obgleich schwerfällig, doch, glaube ich, selbstverständlich sind.

Die Ungleichungen (20), bzw, (21) sollen hier noch in der Form geschrieben werden, in der sie weiter zum wirklichen Gebrauch kommen. Es sei G das Symbol einer geraden Funktion von der oben angegebenen Art. Wendet man die Ungleichungen (20) und (21) an die zufällige Variable $| x - c |$ an, so hat man unmittelbar

$$(24)\ldots \qquad \underset{a}{\overset{(b)}{E}} G(x - c) > G(a) \underset{a}{\overset{(b)}{P}} | x - c |$$

und

$$(25)\ldots \qquad \underset{a}{\overset{(b)}{E}} G(x - c) \leqq G(b) \underset{a}{\overset{(b)}{P}} | x - c | \leqq G(b)$$

wo das zweite Gleichungszeichen in (25) wegfällt, wenn die partielle mathematische Erwartung nicht mit der vollen identisch ist.

10. — Die mathematischen Erwartungen von einigen bestimmten Typen nennt man gewöhnlich *Momente.* Diesen Namen werde ich in einem erweiterten Sinne gebrauchen.

Unter einer G, bzw., *U-Funktion* verstehe man eine (1)′ *gerade, bzw.,* (1)″ *ungerade Funktion, die* (2) *im ganzen reellen Wertgebiet ihres Argumentes definiert,* (3) *überall stetig,* (4) *eindeutig und* (5) *zunehmend ist, und die* (6) *für den Nullwert ihres Arguments selbst zu Null wird.* Die mathematische Erwartung

$$E\, G(x - v) \ , \ bzw. \ , \ E\, U(x - v)$$

nenne ich den G, bzw., *U Moment,* die Funktionen $G\ (x\text{-}v)$, bzw., $U(x\text{-}v)$ ihre *Grundfunktionen* und die Grösse v den *Bezugswert.* Genügt ein Bezugswert (v_0) eines geraden Moments der Bedingung

$$(26)\ldots \qquad E\, G(x - v_0) \leqq E\, G(x - v)$$

bei jedem beliebigen Werte von v, bzw., genügt ein Bezugswert eines ungeraden Moments der Bedingung

$$(27)\ldots \qquad E\, U(x - v_0) = o \ ,$$

so nenne ich v_0 den *zentralen Bezugswert* und den darauf bezogenen Moment den *wohlzentrierten G,* bzw., *U*-Moment.

Es versteht sich von selbst, dass diese Bezeichnungen auch auf Fälle anzuwenden sind, wo Grundfunktionen einigen von obenaufgezählten Requisiten nicht genügen (so sehe z. B. unten den Moment $u_{,)}$), doch werden bei einer allgemeineren Rede von G, bzw., U - Momenten keine solchen, so zu sagen, ausgearteten Typen mitgemeint.

Gehen wir jetzt zu bestimmteren Gebilden über. Man betrachtet gewöhnlich gerade, bzw., ungerade Momente von der Art:

$$m_{2r} = E(x - v)^{2r} \textit{ und } m_{2r+1} = E(x - v)^{2r+1} \ ,$$

wozu noch manchmal sich die die Reihe der geraden Momente vervollständigenden Grössen

$$E \mid x - v \mid^{2r+1}$$

hinzugesellen; die Reihe der ungeraden Momente aber bleibt zur Zeit noch, *m. W.*, lückenhaft.

Unter einem *geraden Moment vom Grade* r, wo r eine beliebige reelle nicht negative Zahl ist, verstehe ich eine Grösse

$$(28)\ldots \qquad g_r = E \mid x - v \mid^r \ ,$$

die auch $g_r\ (x,\ v)$ nötigenfalls geschrieben wird, und unter einem *ungeraden Moment vom Grade* r, wo r wieder den obigen Bedingungen genügt, die Grösse

$$(29)\ldots \qquad u_r = E\, \underset{0}{sgn}\,(x - v) \mid x - v \mid^r$$

oder in ausführlicherer Schreibweise $u_r\ (x,\ v)$. Dabei verstehe ich unter $sgn_0\,(z)$, mit einer Abweichung von der üblichen von KRONECKER stammenden Definition des Symbols sgn (z), die Grösse:

$$(30)\ldots \qquad \underset{0}{sgn}\,(z) = \begin{cases} -1 & , \quad \textit{bei } z < 0 \\ -1 \leqq \theta \leqq +1 & \qquad\quad z = 0 \\ +1 & \qquad\quad z > 0 \end{cases}$$

In der graphischen Darstellung stellt $sgn_0\,(z)$ einen ununterbrochenen Linienzug dar, der aus drei geradlinigen Stücken besteht, deren zwei der z-Axe parallel von $-\infty$ bis 0 und von 0 bis ∞ auf der Höhe -1, bzw., $+1$ verlaufen, der dritte aber sie durch einen zur z-Axe senkrechten von -1 bis zu $+1$ durch 0 gehenden ebenso geradlinigen Zug verbindet. Das KRONECKER'sche Symbol sgn (z) aber hat bekanntlich die Bedeutungen -1, 0 und $+1$, je nachdem $z < 0$, bzw., $= 0$, oder > 0 ist.

Diese Abweichung wird uns durch die Ausnahmestellung des ungeraden Moments vom Grade 0

$$(31)\ldots \qquad u_0\,(x\ ,\ v) = E\, \underset{0}{sgn}\,(x - v)$$

aufgedrängt, die dadurch gewissermassen ausgemerzt wird. Wir müssen darum dieser Grösse einige Betrachtungen widmen.

Definitionsmässig hat man

$$u_0\,(x\ ,\ v) = E\, \underset{0}{sgn}\,(x - v) \mid x - v \mid^0$$

Bei $x \lessgtr v$ ist $|x-v|^0 = 1$, bei $x = v$ hat man aber eine Unbestimmtheitsstelle 0^0, die dadurch unschädlich gemacht wird, dass die Grösse $sgn_\theta\,(0)$ ebenfalls mehrdeutig (von der Mächtigkeit des Kontinuums) ist. So darf also $(x-v)^0$ für den ganzen reellen Wertbereich gleich 1 gesetzt werden und so hat man als die Definition von $u_0\,(x, v)$ die Formel (31).

Man bemerkt leicht weiter, dass jeder ungerade Moment $u_r\,(x, v)$, bei $r > 0$ eine stetige, dabei monotone und zwar abnehnemende Funktion von v ist. Darum hat die Gleichung

$$u_r(x, v) = 0$$

stets eine und dabei lediglich eine reelle Wurzel, den Fall ausgenommen wo u_r bei jedem Werte von v unendlich ist. Was die Gleichung $u_0\,(x, v) = o$ betrifft, so kann sie bei der Definition mittelst des KRONECKER'schen Symbols sgn in einigen Fällen der unstetigen Verteilung gar keine Wurzel haben. In der Tat, ist

$$(32)\ldots\quad E\,sgn\,(x-v)\,|x-v|^0 = -\overset{0}{\underset{-\infty}{P}}(x-v) + o\,.\,p_v + \overset{\infty}{\underset{0}{P}}(x-v) =$$

$$= -\overset{v}{\underset{-\infty}{P}}(x) + \overset{\infty}{\underset{v}{P}}(x)\;.$$

Ist die Wahrscheinlichkeit p_k von x_k von Null verschieden, so vermindert sich Esgn $(x-v)\,|x-v|^0$ beim Uebergang von $v < x_k$ zu $v = x_k$ sprunghaft um p_k und weiter macht sie noch einmal den gleichen Sprung beim Uebergang von $v = x_k$ zu $v > x_k$. So kann sie auch von positiven zu negativen Werten überspringen, ohne dass sie bei irgend einem Werte von v gleich Null wird. Definiert man aber u_0 mittelst unseres Symbols sgn_θ, so hat man

$$(33)\ldots\quad \underset{0}{E\,sgn}\,(x-v)\,|x-v|^0 = \underset{\theta}{E\,sgn}\,(x-v) =$$

$$= -\overset{v}{\underset{-\infty}{P}}(x) + \theta\,p_v + \overset{\infty}{\underset{v}{P}}(x)\;.$$

Die Sprünge bleiben jetzt auf denselben Unstetigkeitsstellen, wie vorher, doch läuft die Grösse $E\,sgn_\theta\,(x-v)$ jedesmal einen ununterbrochenen Zug durch von einem Niveau bis zu dem anderen. Geht sie dabei von positiven zu negativen Werten über, so nimmt sie unterwegs auch den Nullwert an. Darum muss also, bei unserer Definition, die Gleichung

$$u_0\,(x, v) = o$$

stets mindestens eine Wurzelh aben. Ich nenne sie, wie üblich, die (theoretische) *Mediane* von x. Bei unstetigen Verteilungen kann es vorkommen, dass diese Rolle allen Werten in einem Intervall $x_{i-1} < x < x_i$ zukommt. Dieser Umstand erweist sich für uns weiter ohne Belang (siehe Kap. III, Satz 6).

11. — Betrachtet man die Momente g_r und u_r als Funktionen von v, so findet man, dass sie mindestens bei $r > 1$ differenzierbar sind. Ihre Ableitungen findet man, wie folgt.

Bei einer diskreten Verteilung, bei $v = x_k$ und $x_k < v + \Delta v < < x_{k+1}$, hat man

$$(34)\ldots \begin{cases} g_r(x, v) = \sum\limits_{i=1}^{k-1} p_i (v - x_i)^r + \sum\limits_{i=k+1}^{n} p_i (x_i - v)^r \\ g_r(x, v + \Delta v) = \sum\limits_{i=1}^{k-1} p_i (v + \Delta v - x_i)^r + \\ + \sum\limits_{i=k+1}^{n} p_i (x_i - v - \Delta v)^r + p_k (\Delta v)^r , \end{cases}$$

woher, bei $r > 1$,

$$(35)\ldots \quad \lim_{\Delta v \to o} \frac{g_r(x, v + \Delta v) - g_r(x, v)}{\Delta v} = r \sum_{i=1}^{k-1} p_i (v - x_i)^{r-1} - \\ - r \sum_{i=k+1}^{n} p_i (x_i - v)^{r-1}$$

oder

$$(36)\ldots \quad \frac{dg_r}{dv} = - r\, u_{r-1} \quad .$$

In ganz ähnlicher Weise findet man

$$(37)\ldots \quad \frac{du_r}{dv} = - r\, g_{r-1} \quad .$$

Durch Betrachtung der entsprechenden Grenzwerte (siehe oben § 7) überzeugt man sich, dass diese Formeln auch in allgemeineren Fällen richtig sind (1).

Für g_1 findet man durch dasselbe Verfahren, im Fall einer

(1) Mein geehrter Kollege Prof. M. KRAWTSCHUK weist mich auf den Umstand hin, dass die Formeln (36) und (37) sich noch dadurch beweisen lassen, dass man in dem Ausdruck entsprechender Momente den Exponent r als den Grenzwert eines rationellen Bruches darstellt und zwar für einen ungeraden Moment von der Form $r_1 = \frac{2s+1}{2t+1}$ und für einen geraden von der Form $r_2 = \frac{2s}{2t+1}$. In der Tat gehen dann unsere Momente in $u_{r_1} = E(x - v)^{r_1}$ und $g_{r_2} = E(x - v)^{r_2}$ über, woher das gesagte unmittelbar einleuchtet.

diskreten Verteilung und bei $v = x_k$, die rechte Ableitung

$$(38) \ldots \quad \frac{dg_1}{dv} = \frac{d}{dv} E \mid x - v \mid = \sum_{i=1}^{k-1} p_i + p_k - \sum_{i=k+1}^{n} p_i$$

und die linke Ableitung

$$(39) \ldots \quad \frac{dg_1}{dv} = \frac{d}{dv} E \mid x - v \mid = \sum_{i=1}^{k-1} p_i - p_k - \sum_{i=k+1}^{n} p_i \; .$$

Wollten wir die Cronecker'sche Funktion (sgn) der Definition unserer Momente zugrunde legen, so hätten wir für die rechte, bzw., linke Ableitung die Ausdrücke

$$\frac{dg_1}{dv} = - u_0 + p_k \quad ,$$

bzw.,

$$\frac{dg_1}{dv} = - u_0 - p_k$$

gehabt, was wieder eine empfindliche Ausnahme aus der allgemeinen Regel bilden würde. Unsere Definition lässt aber auch diese Ausnahme wegfallen, doch ist dazu noch ein Kunstgriff nötig. Und zwar, wollen wir die Ableitung von g_1 in jedem Punkte, wo sie einen Sprung macht, vervollständigen, indem wir uns diesen lezteren, als einen ununterbrochenen denken und so zwischen der rechten und der linken Ableitung die ganze Menge aller Uebergangswerte als die Werte der Ableitung in diesem Punkte einreihen. So gehen die Formeln (38) und (39) in eine einzige

$$(40) \ldots \quad \frac{dg_1}{dv} = \sum_{i=1}^{k-1} p_i + \theta\, p_k - \sum_{i=k+1}^{n} p_i \, , \; (-1 \leqq \theta \leqq +1) \; ,$$

über, was mit Hilfe des Symbols sgn_θ als

$$(41) \ldots \quad \frac{dg_1}{dv} = - E\, sgn_\theta\, (x - v) = - u_0$$

geschrieben wird. Bemerkt man noch, dass

$$(42) \ldots \quad \frac{du_1}{dv} = \frac{d}{dv} E (x - v) = - E \mid x - v \mid^0 = - g_0 = - 1$$

ist, so sieht man, dass die Formeln (36) und (37) für alle g_r, bzw., u_r gelten, sofern $r \geqq 1$ ist.

Die Grundfunktion des geraden Momentes $g_r(x, v)$ ist $\mid x - v \mid^r$. Da sie als eine eindeutige definiert wird, indem wir lediglich ihre reellen und positiven Werte in Betracht ziehen, so genügt sie allen

Anforderungen, die oben in Betreff einer G-Funktion aufgestellt wurden, wodurch uns die Grundlage für allgemeinere Betrachtungen des nächsten Kapitels geschaffen wird.

Den zentralen Bezugswert des geraden Moments $g_r(x, v)$ nenne ich den *Zentralwert* vom Grade r und bezeichne mit C_r (1). Die Zentralwerte zerfallen in zwei Klassen, je nachdem $r \geqq 1$, oder $o < r < 1$ ist. In obigen Ausführungen wurden nur die Grundlagen zur Behandlung der Zentralwerte der ersten Klasse gelegt. Es wurde nämlich gezeigt, dass der Zentralwert von g_r, bei $r \geqq 1$, sich aus der Gleichung $u_{r-1} = o$ findet, woher es kommt, dass er mit dem Zentralwerte von u_{r-1} zusammenfällt und g_r zum Minimum macht. Allgemeinbekannt sind: C_1, die Mediane, und C_2, die mathematische Erwartung, die sich aus den Gleichungen: $E\, sgn_0\, (x - C_1) = o$, bzw., $E\,(x - C_2) = o$ bestimmen lassen und die geraden Momente: $E \mid x - v \mid$, bzw., $E\,(x - v)^2$ minimalisieren. Was die letzte Grösse betrifft, so bezeichnet man sie gewöhnlich, bei $v = E\,x = C_2$, mit $\mu_2 = E\,(x - Ex)^2$ und nennt sie *Streuung*. In der Reihe der (nach K. Pearson) sog. zentralen Momente $\mu_r = E\,(x - E\,x)^r$, $(r = 1, 2, \ldots)$, ist die Streuung der einzige in unserem Sinne wohlzentrierte Moment. Es wird sich im Weiteren herausstellen, dass die Momente dieser lezten Art für die Formulierung und Begründung mancher allgemeineren Sätze unersetzlich sind.

Was die Zentralwerte der zweiten Klasse $(o < r < 1)$ betrifft, so werden sie in dieser Schrift nicht näher betrachtet, was zur Folge hat, dass einigen Sätzen des nächsten Kapitels entsprechende Einschränkungen auferlegt werden.

Drittes Kapitel.

Ueber hinreichende und notwendige Bedingungen des Bestandes der stochastischen Asymptoten und Grenzwerte.

12. — In der folgenden Darstellung wird stets mit x eine eindimensionelle zufällige Variable und mit φ eine unabhängige Variable be-

---

(1) Vergl. D. Jackson, *Note on the median of a set of numbers*, « Bull. Amer. Mathem. Soc. », Vol. XXVII, n. 4, 1921, und von demselben Verfasser, *Note on quartiles and allied measures*, ibid., Vol. XXIX, n. 1, 1923. Der Verfasser beweist, dass in Fällen, wo C_1 einen einzigen Wert hat, $\lim\limits_{\alpha \to 0} C_{1+\alpha} = C_1$ ist. Daher sein Vorschlag: für die Fälle, wo C_1 mehrdeutig wird, die Mediane als $\lim\limits_{\alpha \to 0} C_{1+\alpha}$ zu definieren. In dieser Arbeit habe ich keine Veranlassung gehabt diesen Schritt

zeichnet, die mit der ersteren stochastisch verbunden ist. Uebersichtlichkeitswegen wiederhole ich die Definition des § 2, um ihr weiter einige keines besonderen Beweises bedürftigen Sätze folgen zu lassen.

Definition 1. *Ist x eine zufällige Variable, die mit einer unabhängigen Variable φ stochastisch verbunden ist, und ist bei jedem gegebenen beliebig kleinen* ε

$$(1)\ldots \qquad \lim_{\varphi_1,\ \varphi_2,\ldots} \overset{(\varepsilon)}{P_{(o)}} \mid x - f(\varphi) \mid = 1 ,$$

so ist $f(\varphi)$ eine stochastische Asymptote von x, was mit $as_B(x) = f(\varphi)$ bezeichnet wird.

Definition 2. *Ist $as_B(x) = f(\varphi) = c$ (onst.), so ist c der stochastische Grenzwert von x. Man schreibe dann $lim_B(x) = c$.*

Folgesatz A. Ist $\underset{\varphi_1,\ \varphi_2\ldots}{as_B}(x) = f(\varphi)$, und hat $f(\varphi)$ einen Grenzwert $\lim\limits_{\varphi_1,\ \varphi_2,\ldots} f(\varphi) = c$, so ist $c = lim_B(x)$.

Folgesatz B. Die Bedingung (1) ist der Bedingung

$$(2)\ldots \qquad \lim_{\varphi_1,\ \varphi_2,\ldots} \overset{\infty}{P_{\varepsilon}} \mid x - f(\varphi) \mid = 0$$

äquivalent.

Folgesatz C. Ist $\underset{\varphi_1,\ \varphi_2,\ldots}{as_B}(x) = f(\varphi)$, so ist $\underset{\varphi_1,\ \varphi_2,\ldots}{lim_B} \{x - f(\varphi)\} = 0$ und vice versa.

Bezeichnung. Besteht zwischen zwei Funktionen einer und derselben unabhängigen Variable die Beziehung $\lim\limits_{\varphi_1,\ \varphi_2,\ldots} \{f(\varphi) - F(\varphi)\} = 0$, so sind sie für einander *Asymptoten*. Sodann schreibe man: $f(\varphi) = as\, F(\varphi)$.

Folgesatz D. Ist $v = as_B(x)$ und sind $v_1 = as(x)$, $v_2 = as(x)$, ... $v_k = as(x)$, so werden auch v_1, v_2, ... v_k stochastische Asymptoten von x sein. Ich nenne sie die *äquivalenten stochastischen Asymptoten*. Hat also eine zufällige Variable *eine* stochastische Asymptote, so hat sie auch eine unbegrenzte Menge von solchen, die ihr und zueinander äquivalent, keine aber die ihr nicht äquivalent sind.

Folgesatz E. Sind $v_1 = \underset{\varphi_1,\ \varphi_2\ldots}{as_B}(x)$ und $v_2 = \underset{\varphi_1,\ \varphi_2,\ldots}{as_B}(x)$, so ist $\lim\limits_{\varphi_1,\ \varphi_2,\ldots} (v_1 - v_2) = 0$, woher die Gleichung $\lim v_1 = \lim v_2$ folgt, wenn irgend eine von diesen Grössen einen Grenzwert hat.

---

mitzumachen. D. JACKSON'S Standpunkt ist übrigens von dem hier vertretenen recht verschieden; doch ziehe ich es vor das Urteil darüber dem kompetenten Leser selbst zu überlassen.

13. — Gehen wir nun zur Aufstellung der Bedingungen über, bei deren Erfüllung eine zufällige Variable (x) stochastische Asymptoten besitzen muss, bzw., besitzen kann. Das Problem wird dabei nur in seinen allgemeinsten, mathematisch relevanten Zügen aufgefasst.

Die Darstellung fängt mit einigen allgemeinen Sätzen an, die weiter in gehöriger Weise spezifiziert werden. Unter einer G-Funktion wird (wie schon oben in § 10) eine (1) im ganzen reelen Wertgebiet definierte, (2) gerade, (3) eindeutige, (4) stetige und (5) zunehmende Funktion gemeint, von der Art, dass (6) $G(0) = 0$ ist.

Satz 1. *Ist eine zufällige Variable* (x) *mit einer unabhängigen Variable* (φ) *stochastisch verbunden, und ist*

$$(3)\ldots \qquad \lim_{\varphi_1,\ \varphi_2,\ldots} E\, G(x - v) = 0\ ,$$

wo $v = f(\varphi)$ *ist, so ist* $as_B\,(x) = v$.

Da bei jedem gegebenen beliebig kleinen ε

$$E\, G(x - v) \geqq \overset{\infty}{\underset{\varepsilon}{E}}\, G(x - v)$$

ist, so hat man der Formel (24) des 2 ten Kapitels zufolge, dass à fortiori

$$E\, G(x - v) > G(\varepsilon) \overset{\infty}{\underset{\varepsilon}{P}} \mid x - v \mid$$

ist, woher sich die Ungleichung

$$(4)\ldots \qquad \overset{\infty}{\underset{\varepsilon}{P}} \mid x - v \mid < \frac{E\, G(x - v)}{G(\varepsilon)}$$

ergibt (1). Da $G(\varepsilon)$ bei jedem von 0 verschiedenen ε selbst von Null

(1) Ersetzt man in MARKOFF'S Ungleichung

$$P \left\{ U \leqq t\, E\, U \right\} > 1 - \frac{1}{t}\ ,$$

wo $U > 0$ ist, die Grösse U durch $G(x - v)$, und $t\, E\, U$ durch $t\, E\, G(x - v) = G(\varepsilon)$, so kommt man leicht zu Ungleichung (4); ersetzt man in Ungleichung (4) die Grösse $G(x - v)$ durch U, so kommt man wieder zu MARKOFF'schen Ungleichung. Ebenso sind die Ungleichungen von TSCHEBYSCHEFF, von MEDOLAGHI, K. PEARSON, von P. CANTELLI, C. LURQUIN, und A. GULDBERG aus (4) abzuleiten. A. MARKOFF, *Wahrscheinlichkeitsrechnung*, 4 Aufl. 1924, p. 86, MEDOLAGHI, *Verhandlungen des IV Congress der Aktuarien*, Wien 1909, Vol. I (zitiert nach F. VINCI, « Giorn. degle Economisti », 1921, XII), K. PEARSON, « Biometrika », XII, p. 284, P. CANTELLI, *Rendiconti della R. Accademia dei Lincei*, 1916, C. LURQUIN, *Comptes rendus*, t. 175, 1922, p. 1035, (drei letzte Abhandlungen zitiere ich nach der folgenden), A. GULDBERG, *Comptes rendus*, t. 175, 1922, p. 1382, DERSELBE, « Metron », Vol. III, N. 1, 1923.

verschieden ist und $E\,G\,(x - v)$ voraussetzungsgemäss den Grenzwert Null hat, so hat man

$$\lim_{\varphi_1,\,\varphi_2,\,\ldots} \overset{\infty}{P}_{\varepsilon} \mid x - v \mid = 0 .$$

und

$$v = as_B(x) \;,$$

was zu beweisen war.

Folgesatz 1 a. Ist $\lim\limits_{\varphi_1,\,\varphi_2,\ldots} E\,G\,(x - v) = 0$, *so ist auch der zentrale Bezugswert* $v_0 = as_B\,(x)$. *Die Grösse* v *ist dabei eine Asymptote von* v_0 *so, dass* $\lim\limits_{\varphi_1,\,\varphi_2,\ldots} (v - v_0) = 0$ *ist.*

Diese Folgerung ist evident, da nach der Definition von v_0 (§ 10)

$$E\,G\,(x - v_0) \leqq E\,G\,(x - v)$$

ist. Konvergiert $E\,G\,(x - v)$ gegen Null, so tut es auch $E\,G\,(x - v_0)$ Das Uebrige geht aus dem Satz 1 und dem Folgesatz E (§ 12) hervor.

Folgesatz 1 b. Konvergiert ein gerader Moment $g_r\,(x, v)$ *gegen Null, so sind sowohl sein Bezugswert* (v), *wie auch der Zentralwert* (C_r) *stochastische Asymptoten von* x.

Erinnert man sich, dass $g_r\,(x, v) = E \mid x - v \mid^r$, bei $r > 0$, ein G. Moment ist, so folgt aus der Bedingung: $\lim\limits_{\varphi_1,\,\varphi_2\,\ldots} g_r\,(x, v) = 0$, dem Satz 1 zufolge, dass $v = as_B\,(x)$ ist, und, nach dem Folgesatz 1a, dass auch $C_r = as_B\,(x)$ sein muss.

Es seien zwei spezielle Fälle hervorgehoben. (1) Ist $\lim\limits_{\varphi_1,\,\varphi_2,\ldots} E\,(x - v)^2 = 0$, so ist auch $\lim\limits_{\varphi_1,\,\varphi_2,\ldots} E\,(x - Ex)^2 = 0$. Sowohl v, wie auch $C_2 = Ex$ sind dann stochastische Asymptoten von x. (2) Ist $\lim\limits_{\varphi_1,\,\varphi_2,\ldots} E \mid x - v \mid = 0$, so ist auch $\lim\limits_{\varphi_1,\,\varphi_2,\ldots} E \mid x - C_1 \mid = 0$, woher $v = as_B\,(x)$ und $C_1 = $ $=$ Med. $(x) = as_B\,(x)$. Besonders wichtig ist das Kriterium

$$(5)\ldots \qquad \lim_{\varphi_1\,,\,\varphi_2\,,\ldots} E\,(x - E\,x\,)^2 = 0\,.$$

Folgesatz 1 c. Konvergiert ein gerader Moment vom Grade r gegen Null und hat ein Zentralwert C_r *einen Grenzwert (bzw., ist* C_r *konstant), so ist* $\lim C_r$ *(bzw.,* C_r *) der stochastische Grenzwert. Der Bezugswert von* g_r *muss dann entweder mit* $\lim C_r$ *(bzw.,* C_r *) zusammenfallen oder gegen diese Grösse konvergieren.*

Nach allem früheren ist der Satz keines Beweises bedürftig.

Dass die Bedingung des Satzes 1 eine lediglich hinreichende, nicht aber eine notwendige ist besagt der

Satz 2. Eine zufällige Variable kann stochastische Asymptote haben, wenn auch alle darauf bezogenen geraden Momente beliebigen positiven Grades $(r > 0)$ *unbegrenzt zunehmen.*

Zum Beweise genüge ein Beispiel. Es sei die Wahrscheinlichkeitsverteilung durch das Schema

$$(6)\ldots \qquad \left\{ \begin{array}{llllll} x: & -n\,, & -\frac{1}{n}\,, & \frac{1}{n}\,, & n & \\ p: & \frac{1}{2\,lgn}\,, & \frac{1}{2}-\frac{1}{2\,lgn}\,, & \frac{1}{2}-\frac{1}{2\,lgn}\,, & \frac{1}{2\,lgn} & \end{array} \right\}$$

gegeben. Es ist klar, dass x einen stochastischen Grenzwert und zwar $\lim\limits_{n \to \infty} {}_B (x) = 0$ hat. Doch wächst der darauf bezogene gerade Moment

$$E \mid x - 0 \mid^{\alpha} = \frac{n^{\alpha}}{lgn} + \frac{1}{n^{\alpha}}\left(1 - \frac{1}{lgn}\right)$$

bei unbegrenzter Zunahme von n über alle Grenzen hinaus, da n^{α}/lgn bekanntlich mit n unbegrenzt zunimmt, wie klein α auch sein möge (1).

Satz 3. Ist eine G — *Funktion von der Art, dass* $\lim\limits_{z \to \infty} G(z) = G(\infty)$ *eine endliche Grösse ist, so wird die Bedingung*

$$(7)\ldots \qquad \lim_{\varphi_1\,,\,\varphi_2\,,\,\ldots} E\, G(x - v_0) = 0\,,$$

wo v_0 *ein zentraler Bezugswert ist, nicht lediglich eine hinreichende sondern auch eine notwendige dafür, dass* x *stochastische Asymptoten habe. Wird sie erfüllt, so wird* $v_0 = as_B(x)$. *Wird sie nicht erfüllt, so hat* x *keine stochastischen Asymptoten überhaupt.*

Dass diese Bedingung hinreichend ist, wurde schon im Satz 1 bewiesen. Es soll nur also dargetan werden, dass sie auch notwendig ist. Schreibt man

$$E\, G(x - v_0) = \mathop{E}_{(o)}^{(\varepsilon)} G(x - v_0) + \mathop{E}_{\varepsilon}^{\infty} G(x - v_0)\,,$$

(1) Die Sätze 1 und 2 sind als in verschiedenen Richtungen vorgenommene Verallgemeinerungen einiger schon bekannten entstanden. Vergl., z. B., A. MARKOFF, *op. cit.*, p. 116-17, 120. Die Erweiterung dieser Sätze auf beliebige Bezugswerte halte ich besonders von Wichtigkeit, da es damit der Entscheidung der Frage vorgearbeitet wird, ob die mathematische Erwartung, die üblicher Weise allein in diesem Zusammenhange berücksichtigt wird, stets eine stochastische Asymptote sein muss, wenn eine zufällige Variable eine solche überhaupt hat.

so hat man nach Massgabe der Formel 25 (Kap. II)

$$E\,G(x - v_0) \leqq G(\varepsilon) + G(\infty) \overset{\infty}{\underset{\varepsilon}{P}} \mid x - v_0 \mid,$$

woher sich die Ungleichung

$$(8)\ldots \qquad \overset{\infty}{\underset{\varepsilon}{P}} \mid x - v_0 \mid \geqq \frac{E\,G(x - v_0) - G(\varepsilon)}{G(\infty)}$$

ergibt. Ist die Bedingung

$$\lim_{\varphi_1,\,\varphi_2,\,\ldots} E\,G(x - v_0) = 0$$

nicht erfüllt, so wird es in jeder unbegrenzten Teilfolge $\varphi_{n_0}, \varphi_{n_0+1}\ldots$, eine unbegrenzte Menge von solchen Werten φ_n geben, dass $E\,G\,(x - v_0)$ nicht kleiner als eine konstante von 0 verschiedene positive Grösse c sein wird. Da aber bei genügend kleinem ε die Grösse $G\,(\varepsilon)$ beliebig klein und somit auch kleiner als c werden kann, so wird also die Wahrscheinlichkeit $\overset{\infty}{\underset{\varepsilon}{P}} \mid x - v_0 \mid$ nie dauernd unter eine positive Grösse

$$\frac{c - G(\varepsilon)}{G(\infty)}$$

fallen können. Die Grösse v_0 ist daher keine stochastische Asymptote von x, und so hat x, dem Satz 1a zufolge, keine stochastischen Asymptoten überhaupt.

Es sei nun erlaubt sich einer Bemerkung MARKOFF's (1) zu erinnern, dass es schwerlich je gelingen werde in dieser, ebenso wie auch in manchen anderen Fragen, eine hinreichende und notwendige Bedingung aufzustellen. So kann dem Satz 3 eine prinzipielle Bedeutung nicht gänzlich abgesprochen werden. Vielleicht wird er in mathematisch mehr geschulten Händen auch zu irgend welchen weiteren Ergebnissen führen können.

14. — Betrachten wir nun den Fall einer zufälligen Variablen, deren mögliche Werte in einem Intervall enthalten sind, dessen Länge eine feste obere Grenze hat. Da gelten folgende Sätze.

Satz 4. Es sei x eine zufällige Variable, die mit einer unabhängigen Variable φ stochastisch verbunden ist, und es sei $v = f(\varphi)$. Hat nun die Differenz möglicher Werte von x und der Grösse v, ihrem absoluten

(1) *Op. cit.*, p. 116.

Werte nach, eine feste obere Grenze (H), so wird die Bedingung

$$(9)\ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots} E\, G(x - v) = 0$$

nicht lediglich hinreichend, sondern auch notwendig dafür, dass $v = as_B\,(x)$ ist.

Der Beweis gestaltet sich ganz ähnlich dem des Satzes 3. Da

$$E\, G(x - v) = \overset{(\varepsilon)}{\underset{(o)}{E}}\, G(x - v) + \overset{\infty}{\underset{\varepsilon}{E}}\, G(x - v)$$

ist, so hat man nach Massgabe der Formel 25 (Kap. II)

$$E\, G(x - v) \leqq G(\varepsilon) + G(H)\, \overset{H}{\underset{\varepsilon}{P}}\, |\, x - v\, | ,$$

woher sich die Ungleichung (1)

$$(10)\ldots \qquad \overset{\infty}{\underset{\varepsilon}{P}}\, |\, x - v\, | = \overset{H}{\underset{\varepsilon}{P}}\, |\, x - v\, | \geqq \frac{E\, G(x - v) - G(\varepsilon)}{G(H)}$$

ergibt, die uns in derselben Schlussweise, wie oben, zum Ergebnis führt, dass, wenn die Bedingung (9) nicht erfüllt ist, auch die Bedingung

$$\lim_{\varphi_1, \varphi_2 \ldots} \overset{\infty}{\underset{\varepsilon}{P}}\, |\, x - v\, | = 0$$

nicht gelten kann. Unsere Bedingung (9) ist also eine notwendige, nach dem Satz 1 aber ist sie auch hinreichend.

Als eine unmittelbare Folgerung des bewiesenen Satzes haben wir nun den wichtigen.

Satz 5. Sind alle möglichen Werte einer zufälligen Variable (x) in einem Intervall enthalten, dessen Länge eine feste obere Grenze (H) hat, so wird die Gleichung

$$(11)\ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots} E\, G(x - v_0) = 0 \ ,$$

wo v_0 ein zentraler Bezugswert ist, eine sowohl hinreichende wie auch notwendige Bedingung dafür, dass x stochastische Asymptoten hat.

Es ist wieder nur die Notwendigkeit der Bedingung (11) festzustellen. Es sei angenommen, dass x eine stochastische Asymptote habe, die wir mit v bezeichnen wollen. Dann wird sich solch ein Wert der unabhängigen Variable φ_k finden, dass bei jedem φ_n $(n > k)$

---

(1) A. A. Markoff, *op. cit.*, p. 119-20; F. P. Cantelli, *Sulla legge dei grandi numeri*, loc. cit., p. 332.

die Wahrscheinlichkeit $\overset{(\varepsilon)}{\underset{(0)}{P}} \mid x - v \mid$ der Einheit beliebig nahe wird, wie klein ε auch sein möge. Darum muss es also dann stets mindestens *einen* solchen möglichen Wert von x geben, dass die Differenz, zwischen ihm und v, ihrer absoluten Grösse nach, eine beliebig kleine positive Grösse ε nicht übertreffe; darum wird auch die Differenz $\mid x - v \mid$, für alle möglichen Werte von x, voraussetzungsgemäss nicht grösser als eine endliche Grösse $H' = H + h$ sein können, wo h eine beliebige positive endliche Grösse ist. Es folgt daher nach dem vorigen Satz 4, dass

$$\lim_{\varphi_1, \varphi_2, \ldots} E\, G(x - v) = 0$$

und nach dem Satz 1 a, dass auch

$$\lim_{\varphi_1, \varphi_2, \ldots} E\, G(x - v_0) = 0$$

sein muss, was zu beweisen war.

Folgesatz 5 a. Sind alle möglichen Werte einer zufälligen Variable (x) *in einem Intervall enthalten, dessen Länge eine feste obere Grenze hat, so ist für den Bestand stochastischer Asymptoten sowohl hinreichend, wie notwendig, dass*

$$(12) \ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots} E \mid x - C_k \mid^r = 0$$

ist.

Folgesatz 5 b. Hat eine zufällige Variable, deren mögliche Werte den Bedingungen des vorigen Satzes genügen, stochastische Asymptoten, so gilt die Beziehung

$$(13) \ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots} (C_r - C_s) = 0 \ ,$$

die sich in die Gleichung

$$(13_a) \ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots} C_r = \lim_{\varphi_1, \varphi_2, \ldots} C_s$$

verwandelt, wenn irgend einer von den Zentralwerten einen Grenzwert hat.

Nach dem Satz 1 ist die Bedingung (12) hinreichend. Nach dem Satz 5 ist jede Bedingung von der Art:

$$\lim_{\varphi_1, \varphi_2, \ldots} g_r(x \ , \ C_r) = 0$$

notwendig. So sind also entweder alle Zentralwerte stochastische Asymptoten, oder es ist keiner von ihnen eine solche. Besitzt unsere zufällige Variable stochastische Asymptoten, so hat man: $C_r = as_B(x)$ und $C_s = as_B(x)$, woher nach dem Folgesatz E (§ 12) sich die Gleichung (13), bzw., (13a) ergibt. Da aber die Abweichungen möglicher Werte von einer stochastischen Asymptote, ihren absoluten Grössen nach, bei den Voraussetzungen unseres Satzes eine obere Grenze haben müssen, so gilt, nach dem Satz 4, auch die Gleichung (12).

Den besonderen Fall der bewiesenen Sätze bilden die Bedingungen:

$$\lim E \mid x - Ex \mid = 0 \;,\; \lim E(x - Ex)^2 = 0 \;,$$
$$\lim E \mid x - Ex \mid^3 = 0 \;,\; \ldots$$

sowohl auch:

$$\lim E \mid x - C_1 \mid = 0 \;,\; \lim E(x - C_1)^2 = 0 \;,$$
$$\lim E \mid x - C_1 \mid^3 = 0 \;,\; \ldots$$

von denen jede für sich bei obigen Voraussetzungen hinreichend und notwendig ist.

15. — Haben die absoluten Grössen der Abstände zwischen möglichen Werten einer zufälligen Variable keine feste obere Grenze, so sind die Verhältnisse zwischen stochastischen Asymptoten und Momenten viel schwerer zu entwirren. Ich fange aber mit einigen fast auf der Oberfläche liegenden Beziehungen an.

Satz 6. Die Mediane ist eine unvermeidliche Asymptote, d. h., hat eine zufällige Variable stochastische Asymptoten überhaupt, so muss auch die Mediane (C_1) eine solche sein.

Es sei $v = as_B(x)$, und es sei angenommen, dass C_1 keine stochastische Asymptote ist. Sodann wird es in jeder unbegrenzten Teilfolge: φ_k, $\varphi_{k+1}, \ldots$ eine unbegrenzte Menge von solchen Werten φ_n, $(n > k)$, geben, dass $(C_1 - v)$ nicht kleiner als eine konstante, von Null verschiedene, positive Grösse c sein wird. Dabei wird die Wahrscheinlichkeit jeder beliebigen konstanten Umgebung von v: $\overset{(\varepsilon)}{\underset{(o)}{P}} \mid x - v \mid$, die der Bedingung $\varepsilon \leq c$ genügt, nach der Grundeigenschaft der Mediane, nicht grösser als 1/2 sein können; nach der Eigenschaft der stochastischen Asymptote aber wird sie bei passend gewähltem φ_k beliebig wenig von 1 unterscheiden müssen. Also ist die obige Annahme unmöglich und der Satz b richtig.

Die Eigenschaft einer «unvermeidlichen Asymptote» hat die Mediane evidenter Weise mit dem dichtesten Werte (sog. «Mode»), vielleicht mit noch anderen Grössen gemein, doch nicht mit den übrigen Zentralwerten der ersten Klasse. Das besagt

Satz 7. Besitzt eine zufällige Variable (x) *stochastische Asymptoten, und hat dabei ein wohlzentrierter gerader Moment* $E \mid x - C_r \mid^r$ *vom Grade* $r > 1$ *keine feste obere Grenze, so ist der Zentralwert* C_r *obgleich eine mögliche doch aber keine unvermeidliche stochastische Asymptote.*

Ich führe zuerst einige Beispiele an. Es sei:

$$\text{(I)} \dots \quad \left\{ \begin{matrix} x: -n & , & 0 & , & n \\ p: \dfrac{1}{n} & , & 1 - \dfrac{2}{n} & , & \dfrac{1}{n} \end{matrix} \right\}.$$

Dann hat man: $\lim\limits_{n \to \infty} {}_B (x) = 0,\ Ex = 0,$

$\mu_2 = g_2(x, C_2) = E(x - Ex)^2 = 2n \rightarrowtail \infty$. So kann also der wohlzentrierte gerade Moment zweiten Grades unbegrenzt zunehmen, der Zentralwert zweiten Grades aber (Ex) kann dabei eine stochastische Asymptote sein.

Bei der Wahrscheinlichkeitsverteilung

$$\text{(II)} \dots \quad \left\{ \begin{matrix} x: -n & , & -1 & , & n \\ p: \dfrac{1}{n} & , & 1 - \dfrac{4}{n} & , & \dfrac{3}{n} \end{matrix} \right\}$$

ist die Sache schon anders:

$$\lim_{n \to \infty} {}_B (x) = -1; Ex = 1 + \frac{4}{n} \to 1; E(x - Ex)^2 = 4n - \frac{4}{n}\left(3 + \frac{4}{n}\right) \to \infty,$$

ebenso wie in beiden folgenden Beispielen.

$$\text{(III)} \dots \quad \left\{ \begin{matrix} x: -n^2 & , & -a & , & n^2 \\ p: \dfrac{1}{n} & , & 1 - \dfrac{3}{n} & , & \dfrac{2}{n} \end{matrix} \right\}$$

$$\lim_{n \rightarrowtail \infty} {}_B (x) = -a; E(x) = n - a\left(1 - \frac{3}{n}\right) \rightarrowtail n \rightarrowtail \infty; E(x - Ex)^2 \rightarrowtail \infty.$$

$$\text{(IV)} \dots \quad \left\{ \begin{matrix} x: -n^2 & , & -n & , & n^2 \\ p: \dfrac{1}{n} & , & 1 - \dfrac{4}{n} & , & \dfrac{3}{n} \end{matrix} \right\}$$

$$\underset{n \rightarrowtail \infty}{as_B} (x) = -n \rightarrowtail -\infty \;;\; Ex = n + 4 \rightarrowtail \infty \;;\; E(x - Ex)^2 \rightarrowtail \infty.$$

Diese Beispiele könnten leicht an einem Spielschema interpretiert werden, wo in Fällen II — IV eine Spielbank die mathematische Erwartung des Gewinns zu ihren Gunsten, und doch bei genügend grossem n fast keine Gewinnschancen hätte.

Um den Satz auch allgemein zu beweisen, nehmen wir das folgende Schema:

$$\text{(V)} \ldots \qquad \left\{ \begin{array}{lll} x: & 0, & n \\ p: & 1 - \dfrac{1}{n^k}, & \dfrac{1}{n_k} \end{array} \right\}$$

und finden den Zentralwert C_r aus der Gleichung $u_{r-1} = 0$, oder

$$(n - C_r)^{r-1} \frac{1}{n^k} - C_r^{r-1} \left(1 - \frac{1}{n^k}\right) = 0,$$

Es ergibt sich daraus

$$\text{(14)} \ldots \qquad C_r = \alpha\, n^{\frac{r-k-1}{r-1}},$$

wo α zur Abkürzung statt einer Grösse eingesetzt ist, die bei genügend grossem n der Einheit beliebig nahe kommt. Weiter findet man

$$\text{(15)} \ldots \qquad g_r(x, C_r) = \beta\, n^{r-k},$$

wo β wieder eine Grösse ist, die bei $n \to \infty$ gegen 1 konvergiert.

Ist $r > k$, so nimmt g_r mit n unbegrenzt zu. Die Grösse C_r aber konvergiert dabei entweder gegen 0, oder gegen 1, oder wächst mit n ins Unendliche, je nachdem

$$k > r - 1 \quad ; \quad k = r - 1 \quad ; \quad k < r - 1$$

ist. In beiden letzten Fällen ist C_r keine stochastische Asymptote. So ist unser Satz für alle Werte von r, bei $r > 1$, bewiesen. Kein Zentralwert erster Klasse ausser der Mediane ist also eine unvermeidliche Asymptote. Auch findet man leicht, dass, wenn in obigem Schema $\dfrac{1}{n^k}$ durch $\dfrac{1}{\log n}$ ersetzt wird, $\lim_{n \to \infty} C_r = \infty$ wird, und zwar bei jedem $r > 1$, wie klein $r - 1$ auch sein möge. Kein Zentralwert der ersten Klasse ausser C_1 ist dann also eine stochastische Asymptote.

16. — Um weiter gehen zu können, beweisen wir hier zwei Hilfsätze.

Hilfsatz A. Hat ein um einen beliebigen Bezugswert (l) *zentrierter gerader Moment* $g_r(x, l)$ *eine feste obere Grenze* (A), *ist weiter eine Grösse* $v \doteq as_B(x)$, *so müssen sowohl die Grösse* $| C_r - v |$, *wie auch der um die stochastische Asymptote zentrierte Moment* $g_r(x, v)$ *auch fest von oben begrenzt sein.*

Ist irgend ein Moment $E\,|x - l|^r < A$, so muss auch $E\,|x - C_r|^r < A$ sein. Bezeichnet man einen partiellen Moment mit dem Symbol E', so hat man à fortiori:

$$E'\,|x - C_r|^r < A.$$

Nun definiere ich die Grenzen für E' durch die Ungleichung

$$| x - v | \leqq \alpha,$$

wo α beliebig klein ist. Die Wahrscheinlichkeit dieses Intervalls ist $\overset{(\alpha)}{\underset{(0)}{P}} | x - v |$, die Intervalsgrenzen für $| x - C_r |$ aber sind

$$| v - C_r | - \alpha \leqq x - C_r \leqq | v - C_r | + \alpha.$$

Man hat daher, der Formel (24) von Kap. II zufolge:

$$(16) \quad A > E' | x - C_r |^r \geqq | (| v - C_r | - \alpha) |^r \overset{(\alpha)}{\underset{(0)}{P}} | x - v | .$$

Da $\overset{(\alpha)}{\underset{(0)}{P}} | x - v |$ bei beliebig kleinem α beliebig nahe zu 1 kommt, so sieht man aus den vorstehenden Ungleichungen, dass $| v - C_r |$ weder unendlich sein, noch unbegrenzt ins Unendliche wachsen kann.

Beachtet man weiter, dass die Grösse $E \; | x - l |^r$ eine stetige Funktion von l ist, so ergibt sich, dass einer endlichen Differenz der Bezugswerte $| v - C_r |$ auch eine endliche Differenz der Momente: $g_r (x, v)$ und $g_r (x, C_r)$ entsprechen muss. Ist, bei den Voraussetzungen unseres Satzes, $g_r (x, C_r)$ fest von oben begrenzt, so muss auch $g_r (x, v)$ eine feste obere Grenze haben.

Hilfsatz B. Hat ein gerader um eine stochastische Asymptote zentrierter Moment $g_r (x, v)$ vom Grade $r > 1$ eine feste obere Grenze und konvergiert dabei $u_{r-1} (x, v)$ gegen Null, so ist der Zentralwert C_r eine stochastische Asymptote.

Da bei jedem $r > 1$ nicht nur die Momente g_r, sondern auch ihre erste Ableitungen stetig sind, so kann man also $g_r = E \; | x - v |^r$ als Funktion von v in die Taylor'sche Reihe entwickeln. Es sei bezeichnet

$$g_r (x, v) = F (v)$$

$$\frac{d}{dv} g_r (x, v) = - r \, u_{r-1} (x, v) = F' (v)$$

$$C_r = v_0$$

Dann hat man

$$(17) \; \ldots \; F (v) = F(v_0) + (v_0 - v) \, F' \left\{ v_0 + \vartheta (v - v_0) \right\},$$

(wo $o < \vartheta < 1$ ist), woher, da F' monoton, $F (v_0)$ Minimum, $F' (v_0) = 0$ ist, sich die Ungleichungen

$$(18) \; \ldots \; F (v_0) < F (v) < F (v_0) + (v - v_0) \, F' (v)$$

ergeben. Da nach dem vorstehenden Satze $v - v_0$ endlich ist, die Grösse $F'(v)$ aber voraussetzungsgemäss gegen Null konvergiert, so hat man

$$\lim_{\varphi_1, \varphi_2, \dots} \left\{ F(v) - F(v_0) \right\} = o\ ,$$

woher sich

$$\lim_{\varphi_1, \varphi_2, \dots} (v - v_0) = o$$

ergibt. Es ist also $v_0 = C_r$ die stochastische Asymptote von x, was zu beweisen war.

Nun können wir den folgenden wichtigen Satz beweisen.

Satz 8. *Hat ein um einen beliebigen Bezugswert zentrierter gerader Moment* $g_r\ (x,\ l)$ *vom beliebigen positiven Grade* $(r > 0)$ *eine feste obere Grenze, und hat dabei die zufällige Variable stochastische Asymptoten, so müssen* (I) *alle um irgend eine Asymptote zentrierten Momente niedrigeren Grades den Grenzwert Null haben und* (II) *alle Zentralwerte niedrigeren Grades (mit Einschluss von* C_r , *wenn* $r \geqq 1$ *ist) stochastische Asymptoten sein.*

Es sei v irgend eine stochastische Asymptote. Man zerlege den Moment $g_{r-s}(x,\ v)$ $(o < s < r)$ wie folgt:

$$(19) \dots \quad E\,|\,x - v\,|^{\,r-s} = \mathop{E}_{(o)}^{(\varepsilon)} |\,x - v\,|^{\,r-s} + \\ + \mathop{E}_{\varepsilon}^{(H)} |\,x - v\,|^{\,r-s} + \mathop{E}_{H}^{\infty} |\,x - v\,|^{\,r-s}.$$

Hat $g_r\ (x,\ l)$ eine obere Grenze, so muss nach dem Hilfsatz A auch $g_r\ (x,\ v)$ fest von oben begrenzt sein. Es sei also

$$E\,|\,x - v\,|^{\,r} < A,$$

woher à fortiori

$$A > \mathop{E}_{H}^{\infty} |\,x - v\,|^{\,r} = \mathop{E}_{H}^{\infty} |\,x - v\,|^{\,s}\ |\,x - v\,|^{\,r-s} \\ \geqq H^s \mathop{E}_{H}^{\infty} |\,x - v\,|^{\,r-s}$$

und endlich

$$(20) \dots \quad \mathop{E}_{H}^{\infty} |\,x - v\,|^{\,r-s} < \frac{A}{H^s}\ .$$

Weiter hat man der Formel 25 (Kap. II) zufolge:

$$(21) \dots \quad \mathop{E}_{(o)}^{(\varepsilon)} |\,x - v\,|^{\,r-s} \leqq \varepsilon^{r-s}$$

und

$$(22)\ldots \qquad \overset{(H)}{\underset{\varepsilon}{E}} \mid x - v \mid^{r-s} \leqq H^{r-s} \overset{\infty}{\underset{\varepsilon}{P}} \mid x - v \mid .$$

Es sei λ eine gegebene beliebig kleine Grösse. Dann nehme man (a) H so gross, dass $A/H^s < \frac{1}{3}\lambda$ (was bei jedem $s > o$ möglich ist), (b) ε so klein, dass $\varepsilon^{r-s} < \frac{1}{3}\lambda$ (was seinerseits bei jedem $r > o$ und $s < r$ möglich ist) und (c) wähle die unabhängige Variable φ so weit in der Folge: $\varphi_1, \varphi_2, \ldots$, dass

$$H^{r-s} \overset{\infty}{\underset{\varepsilon}{P}} \mid x - v \mid < \frac{1}{3}\lambda$$

wird. Dann ergibt sich durch Summieren von Ungleichungen (20) (21) und (22), dass

$$(23)\ldots \qquad E \mid x - v \mid^{r-s} < \lambda$$

ist. So ist also, wenn $r - s = k$ gesetzt wird,

$$(24)\ldots \qquad \lim_{\varphi_1,\, \varphi_2 \ldots} E \mid x - v \mid^k = 0 \;,\; (0 < k < r)$$

woher, nach dem Satz $1a$,

$$(25)\ldots \qquad C_k = as_B(x) \;,\; (0 < k < r) \;.$$

Da aber stets $u_k \leqq g_k$ ist, so hat man auch

$$(26)\ldots \qquad \lim_{\varphi_1,\, \varphi_2, \ldots} u_k(x \,,\, v) = 0 \quad , \quad (0 < k < r) \;,$$

woher bei $r > 1$ die Gleichung

$$\lim_{\varphi_1,\, \varphi_2, \ldots} u_{r-1}(x \,,\, v) = 0$$

folgt, die uns nach dem Hilfsatz B zu schliessen gibt, dass auch

$$(27)\ldots \qquad C_r = as_B\,(x)$$

ist.

Ehe wir weiter gehen, scheint es mir von einigem Interesse für denselben Satz noch einen anderen Beweis zu liefern. Schreibt man die wohlbekannte Liapunoff'sche Ungleichung (1) in der Form

$$(28)\ldots \qquad g_k^{\;r-h} < g_h^{\;r-k}\, g_r^{\;k-h} \;,\; (r > k > h \geqq 0)$$

---

(1) A. Liapounoff, *Nouvelle forme du théorème sur la limite de probabilité*, « Mém. de l'Académie Imp. de science de St. Petersbourg », VIII Serie. Classe Phys.-Math., 1902. Vol. XII, N. 5, p. 2-3.

und nimmt man in Betracht, dass sie nicht lediglich für die vollen, sondern auch für die partiellen mathematischen Erwartungen gilt, so hat man, bei $h = o$, für partielle Momente im beliebigen Intervall I:

$$(29)\ldots \quad (E_I \mid x - v \mid^k)^r < P_I^{\,r-k} (E_I \mid x - v \mid^r)^k \,, \quad (r > k > 0)$$

woher die Ungleichung

$$E_I \mid x - v \mid^k < P_I^{\frac{r-k}{r}} (E_I \mid x - v \mid^r)^{\frac{k}{r}} \,.$$

und à fortiori die Ungleichung

$$(30)\ldots \qquad E_I \mid x - v \mid^k < P_I^{\frac{r-k}{r}} (E \mid x - v \mid^r)^{\frac{k}{r}}$$

hervorgeht.

Ist nun $v = as_B\,(x)$, so hat man (Formel 25 Kap. II)

$$(31)\ldots \qquad \overset{(\varepsilon)}{\underset{(o)}{E}} \mid x - v \mid^k \leqq \varepsilon^k$$

und nach der Formel (30)

$$(32)\ldots \qquad \overset{\infty}{\underset{\varepsilon}{E}} \mid x - v \mid^k < \left(\overset{\infty}{\underset{\varepsilon}{P}} \mid x - v \mid\right)^{\frac{r-k}{r}} (E \mid x - v \mid^r)^{\frac{k}{r}} \,.$$

Addiert man (31) und (32) und schreibt man $g_r = g_r\,(x, v)$, bzw., $g_k = g_k\,(x, v)$, so hat man

$$(33)\ldots \qquad g_k < \varepsilon^k + \left(\overset{\infty}{\underset{\varepsilon}{P}} \mid x - v \mid\right)^{\frac{r-k}{r}} g_r^{\frac{k}{r}} \,.$$

Da ε beliebig klein gewählt werden kann und

$$\lim_{\varphi_1,\, \varphi_2, \ldots} \overset{\infty}{\underset{\varepsilon}{P}} \mid x - v \mid = 0$$

ist, so muss (bei $r > k > o$) auch

$$\lim_{\varphi_1,\, \varphi_2, \ldots} g_k = 0$$

sein, wenn nur g_r eine feste obere Grenze hat. Weiteres ganz wie oben.

Aus dem bewiesenen Satze können noch einige wichtige Folgerungen abgeleitet werden.

Satz 9. Hat eine zufällige Variable stochastische Asymptoten, konvergiert aber ein gerader Moment vom Grade r, der um einen beliebigen Zentralwert vom Grade $k < r$ $(o < r < 1)$, *bzw., vom Grade* $k \leqq r$

$(r \geqq 1)$ *zentriert ist, nicht gegen 0, so können alle geraden Momente höheren Grades $g_{r+\alpha}$, $(\alpha > 0)$, keine obere Grenze haben* (1).

Denn hätte ein gerader Moment $g_{r+\alpha}$ eine obere Grenze, so müsste der auf eine Asymptote (v) bezogene Moment $E \mid x - v \mid^r$, nach dem Satz 8, den Grenzwert 0 haben. Da aber dabei alle C_k $(k < r$, bzw., $k \leqq r)$ stochastische Asymptoten sein müssten, so müsste auch jeder gerade Moment $g_r\,(x, C_k)$ gegen 0 konvergieren, was den Voraussetzungen des Satzes widerspricht.

Satz 10. Konvergiert ein gerader Moment $g_r\,(x, C_k)$, $(o < k < r$, bzw, $o < k \leqq r$, wenn $r > 1$ ist) nicht gegen 0, und hat ein anderer gerader Moment von einem höheren Grade $g_s\,(x, C_k)$, $(s > r)$, eine feste obere Grenze, so besitzt die zufällige Variable keine stochastischen Asymptoten.

Der Satz ist eine unmittelbare Folge des vorhergehenden und so bedarf er keines besonderen Beweises. Er ist insofern von Wichtigkeit, dass er eine zwar nicht hinreichende, doch eine notwendige Bedingung des Bestandes der stochastischen Asymptoten ist. Man kann ihm aber für einen bestimmten Fall auch eine sowohl notwendige wie hinreichende Bedingung ablesen.

Satz 10a. Konvergiert ein gerader Moment $g_r\,(x, C_k)$, $(o < k \underset{(=)}{<} r)$, nicht gegen 0 und hat er dabei eine feste obere Grenze, so ist es eine für den Bestand stochastischer Asymptoten sowohl hinreichende, wie notwendige Bedingung:

$$\lim_{\varphi_1, \varphi_2, \ldots} g_s\,(x, C_i) = 0\,, \quad \left\{ \begin{array}{r} 0 < s < r;\ 0 < i < r;\ \textit{bei}\ r < 1 \\ 0 < i \leqq r;\ \textit{bei}\ r \geqq 1 \end{array} \right\}.$$

Nach allem vorhergehenden ist der Satz evident. Gleich werden wir sehen, dass er sich auf ein bisher noch unaufgeklärtes Problem anwenden lässt.

17. — Es ist wohlbekannt, dass der arithmetische Durchschnitt $(x_{(n)})$ einer unbegrenzt wachsender Anzahl stochastisch von einander unabhängiger zufälliger Variablen $(x_1, x_2, \ldots x_n, \ldots)$ dem Gesetz der grossen Zahlen mindestens in dem Falle untersteht, wenn alle Streuungen $(\overset{(1)}{\mu_2}, \overset{(2)}{\mu_2}, \ldots \overset{(n)}{\mu_2}, \ldots)$ eine feste obere Grenze haben. Aus der bekannten Formel für die Streuung des Durchschnitts (siehe oben Form. 8 Kap. I)

$$\mu_{2\,(n)} = \frac{1}{n} \left\{ \frac{1}{n} \sum_{i=1}^{n} \mu_2^{(i)} \right\}$$

(1) Die Unbestimmtheit in Betreff von C_r bei $r < 1$ hat ihren Grund in der Unvollständigkeit obiger Untersuchungen in Betreff der Zentralwerte zweiter Klasse.

ist noch ein weiter gehender Schluss zu ziehen, und zwar, dass *lim* $\mu_{2\,(n)} = 0$ auch in *den* Fällen ist, wo

$$\frac{1}{n} \sum_{i=1}^{n} \mu_2^{(i)}$$

zwar unbegrenzt zunehmend, doch von einer niedrigeren Unendlichkeitsordnung als n ist. Obwohl diese Bedingung lediglich als eine hinreichende eingeführt wird, doch ist die bisher noch bestehende Lücke in den meisten Fragen der theoretischen Statistik keine sehr empfindliche, da die Unbestimmtheit, inbetreff der Anwendbarkeit des Gesetzes der grossen Zahlen lediglich auf die überaus exzeptionellen Fälle der ins Unendliche wachsenden mittleren Fehler begrenzt wird.

Sind aber zufällige Variablen nicht unabhängig, so nimmt die Sache insofern ein ganz anderes Aussehen an, dass die Unbestimmtheit sich nun auf alle Fälle erstreckt, wo zufällige Variablen beliebig grosse Werte annehmen können und im Durchschnitt positiv korreliert sind. Konvergiert nämlich die durchschnittliche Grösse von Momenten $\mu_{11}^{(i,j)} = E\,(x^{(i)} - Ex^{(i)})\,(x^{(j)} - Ex^{(j)})$, d. h. die Grösse

$$\mu_{[1,1;\,n]} = \frac{1}{n\,(n-1)} \sum_{i \neq j} \sum \mu_{1,1}^{(i,j)}$$

bei unbegrenzt zunehmender Versuchszahl (n) nicht gegen Null, so geht sie bekanntlich (siehe gleich unten) in den Ausdruck für die Streuung des Durchschnittswertes als ein von n unabhängiges Glied ein, und so bleibt $\mu_{2(n)}$, auch bei einer beliebig grossen Versuchszahl von Null verschieden. Somit versagt aber das einzige leicht anwendbare Kriterium und, da es lediglich ein hinreichendes ist, so bleiben wir vom Standpunkt eines strengen mathematischen Raisonement vollständig im Dunkel. Einem reinen Mathematiker kann die Frage damit bisweilen für erledigt gelten (1), einem Theoretiker mit weiterem Sehfelde soll aber die Sache doch in einem etwas anderen Lichte erscheinen. Man kann sich schwerlich den intuitiven Einblick in stochastische Zusammenhänge verbieten, der zur Einsicht führt, dass in solchen Fällen das Gesetz der grossen Zahlen keine Anwendung findet (2). Dass sich eine solche Einsicht vielleicht nur auf etwas enger begrenzte Sachverhalte bezieht, sei in vollem Masse eingestanden,

---

(1) A. A. Markoff, *Wahrscheinlichkeitsrechnung*, 4 Aufl., p. 116-134.
(2) Al. A. Tschuprow, *Zur Theorie der Stabilität statistischer Reihen*, loc. cit., p. 208.

dass aber, z. B., bei einem uniformen korrelativen Zusammenhange (1) das Gesetz der grossen Zahlen nicht vielleicht eine seltene auf künstliche Voraussetzungen begrenzte Ausnahme, sondern eine Regel bilden kann, wird kein theoretischer Statistiker glauben. Doch bleibt die Frage insofern prinzipiell wichtig, dass es keine normale Sache ist strenge Beweisführungen durch mehr oder minder unklare Intuitionen zu ersetzen. Die obenaufgestellten Sätze geben uns nun die Möglichkeit den ersten Schritt zur endgültigen Klärung der Frage vorzunehmen.

Es seien: $x_1, x_2, \dots x_n, \dots$ zufällige, in beliebiger Weise stochastisch miteinander verbundene Variablen, von denen jede ihre eigene Wahrscheinlichkeitsverteilung hat, und es sei angenommen, dass an ihnen der Reihe nach n Versuche ausgeführt werden. Es sei (2).

$$\overset{(i)}{\mu_h} = E\,(x_i - E\,x_i)^h \; ; \; \mu_{[h\,;\,n]} = \frac{1}{n} \sum_{i=1}^{n} \overset{(i)}{\mu_h} \; ;$$

$$\overset{(i_1,\, i_2 \dots i_j)}{\mu_{h_1,\, h_2, \dots h_j}} = E\,(x_{i_1} - E\,x_{i_1})^{h_1} (x_{i_2} - E\,x_{i_2})^{h_2} \dots (x_{i_j} - E\,x_{i_j})^{h_j} \; ;$$

$$\mu_{[h_1,\, h_2, \dots h_j\,;\,n]} = \frac{1}{n\,(n-1) \dots (n-j+1)} S_j\, \overset{(i_1,\, i_2, \dots i_j)}{\mu_{h_1,\, h_2, \dots h_j}} \,,$$

wo S_j eine j-fache Summe bezeichnet, die auf alle ungleiche Werte von $i_1, i_2, \dots i_j$ erstreckt wird. Es sei ferner

$$x_{(n)} = \frac{1}{n} \sum_{i=1}^{n} x'_i$$

und

$$\mu_{h(n)} = E\,(x_{(n)} - E\,x_{(n)})^h \,.$$

Da hat man für den zweiten und den vierten auf mathematische Erwartung bezogenen Momente des Durchschnittswertes $x_{(n)}$ folgende Ausdrücke (3):

$$(34) \; \dots \quad \mu_{2(n)} = \frac{1}{n} \mu_{[2;\,n]} + \frac{n-1}{n} \mu_{[1,\,1;\,n]}$$

und

$$(35) \quad \mu_{4(n)} = \frac{1}{n^3} \mu_{[4;\,n]} + \frac{3\,(n-1)}{n^3} \mu_{[2,\,2;\,n]} + \frac{4\,(n-1)}{n^3} \mu_{[3,\,1;\,n]} + \\ + \frac{6(n-1)(n-2)}{n^3} \mu_{[2,\,1,\,1;\,n]} + \frac{(n-1)\,(n-2)\,(n-3)}{n^3} \mu_{[1,\,1,\,1,\,1;\,n]} \,.$$

(1) Al. A. Tschuprow, *On the mathematical expectation of the moments of frequency distributions in the case of correlated observations*, « Metron », Vol. II, N. 3 4, 1923, p. 469.

(2) Al. A. Tschuprow, *op. cit.*, p. 472.

(3) *Ibid.*, p. 478.

Ist hier $\mu_{[1,1;n]}$ positiv, so kann man sagen, dass unsere n zufälligen Variablen « im Durchschnitt » positiv korreliert sind. Behält nun dieser « durchschnittliche Produktmoment » $\mu_{[1,1;n]}$ auch bei unbegrenzt wachsendem n einen immer positiven und nicht unbegrenzt abnehmenden Wert bei, so kann, wie aus der Formel (34) ersichtlich ist, auch die Streuung des Durchschnitts ($\mu_{2(n)}$) nicht gegen 0 konvergieren. Zwei Fälle sind dabei zu unterscheiden. Haben die Abweichungen möglicher Werte von einander, daher auch von entsprechenden mathematischen Erwartungen, eine feste obere Grenze, so sind auch die Abweichungen $| x_{(n)} - E x_{(n)} |$ fest von oben begrenzt. Ist, wie angenommen, die Bedingung $\lim\limits_{n \to \infty} \mu_{[1,1;n]} = 0$ und somit auch die Bedingung $\lim\limits_{n \to \infty} \mu_{2(n)} = 0$ nicht erfüllt, so kann die Grösse $x_{(n)}$, nach dem Satz 5a, keine stochastischen Asymptoten haben. Das Gesetz der grossen Zahlen findet auf solche Fälle keine Anwendung (1).

Uns interessiert hier aber der andere Fall, wo die möglichen Werte zufälliger Variablen entweder durch keine endlichen Intervalle begrenzt werden, oder wo die sie enthaltenden Intervalle unbegrenzt zunehmen. Da die Bedingung $\lim \mu_{2(n)} = 0$ bei dieser Voraussetzung keine notwendige ist, so kann deren Nichterfüllung über die Geltung des Gesetzes der grossen Zahlen keinen Aufschluss geben.

Nehmen wir an, dass der durchschnittliche vierte Moment $\mu_{[4;n]}$ eine feste obere Grenze hat. Aus den Ungleichungen (18), bzw., (19) (Kap. II) schliesst man leicht, dass wenn alle $\mu_4^{(i)}$ endlich sind und nicht unbegrenzt wachsen können, alle Produktmomente vierten Grades: $\overset{(i,j)}{\mu_{2,2}}$, $\overset{(i,j)}{\mu_{3,1}}$, $\overset{(i,j,k)}{\mu_{2,1,1}}$, und $\overset{(i,j,k,l)}{\mu_{1,1,1,1}}$ auch eine feste obere Grenze haben müssen. Eine weitere Ueberlegung, die ich dem Leser überlassen darf, zeigt sodann, dass wir uns mit einer engeren Voraussetzung begnügen können. Ist nämlich der durchschnittliche vierte Moment $\mu_{[4;n]}$, wie oben angenommen, fest von oben begrenzt, so können auch die durchschnittlichen Produktmomente vierten Grades: $\mu_{[2,2;n]}$ u. s. w. nicht unbegrenzt zunehmen. Der vierte Moment des arithmetischen Durchschnitts $\mu_{4(n)}$ muss also auch eine feste obere Grenze haben.

Erinnert man sich nun des Satzes 10, so findet man, dass bei obigen Voraussetzungen das Gesetz der grossen Zahlen auf den arithmetischen Durchschnitt zufälliger im Durchschnitt positiv korrelierter Grössen keine Anwendung findet. In der Tat, hätte $x_{(n)}$ sto-

---

(1) A. A. Markoff, *Wahrscheinlichkeitsrechnung*, 4 Aufl., p. 119-120.

chastische Asymptoten, so müsste, bei der Nichterfüllung der Bedingung $lim\ \mu_{2(n)} = 0$, der vierte Moment $\mu_{4(n)}$ unbegrenzt wachsen, bzw., bei fest von oben begrenztem vierten Moment, müsste der zweite $\mu_{2(n)}$ gegen 0 konvergieren. Ich formuliere dieses Resultat als

Satz 11. Divergiert der durchschnittliche Produktmoment zweiten Grades $\mu_{[1,1;n]}$ *einer unbegrenzt wachsenden Anzahl zufälliger Variablen nicht gegen Null, und besitzt der durchschnittliche vierte Moment* $\mu_{[4;n]}$ *eine feste obere Grenze, so ist der arithmetische Durchschnitt der zufälligen Variablen dem Gesetz der grossen Zahlen nicht unterworfen* (1).

Für die meisten Fälle, mit denen man in der theoretischen Statistik zu tun hat, ist das Problem durch diesen Satz erledigt. Von der prinzipiellen Seite aber bleibt noch Manches im Dunkel. So erhebt sich, z. B., die Frage, ob vielleicht schon die Nichterfüllung der Bedingung $lim\ \mu_{[1,1;n]} = 0$ allein die Anwendbarkeit des Gesetzes der grossen Zahlen auf den arithmetischen Durchschnitt zufälliger Grossen ausschliesse. Für den Fall, wo $\mu_{[2;n]}$ fest von oben begrenzt ist, gibt unser Satz 10a die Anweisung, dass der Moment

$$g_{2-\alpha} = E \mid x_{(n)} - E\, x_{(n)} \mid^{2-\alpha}, (0 < \alpha < 2)$$

untersucht werden muss. Es ist nämlich die Frage zu entscheiden, ob

(I) bei $\mu_{[2;n]} < A$

und

(II) bei der Nichterfüllung der Bedingung

$$\lim_{n \to \infty} \mu_{[1,1;n]} = 0\ ,$$

$g_{2-\alpha}$ gegen Null konvergieren könne, womit das Problem auf ein rein analytisches zurückgeführt wird. Ueber den zweiten Fall, wo $\mu_{[2;n]}$ keine feste obere Grenze hat, ist zur Zeit, m. W., gar nichts zu sagen.

In diesem nicht recht befriedigenden Zustand muss ich die Sache bisweilen auf sich beruhen lassen.

Viertes Kapitel.

Ueber stochastische Asymptoten und Grenzwerte der zufällig variablen Vektoren.

18. — Durch die Analogie mit den gewöhnlichen Grenzwerten wird uns die Frage nahegelegt, ob nicht auch für stochastische Grenz-

(1) Dass die zufälligen Variablen, *positiv* korreliert sein müssen, ist in der Formulierung des Satzes 11 nicht erwähnt, da es sich bei gemachten Voraus-

werte, bzw., Asymptoten, ähnliche Sätze über Summen, Produkte, überhaupt Funktionen zweier oder mehrerer stochastischen Grenzwerte bestehen, wie für die ersteren (1).

Fangen wir mit der formellen Erweiterung der Begriffe der stochastischen Asymptote und des stochastischen Grenzwerts an. Sind $x_1, x_2, \ldots x_n, \ldots$ zufällige Variablen und ist $\overset{(k_i)}{x_i}$ irgend ein von den möglichen Werten von x_i, so bildet der Inbegriff der Werte $(\overset{(k_1)}{x_1}, \overset{(k_2)}{x_2}, \ldots \overset{(k_n)}{x_n})$ einen Punkt in dem n-dimensionellen Kontinuum und es sei mit $\overline{X}^{(k)}$ oder

$$\left\{ \overset{(k_1)}{x_1}, \overset{(k_2)}{x_2}, \ldots \overset{(k_n)}{x_n} \right\}$$

der Vektor bezeichnet, der ihn mit dem Nullpunkte des Koordinatensystems verbindet und dessen Komponenten: $\overset{(k_1)}{x_1}, \overset{(k_2)}{x_2}, \ldots \overset{(k_n)}{x_n}$ sind. So kommt man zur Idee des zufällig variablen Vektors

$$\left\{ x_1, x_2, \ldots x_n \right\}$$

oder in kürzerer Bezeichnung $\overline{X}$, der verschiedene n-dimensionelle Werte $\left\{ \overset{(k_1)}{x_1}, \overset{(k_2)}{x_2}, \ldots \overset{(k_n)}{x_n} \right\}$ annehmen kann (2). Durch die Bedingungen:

$$\begin{array}{l} a_1 < x_1 \leqq b_1 \\ a_2 < x_2 \leqq b_2 \\ \cdots\cdots\cdots \\ a_n < x_n \leqq b_n \end{array}$$

oder in einer kürzeren Schreibweise

$$(1) \qquad a_i < x_i \leqq b_i \quad , \quad (i = 1, 2, \ldots n)$$

wird ein n-dimensioneller Intervall bezeichnet, dessen Wahrscheinlichkeit sei

$$(2) \ldots \qquad \overset{(b_i)}{\underset{a_i}{P}} (x_i)_{i = 1, 2, \ldots n}$$

---

setzungen von selbst ergibt. Ist in der Tat $\mu[4; n]$ fest von oben begrenzt, so hat auch $\mu[2; n]$, der bekannter Ungleichung $\mu_2^2 < \mu_4$ zufolge, eine feste obere Grenze. Sodann konvergiert $\frac{1}{n}\mu[2; n]$ gegen Null, und da $\mu_2(n)$ eine wesentlich positive Grösse ist, so muss $\mu[1,1; n]$, wenn es negativ ist, gegen Null konvergieren. Siehe A. A. Markoff, *Die Wahrscheinlichkeitsrechnung*, 4 Aufl., p. 125.

(1) F. P. Cantelli, *Sulla legge dei grandi numeri*, loc. cit., p. 349.

(2) R. v. Mises, *Fundamentalsätze der W.-R.*, loc. cit, p. 54-55. 70 ff. Derserlbe, *Grundlagen der W. R.*, loc. cit., p. 66 ff.

Der Inbegriff aller Wahrscheinlichkeiten für alle Intervalle des Kontinuums ist die *Wahrscheinlichkeitsverteilung.* Besteht eine bestimmte Wahrscheinlichkeitsverteilung für einen Vektor $\bar{X}$ so ist er damit als ein zufällig variabler Vektor oder kürzer als ein *zufälliger Vektor* definiert.

Betrachten wir einen zufälligen Vektor $\bar{X}$, der mit unabhängigen Variablen: $\varphi_1, \varphi_2, \ldots \varphi_n$ stochastisch verbunden ist (siehe oben § 1). Es sei $\bar{V}$ ein Vektor, dessen Komponenten Funktionen von $\varphi_1, \varphi_2, \ldots \varphi_n$ sind:

$$v_1 = f_1(\varphi_1) \ , \ v_2 = f_2(\varphi_2) \ , \ldots \ v_n = f_n(\varphi_n) \ .$$

Es seien weiter $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_n$ irgend welche gegebene beliebig kleine Zahlen und

$$(3)\ldots \qquad \overset{(\varepsilon_i)}{\underset{(o)}{P}} \, | \, x_i - v_i \, |_{i=1,2,\ldots n}$$

die Wahrscheinlichkeit des Zusammenbestehens der Bedingungen:

$$(4)\ldots \qquad | \, x_i - v_i \, | \leqq \varepsilon_i \quad , \quad (i = 1, 2, \ldots n) \ .$$

Wird die Bedingung

$$(5)\ldots \qquad \lim_{\varphi'_i, \, \varphi''_i, \ldots} \overset{(\varepsilon_i)}{\underset{(o)}{P}} \, | \, x_i - v_i \, |_{i=1,2,\ldots n} = 1$$

erfüllt, so sage man, dass der Vektor $\bar{V}$ eine stochastische Asymptote des zufälligen Vektors $\bar{X}$ ist, und schreibe

$$(6)\ldots \qquad \bar{V} = as_B(\bar{X}) \ ,$$

bzw.

$$(7)\ldots \qquad \{ v_1 \ , \ v_2 \ , \ldots v_n \} = as_B \ \{ x_1 \ , \ x_2 \ , \ldots x_n \} \ .$$

Ist $\bar{V}$ konstant, so ist es der stochastische Grenzwert von $\bar{X}$.

Weiter kann der Begriff der stochastischen Asymptote, bzw., des stochastischen Grenzwerts noch auf andere zusammengesetzte Gebilde erweitert werden, so, z. B., auf Quaternionen, Tensoren, zufällig variable Linien, Flächen u. s. f.

19. — Weiteren Darlegungen sei ein Hilfsatz vorangestellt. Es seien $\alpha_1, \alpha_2, \ldots \alpha_n$ Ereignisse, die alle zusammenbestehen müssen, von denen aber jedes in zwei unvereinbaren Modalitäten i, bzw. e eintreten kann. Man bezeichne mit $P\ (i_1,\ e_2,\ i_3 \ldots\ e_n)$ die Wahrscheinlichkeit einer bestimmten Kombination der Ereignisse und mit $\sum P\,(e_i\ ,\ \alpha_1,\ \alpha_2, \ldots\ \dot{\alpha}_{i-1}\ ,\ \alpha_{i+1}\ , \ldots\ \alpha_n)$ oder kürzer mit $\sum P\ (e_i\ ,\ \alpha)$ die Summe der Wahrscheinlichkeiten aller Kombinationen, die durch

Einsetzen von i und e auf alle mit α besetzten Stellen zustande kommen. Wird eine Reihe von solchen Summen gebildet, so bedeutet der Strich bei der Summe, dass aus ihr alle in vorangehenden Summen vorkommenden Summanden weggelassen sind. Man bemerkt leicht, dass

$$\Sigma P(e_i, \alpha) = P(e_i)$$

und dass

$$(8)\ldots\ P(i_1, i_2, \ldots i_n) = 1 - \left\{ \Sigma P(e_1, \alpha) + \Sigma' P(e_2, \alpha) + \ldots + \right. \\ \left. + \Sigma' P(e_n, \alpha) \right\}$$

ist, woher sich nach Massgabe der evidenten Beziehung

$$(9)\ldots \qquad \Sigma' P(e_i, \alpha) \leqq \Sigma P(e_i, \alpha) = P(e_i)$$

die *Ungleichung von* G. BOOLE

$$(10)\ldots \quad P(i_1, i_2, \ldots i_n) \geqq 1 - \left\{ P(e_1) + P(e_2) + \ldots + P(e_n) \right\}$$

ergibt (1).

Jetzt kann der *1 Satz über die geometrische Addition stochastischer Asymptoten* bewiesen werden.

Satz 12. Es seien $x_1, x_2, \ldots x_n$ *eine endliche Anzahl zufälliger Variablen, von denen jede eine stochastische Asymptote* $v_i = as_B(x_i)$, $(\varphi_i', \varphi_i'', \ldots)$, *hat. Dann ist auch der Vektor* $\bar{V} = \{v_1, v_2, \ldots v_n\}$ *die stochastische Asymptote des zufälligen Vektors* $\bar{X} = \{x_1, x_2, \ldots x_n\}$.

Nach dem BOOLE'schen Satz (Formel 10) hat man

$$(11)\ldots \quad \overset{(\varepsilon_i)}{\underset{(o)}{P}} \left| x_i - v_i \right|_{i=1,2,\ldots n} \geqq 1 - \sum_{i=1}^{n} \overset{\infty}{\underset{\varepsilon_i}{P}} \left| x_i - v_i \right| .$$

Ist $v_i = as_B(x_i)$, so hat man bei entsprechend gewählten Werten von $\varphi_1, \varphi_2, \ldots \varphi_n$

$$\overset{\infty}{\underset{\varepsilon_i}{P}} \left| x_i - v_i \right| < \frac{1}{n} \lambda \quad , \quad (i = 1, 2, \ldots n) ,$$

wo λ eine gegebene beliebig kleine Grösse ist, und so ergibt sich aus (11) die Ungleichung

$$\overset{(\varepsilon_i)}{\underset{(o)}{P}} \left| x_i - v_i \right|_{i=1,2,\ldots n} > 1 - \lambda ,$$

(1) G. BOOLE, *An Investigation of the Laws of Thought*, London 1854, p. 307 (Cit. nach. F. P. CANTELLI, *Sulla probabilità, come limite della frequenza*, loc. cit. p. 40-41). Da das BOOLE'sche Werk vieilleicht Manchem, wie dem Schreiber dieses unzugänglich ist, so glaubte ich den allgemeinen Beweis seines wichtigen Satzes hier geben zu dürfen.

bzw.,

$$(12) \ldots \quad \lim_{\varphi'_i, \varphi''_i, \ldots} \overset{(\varepsilon_i)}{\underset{(o)}{P}} \left| x_i - v_i \right| = 1 \quad , \quad i = 1, 2, \ldots n$$

was zu beweisen war.

Auf das Gebiet des Gesetzes der grossen Zahlen übertragen besagt der Satz, dass der Komplex einer endlichen Anzahl zufälliger Variablen (bzw., der aus ihnen gebildete Vektor) dem Gesetz der grossen Zahlen unterstehen muss, sobald es für jede einzelne zufällige Variable gilt. Die stochastische Unabhängigkeit von $x_1, x_2, \ldots x_n$ wird dabei nicht vorausgesetzt (1).

Satz 13. Hat ein n-dimensioneller zufälliger Vektor $\overline{X}$ eine stochastische Asymptote $\overline{V}$, so ist jede einer Koordinatenaxe parallele $\overline{V}$-Komponente die stochastische Asymptote der entsprechenden $\overline{X}$-Komponente.

Es ist nach Massgabe der evidenten Beziehung

$$(13) \qquad \overset{(\varepsilon_i)}{\underset{(o)}{P}} \left| x_i - v_i \right|_{i = 1, 2, \ldots k, \ldots n} \leq \overset{(\varepsilon_k)}{\underset{(o)}{P}} \left| x_k - v_k \right|$$

ersichtlich, dass wenn die Grösse rechts nicht den Grenzwert 1 hat, so kann auch die Grösse links diesen Grenzwert nicht haben. Ist also v_k keine stochastische Asymptote von x_k, so ist auch $\overline{V}$ keine stochastische Asymptote von $\overline{x}$. Die Bedingung

$$(14) \ldots \quad v_i = as_B (x_i) \quad , \quad (i = 1, 2, \ldots n)$$

---

(1) Es sind zwei Gedanken auseinanderzuhalten: *erstens* die Idee der Verallgemeinerung des stochastischen Grenzwertsbegriffs (bzw., der Idee des Gesetzes der grossen Zahlen) auf zusammengesetzte Gebilde (z. B. Vektoren) und auf Fälle, wo Wahrscheinlichkeitsverteilung Funktion mehrerer unabhängigen Variablen ist, und, *zweitens*, die Sätze über die Addition der stochastischen Asymptoten und Grenzwerte. Es sei in diesem Zusammenhange erlaubt noch einige Bemerkungen betreffend mancherlei Punkte in F. P. Cantelli's Forschungen machen zu dürfen. Unterscheidet man diese zwei Seiten des Problems, so soll man den Ideen in « Su due applicazioni » etc. (loc. cit., p. 295-300) die Entwickelung nach zwei Richtungen geben. Die Verallgemeinerung der stochastischen Grenzwertsidee, die von F. P. Cantelli gefordert wird, soll in eine organische Verbindung mit dem ganzen System der stochastischen Begriffsbildungen gebracht werden, welcher Versuch oben auch gemacht wurde (siehe §§ 5, 18). Der F. P. Cantelli'sche Satz aber (*op. cit.*, p. 297-8) soll eine weiter gehende Umgestaltung, wie es oben in Text dieses §'s geschah, erfahren. Erstens, ist der Satz nicht lediglich auf die Durchschnittswerte zu beziehen, sondern auf zufällige Variablen überhaupt, und, zweitens, hat man ihn von den seine Geltung einschränkenden Bedingungen zu befreien, mit denen er bei F. P. Cantelli verbunden ist.

erweist sich also als eine nicht lediglich hinreichende sondern auch als eine notwendige für den Bestand der Beziehung

$$(15) \ldots\ldots \quad \{v_1, v_2, \ldots v_n\} = as_B \{x_1, x_2, \ldots x_n\}$$

Aus den bewiesenen Sätzen ergibt sich nun noch der *Satz über die algebraische Addition stochastischer Asymptoten.*

Satz 14. Hat jede zufällige Variable $x_1, x_2, \ldots x_n$ eine stochastische Asymptote $v_i = as_B\,(x_i)$, $(i = 1, 2, \ldots n)$, so ist die (algebraische) Summe dieser letzteren die stochastische Asymptote der Summe der zufälligen Variablen, oder symbolisch:

$$(16) \ldots \sum_{i=1}^{n} as_B\,(x_i) = as_B \sum_{i=1}^{n} x_i\,, \quad (\varphi'_i\,, \varphi''_i\,, \ldots\,;\; i = 1, 2, \ldots n).$$

Da aus dem Zusammenbestehen der Bedingungen

$$(17) \quad |x_1 - v_1| \leqq \varepsilon_1\,, \; |x_2 - v_2| \leqq \varepsilon_2\,, \ldots |x_n - v_n| \leqq \varepsilon_n$$

die Ungleichung

$$(18) \ldots\ldots \quad \left|\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} v_i\right| \leqq \sum_{i=1}^{n} \varepsilon_i$$

folgt, nicht aber umgekehrt, so hat man

$$(19) \ldots \overset{(\lambda)}{\underset{(o)}{P}} \left|\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} v_i\right| \geqq \overset{(\varepsilon_i)}{\underset{(o)}{P}} |x_i - v_i|_{\,i=1,2,\ldots n},$$

wo $\lambda = \sum_{i=1}^{n} \varepsilon_i$ ist. Nimmt man λ beliebig klein und setzt man $\varepsilon_i = \frac{1}{n}\lambda$ ein so ergibt es sich bei jedem endlichen n, dass die Grösse rechts in (19) nach dem Satz 12 den Grenzwert 1 haben muss und somit à fortiori auch die Grösse links.

Es folgt weiter der II Satz über die geometrische Addition stochastischer Asymptoten.

Satz 15. Sind $\bar{X}_1, \bar{X}_2, \ldots \bar{X}_s$ eine endliche Anzahl zufällig variabler, in ein und demselben n-dimensionellen Kontinuum definierter Vektoren und $\bar{V}_1, \bar{V}_2, \ldots \bar{V}_s$ deren stochastische Asymptoten, so ist die geometrische Summe $\bar{V}_1 + \bar{V}_2 + \ldots + \bar{V}_s$ die stochastische Asymptote der geometrischen Summe $\bar{X}_1 + \bar{X}_2 + \ldots \bar{X}_s$.

Sind $v_{i,k}$ und $x_{i,k}$ die der Koordinatenaxe mit dem Index k parallelen Komponenten von $\bar{V}_i$, bzw., $\bar{X}_i$, so hat man nach Satz 13

$$(20) \ldots\ldots \quad v_{i,k} = as_B\,(x_{i,k})\,,$$

woher nach Satz 14

$$(21) \ldots\ldots \quad \sum_{i=1}^{s} v_{i,k} = as_B \sum_{i=1}^{s} x_{i,k}$$

und nach Satz 12

$$(22)\ldots\ldots \quad \left\{ \sum_{i=1}^{s} v_{i,1}, \sum_{i=1}^{s} v_{i,2}, \ldots \sum_{i=1}^{s} v_{i,n} \right\} =$$

$$= as_B \left\{ \sum_{i=1}^{s} x_{i,1}, \sum_{i=1}^{s} x_{i,2}, \ldots \sum_{i=1}^{s} x_{i,n} \right\}.$$

Da der erste Vektor die geometrische Summe von $\overline{V}_1, \overline{V}_2, \ldots \overline{V}_s$ und der zweite von $\overline{X}_1, \overline{X}_2, \ldots \overline{X}_s$ ist, so ist der Satz bewiesen.

Folgende Sätze können nun ohne Beweis angegeben werden.

Satz 16

$$(a)\ldots\ldots \quad as_B(A\,x) = A\; as_B(x).$$

$$(b)\ldots\ldots \quad as_B(A\,\overline{X}) = A\; as_B(\overline{X})$$

$$(c)\ldots\ldots \quad as_B(A_1\overline{X}_1 + A_2\overline{X}_2 + \ldots + A_s\overline{X}_s) = \sum_{i=1}^{s} A_i\; as_B(\overline{X}_i)$$

$$(d)\ldots\ldots \quad as_B\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} as_B(x_i),$$

wo mit A, bzw., $A_1, A_2, \ldots A_s$, ebenso auch mit n und s konstante endliche Zahlen bezeichnet sind.

20. — An diesen Punkt angelangt, befinden wir uns auf dem Scheidewege. Ich gehe zuerst in diesem und zwei nächsten Paragraphen auf einige Fragen über unendliche Summen ein; dann wird im § 23 (Kap. V) noch eine andere Richtung genommen, um einige Resultate über stetige Funktionen zufälliger Variablen überhaupt mitzuteilen. In beiden Gebieten mögen wir uns aber mit den ersten Schritten zufrieden geben.

Es sei x eine zufällige Variable, die mit einer unabhängigen Variable φ stochastisch verbunden ist und eine stochastische Asymptote $v = f(\varphi)$ hat. Einer unbegrenzten Folge bestimmter φ-Werte: $\varphi_1, \varphi_2, \ldots \varphi_n, \ldots$ wird dann eindeutig eine ebenfalls unbegrenzte Folge zufälliger Variablen: $x_1, x_2, \ldots x_n, \ldots$ zugeordnet, wo mit x_n nichts anderes als unsere zufällige Variable x bei $\varphi = \varphi_n$ gemeint ist. Bezeichnet man noch $v_n = f(\varphi_n)$, so hat man

$$\lim_{n \to \infty} \overset{(\varepsilon)}{\underset{(o)}{P}} \mid x_n - v_n \mid = 1$$

oder $v_n = as_B\ (x_n)$.

Betrachtet man nun den Inbegriff der zufälligen Variablen x_n, $x_{n+1}, \ldots x_s$ als einen zufälligen Vektor

$$\overline{X}_{n,s} = \left\{ x_n, x_{n+1}, \ldots x_s \right\}.$$

und den Inbegriff der Werte: $v_n, v_{n+1}, \ldots v_s$ als einen (nicht zufälligen) Vektor

$$\overline{V}_{n,s} = \{ v_n, v_{n+1}, \ldots v_s \} ,$$

und lässt man s unbegrenzt zunehmen, so kommt man zu unendlich-dimensionellen Vektoren

$$\overline{X}_{n,\infty} = \{ x_n, x_{n+1}, \ldots \}$$

und

$$\overline{V}_{n,\infty} = \{ v_n, v_{n+1}, \ldots \}$$

Lässt es sich weiter, wie es sich gleich unten herausstellt, mit der Wahrscheinlichkeit der Erfüllung einer unbegrenzten Reihe der Bedingungen

$$| x_i - v_i | \leqq \varepsilon \quad , \quad (i = n, n+1, \ldots) ,$$

die wir mit

$$\overset{(\varepsilon)}{\underset{(o)}{P}} | x_i - v_i |_{i=n,n+1,\ldots}$$

bezeichnen, einen bestimmten Sinn verbinden und ist

$$\lim_{n \to \infty} \overset{(\varepsilon)}{\underset{(o)}{P}} | x_i - v_i |_{i=n,n+1,\ldots} = 1 ,$$

so sage man, dass

$$\{ v_n , v_{n+1}, \ldots \} = as_B \{ x_n, x_{n+1}, \ldots \}$$

ist. Es fragt sich, ob, bzw., unter welchen Bedingungen, der oben bewiesene Additionssatz auch für diese ∞-dimensionellen Vektoren Geltung habe.

Ich führe zuerst folgende kürzere Bezeichnungen ein. Es sei nämlich:

$$(23)\ldots \quad \left\{ \begin{array}{llll} P_n & = \overset{(\varepsilon)}{\underset{(o)}{P}} | x_n - v_n | \quad , & Q_n & = 1 - P_n \\ P_{n,s} & = \overset{(\varepsilon)}{\underset{(o)}{P}} | x_i - v_i |_{i=n,n+1,\ldots s}, & Q_{n,s} & = 1 - P_{n,s} \\ P_{n,\infty} & = \lim\limits_{s \to \infty} P_{n,s} \quad , & Q_{n,\infty} & = 1 - P_{n,\infty}. \end{array} \right.$$

Die stochastische Unabhängigkeit von $x_1, x_2, \ldots x_n, \ldots$ wird nicht vorausgesetzt.

Man bemerke erst, dass

$$(24)\ldots \quad P_n \geqq P_{n,n+1} \geqq P_{n,n+2}, \geqq \cdots \geqq P_{n,s} \geqq \cdots \geqq 0$$

ist, so, dass diese unbegrenzte, nicht zunehmende, von unten begrenzte Folge einen Grenzwert, den wir oben mit $P_{n,\infty}$ bezeichneten, haben muss (1). Ist dieser Grenzwert gleich 0, so ist $P_{n,\infty}=0$, $P_{n+1,\infty}=0$ u. s. w., was sich mit $\lim\limits_{n\to\infty} P_{n,\infty}=0$ schreiben lässt. $\bar{V}_{n,\infty}$ ist dann keine stochastische Asymptote von $\bar{X}_{n,\infty}$.

Betrachten wir nun den Fall, wo $P_{n,\infty}$ von Null verschieden ist. Da $P_{n,s}\leq P_{n+1,s}$ ist, und zwar bei jedem s, so gilt dieselbe Beziehung auch für die Grenzwerte, woher

(25) . . . $$P_{n,\infty}\leqq P_{n+1,\infty}\leqq P_{n+2,\infty}\leqq \ldots \leqq 1$$

Diese unbegrenzte, nicht abnehmende, von oben begrenzte Folge muss also einen Grenzwert, und zwar einen von Null verschiedenen haben, da die erste Grösse in der Folge ($P_{n,\infty}$) voraussetzungsgemäss >0 ist.

Um die Bedingungen, bei welchen $\lim P_{n,\infty}=1$ ist, kennen zu lernen, schreiben wir auf Grund des Boole'schen Satzes (2)

(26) . . . $$P_{n,s}\geqq 1-\sum_{i=n}^{s} Q_i \quad ,$$

woher sich für die Grenzwerte die Ungleichung

(27) . . . $$P_{n,\infty}\geqq 1-\sum_{i=n}^{\infty} Q_i \quad ,$$

ergibt, wo

(28) . . . $$\sum_{i=n}^{\infty} Q_i=\lim_{s\to\infty}\sum_{i=n}^{s} Q_i$$

ist. Nimmt man an, dass die Reihe $Q_n+Q_{n+1}+\ldots$ in dem engeren Sinne konvergiert, d. h. einen endlichen Grenzwert hat, so hat man bekanntlich

(29) . . . $$\lim_{n\to\infty}\sum_{i=n}^{\infty} Q_i=0 \quad ,$$

woher, der Formel (27) zufolge, sich schliessen lässt, dass

(30) . . . $$\lim_{n\to\infty} P_{n,\infty}=1$$

ist.

Betrachten wir jetzt noch den Fall der stochastischen Unabhängigkeit von $x_1, x_2, \ldots x_n, \ldots$ von einander. Da hat man

(31) . . . $$P_{n,s}=P_n P_{n+1}\ldots P_s=(1-Q_n)(1-Q_{n+1})\ldots(1-Q_s) \ ,$$

---

(1) F. P. Cantelli, *Sulla probabilità come limite della frequenza*, loc. cit. p. 41.
(2) *Ibid.* p. 41.

und da die Grössen Q_n, $Q_{n+1}, \ldots$ wesentlich positiv sind, so wird bekanntlich $\lim\limits_{s \to \infty} P_{n,s}$ einen Grenzwert 0, oder einen von 0 verschiedenen haben, je nachdem die unendliche Reihe: Q_n, $Q_{n+1}, \ldots$ divergent oder konvergent ist. Wir wissen aber, dass bei $P_{n,\infty} = 0$ auch $\lim\limits_{n \to \infty} P_{n,\infty} = 0$ ist, dass aber, bei $P_{n,\infty}$ endlichem und von 0 verschiedenem, $\lim\limits_{n \to \infty} P_{n,\infty} = 1$ ist. So verrät sich die Konvergenz der genannten Reihe im Falle der stochastischen Unabhängigkeit nicht nur als eine hinreichende, sondern auch als eine notwendige Bedingung dafür, das $\bar{V}_{n,\infty} = as_B\, \bar{X}_{n,\infty}$ ist.

Es wurde im Vorigen vorausgesetzt, dass $v_n = \underset{n \to \infty}{as_B} (x_n)$ ist. Lässt man diese Voraussetzung fallen, so folgt daraus, dass die Bedingung $\lim\limits_{n \to \infty} P_n = 1$ nicht erfüllt wird, und da nach der Formel (24)

$$P_{n,\infty} \leqq P_n \leqq 1$$

sein muss, so kann auch die Bedingung

$$\lim_{n \to \infty} P_{n,\infty} = 1$$

nicht erfüllt werden. Sodann ist also $\bar{V}_{n,\infty}$ wieder keine stochastische Asymptote von $\bar{X}_{n,\infty}$.

Fasst man die Resultate obiger Untersuchungen zusammen, so kommt man zu

Satz 17. Es seien $\varphi_1, \varphi_2, \ldots \varphi_n, \ldots$ *eine unbegrenzte Folge von Werten einer unabhängigen Variable* φ *und es sei* x_i *eine zufällige Variable, die mit* φ_i *stochastisch verbunden ist. Ist weiter* $v_i = f(\varphi_i)$, *und*

$$Q_n = \overset{\infty}{\underset{\varepsilon}{P}} \mid x_n - v_n \mid,$$

so wird

$$\{v_n, v_{n+1}, \ldots\} = \underset{n \to \infty}{as_B} \{x_n, x_{n+1}, \ldots\}$$

sein, wenn (*a*) $v_i = \underset{i \to \infty}{as_B} (x_i)$ *ist und* (*b*) *die unendliche Reihe:*

$Q_n + Q_{n+1} + \ldots$ *gegen einen endlichen Wert konvergiert.*

Die Bedingung (*a*) *ist dabei eine notwendige, bei der stochastischen Unabhängigheit aber der zufälligen Variablen* $x_n, x_{n+1}, \ldots$ *von einander, wird auch die Bedingung* (*b*) *notwendig.* Wie es damit im Falle der stochastischen Verbundenheit steht, bleibe dahingestellt.

21. — Wollen wir obiges allgemeines Schema durch ein konkreteres illustrieren. Es seien x_n, $x_{n+1}, \ldots x_s, \ldots$ zufällige Variablen, von denen jede durch eine entsprechende Versuchsserie und zwar zu n, $n+1, \ldots s, \ldots$ Versuchen, die von einander nicht unabhängig vorausgesetzt werden sollen, bestimmt wird. Es werde nur angenommen, dass die Versuchsanordnungen für eine Serie zu n durch die Zahl n vollständig definiert ist, so dass die Wahrscheinlichkeitsverteilung von x_n die Funktion lediglich von n ist. Hat man, z. B., eine Reihe von Urnen mit Wahrscheinlichkeiten der Ziehung einer weissen Kugel: $p_1, p_2, \ldots p_n, p_{n+1}, \ldots$, so kann x_n die Häufigkeit ihres Erscheinens in n Ziehungen sein, die der Reihe nach aus n ersten Urnen vorgenommen werden können. Oder, hat man eine unbegrenzte Reihe von Versuchsanordnungen für zufällige Variablen $\xi_1, \xi_2, \ldots \xi_n, \xi_{n+1}, \ldots$ und wird stets eine Reihe aus n Versuchen auf den n ersten Variablen dieser Reihe vorgenommen, so kann unsere zufällige Variable x_n entweder dem arithmetischen Durchschnitt: $\xi_{(n)} = \frac{1}{n} \sum\limits_{i=1}^{n} \xi'_i$,

oder der empirischen Streuung: $\frac{1}{n} \sum\limits_{i=1}^{n} (\xi'_i - \xi_{(n)})^2$, bzw., einer anderen Funktion von $\xi_1, \xi_2, \ldots \xi_n$ gleichgesetzt werden. Die stochastischen Bedingungen der Versuche können dabei in sehr mannigfaltiger Weise angeordnet werden. So kann man, z. B., als die n-te Urne *die* nehmen, aus welcher schon n — 1 Ziehungen mit Zurücklegen vorgenommen wurden, man kann die Urne durch Los oder durch Resultat des unmittelbar vorangehenden Versuchs bestimmen lassen u. s. w.

Nun ist es und zwar bei jeder gegebenen Versuchsanordnung noch in sehr verschiedener Weise möglich einzelne zufällige Variablen x_n, $x_{n+1}, \ldots$ in ∞ dimensionelle Gesammtheiten zu vereinigen und so verschiedene zufällige Vektoren zu bilden. Ich will hier etwas näher deren zwei Arten betrachten.

(I) An einer unbegrenzten Reihe von zufälligen Variablen: ξ_1, $\xi_2, \ldots \xi_n, \ldots$ sei eine Versuchsserie zu n Versuchen in der Weise vollbracht, dass der erste Versuch an der ersten Variable, der zweite an der zweiten u. s. f. angestellt wird. Es sei damit ein empirischer Wert der zufälligen Variable $x_n = F(\xi'_1, \xi_2', \ldots \xi'_n)$ bestimmt worden. In derselben Weise *mittelst einer neuen und von der ersten unabhängigen Versuchsserie* zu $n+1$ Versuchen sei ein empirischer Wert der zufälligen Variable $x_{n+1} = F(\xi_1'', \xi_2'', \ldots \xi''_{n+1})$ bestimmt. Die durch dieses Verfahren definierte unbegrenzte Reihe von unabhängigen zufälligen Variablen: x_n, $x_{n+1}, \ldots x_s, \ldots$ wird einen ∞-dimensionellen Vektor bilden, den ich mit $\{x_n, x_{n+1}, \ldots\}^{I}$ oder kürzer mit $\bar{X}^{I}_{n,\infty}$ bezeichne.

(II) Man kann aber aus ein und derselben unbegrenzten Versuchsfolge die ersten n, dann die ersten $n+1$, u. s. f. heraussondern um dadurch eine unbegrenzte Folge der zufälligen Variablen: x_n, $x_{n+1}, \ldots$ bestimmen zu lassen. So wird wieder ein ∞-dimensioneller Vektor: $\{x_n, x_{n+1}, \ldots\}^{II}$, kürzer $\bar{X}^{II}_{n,\infty}$ definiert, den man auch den (verallgemeinerten) Cantelli'schen Vektor nennen kann (1).

Beginnt man mit dem ersten Fall, so findet man aus der Formel (31), indem man sich der bekannten Ungleichung

$$lg(1-x) < -x$$

erinnert, dass

$$(32)\ldots \qquad lg P^{I}_{n,s} < -\sum_{i=n}^{s} Q_i$$

und

$$(33)\ldots. \qquad P^{I}_{n,s} < e^{-\sum\limits_{i=n}^{s} Q_i}$$

ist. Da aber

$$e^{-x} < 1 - x + \frac{1}{2}x^2$$

ist, so kommt man unmittelbar zu Ungleichung

$$(34)\ldots \qquad P^{I}_{n,s} < 1 - \sum_{i=n}^{s} Q_i + \frac{1}{2}\left(\sum_{i=n}^{s} Q_i\right)^2,$$

woher und unter der Rücksichtnahme auf die Formeln (23) und (26) sich die Ungleichungen

$$(35)\ldots \qquad \sum_{i=n}^{s} Q_i \left(1 - \frac{1}{2}\sum_{i=n}^{s} Q_i\right) < Q^{I}_{n,s} \leqq \sum_{i=n}^{s} Q_i$$

und endlich, nachdem man zu den Grenzwerten übergeht, die Ungleichungen

$$(36)\ldots \qquad \sum_{n}^{\infty} Q_i \left(1 - \frac{1}{2}\sum_{n}^{\infty} Q_i\right) < Q^{I}_{n,\infty} \leqq \sum_{n}^{\infty} Q_i$$

---

(1) F. P. Cantelli betrachtet (loc. cit., p. 41 ff.) eine unbegrenzte Reihe unabhängiger zufälliger Variablen: $\xi_1, \xi_2, \ldots \xi_n, \ldots$ und bildet dann einen (in unserer Sprache) ∞ — dimensionellen Vektor aus den Durchschnittswerten $\xi_{(n)}$, $\xi_{(n+1)}, \ldots$ nach dem Modus II. Das Wesen der Sache wird dabei aber nicht geändert, wenn man auch in Betreff von $\xi_1, \xi_2, \ldots$ die Voraussetzung der Unabhängigket fallen lässt.

ergeben. Hat nun die Summe $Q_n + Q_{n+1} + \ldots$ keinen eigentlichen Grenzwert, so liefert uns die Formel (36) die triviale Ungleichung $-\infty < Q^{\mathrm{I}}_{n,\infty} < \infty$; ist aber die Reihe $Q_n + Q_{n+1} + \ldots$ konvergent, so hat man

$$\lim_{n \to \infty} Q^{I}_{n,\infty} = 0$$

und

$$(37)\ldots \qquad \lim_{n \to \infty} \frac{Q^{I}_{n,\infty}}{\sum\limits_{n}^{\infty} Q_i} = 1 ,$$

was auch in der Form $Q^{\mathrm{I}}_{n,\infty} \sim \sum\limits_{n}^{\infty} Q_i$, $(n \to \infty)$ geschrieben werden kann. Bei genügend grossen Werten von n können daraus die Werte von $Q^{I}_{n,\infty}$ und $P^{I}_{n,\infty}$ mieiner grossen Approximation berechnet werden.

Ich will diese Berechnung für den Fall unabhängiger zufälliger Variablen: $\xi_1, \xi_2, \ldots \xi_n, \ldots$ von denen jede dieselbe Wahrscheinlichkeitsverteilung hat, auch wirklich durchführen. Es sei

$$M = E\,\xi_1 = E\,\xi_2 = \ldots$$
$$\sigma^2 = \mu_2 = E(\xi_1 - M)^2 = E(\xi_2 - M)^2 = \ldots$$
$$\xi_{(n)} = \frac{1}{n} \sum_{i=1}^{n} \xi'_i$$

$$\sigma^2_{(n)} = \mu_{2\,(n)} = E(\xi_{(n)} - E\,\xi_{(n)})^2 = \frac{\sigma^2}{n} .$$

Ist noch der vierte Moment endlich, und fügt sich also $\xi_{(n)}$ dem Gauss-Laplace'schen Grenzgesetze welches uns bei grossem n als eine gute approximative Formel gelten darf, so kann man bekanntlich, ebenso bei grossem n, auch die approximative Formel

$$(38) \qquad \sqrt{\frac{2}{\pi}} \int_{\varepsilon}^{\infty} e^{-\frac{t^2}{2\sigma^2_t}} dt = \sqrt{\frac{2}{\pi}}\, \frac{\sigma_t}{\varepsilon}\, e^{-\frac{\varepsilon^2}{2\sigma^2_t}}$$

anwenden. Setzt man hierin statt t und σ_t die Grössen $(\xi_{(n)} - M)$, bzw., $\sigma_{(n)} = \frac{\sigma}{\sqrt{n}}$ ein, so findet man:

$$(39)\ldots\ldots \qquad \dot{Q}_n = \sqrt{\frac{2}{\pi}}\, \frac{\sigma}{\varepsilon\sqrt{n}}\, e^{-\frac{\varepsilon^2 n}{2\sigma^2}}.$$

Führt man weiter zur Abkürzung die Bezeichnung

$$(40) \ldots\ldots \qquad u = e^{-\frac{\varepsilon^2}{2\sigma^2}}$$

ein, so findet man (wenn das Gleichungszeichen im Sinne der approximativen Gleichung geschrieben wird):

$$(41) \ldots Q^I_{n,\infty} = \sqrt{\frac{2}{\pi}} \frac{\sigma u^n}{\varepsilon \sqrt{n}} \left(1 + \frac{u}{\sqrt{1+\frac{1}{n}}} + \frac{u^2}{\sqrt{1+\frac{2}{n}}} + \ldots\right)$$

$$= Q_n S ,$$

woher den Beziehungen

$$(42)\ 1 + \frac{u}{\sqrt{1+\frac{1}{n}}} + \left(\frac{u}{\sqrt{1+\frac{1}{n}}}\right)^2 + \ldots < S < 1 + u + u^2 + \ldots ,$$

bzw.,

$$(43) \ldots\ldots \qquad \frac{1}{1 - \frac{u}{\sqrt{1+\frac{1}{n}}}} < S < \frac{1}{1-u}$$

zufolge sich die Ungleichungen

$$(44) \ldots\ldots \qquad \frac{Q_n}{1-u} \cdot \frac{1-u}{1 - u\sqrt{\frac{n}{n+1}}} < Q^I_{n,\infty} < \frac{Q_n}{1-u}$$

ergeben. Bei grossem n kann man also *fast* genau

$$(45) \ldots\ldots \qquad Q^I_{n,\infty} = \frac{1}{1-u} Q_n$$

setzen. Ist noch die Grösse $\frac{\varepsilon^2}{\sigma^2}$ klein genug, so hat man eine approximative Gleichung

$$u = e^{-\frac{\varepsilon^2}{2\sigma^2}} = 1 - \frac{\varepsilon^2}{2\sigma^2} ,$$

was, in (45) eingesetzt, uns eine einfache Formel

$$(46) \ldots\ldots \qquad Q^I_{n,\infty} = 2\left(\frac{\sigma}{\varepsilon}\right)^2 Q_n$$

gibt, die eine sehr gute Approximation sein kann, wenn n sehr gross, ε/σ sehr klein und $Q_{n,\infty}$ ebenso sehr klein ausfällt. In entgegengesetzten Fällen soll man die Rechnung mit anderen, ebenfalls gut bekannten, Mitteln strenger durchführen, was keine prinzipiellen Hindernisse bietet, und für uns von keinem weiteren Interesse ist.

Für das BERNOULLI'sche Schema ist σ unserer Formel gleich $\sqrt{p\,q}$, und so verwandelt sich die Formel (46) in

$$(47)\ldots\ldots \qquad Q^{I}_{n,\infty} = \frac{2\,p\,q}{\varepsilon^2}\,Q_n$$

wo für Q_n, nach (39),

$$(48)\ldots\ldots \qquad Q_n = \sqrt{\frac{2}{\pi}}\;.\;\frac{\sqrt{p\,q}}{\varepsilon\sqrt{n}}\;e^{-\frac{\varepsilon^2 n}{2pq}}$$

approximativ gelten kann.

Z. B., bei $p = \frac{1}{2}$, $q = \frac{1}{2}$, $n = 10.000.000.000$, $\varepsilon = 0{,}01$, hat man

$$Q_n = \sqrt{\frac{2}{\pi}}\;\frac{\sqrt{\frac{1}{2}\cdot\frac{1}{2}}}{0{,}01\sqrt{10^{10}}}\;e^{-\frac{10^{-4}\,10^{10}}{2.\frac{1}{2}\cdot\frac{1}{2}}} = 4{,}0.10^{-4}\,.\,e^{-2.10^{6}} = 4.10^{-868593},$$

woher

$$Q^{(I)}_{n,\infty} = \frac{2\,.\,{}^1/_2\,.\,{}^1/_2}{(0{,}01)^2}\,Q_n = 5000\;Q_n = 2.10^{-868589}.$$

Sogar bei $p = {}^1/_2$, $q = {}^1/_2$, $n = 10^{10}$, $\varepsilon = 0{,}0001$ hat man noch

$$Q_n = 6.10^{-89} \text{ und } Q^{I}_{n,\infty} = \frac{1}{2}\,.\,10^8\;Q_n = 3.10^{-81}\;.$$

Hat man also (unter den Voraussetzungen des BERNOULLI'schen Falles und bei $p = {}^1/_2$) eine unbegrenzte Folge unabhängiger Versuchserien zu $10^{10}, 10^{10}+1, 10^{10}+2$, u. s. w. Versuchen, so ist es fast gänzlich ausgeschlossen ($Q^{I}_{n,\infty} = 3.10^{-81}$), dass die Häufigkeit eines Ereignisses von seiner Wahrscheinlichkeit (${}^1/_2$) irgend jemals mehr als um $0{,}02\,{}^0/_0$ $\left(100\,.\,\frac{0{,}0001}{0{,}5}\right)$ abweichen wird.

Was den CANTELLI'schen Vektor $\overline{X}^{II}_{n,\infty}$ betrifft, so bemerkt man leicht, dass die Berechnung von $Q^{II}_{n,\infty}$ für ihn sehr verschieden ausfallen kann, je nach der Art des stochastischen Zusammenhanges zwischen zufälligen Variablen. Obige Ausführungen aber geben uns übrigens die niedere und obere Grenzen für $Q^{II}_{n,\infty}$, da stets die

Ungleichungen

$$(49) \ldots\ldots \qquad Q_n \leqq \overset{II}{Q}_{n,\infty} \leqq \sum_{n}^{\infty} Q_i$$

gelten müssen. Bei den Voraussetzungen des letzten Beispiels hat man darum $3.10^{-89} \leqq Q^{II}_{n,\infty} \leqq 3.10^{-81}$, was manchmal zur Abschä tzung der Grössenordnung genügen kann.

22. — Der Satz 17 kann uns noch verschiedene andere abgeleitete Kriterien für die Beurteilun geines ∞-dimensionellen Vektors abgeben. So wird, z. B., bekanntlich die unendliche Reihe $Q_n + Q_{n+1} + \ldots$ konvergent sein, wenn

$$\lim_{n \to \infty} \left(Q_n \, n^{1+\alpha}\right)$$

bei $\alpha > 0$ einen endlichen Wert hat, und divergent, wenn

$$\lim_{n \to \infty} \left(Q_n \, n\right)$$

nicht Null ist. Bei $\lim\, (Q_n\, n) = 0$, kann sie aber ebenso konvergent, wie auch divergent sein. Setzt man in die Formel 4 (Kap. III) statt $E\, G\,(x - v)$ die Grösse $g_r\,(x, Ex)$ und statt $G\,(\varepsilon)$ die Grösse ε^r ein, so hat man eine bekannte Ungleichung

$$(50) \ldots\ldots \qquad \overset{\infty}{\underset{\varepsilon}{P}} \mid x - E\,x \mid < \frac{g_r\;(x, E\,x)}{\varepsilon^r} \,,$$

wo die Wahrscheinlichkeit links mit obigem Q_n identisch ist. So kommt man zum Kriterium der Konvergenz der Reihe $Q_n + Q_{n+1} + \cdots$, das darin besteht, dass der Grenzwert

$$(51) \ldots \qquad \lim_{n \to \infty} \frac{g_r\,(x, E\,x).\; n^{1+\alpha}}{\varepsilon^r} \qquad ,(\alpha > o)$$

oder, bei $\varepsilon = \text{const}$, der Grenzwert

$$(52) \ldots\ldots \qquad \lim\, g_r\;(x, E\,x)\;.\; n^{1+\alpha} \qquad , \quad (\alpha > o)$$

endlich sein muss. Man beachte dabei, dass diese Bedingung eine lediglich hinreichende ist.

Will man dieses Kriterium an die unendliche Folge arithmetischer Durchschnitte, d. h. an den Vektor

$$\bar{X}^I = \{x_{(n)}\;,\; x_{(n+1)}\,, \ldots\}^I ,$$

bzw., an den Vektor

$$\bar{X}^{II} = \{x_n\,,\; x_{(n+1)}\,, \ldots\}^{II}$$

anwenden, so hat man die Momente von $x_{(n)}$ in Betracht zu ziehen. Die Unabhängigkeit der betreffenden zufälligen Variablen von einander vorausgesetzt, hat man, z. B., für den vierten zentralen Moment von $x_{(n)}$ folgenden Ausdruck (1)

$$(53) \ldots \quad \mu_{4\,(n)} = \frac{3}{n^2}\,\mu^2_{[2\,;\,n]} + \frac{1}{n^3}\,\mu_{[4\,;\,n]} - \frac{3}{n^3}\left\{\frac{1}{n}\sum_{i=1}^{n}(\mu_2^{(i)})^2\right\}.$$

Man findet sofort, dass

$$n^{1+\alpha}\mu_{4\,(n)}$$

gegen Null konvergieren wird, wenn

$$(54) \ldots \quad \begin{cases} \lim\limits_{n \to \infty} \dfrac{\mu^2_{[2\,;\,n]}}{n^{1-\alpha}} = o \\ \text{und} \\ \lim\limits_{n \to \infty} \dfrac{\mu_{[4\,;\,n]}}{n^{2-\alpha}} = o \end{cases}$$

ist: $\mu_{[2;n]}$ und $\mu_{[4;n]}$ müssen entweder von oben begrenzt sein, oder von niedrigerer Unendlichkeitsordnung als $n^{1/2}$, bzw., n^2 sein. Da bei der Erfüllung dieser Bedingungen noch stets auch die Bedingung

$$(55) \ldots \quad \lim_{n \to \infty} \mu_{2\,(n)} = \lim_{n \to \infty} \frac{1}{n}\,\mu_{[2\,;\,n]} = o$$

gilt, was den Bestand stochastischer Asymptoten von $x_{(n)}$ zur Folge hat, so ergibt sich daraus nach Satz 17 auch die stochastische Konvergenz beider Vektoren $\overline{X}^{\mathrm{I}}$ und $\overline{X}^{\mathrm{II}}$ (2).

Es sei noch bemerkt, dass wenn $\mu^2_{[2;n]}$ und $\mu_{[4;n]}$ feste obere Grenzen haben, oder wenn die Unendlichkeitsordnung von $\mu_{[4;n]}$ die Unendlichkeitsordnung von $\mu^2_{[2;n]}$ weniger als um 1 übertrifft, so wird

$$\lim_{n \to \infty} \frac{\mu_{[4\,;\,n]}}{n\,\mu^2_{[2\,;\,n]}} = o$$

sein, was einen speziellen Fall der Liapunoff's Bedingung darstellt (3).

---

(1) AL. A. TSCHUPROW, *On the mathematical expectation of the moments of frequency distributions*, Part. II, « Biometrika », Vol. XIII, 1921, p. 286.

(2) Cf. F. P. CANTELLI, *Sulla probabilità come limite della frequenza*, loc. cit., p. 43.

(3) Bezeichnet man $g_r^{(i)} = E \mid x_i - E\,x_i \mid^r$ und $g_{[r;\,n]} = \frac{1}{n}\sum_{i=1}^{n} g_r^{(i)}$, so kann LIAPUNOFF's Bedingung am bequemsten in der Form

$$\lim_{n \to \infty} \frac{g^2_{[k;\,n]}}{n^{k-2}\,g^k_{[2\,;\,n]}} = o \quad , \quad (k > 2)$$

geschrieben werden. A. LIAPUNOFF, *op. cit.*, p. 3.

Sodann wird sich der arithmetische Durchschnitt dem Gauss-Laplace'schen Grenzgesetz fügen und die Wahrscheinlichkeiten $P^{\mathrm{I}}_{n,\infty}$ und $P^{\mathrm{II}}_{n,\infty}$ werden sich bei genügend grossen n nach den Vorschriften des vorigen Paragraphs ermitteln, bzw., abschätzen lassen.

Fünftes Kapitel.

Ueber stochastische Grenzwerte der zufälligen Variablen, die Funktionen anderer zufälligen Grössen sind.

23. — Der Satz über algebraische Addition stochastischer Asymptoten, den wir im § 19 bewiesen haben, ist lediglich ein spezieller Fall allgemeinerer Sätze über stochastische Asymptoten der zufälligen Variablen, die Funktionen anderer zufälligen Grössen sind. Nicht jede Funktion einer zufälligen Variable, die stochastische Asymptoten hat, muss auch selbst solche haben. So besitzt die zufällige Variable, deren Verteilung durch das Schema:

$$\left\{\begin{array}{llll} x: & n-\dfrac{1}{n}, & n, & n+\dfrac{1}{n} \\ p: & \dfrac{1}{2}-\dfrac{1}{2n}, & \dfrac{1}{n}, & \dfrac{1}{2}-\dfrac{1}{2n} \end{array}\right\}$$

gegeben ist, eine stochastische Asymptote: $\underset{n\to\infty}{as_B}(x)=n$. Das Quadrat von x hat aber keine stochastischen Asymptoten, was aus dem Verteilungschema

$$\left\{\begin{array}{llll} x^2: & n^2-2+\dfrac{1}{n^2}, & n^2, & n^2+2+\dfrac{1}{n^2} \\ p: & \dfrac{1}{2}-\dfrac{1}{2n}, & \dfrac{1}{n}, & \dfrac{1}{2}-\dfrac{1}{2n} \end{array}\right\}$$

ersichtlich ist. Noch ein Beispiel. Die Häufigkeit $\left(\dfrac{m}{n}\right)$ eines Ereignisses hat bekanntlich unter gewissen Voraussetzungen einen stochastischen Grenzwert. Das Produkt aber aus der Häufigkeit auf die Versuchszahl $\left(\dfrac{m}{n}\cdot n\right)$ ist die Ereigniszahl, eine zufällige Variable, die weder einen stochastischen Grenzwert, noch überhaupt eine stochastische Asymptote hat.

Der Leser, dem die heutige Lehre von reellen Funktionen bekannt ist, bemerkt sofort, dass man hier ein Thema berührt, für dessen Behandlung uns mathematische Instrumente von ausgezeichneter Qualität und Schärfe zur Verfügung stehen. Doch erheischt die systematische Betrachtung betreffender Fragen eine Reihe von Untersuchungen, die eine besondere monographische Bearbeitung erfordern und im Rahmen dieser Arbeit nicht ausgeführt werden können. Das folgende soll also nur als ein lediglich sehr bescheidener Beitrag zu dem hier gestreiften Problem betrachtet werden.

Satz 18. Es sei: $x_1, x_2, \ldots x_n$ *eine endliche Anzahl (eindimensioneller) zufälliger Variablen, von denen jede* (x_i) *mit einer unabhängigen Variable* (φ_i) *stochastisch verbunden ist und einen stochastischen Grenzwert* $c_i = \lim_B\limits_{\varphi'_i\,,\,\varphi''_i\,,\ldots} (x_i)$ *hat. Ist* $F\ (x_1,\ x_2, \ldots\ x_n)$ *eine von* $\varphi_1, \varphi_2, \ldots \varphi_n$ *direkt unabhängige und in der Nähe von dem Punkte* ($c_1, c_2, \ldots c_n$) *stetige Funktion, so besitzt diese letztere einen stochastischen Grenzwert, welcher der Funktion der betreffenden Grenzwerte gleich ist, oder in Symbolen:*

$$(1)\ldots \qquad \lim_{(\varphi'_i\,\varphi''_i\,,\ldots;\,i=1\,,2\,,\ldots n)} F\,(x_1\,,x_2\,,\ldots x_n) = F\,(c_1\,,c_2\,,\ldots c_n)\ ,$$

Der Beweis gestaltet sich folgendermassen. Nach dem Satz 12 muss es in jeder unbegrenzten Folge $\varphi'_i\ ,\ \varphi''_i\ ,\ldots$ solch einen Wert $\varphi_i^{(k_i)}$ geben, dass, bei beliebigen Werten von

$$\varphi_i^{(h_i)}\ (h_i > k_i\ ;\ i = 1,2,\ldots n),$$

die Wahrscheinlichkeit des Zusammenbestehens der Ungleichungen

$$(2)\ldots \qquad |\ x_i - c_i\ | \leqq \varepsilon_i\ \ ,\quad (i = 1\,,2\,,\ldots n)$$

der Bedingung

$$(3)\ldots \qquad \overset{(\varepsilon_i)}{\underset{(o)}{P}}\ |\ x_i - c_i\ |_{\,i=1\,,2\,,\ldots n} > 1 - \eta$$

genügen wird, wo $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_n$ und η beliebig kleine vorgegebene Zahlen sind. Bezeichnet man mit λ die Differenz zwischen dem grössten und dem kleinsten Werte der Funktion F in dem durch (2) definierten Intervall, so bemerkt man, dass die Erfüllung von (2) die Erfüllung der Ungleichung

$$(4)\ldots \qquad |\ F(x_1\,,x_2\,,\ldots x_n) - F(c_1\,c_2\,,\ldots c_n)\ | \leqq \lambda$$

zur Folge hat, dass aber die Erfüllung von (4) die Geltung von (2) nicht voraussetzt. Es folgt daher, dass

$$(5)\ldots \qquad \overset{(\lambda)}{\underset{(o)}{P}}\ |\ F(x_1\,,x_2\,,\ldots x_n) - F(c_1\,,c_2\,,\ldots c_n)\ | \geqq$$
$$\geqq \overset{(\varepsilon_i)}{\underset{(o)}{P}}\ |\ x_i - c_i\ |_{\,i=1\,,2\,,\ldots n} > 1 - \eta$$

ist. Da aber für jede in der Nähe von $(c_1, c_2, \ldots c_n)$ stetige und von $\varphi_1, \varphi_2, \ldots \varphi_n$ direkt unabhängige Funktion die Intervallsgrenzen $\pm \varepsilon_1, \pm \varepsilon_2, \ldots \pm \varepsilon_n$ so gewählt werden können, dass λ einer vorgegebenen beliebig kleinen konstanten Zahl gleich wird, so ist die Ungleichung (5) der Gleichung

$$(6)\ldots \quad \lim_{\varphi'_i\, \varphi''_i\,, \ldots (i=1,2,\ldots n)} \overset{(\lambda)}{\underset{(o)}{P}} \mid F(x_1, x_2, \ldots x_n) - F(c_1, c_2, \ldots c_n) \mid = 1$$

äquivalent, wo λ eine gegebene beliebig kleine Zahl ist. Damit ist aber der Satz 18 bewiesen.

24. — Gehen wir nun zur Auffindung einiger wichtigen Folgerungen des bewiesenen Satzes über. Man erinnere sich, dass, wenn eine zufällige Variable einen stochastischen Grenzwert hat, jede stochastische Asymptote gegen diesen letzteren konvergieren muss. Man hat also für die Variablen des vorigen Paragraphen

$$\lim as_B\, F(x_1, x_2, \ldots x_n) = \lim_B F(x_1, x_2, \ldots x_n)$$

und

$$\lim as_B(x_i) = \lim_B (x_i) = c_i \ ,$$

woher, der Formel (1) zufolge,

$$(7)\ldots \quad \lim as_B F(x_1, x_2, \ldots x_n) = F\{\lim as_B(x_1), \lim as_B(x_2), \ldots \lim as_B(x_n)\}.$$

Setzt man zur Abkürzung

$$y = F(x_1, x_2, \ldots, x_n)$$

und erinnert man sich, dass die Mediane (C_1) eine unvermeidliche Asymptote ist, so hat man

$$(8)\ldots \quad \lim C_1^{(y)} = F(\lim C_1^{(1)}, \lim C_1^{(2)}, \ldots \lim C_1^{(n)}),$$

und ähnlich auch für jede andere unvermeidliche stochastische Asymptote, so, z. B., für die « Mode ».

Was die Zentralwerte betrifft, so müssen dazu noch einige andere bestimmte Bedingungen hinzukommen. Und zwar, *erstens*, müssen die betreffenden mathematischen Erwartungen überhaupt bestehen, damit man von entsprechenden Zentralwerten reden könnte, und, *zweitens*, müssen diese letzteren stochastische Asymptoten sein. Eine bei den Voraussetzungen des Satzes 18 hinreichende, obwohl noch keine notwendige Bedingung dafür ist die, dass betreffende Momente feste obere Grenzen haben (siehe oben Satz 8). Sind diese Bedingungen erfüllt, so ist die Formel (7) auf entsprechende Zentralwerte anwendbar. Z. B.:

1) bestehen die Momente: $\mu_2^{(1)}, \mu_2^{(2)}, \ldots \mu_2^{(n)}; \mu_2^{(y)}$ und die mathematischen Erwartungen: $Ex_1, Ex_2, \ldots Ex_n; Ey$;

2) haben $\mu_2^{(1)}, \mu_2^{(2)}, \ldots \mu_2^{(n)}; \mu_2^{(y)}$ feste obere Grenzen, und

3) sind die Voraussetzungen des Satzes (18) erfüllt, so müssen $Ex_i = as_B\ (x_i)$, $Ey = as_B\ (y)$ sein, und es gilt darum die Gleichung

$$(9) \ldots \lim E\,F(x_1, x_2, \ldots x_n) = F(\lim Ex_1, \lim Ex_2, \ldots \lim Ex_n).$$

a) Liegen mögliche Werte unserer zufälligen Variablen: x_1, $x_2, \ldots x_n$ in Intervallen, deren Längen von oben begrenzt werden, und

b) ist die Funktion F überall stetig, so sind die Bedingungen (1) und (2) von selbst erfüllt, ebenso auch alle ähnlichen, die andere Zentralwerte betreffen. So müssen dann zu (*a*) und (*b*) nur die entsprechenden Veraussetzungen des Satzes (18) hinzukommen. Sie werden erfüllt, z. B.,

c) wenn alle Streuungen: $\mu_2^{(1)}, \mu_2^{(2)}, \ldots \mu_2^{(n)}$ gegen Null konvergieren, und

d) wenn jede mathematische Erwartung Ex_i $(i = 1, 2, \ldots n)$ einen endlichen Grenzwert hat. Dann wird die Gleichung (9) für alle stetigen Funktionen und für alle Wahrscheinlichkeitsverteilungen gelten müssen.

Ist die Bedingung (*a*) nicht erfüllt, so muss man direkter Weise sich überzeugen, dass sowohl $\mu_2^{(1)}, \mu_2^{(2)}, \ldots \mu_2^{(n)}$, wie auch $\mu_2^{(y)}$ feste obere Grenzen haben, was besonders in Betreff der letzteren Grösse nicht stets eine leichte Aufgabe ist.

Lassen wir nun auch die Bedingung (*b*) fallen, so kann es sich manchmal treffen, dass die Funktion $F\ (x_1, x_2, \ldots x_n)$ überhaupt keine, bzw., keine uns interessierenden Momente haben wird. Und doch kann dabei der Satz (18) volle Geltung haben, wenn nur die Funktion F in unmittelbarer Nähe von $c_1, c_2, \ldots c_n$ stetig bleibt. Statt voller mathematischer Erwartungen kann man in solchen Fällen die partiellen, bzw., die bedingten betrachten, die durch Ausscheidung aller Unbestimmtheits, bzw., Unendlichkeitsstellen (ev. mit ihren nächsten Umgebungen) definiert werden. Bezeichnet man die Wahrscheinlichkeit der Ausgeschiedenen Stellen mit Q, und aller übrigen mit $P = 1 - Q$, und führt man für die partielle mathematische Erwartung die Bezeichnungen: $E'_{[0/0]}\,x$, bzw., $E'_{[\infty]}\,x$, oder allgemein $E'_{[J]}\,x$ ein, so können die entsprechenden bedingten mathematischen Erwartungen mit $E_{[0/0]}\,x = \frac{1}{P}\,E'_{[0/0]}\,x$, bzw., mit $E_{[\infty]}\,x = \frac{1}{P}\,E'_{[\infty]}\,x$, oder allgemein mit $E_{[J]}\,x$ bezeichnet werden (1). Lässt man nicht ausser Acht, dass es sich um Fälle handelt, wo es überhaupt keine, bzw., keine endliche mathematische Erwartung gibt, so ersieht man sofort, dass die Frage, welche von beiden Ersatzwerten « richtiger » ist, keinen Sinn hat, und dass die Wahl lediglich nach Zweckmässigkeitsgründen zu treffen ist. Dabei muss man aber sich hüten die begrifflichen Grenzen zwischen der mathematischen Erwartung und diesen Ersatzwerten zu verwischen. Es soll volle Klarheit geschaffen werden sowohl über die Zwecke, für welche sie zu Verwendung kommen können, wie auch über ihre Eigenschaften, die sie mit vollen mathematischen Erwartungen gemein haben mögen und derenthalber sie die Rolle der mathematischen Erwartungen übernehmen dürfen.

Scheidet man von möglichen Werten (bzw., Intervallen) einer zufälligen Variable (x) einige (J) aus, und dividiert man die Wahrscheinlichkeiten übriggebliebener Werte durch ihre Gesammtwahrscheinlichkeit (P), so erhält man eine neue zufällige Variable ($x_{[J]}$), die man die *bedingte zufällige Variable* nennen kann. Die bedingte mathematische Erwartung der ersten (x) ist dabei die unbedingte mathematische Erwartung der zweiten ($x_{[J]}$). So kann man auf diese letztere alle Sätze des III Kapitels anwenden. Ihre Ausnützung aber zur Beurteilung der ursprünglichen zufälligen Variable beruht dabei auf folgendem Satze, den ich hier ohne Beweis aufstellen darf.

Satz 19. Konvergiert die Gesammtwahrscheinlichkeit (P) *der bedingten zufälligen Variable* ($x_{[J]}$) *gegen 1, und hat die ursprüngliche zufällige Variable* (x) *stochastische Asymptoten, so hat stochastische Asymptoten auch die bedingte und vice versa. Jede stochastische Asymptote von* x *ist dabei eine stochastische Asymptote von* $x_{[J]}$ *uvd umgekehrt.*

Die bedingte zufällige Variable, die dieselben stochastischen Asymptoten, wie die ursprüngliche hat, sei die *äquivalente zufällige Variable* genannt und, ebenso wie auch ihre Momente und Zentralwerte, mit demselben, nur fett gedruckten, Buchstaben bezeichnet. Nach dem Satz 19 besteht die hinreichende und notwendige Bedingung dieser Aequivalenz einfach darin, dass die Gesammtwahrscheinlichkeit aller ausgeschiedenen Werte gegen Null konvergieren muss. Sind

---

(1) Das Ersetzen der *vollen* mathematischen Erwartung durch eine *partielle* stammt, m. W., von G. Bohlmann her (*op. cit.*, p. 386-387), der die Grösse $E'_{[J]}\,x$ den Hauptwert nennt. A. A. Tschuprow schlägt dagegen vor die *bedingte* mathematische Erwartung als einen Ersatz zu gebrauchen (*Ueber die mathematische Erwartung des Quotienten.* « Forschungen der russischen Gelehrten im Auslande », Bd. I, Berlin 1922, p. 242-243. russisch).

$x_1, x_2, \ldots x_n$ die equivalenten zufälligen Variablen für $x_1, x_2, \ldots x_n$; $Q_1, Q_2, \ldots Q_n$ ihre ausgeschiedenen Wahrscheinlichkeiten; $y = F(x_1\ x_2, \ldots x_n)$ eine Funktion von $x_1, x_2, \ldots x_n$; und Q_y die ausgeschiedene Wahrscheinlichkeit für $y_{[J]} = F(x_1, x_2, \ldots x_n)$, so ist nach dem BOOLE'schen Satz (siehe oben § 19)

$$(10)\ldots \qquad Q_y \leqq \sum_{i=1}^{n} Q_i$$

und da jedes Q_i gegen Null konvergiert, so ist auch $lim\ Q_y = 0$. Die bedingte zufällige Variable $y_{[J]}$, die wir jetzt auch y (bzw. $y_{[J]}$) schreiben dürfen, ist also der zufälligen Variable y äquivalent und hat alle stochastischen Asymptoten mit ihr gemein.

So kommt man zu folgendem Ergebniss: Sind die Voraussetzungen des Satzes (18) erfüllt und sind die Streuungen von $x_1, x_2, \ldots x_n$ fest von oben begrenzt, so ist

$$lim_B F(x_1, x_2, \ldots x_n) = F(lim\ E\ x_1, lim\ E\ x_2, \ldots lim\ E\ x_n).$$

Ist noch $lim \sum_{i=1}^{n} Q_i = 0$ und hat die bedingte Streuung

$$\mu_{2[J]}^{(y)} = E_{[J]}(y - E_{[J]}\ y)^2$$

eine feste obere Grenze, so ist $E_{[J]}\ y = E\ y$ eine stochastische Asymptote von y, woher

$$(11)\ldots \qquad lim\ E\ y = lim_B(y)\ .$$

Man hat also

$$(12)\ldots\ lim\ E\ F(x_1, x_2, \ldots x_n) = F(lim\ E\ x_1, lim\ E\ x_2, \ldots lim\ E\ x_n),$$

wo links auch $E_{[J]}\ F(x_1, x_2, \ldots x_n)$ geschrieben werden kann, bzw., auch $E_{[c/_0]}\ F(x_1, x_2, \ldots x_n)$, wenn die ausgeschiedenen Stellen lediglich Unbestimmtheitsstellen sind. Statt $E_{[J]}\ F$ kann man auch die entsprechende partielle mathematische Erwartung $E'_{[J]}\ F$ darin einsetzen, da sie sich von der ersten Grösse nur durch den Faktor P unterscheidet, der bei gemachten Voraussetzungen gegen Eins konvergiert.

Fasst man nun alle Ergebnisse dieser Untersuchungen zusammen, so kommt man zu

Satz 20. Es seien $x_1, x_2, \ldots x_n$ zufällige Variablen die mit unabhängigen Variablen $\varphi_1, \varphi_2, \ldots \varphi_n$ stochastisch verbunden sind und stochastische Grenzwerte:

$$c_i = \underset{\varphi'i\ \varphi''i\ ,\ldots}{lim_B}(x_i)\quad , \quad (i = 1, 2, \ldots n)$$

haben. Es sei $y = F(x_1\ x_2, \ldots x_n)$ eine Funktion, die von φ_1,

$\varphi_2, \ldots \varphi_n$ *direkt unabhängig und in der Nähe von* $(c_1, c_2, \ldots c_n)$ *stetig ist.*

(I) **Bezeichnet man mit W_i eine unvermeidliche Asymptote von x_i und mit W_y eine solche von y, so hat man**

$$(13)\ldots \qquad \lim W_y = F(\lim W_1, \lim W_2, \ldots \lim W_n), \quad (\varphi'_i, \varphi''_i, \ldots; i = 1, 2, \ldots n)$$

(II) *Gibt es für alle zufälligen Variablen:* $x_1, x_2, \ldots x_n$ *und* y *die geraden Momente vom Grade* r *und haben diese letzteren feste obere Grenzen, so bestehen noch die Beziehungen*

$$(14)\ldots \qquad \lim C_h^{(y)} = F(\lim C_{k_1}^{(1)}, \lim C_{k_2}^{(2)}, \ldots \lim C_{k_n}^{(n)}), \quad (o < h \leqq r; 0 < k_i \leqq r; i = 1, 2, \ldots n),$$

wo alle k_i *und* h *sowohl dieselben, ebenso wie auch verschiedene Werte annehmen können.*

(III) *Es seien* $x_1, x_2, \ldots x_n$ *und*

$$y = F(x_1, x_2, \ldots x_n)$$

bedingte zufällige Variablen, die durch Ausscheidung von Unbestimmtheits, bzw., Unendlichkeitsstellen der Funktion

$$y = F(x_1, x_2, \ldots x_n)$$

entstehen. Konvergieren dann die Wahrscheinlichkeiten der ausgeschiedenen Stellen von $x_1, x_2, \ldots x_n$ *gegen 0, und genügen die geraden Momente vom Grade* r *der Variablen:* $x_1, x_2, \ldots x_n$ *und* y *den Erfordernissen von N II, so gilt die Gleichung*

$$(15)\ldots \qquad \lim C_h^{(y)} = F(\lim C_{k_1}^{(1)}, \lim C_{k_2}^{(2)}, \ldots \lim C_{k_n}^{(n)}), \quad (o < h \leqq r; 0 < k_i \leqq r; i = 1, 2, \ldots n),$$

wo $C_h^{(y)}$ *der Zentralwert* h *— ten Grades von* y*, d. h. der bedingte Zentralwert* h *— ten Grades von* y *ist. Er kann in (15) auch durch betreffenden partiellen Zentralwert ersetzt werden* (1).

(1) Der Satz 20 ist eine Umformung und eine partielle Verallgemeinerung des G. BOHLMANN'schen «ersten Hilfsatzes der mathematischen Statistik» (siehe *op. cit.*, p. 342-347, 362-370). Doch bestehen hier sehr wichtige Divergenzpunkte. (1) Der BOHLMANN'sche Satz bezieht sich lediglich auf mathematische Erwartungen, was natürlich der wichtigste spezielle Fall auch des Satzes 20 ist. (2) Die Grenzbeziehung, die in dem BOHLMANN'schen Satze behauptet wird, ist nicht ein Limes, sondern ein Aequivalenzverhältniss:

$$\lim \left| \frac{E\,F(x)}{F(Ex)} \right| = 1 .$$

25. — Ich füge nun einige Beispiele hinzu.

A) Es sei $x = \frac{m}{n}$ eine (relative) Häufigkeit in einer Reihe unabhängiger Versuche bei Ereigniszahl m und Versuchszahl n. Da hat man $E\,x = p$, wo mit p eine Durchschnittswahrscheinlichkeit beim einzelnen Versuch bezeichnet wird, und es sei vorausgesetzt, dass sie entweder konstant ist oder einen Grenzwert $p_0 = \lim_{n \to \infty} p$ hat. Für jede von n direkt unabhängige und überall stetige Funktion gilt dann

$$\lim_{n \to \infty} E\,F\left(\frac{m}{n}\right) = F(p_0)$$

und ähnlich für jede in der Nähe des Punktes p_0 stetige, wenn Unendlichkeits — und Unbestimmtheitsstellen, laut Vorschrift von n^0 *III* des Satzes (20), ausgeschieden werden können.

So hat man, z. B., für die Funktion $y = e^{A\frac{m}{n}}$, wo A eine konstante ist,

$$\lim_{n \to \infty} E\,e^{A\frac{m}{n}} = e^{A\,p_0}, \tag{16}$$

was beim Bernoullischen Schema ($p_0 = p$) sich auch direkt berechnen lässt. Für jedes k nämlich hat man

$$E\,e^{k\,m} = \sum_{m=0}^{n} \frac{n!}{m!\,(n-m)!}\,p^m\,q^{n-m}\,e^{k\,m} = (q + p\,e^k)^n$$

---

(3) Von der zufälligen Variable in BOHLMANN'S Satze wird nicht vorausgesetzt, dass sie einen stochastischen Grenzwert hat. So gelten BOHLMANN'S Sätze auch für Ereigniszahl, die mit keiner Grösse in stochastischer Asymptoten-sondern in einer stochastischen Aequivalenzbeziehung zu ihrer mathematischen Erwartung steht. Für diese zufällige Variable gilt unser Satz überhaupt nicht. Da in Anwendungen meistens solche Funktionen der Ereigniszahl vorkommen, wo diese letztere durch die Versuchszahl dividiert werden kann, so ist dieser Punkt praktisch keine wichtige Begrenzung des Geltungsbereichs unseres Satzes. Prinzipiell aber soll er stets im Auge behalten werden. (4) Im engsten Zusammenhange mit den vorigen Punkten (2) und (3) befindet sich der Umstand, dass BOHLMANN'S Satz nicht für alle Wahrscheinlichkeitsverteilungen und nicht für alle sogar überall stetige Funktionen gilt. Unser Satz dagegen gilt für alle Wahrscheinlichkeitsverteilungen, bei denen stochastische Grenzwerte bestehen, und bei gewissen Klassen von zufälligen Variablen (so z. B. bei Häufigkeiten) für alle in der Nähe des betreffenden Punktes stetige Funktionen. (5) Mit der veränderten Problemstellung hängt bei uns auch die grössere Durchsichtigkeit der Begründung am engsten zusammen.

Setzt man für k die Grösse $\frac{A}{n}$ ein, so wird

$$E\, e^{A\frac{m}{n}} = (q + p\, e^{\frac{A}{n}})^n = \left[1 + p\,(e^{\frac{A}{n}} - 1)\right]^n,$$

woher sich

$$(17)\ldots \qquad \lim_{n \to \infty} e^{A\frac{m}{n}} = e^{p \lim [n\,(e^{A/n} - 1)]} = e^{Ap}$$

ergibt (1).

B) Die Funktion $\frac{m}{n - m}$ kann in der Form $\frac{m/n}{1 - m/n}$ geschrieben werden. Bei $m = n$ hat sie eine Unendlichkeitsstelle, deren Wahrscheinlichkeit bei vorausgesetzten Unabhängigkeit der Versuche sich gleich $Q = p_1\, p_2 \ldots p_n$ ergibt. Ist $\lim\limits_{n \to \infty} \frac{p_1 + p_2 + \ldots + p_n}{n} = p_0$ von 1 verschieden, so ist $\lim\, Q = 0$. Sodann ist die Grösse

$$E[\infty] \frac{m/n}{1 - m/n} = \frac{1}{\sum\limits_{m=0}^{n-1} P_{m/n}} \sum_{m=0}^{n-1} P_{m/n} \frac{m/n}{1 - m/n},$$

wo $P_{m/n}$ die Wahrscheinlichkeit des m — maligen Eintretens des betreffenden Ereignisses in n Versuchen ist, eine stochastische Asymptote und folglich

$$(18)\ldots \quad \lim E[\infty] \frac{m/n}{1 - m/n} = \frac{p_0}{1 - p_0} = \lim{}_B \frac{n - m}{m}.$$

C) Es sei, dass an zwei stochastisch mit einander verbundenen Variablen x und y n Versuche vorgenommen werden. Man bezeichne mit $n_{i\,|}$, bzw., $n_{|\,j}$ die Zahl der Versuche, wo x den Wert x_i, bzw., y den Wert y_j angenommen haben, und mit $n_{i\,|\,j}$ die Zahl der Versuche, bei welchen sowohl x den Wert x_i, wie auch y den Wert y_j zeigten. Es seien die entsprechenden Wahrscheinlichkeiten mit $p_{i\,|}$, $p_{|\,j}$ und $p_{i\,|\,j}$ bezeichnet. Die Grösse

$$\varphi^2 = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{[p_{i\,|\,j} - p_{i\,|}\; p_{|\,j}]^2}{p_{i\,|}\; p_{|\,j}}$$

(1) Vergl. G. Bohlmann, *op. cit.*, p. 399.

ist the mean square contingency der Engländer und die Grösse

$$\varphi^2_a = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{\left[n_{i|j} - \frac{1}{n} n_{i|} \, n_{|j}\right]^2}{n_{i|} \, n_{|j}}$$

the approximate value of the mean square contingency (K. Pearson). Diese Grösse hat $k+l$ Unbestimmtheitsstellen, da bei $n_{i|}=0$ und bei $n_{|j}=0$ sowohl der Nenner, wie auch der Zähler gleich Null werden. Bezeichnet man $q_{i|}=1-p_{i|}$ und $q_{|j}=1-p_{|j}$, so wird die Wahrscheinlichkeit von $\{ n_{i|} = 0 \} \; Q_{i|} = q^n_{i|}$, und ähnlich

$$Q_{|j} = q^n_{|j} \quad .$$

Da alle q von 1 verschieden sind (im entgegengesetzten Falle würde sich unsere Tabelle auf eine mit wenigeren Reihen reduzieren), so hat man, der Formel (10) zufolge, für die Gesamtwahrscheinlichkeit aller auszuscheidenden Stellen (Q):

$$\lim_{n \to \infty} Q \leqq \lim_{n \to \infty} \left\{ \sum_{i=1}^{k} q^n_{i|} + \sum_{j=1}^{l} q^n_{|j} \right\} = o$$

An $\varphi_a{}^2$ lässt sich also n^0 *III* des Satzes 20 anwenden. Schreibt man es in der Form

$$\varphi^2_a = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{\left[\frac{n_{i|j}}{n} - \frac{n_{i|}}{n} \cdot \frac{n_{|j}}{n}\right]^2}{\frac{n_{i|}}{n} \cdot \frac{n_{|j}}{n}}$$

und nimmt man die Beziehungen

$$lim_B \left(\frac{n_{i|j}}{n}\right) = p_{i|j} \; ; \; lim_B \left(\frac{n_{i|}}{n}\right) = p_{i|} \; ; \; lim_B \left(\frac{n_{|j}}{n}\right) = p_{|j}$$

in Betracht, so findet man

(19) ... $$\lim_{n \to \infty} E_{[\%]} \varphi_a{}^2 = \varphi^2 \quad . \text{ (1)}$$

Damit wird auch bewiesen, dass die Grösse φ^2_a, obgleich sie keine mathematische Erwartung im eigentlichen Sinne hat, doch einen stochastischen Grenzwert besitzt und zwar

$$\lim_{B \; n \to \infty} \varphi^2_a = \lim_{n \to \infty} E_{[\%]} \varphi^2_a = \varphi^2 \; .$$

---

(1) Vergl. AL. A. TSCHUPROW, *Ueber die mathematische Erwartung des Quotienten* (russisch), loc. cit., p. 257-261.

E) Wollen wir noch den stochastischen Grenzwert eines *geometrischen Durchschnitts* auffinden. Es sei eine unbegrenzt fortsetzbare Folge unabhängiger Versuche gegeben, die der Reihe nach an zufälligen Variablen $x_1, x_2, \ldots x_n, \ldots$ vorgenommen werden, und es seien $x'_1, x'_2, \ldots x'_n, \ldots$ die zufälligen Werte, die dabei zu Stande kommen. Es sei vorausgesetzt, dass unsere zufälligen Variablen lediglich positive Werte annehmen können, die eine feste von 0 verschiede positive untere Grenze und eine feste obere Grenze haben. Weiter sei angenommen, dass

$$\frac{1}{n} \sum_{i=1}^{n} E \log x_i$$

einen Grenzwert hat. Es sei

$$y = \sqrt[n]{x'_1 \, x'_2 \ldots x'_n}$$

der geometrische Durchschnitt von $x'_1, x'_2, \ldots x'_n$.

Führt man die Bezeichnungen

$$\eta = \log y \quad \text{und} \quad \xi_i = \log x_i$$

ein, so ist

$$\eta = \frac{1}{n} \sum_{i=1}^{n} \xi'_i \,,$$

und wenn η einen stochastischen Grenzwert η_0 hat, so wird auch $y = e^{\eta}$, nach dem Satz 18, einen stochastischen Grenzwert haben, und zwar

$$(20) \ldots \qquad \lim_B (y) = e^{\eta_0}$$

Bezeichnet man die Streuung von ξ_i mit $\mu_2^{(i)}$, den Produktmoment von ξ_i und ξ_j mit $\mu_{1,1}^{(i,j)}$ und die Streuung von η mit $\mu_{2(n)}$, so hat man nach der uns schon bekannten Formel

$$\mu_{2(n)} = \frac{1}{n} \mu_{[2;\,n]} + \frac{n-1}{n} \mu_{[1,\,1;\,n]} \,.$$

Da bei der Unabhängigkeit der Versuche

$$\begin{aligned} \mu_{1,1}^{(i,j)} &= E\,(\log x_i - E \log x_i)\,(\log x_j - E \log x_j) \\ &= E\,(\log x_i - E \log x_i)\; E\,(\log x_j - E \log x_j) = 0 \end{aligned}$$

ist, so ist auch der arithmetische Durchschnitt aller $\mu_{1,1}^{(i,j)}$, d. h. die Grösse $\mu_{[1,\,1;\,n]} = 0$. Die Grösse $\mu_{[2;\,n]}$, d. h. der arithmetische Durchschnitt aller $\mu_2^{(i)}$, muss bei gemachten Voraussetzungen eine feste obere Grenze haben, und darum wird die Streuung von η, mit unbegrenzter Zunahme der Versuchszahl (n), gegen Null konvergieren. Nach dem Satz 1 hat also η stochastische Asymptoten, und da

$$E\,\eta = \frac{1}{n}\sum_{i=1}^{n} E\,log\,x_i$$

voraussetzungsgemäss einen Grenzwert hat, so ist

$$(21)\ldots \qquad \eta_0 = lim_B(\eta) = \lim_{n \to \infty} \left\{ \frac{1}{n}\sum_{i=1}^{n} E\,log\,x_i \right\}.$$

Die Grösse rechts kann in einer anderen Form dargestellt werden. Es seien mögliche Werte von x mit a_1, $a_2, \ldots a_s$ und ihre Wahrscheinlichkeiten mit p_1, $p_2, \ldots p_s$ bezeichnet. Dann hat man

$$E\,log\,x = \sum_{i=1}^{s} p_i \;\; log\,a_i = log\,(a_1^{p_1}\,a_2^{p_2}\ldots a_s^{p_s}) \;.$$

Die Grösse $a_1^{p_1}\,a_2^{p_2}\ldots a_s^{p_s}$ ist der geometrische Durchschnitt möglicher Werte mit betreffenden Wahrscheinlichkeiten als Gewichten. Nach Analogie mit der mathematischen Erwartung kann sie die *geometrische Erwartung* genannt werden. Ich bezeichne sie mit $\overline{E}\,x$. Es ist also

$$\overline{E}\,x = a_1^{p_1}\,a_2^{p_2}\ldots a_s^{p_s},$$

woher

$$(22)\ldots \qquad E\,log\,x = log\,\overline{E}\,x \text{ (1)}.$$

Sodann hat man

$$(27)\ldots. \qquad \frac{1}{n}\sum_{i=1}^{n} E\,log\,x_i = \frac{1}{n}\sum_{i=1}^{n} log\,\overline{E}x_i$$

$$= log\sqrt[n]{\overline{E}\,x_1\,\overline{E}\,x_2\ldots\overline{E}_n}$$

und

$$\eta_0 = log\left\{ \lim_{n \to \infty} \sqrt[n]{\overline{E}\,x_1\,\overline{E}\,x_2\ldots\overline{E}\,x_n} \right\},$$

---

(1) Aus dieser Formel hat man

$$(23)\ldots \qquad \overline{E}\,x = exp.\,(E\,log\,x) \quad ,$$

woher sich folgende Grundeigenschaften von $\overline{E}\,x$ sofort ergeben:

$$(24)\ldots. \qquad \overline{E}\,A\,x = A\,\overline{E}\,x$$

$$(25)\ldots. \qquad \overline{E}\,x^k = (\overline{E}\,x)^k$$

$$(26)\ldots. \qquad \overline{E}\,x\,y\ldots w = \overline{E}\,x\,\overline{E}\,y\ldots\overline{E}\,w \quad ,$$

wo k eine beliebige reele Zahl und x, $y, \ldots w$ beliebige zufällige Variablen sind, deren Unabhängigkeit dabei nicht vorausgesetzt wird.

woher endlich, der Formel (20) zufolge, sich die schliessliche Gleichung

$$(28)\ldots \quad \lim_{B\, n \rightarrowtail \infty} \sqrt[n]{x'_1\, x'_2 \ldots x'_n} = \lim_{n \rightarrowtail \infty} \sqrt[n]{\bar{E}\, x_1\, \bar{E}\, x_1 \ldots \bar{E}\, x_n}$$

ergibt. Nimmt man nun in Betracht die Formeln (25), und (26) so kann diese Gleichung auch in der Form

$$(29)\ldots \quad \lim_{B\, n \rightarrowtail \infty} \sqrt[n]{x'_1\, x'_2 \ldots x'_n} = \lim_{n \rightarrowtail \infty} \bar{E} \sqrt[n]{x_1\, x_2 \ldots x_n}$$

geschrieben werden. Sind die Wahrscheinlichkeitsverteilungen aller Variablen identisch, so sei

$$\bar{E}\, x_1 = \bar{E}\, x_2 = \ldots = \bar{E}\, x\ ,$$

woher

$$(30)\ldots \quad \lim_{B\, n \rightarrowtail \infty} \sqrt[n]{x'_1\, x'_2 \ldots x'_n} = \bar{E}\, x\ .$$

Für den geometrischen Durchschnitt einer *endlichen* Anzahl zufälliger Variablen hat man dagegen direkt nach den Sätzen 18 und 20:

$$(31)\ldots \quad \left.\begin{aligned} \lim_B \sqrt[s]{x_1\, x_2 \ldots x_s} &= \lim \sqrt[s]{E x_1\, E\, x_2 \ldots E\, x_s} \\ &= \lim E \sqrt[s]{x_1\, x_2 \ldots x_s}\ , \end{aligned}\right\}\begin{matrix} \varphi'_i\ , \varphi''_i\ ,\ldots \\ (i = 1, 2, \ldots s), \end{matrix}$$

wenn die betreffenden Voraussetzungen erfüllt sind.

26. — Die Untersuchung der sich an den Satz 18, bzw., 20 anknüpfenden Fragen geht evident über die Grenzen dieser Arbeit hinaus. Da bestehen zuförderst verschiedene Berührungspunkte mit den Problemen der mehrmals zitierten G. Bohlmann'schen Untersuchung, doch können in dieser Richtung laufende Fäden hier nicht weiter verfolgt werden. Andererseits gibt uns der gewonnene Standpunkt noch neue Möglichkeiten für die Aufsuchung der Bedingungen des Bestandes der stochastischen Asymptoten, wovon ich hier eine kleine Probe geben will, um einen Ausblick noch in dieser Richtung zu gewinnen.

Satz. 21. Es sei x *eine zufällige Variable, die mit einer unabhängigen Variable* φ *stochastisch verbunden ist. Es sei* $y = F(x)$ *eine den Bedingungen des Satzes (20) genügende Funktion, und es sei noch angenommen, dass* $E\, x$ *einen endlichen Grenzwert* $x_0 = \lim E\, x$ *hat, und die Streuungen* $\mu_2^{(x)}$, $\mu_2^{(y)}$ *fest von oben be-*

grenzt sind. Wird dabei die Bedingung

$$\lim_{\varphi', \varphi'' \ldots} E\, F(x) = F(x_0)$$

nicht erfüllt sein, so kann x keine stochastische Asymptoten haben.

In der Tat, wenn x stochastische Asymptoten hätte, so würden dann alle Voraussetzungen des Satzes 20 erfüllt sein, und so müsste auch die obenangegebene Grenzbeziehung gelten, was den Bedingungen des Satzes widerspräche. Es würde sehr leicht sein dieser Bedingung (die eine notwendige, doch aber keine hinreichende ist) eine allgemeinere Formulierung zu geben, doch dürfen wir es hier unterlassen.

Mit Hilfe des Satzes 20 lässt sich noch eine andere notwendige Bedingung begründen. Es sei eine Funktion $y = \Phi(x, H)$ durch die Bedingungen definiert, dass bei $|x| \leqq H$, sie gleich x, bei $|x| > H$, sie gleich Null ist, oder in Symbolen:

$$(32) \ldots \quad \Phi(x, H) = \begin{cases} x, & |x| \leqq H \\ o, & |x| > H \end{cases}$$

Da sie im Intervall $-H < x < H$ stetig ist, so lässt sich darauf der Satz 20 (n^0 *II*) anwenden, wenn nur $\lim_B(x)$ im demselben Intervall liegt.

Es sei x eine zufällige Variable, deren Streuung (μ_2) fest von oben begrenzt ist. Hat sie stochastische Asymptoten, so ist

$$E\, x = as_B(x) \quad .$$

Führt man noch eine zufällige Variable $Z = x - E\, x$ ein, so hat man

$$E\, Z = o, \ \lim_B(Z) = o, \ E(Z - E\, Z)^2 = \mu_2 \, .$$

Betrachten wir eine Funktion

$$y = \Phi(Z^2, H^2),$$

die in der Nähe des Punktes $Z = 0$ stetig ist und lediglich im Endlichen liegende mögliche Werte hat. Besitzt x stochastische Asymptoten, hat also Z einen stochastischen Grenzwert ($\lim_B Z = 0$), so muss nach Satz 20 die Gleichung

$$(33) \ldots \lim E\, \Phi(Z^2, H^2) = \Phi[(\lim E\, Z)^2, H^2] = \Phi(o, H^2) = o$$

gelten, was, nach der Definition von Φ, mit der Bedingung

$$(34) \ldots \quad \lim_{\varphi_1, \varphi_2 \ldots} \overset{(H)}{\underset{(o)}{E}} (x - E\, x)^2 = o$$

identisch ist. Ist sie nicht erfüllt, so hat x keine stochastischen Asymptoten.

So haben wir

Satz 22. Hat die Streuung einer zufälligen Variable (x) *eine feste obere Grenze, und konvergiert die partielle Streuung*

$$\overset{(H)}{\underset{(o)}{E}} \mid x - E\, x \mid^2,$$

wo H eine beliebige konstante positive Grösse ist, nicht gegen Null, so kann x keine stochastischen Asymptoten haben.

Bemerkt man, dass, wenn w eine unvermeidliche Asymptote ist, $lim_B\, (x - w) = 0$ und $\overset{(H)}{\underset{(o)}{E}}\, (x - w) = as_B\, (x - w)$ sein muss, so erhält man durch ein ähnliches Raisonement

$$(35)\ldots \qquad lim\, E\, \Phi\, (Z^k, H^k) = \Phi\, [(lim\, \overset{(H)}{\underset{(o)}{E}} Z)^k, H^k] = \\ = \Phi\, (o, H^k) = o\;,$$

wo $Z = x - w$ ist, woher

$$(36), \ldots \qquad \lim_{\varphi_1, \varphi_2, \ldots} \overset{(H)}{\underset{(o)}{E}} \mid x - w \mid^k = o\;,$$

wenn x überhaupt stochastisch konvergent ist. Ist (36) nicht erfüllt, so kann x keine stochastischen Asymptoten haben. Dieses Ergebniss formulieren wir wie folgt:

Satz 23. Konvergiert ein partieller gerader Moment

$$\overset{(H)}{\underset{(o)}{E}} \mid x - w \mid^k\;,$$

wo k und H beliebige positive Zahlen sind und w eine unvermeidliche Asymptote ist, nicht gegen Null, so hat x keine stochastischen Asymptoten.

Beide Sätze könnten auch mit der früheren Methode der Zerlegung in partielle mathematische Erwartungen abgeleitet werden, doch ist das im obigen angewandte Verfahren ein kürzeres. Vielleich wird es sich noch in einigen anderen Fällen bewähren können.

R. A. Fisher, M. A.
Rothamsted Experimental Station

Applications of "Student's„ Distribution.

1. - Introductory.

The Theory of Errors may be said to have taken its origin in the fact that the accuracy of the mean of a number of observations may be estimated from the discrepancies observed among the individual values used in obtaining the mean. In the simple theory appropriate to samples drawn from a normal population, of which the variance (mean squared deviation) is σ^2, it is easy to show that the mean of n' observations will be distributed normally with variance

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n'}$$

consequently, if σ^2 were known *a priori*, the sampling distribution of the mean would be fully known. To test any hypothesis respecting the mean of the population, as for example that the mean of the population was any assigned value m, we should merely need to calculate

$$t = (\bar{x} - m)\frac{\sqrt{n'}}{\sigma}$$

and the probability integral

$$P = \frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-\frac{1}{2}t^2}\,dt$$

would give, with any degree of accuracy required, the probability, on that hypothesis, that a greater discrepancy should occur than that actually observed. If the value of P so calculated turned out to be a small quantity such as 0·01, we should conclude with some confidence that the hypothesis was not in fact true of the population actually sampled.

2. - " Student's „ Distribution.

In the majority of cases in which such tests are required we have no *a priori* knowledge of the variance of the population, or indeed of whether its distribution is normal or not. The first point (which alone we shall consider in detail) was met by estimating the variance of the population from the sample itself; if $x_1, x_2, \ldots . x_{n'}$, be our observations, we may take

$$s^2 = \frac{S(x - \bar{x})^2}{n' - 1}$$

as an estimate of the unknown variance, σ^2.

s, so calculated, is a perfectly good estimate of σ, but it is seldom or never equal to σ, and, as was first pointed out by « Student » in his fundamental paper of 1908, (1), if in testing significance we substitute s for σ, and calculate

$$t = (x - m)\frac{\sqrt{n'}}{s}$$

we have no right to assume that t will still be distributed in the normal curve, or that the significance of an observation can be accurately tested by the normal probability integral. In order to obtain an accurate test « Student » investigated the distribution of s^2 in random samples; from the relation connecting the first four moments he inferred that it was probably of the Pearsonian Type III, and so obtained the exact distribution, which may be written

$$df = \frac{1}{\left(\frac{n'-3}{2}\right)!}\left(\frac{n'-1}{2\sigma^2}\right)^{\frac{n'-1}{2}} \cdot (s^2)^{\frac{n'-3}{2}} e^{-\frac{(n'-1)s^2}{2\sigma^2}} d(s^2)$$

$$= \frac{1}{\left(\frac{n-2}{2}\right)!}\left(\frac{n}{2\sigma^2}\right)^{\frac{1}{2}n} (s^2)^{\frac{n-2}{2}} e^{-\frac{ns^2}{2\sigma^2}} d(s^2)$$

where $n = n' - 1$.

Assuming that the distribution of s^2 is independent of that of $\bar{x}$ this expression may be used to deduce the exact distribution of

$$t = (\bar{x} - m)\frac{\sqrt{n'}}{s} \quad ;$$

for the distribution of $\bar{x}$ is given by

$$df = \frac{1}{\sqrt{2\pi}} e^{-\frac{n'(\bar{x}-m)^2}{2\sigma^2}} \cdot \frac{\sqrt{n'}}{\sigma} d\bar{x} \quad ,$$

or, for a given value of s by,

$$df = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2 s^2}{2\sigma^2}} \cdot \frac{s}{\sigma} dt \quad ;$$

so that for all values of s

$$df = \frac{1}{\left(\frac{n-2}{2}\right)! \sqrt{2\pi}} \left(\frac{n}{2\sigma^2}\right)^{\frac{1}{2}n} \cdot \frac{dt}{\sigma} \int_0^\infty (s^2)^{\frac{n-1}{2}} e^{-\frac{(n+t^2)s^2}{2\sigma^2}} d(s^2)$$

$$= \frac{\left(\frac{n-1}{2}\right)!}{\left(\frac{n-2}{2}\right)! \sqrt{\pi n}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt \, .$$

This is equivalent to the form given by « Student » in 1908 (1). The result obtained at that time was partly intuitive, since on two points the demonstration was incomplete; (i) the distribution obtained for s^2 agrees with the true distribution in the first four moments, but might conceivably have differed from it in the higher moments. (ii) « Student » demonstrated that s^2 was not correlated with $(\bar{x} - m)^2$, but did not show that the two distributions were entirely independent.

3. - Proof of the exactitude of " Student's „ Distribution for Normal Samples.

One method of proving these two points, which has been found to be valuable in other sampling problems, is to consider the observations $x_1, \ldots, x_{n'}$ as rectangular coordinates in Euclidian space of n' dimensions; the volume element in this space will be

$$dv = dx_1 \, dx_2 \, \ldots \, , \, dx_{n'} \, ,$$

and since the individual observations are independently distributed so that

$$df = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \; ,$$

the frequency element for the sample will be

$$df = \frac{1}{(\sigma \sqrt{2\pi})^{n'}} e^{-\frac{1}{2\sigma^2} S(x-m)^2} dv .$$

The density with which samples occur in any region is therefore proportional to

$$e^{-\frac{1}{2\sigma^2} S(x-m)^2} = e^{-\frac{n'}{2\sigma^2}(\bar{x}-m)^2} . e^{-\frac{(n'-1)s^2}{2\sigma^2}} .$$

To find the simultaneous distribution of $\bar{x}$ and s^2 it is necessary to express the element of volume dv in terms of these two statistics. Observing that $\bar{x}$ is proportional to the distance of the sample point from a fixed « plane » region,

$$S(x) = 0 \quad ,$$

and that s is proportional to the distance from the fixed line,

$$x_1 = x_2 = x_3 = \; . \; . \; . \; . \; . = x_{n'} \quad ,$$

it follows that the element of volume, dv, varies as

$$s^{n'-2} \; ds \; d\bar{x},$$

so that df varies as

$$e^{\frac{n'}{2\sigma^2}(\bar{x}-m)^2} d\bar{x} \; . \; s^{n'-2} \; e^{-\frac{(n'-1)s^2}{2\sigma^2}} ds .$$

Since this falls into two factors involving $\bar{x}$ and s respectively, the two distributions must be wholly independent, and completing each with its necessary constant factors, we have

$$df = \frac{1}{\sqrt{2\pi}} e^{-\frac{n'}{2\sigma^2}(\bar{x}-m)^2} \frac{\sqrt{n'}}{\sigma} d\bar{x}$$

$$\times \frac{1}{\left(\frac{n'-3}{2}\right)} \left(\frac{n'-1}{2\sigma^2}\right)^{\frac{n'-1}{2}} (s^2)^{\frac{n'-3}{2}} e^{-\frac{(n'-1)s^2}{2\sigma^2}} d(s^2) \quad ,$$

the simultaneous distribution from which the distribution of t has already been derived.

4. - Conditions for the wider application of "Student's" distribution.

Although, for the solution of the specific problem attacked by « Student », the simultaneous distribution of $\bar{x}$ and s^2 is all that is required, yet the fact that this distribution is compounded of two *independent* distributions, (i) that of $\frac{(\bar{x}-m)\sqrt{n'}}{\sigma}$, distributed normally about zero with unit standard deviation, and (ii) that of

$$\frac{(n'-1)\,s^2}{\sigma^2} \quad (=\chi^2),$$

in the distribution

$$df = \frac{1}{\frac{n'-3}{2}\,!}\left(\frac{\chi^2}{2}\right)^{\frac{n'-3}{2}} e^{-\frac{1}{2}\chi^2}\,d(\chi^2),$$

in such a way that

$$t = \frac{(\bar{x}-m)\sqrt{n'}}{\sigma} \div \sqrt{\frac{\chi^2}{n'-1}},$$

shows that « Student's » formula for the distribution of t is applicable to all cases which can be reduced to a comparison of the deviation of a normal variate, with an independently distributed estimate of its standard deviation, derived from the sums of squares of homogeneous normal deviations, either from the true mean of the distribution, or from the means of samples.

5. - Significance of differences between means.

This statistical situation occurs very frequently in connection with experimental work; and, consequently. « Student's » distribution affords the solution of a variety of problems beyond that for which it was originally prepared. Of these, one that appears continually under one form or another is the comparison of two mean values. If $\bar{x}_1$ and $\bar{x}_2$ are the means of two samples of n_1 and n_2 values respectively, and we wish to test if the two means are sufficiently alike to warrant the belief that the samples are drawn from the same population, or, on the other hand if the means are significantly different, we may suppose the hypothetical population to have a standard deviation, σ. Then

$$\frac{\bar{x}_1-\bar{x}_2}{\sigma}\;\cdot\;\sqrt{\frac{n_1 n_2}{n_1+n_2}}$$

will be normally distributed with unit standard deviation; further,

$$\frac{\overset{n_1}{\underset{1}{S}}(x_1 - \bar{x}_1)^2 + \overset{n_2}{\underset{1}{S}}(x_2 - \bar{x}_2)^2}{\sigma^2} = \frac{S_1 + S_2}{\sigma^2}$$

will be distributed in the χ^2 distribution for $n = n_1 + n_2 - 2$, (or $n' = n_1 + n_2 - 1$); moreover, these two distributions will be wholly independent. Consequently, if we write

$$t = \frac{(x_1 - x_2)\sqrt{n_1 + n_2 - 2}}{\sqrt{S_1 + S_2}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

then t will be distributed in « Student's » distribution, specified by

$$n = n_1 + n_2 - 2 \ .$$

Example 1.

As an example of the application of this method in experimental work we may take a portion of the data of an electro-culture experiment carried out at Rothamsted in 1922. Eight pots growing three barley plants each, were exposed to the action of a high tension discharge, while nine similar pots were enclosed in an earthed wire cage. The numbers of tillers in each pot were as follows: —

Electrified	16, 16, 20, 16, 20, 17, 15, 21	mean $17.625 = \bar{x}_1$
Caged	17, 27, 18, 25, 27, 29, 27, 23, 17	mean $23.333 = \bar{x}_2$

The difference between the means is therefore 5.708; also

$$S_1 = 37.875, \ S_2 = 184, \ S_1 + S_2 = 221.875$$

multiplying $(S_1 + S_2)$ by 17 and dividing successively by 15, 8 and 9, we have as the estimated variance of the difference between the means, 3.4925; and for the estimated standard deviation, 1.8688. The value of t, which is the ratio of the difference to its estimated standard deviation is therefore 3.054. For $n = 15$, Table I shows that this value will be exceeded by chance about 41 times in 10,000. In other words a difference, positive or negative, greater than that observed will occur by chance only about 8 times in a thousand trials. The difference must therefore be judged significant. The two series are definitely unlike in their tillering; the possibility, however, that unlikeness in the variability as well as unlikeness in the mean, has contributed to the result is not excluded by this test. The possibility that samples from populations alike in their mean, but unlike in their variability should give significant values of t is scarcely of importance in relation to the practical use of the test in experimental work.

In cases where the two samples are equal in number, and in which each individual of one sample corresponds in some way to a particular individual of the second sample, we may test the significance of the difference of the means, either by the above method, or by the method as originally set forth by « Student ». In the latter case the differences between corresponding values are considered as a single sample, and the test shows if their mean differs significantly from zero. When both methods are available, sometimes the one and sometimes the other is the more sensitive; if either shows a significant deviation its testimony cannot be ignored. If, as frequently happens in experimental work the corresponding values of the two samples are positively correlated, the standard deviation of the differences will be reduced by this circumstance; against this advantage we must set off the fact that in treating the results as a single sample, the value of n is only half as great as if the two samples had been treated separately. The results of applying both tests to the same data supply a direct statistical measure of the efficacy of the system of « controls » which has been utilised.

6. - Significance of regression coefficients.

The second class of tests for which « Student's » distribution provides an exact solution, lies in testing the significance of the large class of statistics known as regression coefficients; and also in testing the significance of differences between regression coefficients obtained in different samples.

Consider a simple linear regression formula

$$Y = a + b\,(x - \bar{x})$$

in which the coefficients a and b have been calculated by the equations

$$a = y \qquad\qquad b = \frac{S\,\{\,y\,(x - \bar{x})\,\}}{S\,(x - \bar{x})^2}$$

if, for a given value of x, y is distributed normally with variance σ^2 then *confining attention to samples having the same values of* x, it is evident that b will be distributed normally with variance

$$\sigma^2{}_b = \frac{\sigma^2}{S\,(x - \bar{x})^2}$$

moreover, the mean of this distribution will be the value of the regression, say, β, obtained from an infinitely large sample. In other words

$$\frac{(b-\beta)\sqrt{S(x-\bar{x})^2}}{\sigma}$$

will be normally distributed about zero with unit standard deviation.

This expression involves the unknown parameter, σ, and it is by substituting for σ an estimate of its value, s, derived from the observations, that the distribution is changed from the normal form to that of « Student's » distribution. The estimate which we shall consider is,

$$s^2 = \frac{1}{n'-2} S(y-Y)^2$$

where n' is the number of observations. We shall prove that $\frac{1}{\sigma^2} S(y-Y)^2$ is distributed as is the sum of the squares of $n'-2$ quantities distributed independently and normally with unit standard deviation. It is perhaps worth while to give, at length, an algebraical method of proof, since analogous cases have hitherto been demonstrated only geometrically, by means of a construction in Euclidian hyperspace, and the validity of such methods of proof may not be universally admitted.

If $x_1, x_2, \ldots\ldots x_{n'}$ be distributed normally and independently with unit standard deviation and if,

$$\begin{aligned}
\zeta_1 &= p_{11}\, x_1 + p_{12}\, x_2 + \ldots\ldots\ldots\ldots\ldots p_{1n'}\, x_{n'} \\
\zeta_2 &= p_{21}\, x_1 + p_{22}\, x_2 + \ldots\ldots\ldots\ldots\ldots p_{2n'}\, x_{n'} \\
\zeta_{n'} &= p_{n'1}\, x_1 + p_{n'2}\, x_2 + \ldots\ldots\ldots\ldots\ldots p_{n'n'}\, x_{n'}
\end{aligned}$$

then $\zeta_1, \zeta_2, \ldots\ldots \zeta_{n'}$, will be distributed normally and independently with unit standard deviation, provided for all values of i

$$p^2_{i1} + p^2_{i2} + \ldots\ldots\ldots p^2_{in'} = 1$$

and for all unequal values of i and j

$$p_{i1}\, p_{j1} + p_{i2}\, p_{j2} + \ldots\ldots + p_{in'}\, p_{jn'} = 0$$

Now, if $\zeta_1, \zeta_2, \ldots\ldots \zeta_h$ be any h linear functions of $x_1, x_2 \ldots\ldots x_{n'}$ fulfilling these conditions, the set of h homogeneous equations,

$$\begin{aligned}
p_{11}\, p_{i1} + p_{12}\, p_{i2} + \ldots\ldots\ldots\ldots\ldots p_{1n'}\, p_{in'} &= 0, \\
p_{h1}\, p_{i1} + p_{h2}\, p_{i2} + \ldots\ldots\ldots\ldots\ldots p_{nh'}\, p_{in'} &= 0
\end{aligned}$$

involving the n' unknowns $p_{i_1}, p_{i_2} \ldots\ldots p_{in'}$, can always be solved if h does not exceed $n' - 1$, and every such solution will yield a solution of

$$p^2_{i_1} + p^2_{i_2} + \ldots\ldots p^2_{in'} = 1\ ;$$

consequently we can always find in succession, variates

$$\zeta_{h+1},\ \zeta_{h+2}\ \ldots\ldots\ \zeta_{n'}$$

fulfilling the required conditions.

From this it follows that

$$\overset{n'}{\underset{1}{S}}(x^2) - \overset{h}{\underset{1}{S}}(\zeta^2)$$

can always be expressed in the form

$$\overset{n'}{\underset{h+1}{S}}(\zeta^2)$$

and is therefore distributed as is the sum of the squares of $n' - h$ quantities normally and independently distributed with unit standard deviation. Moreover this distribution will be wholly independent of $\zeta_1, \ldots\ldots \zeta_h$

Now if $$\alpha + \beta(x - \bar{x})$$

is the regression line of the population sampled, the quantities

$$\frac{1}{\sigma}\left\{ y - \alpha - \beta(x - \bar{x}) \right\}$$

are normally and independently distributed with unit standard deviation; and, since a and b have been chosen to make

$$\frac{1}{\sigma^2} S\left\{ y - a - b(x - \bar{x}) \right\}^2$$

a minimum, it may be reduced to the form

$$\frac{1}{\sigma^2} S\left\{ y - \alpha - \beta(x - \bar{x}) \right\}^2 - \frac{n'}{\sigma^2}(a - \alpha)^2 - (b - \beta)^2 \frac{S(x - \bar{x})^2}{\sigma^2}$$

and will, consequently, be distributed as is the sum of the squares of $(n' - 2)$ quantities, each distributed independently and normally with unit standard deviation, provided that

$$\frac{(a - \alpha)\sqrt{n'}}{\sigma} \text{ and } \frac{(b - \beta)}{\sigma}\sqrt{S(x - \bar{x})^2}$$

are so distributed. Now

$$\frac{(a-\alpha)\sqrt{n'}}{\sigma} = S\left\{\frac{1}{\sqrt{n'}} \cdot \frac{y-\alpha-\beta(x-\bar{x})}{\sigma}\right\}$$

and is of the form required for ζ_1 having

$$p_{11} = p_{12} \quad \ldots\ldots \quad = p_{1n'} = \frac{1}{\sqrt{n'}}$$

satisfying the equation

$$p^2_{11} + p^2_{12} + \quad \ldots\ldots \quad + p^2_{1n'} = 1 \quad ;$$

also

$$\frac{b-\beta}{\sigma} S(x-x)^2 = S\left\{\frac{x-\bar{x}}{S(x-x)^2} \quad . \quad \frac{y-\alpha-\beta(x-\bar{x})}{\sigma}\right\} \quad ,$$

which is of the form required for ζ_2, having

$$p_{2i} = \frac{x_i - \bar{x}}{\sqrt{S(x-\bar{x})^2}} \quad ,$$

satisfying both the equations

$$p^2_{21} + p^2_{22} + \quad \ldots\ldots \quad p^2_{2n'} = 1$$

and

$$p_{11}\,p_{21} + p_{12}\,p_{22} + \ldots\ldots + p_{1n'}\,p_{2n'} = S\left\{\frac{1}{n'} \quad \frac{x_i - \bar{x}}{\sqrt{S(x-\bar{x})^2}}\right\} = 0$$

Consequently

$$\frac{1}{\sigma^2} S(y-Y)^2$$

must be distributed in random samples as is the sum of $(n' - 2)$ quantities, each distributed independently and normally with unit standard deviation; and this distribution will be wholly independent of that of a and b.

Substituting for σ its estimated value s, we see that

$$t = \frac{(b-\beta)\sqrt{n'-2}\,\sqrt{S(x-\bar{x})^2}}{\sqrt{(y-Y)^2}}$$

will be distributed in « Student's » distribution for $n = n' - 2$. The quantity t involves no hypothetical quantities, being calculable wholly from the observations. It is the point of the method, as of « Student's » original treatment of the probable error of the mean, to obtain a quantity of known distribution expressible in terms of the observations only. If we had found the distribution of b for samples varying in the values of x observed, we should have been obliged to express the distribution in terms of the unknown standard deviation σ_x in the population sampled; moreover, since σ_x is unknown, we should have been obliged to substitute for it, an estimate based upon $S\,(x - \bar{x})^2$; the inexactitude of the estimate would have vitiated our solution, and required us to make allowance for the sampling variation of $S\,(x - \bar{x})^2$; finally, this process, when allowance had been accurately made would lead us back to the « Student's » distribution found above. The proof given above has, however, the advantage that it is valid whatever may be the distribution of x, provided that y is normally and equally variable in each array, and that regression of y on x is linear in the population sampled.

7. Non-linear Regression.

The same distribution is adequate for the coefficients of a non linear regression formula, still provided that y is normally and equally variable in each array. For suppose the regression equation is of the form

$$Y = a_0 + a_1 X_1, + a_2 X_2 + \ldots\ldots + a_p X_p$$

where $X_o\,(= 1)$, X_1, X_2, ... X_p are orthogonal functions of x, so that for all unequal values of i and j

$$S\,(X_i\, X_j) = 0$$

the summation being taken over the observed values of x. In the most important case X_i will be a polynomial in x of degree i. In any case the coefficient, a_i, will be estimated by means of the equation

$$a_i = \frac{S\,(y\, X_i)}{S\,(X_i^2)}$$

and

$$\sigma^2{}_{ai} = \frac{\sigma^2}{S\,(X^2{}_i)}$$

consequently, by an easy extension of the reasoning used above,

$$t = \frac{(a_i - \alpha_i)\sqrt{n' - p - 1}\ \sqrt{S\,(X^2{}_i)}}{\sqrt{S\,(y - Y)^2}}$$

will be distributed in « Student's » distribution, for $n = n' - p - 1$.

In practice it is quickest to calculate $S(y - Y)^2$ from the relation

$$S(y - Y)^2 = S(y^2) - n' a^2_0 - a^2_1 S(X^2_1) - \ldots\ldots. a^2_p S(X^2_p).$$

Alternatively this relation may be used to check the values of a_i obtained.

Example 2.

The difference in yield of total grain between the two dunged plots of Broadbalk field for the years 1901-1923 are as follows. The yield of plot 2B exceeded that of plot 2A by

	lb. per acre.		lb. per acre.		lb. per acre.		lb. per acre.		lb. per acre.
1901	124	1906	133	1911	207	1916	118	1921	106
1902	340	1907	315	1912	196	1917	221	1922	185
1903	146	1908	175	1913	211	1918	272	1923	142
1904	36	1909	321	1914	— 8	1919	410		

Fitting a straight line to these differences, of the form

$$Y = a_0 + a_1 x,$$

where x is the date measured in years from the central year, 1912, we find $a_0 = 194$, $a_1 = 1.247$. Are these significant?

We have $S(y - Y)^2 = 259{,}872$, $S(x^2) = 1012$ so that for the significance of the mean

$$t = \frac{194 \sqrt{21}.\sqrt{23}}{\sqrt{259{,}872}} = 8.36\ ;$$

and for the significance of the linear rate of increase

$$t = \frac{1.247 \sqrt{21}\ \sqrt{1012}}{\sqrt{259{,}872}} = .36$$

For each test $n = 21$. It is evident at once that whereas the difference in yield of the plots is clearly significant, there is no sign, during the present century, that the difference has been increasing or decreasing. In view of the latter fact the significance of the mean yields could have been tested satisfactorily without fitting the regression line.

8. Multiple Regression.

The same distribution is equally applicable to the coefficients of a multiple regression surface, plane or curved. In general the surface may be represented by

$$Y - \bar{y} = b_1 x_1 + b_2 x_2 \,.\,.\,.\,.\,.\, + b_p x_p$$

where $x_1, \ldots, x_p$ are either variates, or functions of variates, measured from their means, in terms of which the regression is expressed. Since $x_1, \ldots, x_p$ may now be correlated, we shall require the determinant

$$\Delta = \begin{vmatrix} S(x_1^2) & S(x_1 x_2). & \ldots & S(x_1 x_p) \\ S(x_1 x_2) & S(x_2^2) & \ldots & S(x_2 x)_p \\ \ldots & \ldots & \ldots & \ldots \\ S(x_1 x_p) & S(x_2 x_p) & \ldots & S(x_p{}^2) \end{vmatrix}$$

where the summation is taken over the observed n' values.

Then

$$\sigma^2{}_{b_i} = \frac{\sigma^2 \Delta_{ii}}{\Delta}$$

where Δ_{ii} is the cofactor of $S\,(x_i{}^2)$.

Consequently if

$$t_i = \frac{(b_i - \beta_i)\sqrt{n' - p - 1}\,\sqrt{\Delta}}{\sqrt{S\,(y - Y)^2}\,\sqrt{\Delta_{ii}}}$$

where β_i is the population value corresponding to the observed b_i, then t_i will be distributed in « Student's » distribution for $n = n' - p - 1$.

9. Distributions related to "Student's", as that of χ^2 is to the normal curve.

Finally the probability integral with which we are concerned is of value in calculating the probability integral of a wider class of distributions which is related to « Student's » distribution in the same manner as that of χ is related to the normal distribution. This wider class of distributions appears (i) in the study of intraclass correlations (2) (ii) in the comparison of estimates of the variance, or of the standard deviation from normal samples (3, p. 142) (iii) in testing the goodness of fit of regression lines (4) (iv) in testing the significance of a multiple correlation (5), or (v) of a correlation ratio (3, p. 218).

For example, the distribution in random samples of a multiple correlation, R, obtained by correlating n_1 independent variates with a dependent variate, having no real correlation with them, is

$$df = \frac{\frac{n_1 + n_2 - 2}{2}!}{\frac{n_2 - 2}{2}!\,\frac{n_1 - 2}{2}!} (R^2)^{\frac{n_1 - 2}{2}} (1 - R^2)^{\frac{n_2 - 2}{2}} d(R^2)$$

where $n_1 + n_2 + 1$ stands for the number of the sample. If n_1 is even, the probability that R should exceed any specified value is the partial sum of a binomial expansion,

$$P = (1 - R^2)^{\frac{1}{2} n_2} \Big\{ 1 + \frac{n_2}{2} R^2 + \frac{n_2 (n_2 + 2)}{2.4} R^4 \dots\dots$$
$$\dots \frac{n_2 (n_2 + 2) \dots (n_2 + n_1 - 4)}{2 . 4 \dots (n_1 - 2)} R^{n_1 - 2} \Big\},$$

whereas when n_1 is odd

$$P = \frac{2}{\sqrt{\pi}} \frac{\frac{n_2 - 1}{2}!}{\sqrt{\frac{n_2 - 2}{2}}} \int_{\sqrt{\frac{n_2 R^2}{1 - R^2}}}^{\infty} \left(1 + \frac{t^2}{n_2}\right)^{-\frac{1}{2}(n_2 + 1)} dt$$

$$+ \frac{2}{\pi} \cdot \frac{\frac{n_2 - 1}{2}!}{\frac{n_2 - 2}{2}!} (1 - R^2)^{\frac{1}{2} n_2} \Big\{ R + \frac{n_2 + 1}{3} R^3 + \frac{(n_2 + 1)(n_2 + 3)}{3 . 5} R^5 + \dots$$
$$\dots \frac{(n_2 + 1) \dots (n_2 + n_1 - 4)}{3 . 5 \dots (n_1 - 2)} R^{n_1 - 2} \Big\};$$

the analogy of these expressions with those given by PEARSON (6) for the χ^2 distribution is obvious. They become identical when

$$n_2 \to \infty, \ n_1 = n' - 1.$$

In the second form it will be noticed that the probability integral of the normal curve has been replaced by an integral of « Student's » distribution, of which the approximate value may be obtained from the tables. The multiple correlation must be judged significant only if the value of P obtained is too small to allow us to admit the hypothesis that the dependent variate is really uncorrelated with the independent variates.

REFERENCES.

1. STUDENT, (1908). *The probable error of a mean.* « Biometrika », vol. VI, pp. 1-25.
2. R. A. FISHER, (1921). *On the « Probable Error » of a coefficient of correlation deduced from a small sample.* « Metron », vol. I, Pt. 4, pp. 1-32.
3. R. A. FISHER, (1925). *Statistical Methods for Research Workers.* Oliver & Boyd, Edinburgh.
4. R. A. FISHER, (1922). *The goodness of fit of regression formulae, and the distribution of regression coefficients* « Journal of the Royal Statistical Society », vol. LXXXV, pp. 597-612.
5. R. A. FISHER, (1924). *The influence of rainfall on the yield of wheat at Rothamsted.* « Phil. Trans. », vol. CCXIII. pp. 89-142.
6. K. PEARSON, (1900). *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling.* « Phil. Mag. », Series V, L, pp. 157-175.

" STUDENT ,,

New Tables for Testing the Significance of Observations.

In « Biometrika » VI. pp. 1-25 it was suggested if $z = \frac{\bar{x}}{s}$ where $\bar{x}$ is the distance of the mean of a sample of n from the true mean of a normally distributed population, and s is the standard deviation of the same sample, *i. e.*,

$$\sqrt{\frac{S\,(x-\bar{x})^2}{n}}$$

then the frequency of z is given by the frequency curve

$$y = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\,(1+z^2)^{-\frac{n}{2}}$$

and that consequently the integral

$$p = \frac{1}{2} + \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\int_0^z (1+z^2)^{-\frac{n}{2}}\,dz$$

gives the probability that the mean of a sample of n drawn from a normally distributed population, measured in terms of the standard deviation of the sample shall exceed the value z.

Tables were constructed for values of n from 4 *to* 10, and subsequently, « Biometrika » XI. p. 416, from 2 *to* 30.

It has since been shown as in the preceding paper by Mr. FISHER that the suggestion was in fact justified, and that the integral has a much wider application than was originally supposed.

The tables hitherto published suffer however from two defects: (1) that as n increases the z scale becomes very coarse, and (*ii*) that except in the case for which it was designed, n the number in the sample, is not the best number under which to enter the table, but $n - 1$, the number of degrees of freedom.

The present Tables have, therefore, at Mr. FISHER's suggestion been constructed with argument $t = z\sqrt{n}$ where n is now one less than the number in the sample, which we may call n'. They correspond to SHEPPARD's table, when that is used to test the significance of the mean of a large number of observations.

Table I. extends from $t = 0$ to $t = 6$, at intervals of .1, from $n = 1$ to $n = 20$, inclusive; in each column in which values of more than ·99995 occur, the first of these is written 1.0000, and further values are not given.

Table II. gives values beyond $t = 6$, to six places of decimals, from which values accurate to four places of decimals can be calculated by proportional interpolation. The intervals are, therefore, unequal, and increase as t becomes larger. In this table no values are given under $n = 1$ and $n = 2$, as these can be easily calculated from the ordinary trigonometrical tables by the formulae

$$n = 1 \qquad p = \frac{1}{2} + \frac{\theta}{\pi} \text{ (where } \tan\theta = t\text{)}$$

$$n = 2 \qquad p = \frac{1}{2} + \frac{1}{2}\sin\theta \left(\text{where } \tan\theta = \frac{t}{\sqrt{2}}\right)$$

Table III. gives coefficients for calculating the difference between the value for $n = \infty$ *i. e.*, SHEPPARD's table, and that for n, and p is arrived at by the formula

$$p = p_{\infty} - \frac{C_1}{n} - \frac{C_2}{n^2} - \frac{C_3}{n^3} - \frac{C_4}{n^4}$$

This gives values of p, estimated to be accurate to .000005, when n is greater than 20, and, in fact at 20 and 24 the following differences were found:

Values of t	.5	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
Differences $n = 20$	0	0	0	0	+23	+17	−33	−46	−30	+41	+19	+9	+4
Differences $n = 24$					+8	+9	−14	−18	−4	+20	+7		

The above differences are in the seventh place of decimals, and are between values of p given by the approximation and those derived from the cosine formula using seven-place tables. Mr. FISHER's note explains the basis on which the coefficients were calculated.

The methods of calculating and checking the tables were as follows: —

1. Values of p for $t = 5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0,$

and $n = 1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 15, 20, 24,$

were calculated from the cosine formula (« Biometrika » VI. p. 10), using seven-figure tables; these values, though they are the sum of $n/2$ terms, appear to be accurate within about .0000003, and were checked by recalculation. They were also compared with the values obtained by the use of Table III. which both served as a further check, and also to show within what limits Table III. could be used for constructional purposes.

2. From the values thus calculated under $n = 6, 8, 12, 24$, together with $n = \infty$ (*i. e.,* Sheppard), the remaining frame values under $n = 7, 9, 11, 13, 14, 15, 16, 17, 18, 19$, were interpolated by coefficients calculated by Mr. Fisher for asymptotic interpolation, These were checked by recalculation and cross-differencing, *i. e.*, by comparing the difference $p_n - p_{n-1}$ with $p_{n+1} - p_n$ for the same values of t, and any doubtful values were recalculated by the cosine formula, as also were any values in which the fourth place of decimals was doubtful, *i. e.,* whenever the fifth place of decimals was 4 or 5.

3. Having thus obtained a frame, this was filled in to five places of decimals in three ways: (*a*) by interpolation, using where necessary both four and six-point central interpolation. It was found that over the greater part of table the true values lie between the four-point and the six-point interpolation, but for high values of t it was usually sufficient to use four-point, six-point being required only to locate doubtful values. (*b*). For very low values of n ($= 1, 2$, and 3) the frame was found not to be sufficiently close with low values of t, and alternate values had to be calculated from the cosine formula, the remaining odd values being interpolated by four-or six-point central interpolation. (*c*). As n increased it was found possible to make more and more use of Table III beginning at $n = 4$ with values of t less than 1 and ending at $n = 20$ with the whole table These values were recalculated as a check. Second differences were then taken down the columns and any doubtful figures checked from the cosine formula; as before this was done whenever the fourth figure was in any doubt.

Finally, the whole table was cross-differenced, and a very large number of values were recalculated from the cosine formula. Very few alterations were, however, found to be necessary.

Table II. was altogether calculated from the cosine formula; as it is designed to give an accuracy of four figures by proportional interpolation, it was possible to increase the interval between the t entries as t increases.

Table III was calculated from Mr. FISHER 's formulae, and I have to thank Miss W. A MACKENZIE, M. Sc. of the Rothamsted Statistical Laboratory for kindly checking this part of the work.

R. A. FISHER

Expansion of "Student's" integral in powers of n^{-1}.

The exact distribution of z found by « Student » in 1908, may be written

$$df = \frac{\frac{n-1}{2}\,!}{\frac{n-2}{2}\,!\,\sqrt{\pi n}}\left(1+\frac{t^2}{n}\right)^{-\frac{1}{2}(n+1)} dt \qquad \text{I}$$

Clearly, as n tends to ∞, df will tend to its normal value

$$\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}t^2}\, dt, \qquad \text{II}$$

and the approach to normality may be expressed by expanding (I) in inverse powers of n.

Now

$$-\log\left(1+\frac{t^2}{n}\right) = \frac{(-t^2)}{n} + \,\cdot\,\cdot\,\cdot\,\cdot\,\cdot\,\cdot\,\cdot\,\cdot + \frac{(-t^2)^p}{pn^p} + \cdot\,\cdot\,\cdot\,\cdot$$

hence

$$-\frac{1}{2}\left(n-1\right)\log\left(1+\frac{t^2}{n}\right) = -\frac{1}{2}t^2 + \cdot\,\cdot\,\cdot\,\cdot\,\cdot\,\cdot\,\cdot\,\cdot$$

$$\cdot\,\cdot\,\cdot\,\cdot\,\cdot + \frac{p\,(-t^2)^{p+1} + (p+1)(-t^2)^p}{2\,p\,(p+1)\,n^p} + \cdot\,\cdot\,\cdot\,\cdot\,,$$

moreover

$$\log\left(\frac{\frac{n-1}{2}\,!}{\frac{n-2}{2}\,!}\cdot\sqrt{\frac{2}{n}}\right) = -\frac{1}{4n} + \frac{1}{24n^3} - \frac{1}{20n^5} + \cdot\,\cdot\,\cdot\,\cdot\,\cdot\,\cdot\,,$$

consequently, the logarithm of the ratio which the exact formula (I) bears to its limiting form (II) may be written as,

$$\frac{t^4-2t^2-1}{4n}-\frac{2t^6-3t^4}{12n^2}+\frac{3t^8-4t^6+1}{24n^3}-\frac{4t^{10}-5t^8}{40n^4}+\frac{5t^{12}-6t^{10}-3}{60n^5}-\ldots\ldots \qquad \text{III}$$

The exponential expansion of this series will give the expansion of the actual differences of the ordinates, in powers of n^{-1}. Each term contains a polynomial in t^2, and the factor z, standing for the differential coefficient in II.

The first five terms are

$$\left.\begin{array}{l}
(t^4-2t^2-1)\,\dfrac{z}{4n}\\[2mm]
(3t^8-28t^6+30t^4+12t^2+3)\,\dfrac{z}{96n^2}\\[2mm]
(t^{12}-22t^{10}+113t^8-92t^6-33t^4-6t^2+15)\,\dfrac{z}{384n^3}\\[2mm]
(15t^{16}-600t^{14}+7100t^{12}-26616t^{10}+18330t^8+6360t^6+1980t^4-1800t^2-945)\,\dfrac{z}{92160n^4}\\[2mm]
(3t^{20}+190t^{18}+4025t^{16}-38976t^{14}+103702t^{12}-63444t^{10}-21270t^8-7800t^6+4455t^4+1890t^2-17955)\,\dfrac{z}{368640n^5}
\end{array}\right\}\ \text{IV}$$

These polynomials are of some interest in themselves, but for the present we require the integral of each term from t to ∞; these are as follows,

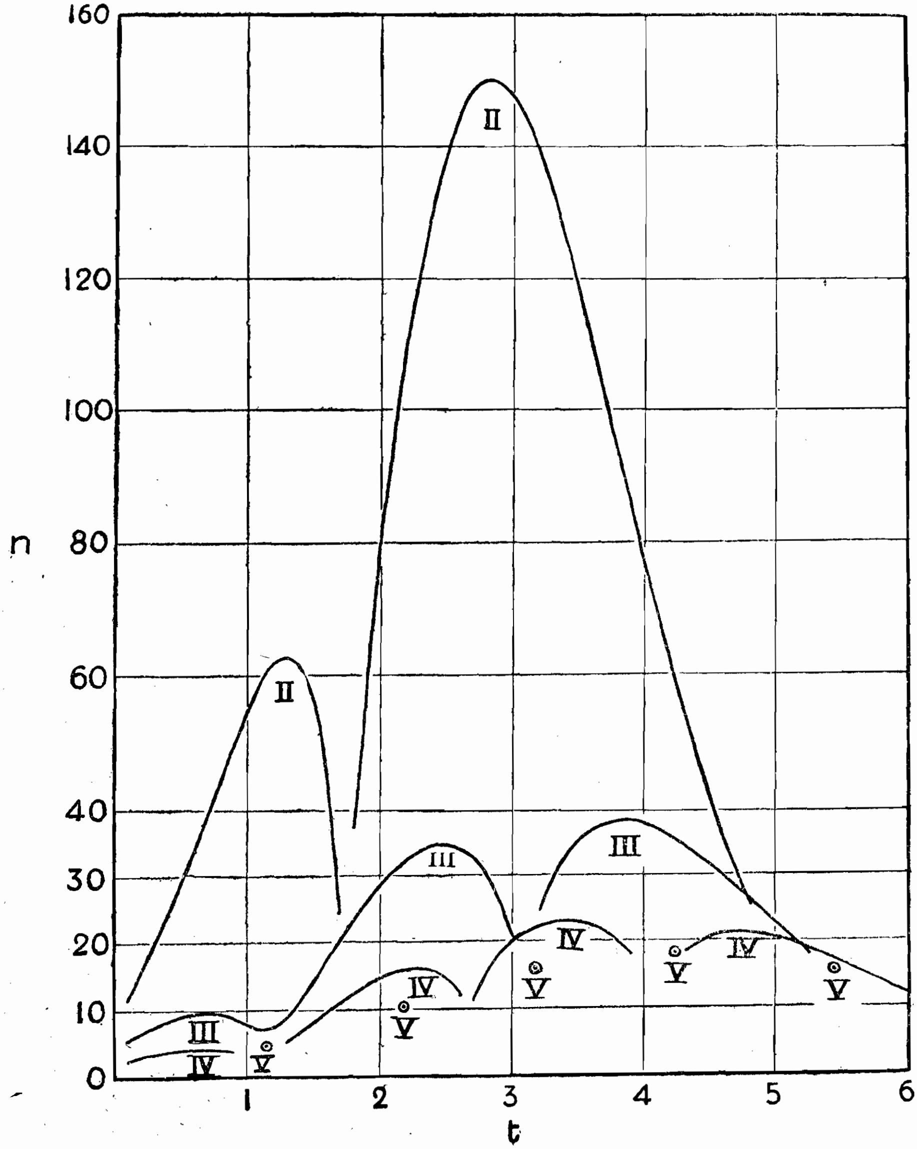
$$\left.\begin{array}{l}
(t^2+1)\,\dfrac{tz}{4n}\\[2mm]
(3t^6-7t^4-5t^2-3)\,\dfrac{tz}{96n^2}\\[2mm]
(t^{10}-11t^8+14t^6+6t^4-3t^2-15)\,\dfrac{tz}{384n^3}\\[2mm]
(15t^{14}-375t^{12}+2225t^{10}-2141t^8-939t^6-213t^4-915t^2+945)\,\dfrac{tz}{92160n^4}\\[2mm]
(3t^{18}-133t^{16}+1764t^{14}-7516t^{12}+5994t^{10}+2490t^8+1140t^6+180t^4+5355t^2+17955)\,\dfrac{tz}{368640n^5}
\end{array}\right\}\ \text{V}$$

The first four of these will be needed in the practical computation of p for values of n exceeding 20; the coefficients of the negative powers of n are given in Table III. For the fifth correction we need only consider the maximum values. These may be found from the zeros of the polynomial in powers of t^2, given in IV; the polynomial is of the $10th.$ degree in t^2, and has five positive roots. For each of these the following Table shows (i) the value of t^2, (ii) the value of t, (iii) the coefficient of n^5 in V (iv) the value of n for which the contribution of this term to p is $\pm$ .00001.

(*i*) t^2	(*ii*) t	(*iii*) C_5	(*iv*) n
1.3133	1.1460	+ .019864	4.567
4.7330	2.1755	— 1.0818	10.158
10.0555	3.1710	+ 10.3695	15.964
17.8547	4.2255	— 19.6603	18.143
29.7308	5.4526	+ 9.1438	15.568

The largest maximum is at $t = 4.2255$; here the value of n, under (*iv*), rises to 18.143. At this point the fifth correction becomes substantial for values of n but little below 20; at $n = 21$, however, the correction is less than 4 in the 6 th. place of decimals. We may conclude that, for values of n exceeding 20, the error involved in neglecting corrections beyond the fourth will not exceed .000005.

Over considerable parts of this region fewer corrections than four will suffice. The diagram shows at a glance how many terms will be needed at each point. The curves pass through those points at which the corresponding correction is $\pm$ ·00001. Only the maxima are shown for the fifth correction. Thus outside the curve II, the first correction alone will suffice, provided we are willing to accept an error of unity in the 5th. place; between the curves II and III, two correction terms should be used, and so on. It will be seen that only in two small regions will the fourth correction be necessary, for values of n exceeding 20, where no more than the minimal accuracy is aimed at; in other regions the expansion is often useful down to the fourth term, if more accurate values of p are required.


160
140
120
100
80
60
40
30
20
10
0
n
1
2
3
4
5
6
t
II
II
III
III
III
IV
IV
IV
IV
V
V
V
V
V

TABLES I, II, III

TABLE I.

t	$n = 1$ $n' = 2$	2 3	3 4	4 5	5 6	6 7	7 8	8 9	9 10	10 11
0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0
.1	.531,7	.535,3	.536,7	.537,4	.537,9	.538,2	.538,4	.538,6	.538,7	.538,8
.2	.562,8	.570,0	.572,9	.574,4	.575.3	.576,0	.576,4	.576,8	.577,0	.517,3
.3	.592,8	.603,8	.608,1	.610,4	.611,9	.612,9	.613,6	.614,1	.614,5	.614,8
.4	.621,1	.636,1	.642,0	.645,2	.647,2	.648,5	.649,5	.650,2	.650,8	.651,2
.5	.647,6	.666,7	.674,3	.678,3	.680.9	.682,6	.683,8	.684,7	.685,5	.686,1
.6	.672,0	.695,3	.704,6	.709,6	.712,7	.714,8	.716,3	.717,4	.718,3	.719,1
.7	.694,4	.721,8	.732,8	.738,7	.742,4	.744,9	.746,7	.748,1	.749,2	.750,1
.8	.714,8	.746,2	.758,9	.765,7	.770,0	.772,9	.775,0	.776,6	.777,8	.778,8
.9	.733,3	.768,4	.782,8	.790,5	.795,3	.798,6	.801,0	.802,8	.804,2	.805,4
1.0	.750,0	.788,7	.804,5	.813,0	.818,4	.822,0	.824,7	.826,7	.828,3	.829,6
1.1	.765,1	.807,0	.824,2	.833,5	.839,3	.843,3	.846,1	.848,3	.850,1	.851,4
1.2	.778,9	.823,5	.841,9	.851,8	.858,1	.862,3	.865,4	.867,8	.869,6	.871,1
1.3	.791,3	.838,4	.857,8	.868,3	.874,8	.879,3	.882,6	.885,1	.887,0	.888,6
1.4	.802,6	.851,8	.872,0	.882,9	.889,8	.894,5	.897,9	.900,5	.902,5	.904,1
1.5	.812,8	.863,8	.884,7	.896,0	.903,0	.907,9	.911,4	.914,0	.916,1	.917,7
1.6	.822,2	.874,6	.896,0	.907,6	.914,8	.919,6	.923,2	.925,9	.928,0	.929,7
1.7	.830,7	.884,4	.906,2	.917,8	.925,1	.930,0	.933,5	.936,2	.938,3	.940,0
1.8	.838,6	.893,2	.915,2	.926,9	.934,1	.939,0	.942,6	.945,2	.947,3	.949,0
1.9	.845,8	.901,1	.923,2	.934,9	.942,1	.946,9	.950,4	.953,0	.955,1	.956,7
2.0	.852,4	.908,2	.930,3	.941,9	.949,0	.953,8	.957,2	.959,7	.961,7	.963,3
2.1	.858,5	.914,7	.936,7	.948,2	.955,1	.959,8	.963,1	.965,5	.967,4	.969,0
2.2	.864,2	.920,6	.942,4	.953,7	.960,5	.964,9	.968,1	.970,5	.972,3	.973,8
2.3	.869.5	.925,9	.947,5	.958,5	.965,1	.969,4	.972,5	.974,8	.976,5	.977,9
2.4	.874,3	.930,8	.952,1	.962,8	.969,2	.973.4	.976,3	.978,4	.980,1	.981,3
2.5	.878,9	.935,2	.956,1	.966.6	.972,8	.976,7	.979,5	.981,5	.983,1	.984,3
2.6	.883,1	.939,2	.959,8	.970,0	.975.9	.979,7	.982,3	.984,2	.985,6	.986,8
2.7	.887,1	.942,9	.963,1	.973,0	.978,6	.982,2	.984,7	.986,5	.987,8	.988,8
2.8	.890,8	.946,3	.966,1	.975,6	.981,0	.984,4	.986,7	.988 4	.989,6	.990,6
2.9	.894,3	.949,4	.968,7	.977,9	.983,1	.986,3	.988,5	.990,1	.991,2	.992,1
3.0	.897,6	.952,3	.971,2	.980,0	.985,0	.988,0	.990,0	.991,5	.992,5	.993,3
3.1	.900,7	.954,9	.973,4	.981,9	.986,6	.989,4	.991,3	.992,7	.993,6	.994,4
3.2	.903,6	.957,3	.975,3	.983,5	.988,0	.990,7	.992,5	.993,7	.994,6	.995,3
3.3	.906,3	.959,6	.977,1	.985,0	.989,3	.991,8	.993,4	.994,6	.995,4	.996,0
3.4	.908,9	.961,7	.978,8	.986,4	.990,4	.992,8	.994,3	.995,3	.996,1	.996,6

t	$n = 11$ $n' = 12$	12 13	13 14	14 15	15 16	16 17	17 18	18 19	19 20	20 21	∞
0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,0	.500,000,0
.1	.538,9	.539,0	.539,1	.539.1	.539,2	.539,2	.539,2	.539,3	.539,3	.539,3	.539,827,8
.2	.577,4	.577,6	.577,7	.577,8	.577,9	.578,0	.578,1	.578,1	.578,2	.578,2	.579,259,7
.3	.615,1	.615,3	.615,5	.615,7	.615,9	.616,0	.616,1	.616,2	.616,3	.616,4	.617,911,4
.4	.651,6	.651,9	.652,2	.652,4	.652,6	.652,8	.652,9	.653,1	.653,2	.653,3	.655,421,7
.5	.686,5	.686,9	.687,3	.687,6	.687,8	.688,1	.688,3	.688,4	.688,6	.688,7	.691,462,5
.6	.719,7	.720,2	.720,6	.721,0	.721,3	.721,5	.721,8	.722,0	.722,2	.722,4	.725,746,9
.7	.750,8	.751,4	.751,9	.752,3	.752,7	.753,0	.753,3	.753,6	.753,8	.754,0	.758,036,3
.8	.779,7	.780,4	.781,0	.781,5	.781,9	.782,3	.782,6	782,9	.783,2	.783,4	.788,144,6
.9	.806,3	.807,1	.807,8	.808,3	.808,8	.809,3	.809,7	.810,0	.810,3	.810,6	.815,939,9
1.0	.830,6	.831,5	.832,2	.832,9	.833,4	.833,9	.834,3	.834,7	.835,1	.835,4	.841,344,7
1.1	.852,6	.853,5	.854,4	.855,1	.855,7	.856,2	.856,7	.857,1	.857,5	.857,8	.864,333,9
1.2	.872,3	.873,4	.874,2	.875,0	.875,6	.876,2	.876,7	.877,2	.877,6	.877,9	.884,930,3
1.3	.889,9	.891,0	.891,9	.892,7	.893,4	.894,0	.894,5	895,0	.895,4	.895,8	.903,199,5
1.4	.905,5	.906,6	.907,5	.908,4	.909,1	.909,7	.910,3	.910,7	.911,2	.911,6	.919,243,3
1.5	.919,1	.920,3	.921,2	.922,1	.922,8	.923,5	.924,0	.924,5	.925,0	.925,4	.933,192,8
1.6	.931,0	.932,2	.933,2	.934,0	.934,8	.935,4	.936,0	.936,5	.937,0	.937,4	.945,200,7
1.7	.941,4	.942,6	.943,5	.944,4	.945,1	.945,8	.946,3	.946,8	.947,3	.947,7	.955,434,5
1.8	.950,3	.951,5	.952,5	.953,3	.954,0	.954,6	.955,2	.955 7	.956,1	.956,5	.964,069,7
1.9	.958,0	.959,1	.960,1	.960,9	.965,6	.962,2	.962,7	963,2	.963,6	.964,0	.971,283,4
2.0	.964,6	.965,7	.966,6	.967,4	.968,0	.968,6	.969,1	.969,6	.970,0	.970,4	.977,249,9
2.1	.970,2	.971,2	.972,1	.972,8	,973,5	.974,0	.974,5	.975,0	.975,3	.975,7	.982,135,6
2.2	.975.0	.975,9	.976,8	.977 4	.978,1	.978,6	.979,0	.979,4	.979,8	.980,1	.986,096,6
2.3	.979,0	.979,9	.980,7	.981,3	.981,9	.982,4	.982,8	983,2	.983,5	.983,8	.989,275,9
2.4	.982,4	.983,2	.984,0	.984,6	.985,1	.985,5	.985,9	986,3	.986,6	.986,9	.991,802,5
2.5	.985,2	.986,0	.986,7	.987,3	.987,7	.988,2	.988,5	.988,8	.989,1	.989,4	.993,790,3
2.6	.987,7	.988,4	.989,0	.989,5	.990,0	.990,3	.990,7	.991,7	.991,2	.991,4	.995,338,8
2.7	.989,7	.990,3	.990,9	.991,4	.991,8	.992,1	.992,4	.992,7	.992,9	.993,1	.996,533,0
2.8	.991,4	.992,0	.992,5	.992,9	.993,3	.993,6	.993,8	.994,1	.994,3	.994,5	.997,444,9
2.9	.992,8	.993,3	.993,8	.994,2	.994,5	.994,5	.995,0	.995,2	.995,4	.995,6	.998,134,2
3.0	.994,0	.994,5	.994,9	.995,2	.995,5	.995,8	.996,0	.996,2	.996,3	.996,5	.998,650,1
3.1	,994,9	.995,4	.995,8	.996,1	.996,3	.996,6	.996,0	.996,9	.997,1	.997,2	.999,032,4
3.2	.995,8	.996,2	.996,5	.996,8	.997,0	.997,2	.997,4	.997,5	.997,6	.997,8	.999,312,9
3.3	.996,5	.996,8	.997,1	.997,4	.997,6	.997,7	.997,9	.998,0	.998,1	.998,2	.999,516,6
3.4	.997,0	.997,4	.997,6	.997,8	.998,0	.998,2	.998,3	.998,4	.998,5	.998,6	.999,663,1

TABLE I. *(continued)*

	$n=1$ $n'=2$	2 3	3 4	4 5	5 6	6 7	7 8	8 9	9 10	10 11	$n=11$ $n'=12$	12 13	13 14	14 15	15 16	16 17	17 18	18 19	19 20	20 21	∞
3.5	.911,4	963,6	980,3	.987,6	.991,4	.993,6	.995,0	.996,0	.996,6	.997,1	.997,5	.997,8	.998,0	.998,2	.998,4	998,5	.998,6	.998,7	.998,8	.998,9	.999,767,4
3.6	.913,8	.965,4	.981,6	.988,6	.992,2	.994,3	.995,6	.996,5	.997,1	.997,6	.997,9	.998,2	.998,4	.998,6	.998,7	.998,8	.998,9	.999,0	.999,0	.999,1	.999,840,9
3.7	.916,0	.967,0	.982, 9	.989,6	.993,0	.995.0	.996,2	.997,0	.997,5	.997,9	.998.2	.998,5	.998,7	.998,8	.998,9	.999,0	.999,1	.999,2	.999,2	.999,3	.999,892,2
3.8	.918,1	.968,6	.984,0	.990,4	.993.7	.995,5	.996,6	.997,4	.997,9	,998,3	.998,5	.998,7	.998,9	.999,0	.999,1	.999,2	.999,3	.999,3	.999,4	,999.4	.999,927,7
3.9	.920,1	.970,1	.985,0	.991,2	.994,3	.996,0	.997,1	.997,7	.998,2	.998,5	.998,8	.998,9	.999,1	.999,2	.999,3	.999,4	.999,4	.999,5	.999,8	.999,6	.999,951,9
4,0	.922,0	.971,4	.986,0	.991,9	.994,8	.996,4	.997,4	.998,0	.998,4	998,7	.999,0	.999,1	.999,2	.999,3	.999,4	.999,5	.999,5	.999,6	.999,6	.999,6	.999,968,3
4.1	.923,9	.972,7	.986,9	.992,6	.995,3	.996,8	99.7,7	.998,3	.998,7	.998,9	.999,1	.999,3	.999,4	.999,5	.999,5	.999,6	.999,6	.999,7	.999,7	.999,7	.999,979,3
4.2	.925,6	.973,9	.987,7	.993,2	.995,8	.997,2	99.8,0	.998,5	.998,8	.999,1	.999,3	.999,4	.999,5	.999,6	.999,6	.999,7	.999,7	.999,7	.999,8	.999,8	.999,986,7
4.3	.927,3	.975,0	.988.4	.993,7	.996,1	.997,5	.998,2	.998.7	.999,0	.999,2	.999,4	.999,5	.999,6	.999,6	.999,7	.999,7	.999,8	.999,8	.999,8	.999,8	.999,991,5
4.4	.928,9	.976,0	.989,1	.994,2	.996,5	.997,7	.998,4	.998,9	.999,1	.999,3	.999,5	.999,6	.999,6	.999,7	.999,7	.999,8	.999,8	.999,8	.999,8	.999,9	.999,994,6
4.5	.930,4	.977,0	.989,8	.994,6	.996.8	.997,9	.998,6	.999,0	.999,3	.999,4	.999,5	.999,6	.999,7	.999,8	.999,8	999,8	.999,8	.999,9	.999,9	.999,9	.999,996,6
4.6	.931,9	.977,9	.990,3	.995,0	.997,1	.998,2	.998,8	.999,1	.999,4	.999,5	.999,6	.999,7	.999,8	.999,8	.999,8	.999,9	.999,9	.999,9	.999,9	.999,9	.999,99 ,9
4.7	.933,3	.978,8	.990,9	.995,3	.997,3	.998,3	,998,9	.999,2	.999,4	.999,6	.999,7	.999,7	.999,8	.999,8	.999,9	.999,9	.999,9	.999,9	.999,9	.999,9	.999,998,7
4,8	.934,6	.979,6	.991,4	.995,7	99.7,6	.998,5	.999,0	.999,3	.999,5	.999,6	.999,7	.999,8	.999,8	.999,9	999,9	.999,9	.999,9	.999,9	.999,9	.999,9	.999,999,2
4.9	.935,9	.980,4	.991,9	.996,0	99.7,8	.998,6	.999,1	.999,4	.999,6	.999,7	.999,8	.999,8	.999,9	.999,9	.999,9	.999,9	.999,9	.999,9	1.000,0	1.000,0	.999,999,5
5.0	.937.2	.981,1	.992,3	.996,3	.997,9	.998,8	.999,2	.999,5	.999,6	.999.7	.999.8	.999,8	.999,9	.999,9	.999,9	999,9	.999,9	1.000,0			.999,999,7
5.1	.938,4	.981,8	.992,7	.996,5	.998,1	.998,9	.999,3	.999,5	.999,7	.999,8	.999,8	.999,9	.999,9	.999,9	.999,9	999,9	1.000,0				.999,999,8
5.2	.939,5	.982,5	.993,1	.996,7	.998,3	.999,0	.999,4	.999,6	.999,7	.999,8	.999,9	.999,9	.999,9	,999,9	.999,9	1.000,0					.999,999,9
5.3	.940,6	.983,1	.993,4	.997,0	.998,4	.999,1	.999,4	.999,6	.999,8	.999,8	.999,9	.999,9	.999,9	,999,9	1.000,0						.999,999,9
5.4	.941,7	.983,7	.993,8	.997,2	.998,5	.999,2	.999,5	.999,7	.999,8	.999,8	.999,9	.999,9	.999,9	1.000,0							1.000,000,0
5.5	.942,8	.984,2	.994,1	.997,3	.998,6	.999,2	.999,5	.999,7	.999,8	.999,9	,999,9	.999,9	.999,9								
5.6	.943,8	.984,8	.994,4	.997,5	.998,7	.999,3	.999,6	.999,7	.999.8	.999,9	.999,9	.999,9	1.000,0								
5.7	.944,7	.985,3	.994,6	.997,7	.998,8	.999,4	.999,6	.999,8	.999,9	.999,9	.999,9	1.000,0									
5.8	.945,7	.985,8	.994,9	.997,8	.998,9	.999,4	.999,7	.999,8	.999,9	.999,9	.999,9										
5.9	.946,6	.986,2	.995,1	.997,9	.999,0	.999,5	.999,7	.999,8	.999,9	.999,9	.999,9										
6.0	.947,4	.986,7	.995,4	.998,1	.999,1	.999,5	.999,7	.999,8	.999,9	.999,9	1.000,0										

TABLE II. — *Supplementary table for high values of t.*

t	$n=3$ $n'=4$	$n=4$ $n'=5$	$n=5$ $n'=6$	$n=6$ $n'=7$	$n=7$ $n'=8$	$n=8$ $n'=9$	$n=9$ $n'=10$	$n=10$ $n'=11$
6.0	.995,364	.998,059	.999,077	.999,518	.999,729	.999,838	.999,899	.999,934
6.5	.996,303	.998.555	999,357	.999,684	.999,833	.999,906		.999,966
7.0	.997,007	.998,904	.999,542	.999,788	.999,894	.999,944	.999,968	
7.5	.997,544	.999,155				.999,965		
8.0	.997,962	.999,338	.999,754	.999,898	.999,954			
8.5	.998,290							
9.0	.998,552	.999,578	.999,859	.999,947				
10.0	.998,936	.999,719	.999,915	.999,971				
11.0	.999,196							
12.0	.999,377	.999,862	.999,965					
14.0	.999,605	.999,924						
16.0	.999,735	.999,955						
20.0	.999,863							
24.0	.999,921							
28.0	.999,950							

Linear interpolation between adjacent entries will give four figure accuracy.

TABLE III.

t	C_1	C_2	C_3	C_4
.1	.010,023,1	— .001,261	— .001,55	+ .000,4
.2	.020,334,2	— .002,616	— .003,08	.000,8
.3	.031,178,5	— .004,177	— .004,53	.001,3
.4	.042,719,3	— .006,087	— 005,86	.001,7
.5	.055,010,2	— .008,509	— .006,97	.002,2
.6	.067,977,8	— .011,595	— .007,72	.002,6
.7	.081,420,2	— .015,432	— .007,96	.002,8
.8	.095,018,8	— .019,991	— .007,74	.002,6
.9	.108,363,2	— .025,066	— .006,51	.002,1
1.0	.120,985,4	— .030,246	— .005,04	.001,1
1.1	.132,399,7	— .034,907	— .003,76	.000,4
1.2	.142,144,2	— .038,248	— .003,75	.001,1
1.3	.149,819,0	— .039,363	— .006,56	.005,8
1.4	.155,117,7	— .037,344	— .014,10	.018,2
1.5	.157,849,6	— .031,399	— .028,41	.043,0
1.6	.157,951,2	— .020,971	— .051,29	.084,9
1.7	.155,486,7	— .005,832	— .083,84	.161,2
1.8	.150,637,0	+ .013,846	— .126,09	.232,3
1.9	.143,682,2	.037,483	— .176,64	.335,1
2.0	.134,977,5	.064,114	— .232,56	.446,6
2.1	.124,924,4	.092,473	— .289,45	.551,2
2.2	.113,944,4	.121,099	— .341,85	.627,7
2.3	.102,451,8	.148,472	— .383,79	.652,2
2.4	.090,832,2	.173,147	— .409,50	.601,0
2.5	.079,425,1	.193,877	— .414,17	.455,3
2.6	.068,512,4	.209,710	— .394,59	+ .205,6
2.7	.058,312,9	.220,052	— .349,67	— .145,4
2.8	.048,980,8	.224,698	— .280,61	— .580,9
2.9	.040,609,7	.223,749	— .190,82	— 1.070,4
3.0	.033,238,9	.217,715	— .085,59	— 1.572,7
3.1	.026,862,2	.207,289	+ .028,49	— 1.368,5
3.2	.021,437,7	.193,351	.144,23	— 1.610,8
3.3	.016,897,1	.176,859	.254,58	— 1.837,2
3.4	.013,155,2	.158,774	.353,27	— 2.034,8
3.5	.010,117,7	.139,969	.435,35	— 2.732.8

TABLE III (*continued*).

t	C_1	C_2	C_3	C_4
3.6	.007,687,9	.121,313	.497,49	— 2.503,1
3.7	.005,772,2	.103,371	.538,13	— 2.123,7
3.8	.004,282,3	.086,649	.557,33	— 1.625,7
3.9	.003,139,7	.071,486	.556,67	— 1.050,1
4.0	.002,275,1	.058,066	.538,80	— .441,7
4.1	.001,629,5	.046,453	.507,08	+ .155,5
4.2	.001,153,8	.036,613	.465,19	.702,6
4.3	.000,807,4	.028,438	.416,81	1.169,2
4.4	.000,558,6	.021,773	.365,31	1.535,3
4.5	.000,382,1	.016,436	.313,56	1.791,6
4.6	.000,258,4	.012,235	.263,86	1.938,7
4.7	.000,172,8	.008,984	.217,87	1.985,3
4.8	.000,114,3	.006,508	.176,63	1.946,2
4.9	.000,074,6	.004,652	.140,70	1.839,4
5.0	.000,048,3	.003,281	.110,17	1.683,9
5.1	.000,030,9	.002,284	.084,85	1.498,6
5.2	.000,019,5	.001,570	.064,29	1.299,5
5.3	.000,012,2	.001,065	.047,96	1.100,4
5.4	.000,007,6	.000,713	.035,27	.910,7
5.5	.000,004,6	.000,472	.025,47	.379,9
5.6	.000,002,8	.000,308	.018,15	.506,5
5.7	.000,001,7	.000,199	.012,74	.437,0
5.8	.000,001,0	.000,127	.008,82	.349,5
5.9	.000,000,6	.000,080	.006,02	.262,6
6.0	.000,000,3	.000,050	.004,05	.193,9

MARCELLO BOLDRINI.

Capacità contributiva e gravame fiscale di alcuni Stati.

1. — La capacità contributiva assoluta di una paese è misurata da tutta quella parte della produzione o reddito privato che supera il consumo minimo necessario per mantenere la produzione allo stesso livello (1).

In questo consumo rientra la spesa strettamente necessaria alla sussistenza, educazione ed istruzione tecnica della popolazione, che, sotto il punto di vista economico, rappresenta il costo di esercizio, di ammortamento, di ricostituzione dei capitali umani. Secondo opinioni autorevoli, non si dovrebbe, anzi, parlare che di reddito al netto delle sussistenze (2).

Dal confronto fra i redditi medi individuali dei vari Stati, dopo avere detratto il minimo medio di esistenza, si può desumere la capacità contributiva relativa di essi.

Ma, oltre e più che la capacità contributiva assoluta e relativa di più Stati, ha importanza la conoscenza del rispettivo gravame fiscale assoluto e comparativo. È evidente, che il gravame fiscale di un paese sarebbe massimo quando l'ammontare delle imposte assorbisse il cento per cento della parte di reddito che misura la capacità contributiva di quel paese e sarebbe nullo se nessuna parte di tale reddito venisse prelevata dal fisco. Cosicchè, è logico che si assuma come misura del gravame fiscale assoluto sopportato da uno

---

(1) I principali risultati delle indagini statistiche svolte nella presente memoria sono stati riassunti nel rapporto: *International Comparisons of the Burden of Taxation.* Roma, Libreria dello Stato, 1925. Per il concetto di capacità contributiva e di gravame fiscale, vedasi: G. F. SHIRRAS, *Taxable Capacity and Burden of Taxation and Public Debt.* « Journal of the Royal Statistical Society », vol LXXXVIII, part. IV, 1925.

(2) Vedasi: C. GINI, *A Comparison of the Wealth and National Income of several important Nations.* Roma, Libreria dello Stato, 1925, p. 36 segg. A. LORIA, *La sintesi economica.* Torino, 1909, p. 239 segg.

Stato la percentuale del reddito nazionale al netto delle sussistenze assorbita dall'ammontare totale delle imposte erariali e locali. Per ragioni di brevità potremo chiamare tale cifra *percentuale*, o, meglio, *indice del gravame fiscale.*

Meno logico è, invece, cercare di farsi una idea comparativa della pressione fiscale sopportata dai vari Stati basandosi sul confronto fra le percentuali di reddito netto assorbito dalle imposte in ciascuno di essi. Come si intuisce facilmente e come si vedrà più tardi, non è affatto necessario che uguali percentuali esprimano lo stesso gravame fiscale e diverse percentuali esprimano un gravame fiscale differente nei vari Stati. Tuttavia, tale criterio può essere accolto come una prima approssimazione, in vista della sua straordinaria semplicità, e, soprattutto, perchè merita, almeno, di essere sostituito all'altro, del confronto fra l'ammontare medio individuale delle imposte pagate nei diversi Stati, come si fa non di rado (1).

2. — Possediamo una valutazione recente del reddito nazionale al netto delle sussistenze, ed informazioni ufficiali o valutazioni attendibili dell'ammontare delle imposte erariali e locali dei cinque Stati che furono i principali alleati durante la guerra mondiale (2).

---

(1) Anche recentemente il Senatore americano BORAH, dal confronto fra il carico medio individuale di imposte nei vari Stati cercava di formarsi una idea comparativa del gravame fiscale da essi sopportato.

(2) Vedasi, per la valutazione del reddito nazionale netto degli Stati: C. GINI. *A Comparison of the Wealth etc.*, cit. Il GINI assegna come minimo di esistenza una detrazione media individuale di 48 dollari agli italiani, di 60 dollari ai francesi ed ai belgi, di 75 dollari agli inglesi e di 100 dollari agli americani. In realtà, queste cifre rappresentano un *limite inferiore del minimo di esistenza*. Aggiungasi che i criteri proposti da qualche autore per valutare questo minimo porterebbero ad assegnare agli italiani molto meno del 48 % della detrazione adottata per gli americani. Così, lo STAMP vorrebbe che il minimo di esistenza calcolato variasse, da un paese all'altro, in proporzione del reddito medio individuale lordo. In tal caso, il minimo necessario alla vita di un italiano risulterebbe meno di 1/6 del minimo necessario alla vita di un cittadino degli Stati Uniti, sol' perchè il reddito medio di un italiano è oggi di dollari 101 e quello di un cittadino americano di dollari 614. Il minimo necessario alla esistenza si potrebbe anche calcolare per i vari paesi in proporzione del salario delle classi operaie. Ma, anche per questa via, si assegnerebbe ad un italiano 1/5 appena del minimo di esistenza di un cittadino degli Stati Uniti, perchè questo è, all'incirca, il rapporto fra il potere reale di acquisto dei salari operai nei due Stati. Evidentemente, il GINI ha stabilito minimi di esistenza bassi ed ha attenuato le differenze che tra essi si verificano nei vari Stati, allo scopo di evitare che i suoi procedimenti potessero essere sospettati di far figurare gli italiani in una condizione più sfavorevole della reale, sospetto che era particolarmente da evitarsi tenuto conto della circo-

Possiamo, perciò, in base ai concetti fissati nel paragrafo precedente, valutare comparativamente la capacità contributiva di tali Stati e il gravame fiscale da essi sopportato.

Nel seguente specchietto è indicato:

a) l'ammontare del reddito nazionale e il reddito medio per testa di abitante nei cinque Stati a metà del 1925, diminuito del minimo di esistenza ed espresso in dollari;

b) l'ammontare delle imposte pagate nei cinque Stati sotto forma di contribuzioni dirette e indirette, erariali e locali, accresciuto dell'ammontare dei prodotti netti dei monopoli e delle imprese di Stato. Le cifre si riferiscono all'anno più recente per il quale si posseggono i risultati consuntivi e sono state tradotte in dollari al cambio medio di tale anno.

c) il rapporto percentuale tra le cifre di cui alla lettera *b*) e le cifre di cui alla lettera *a*);

d) il rapporto delle percentuali italiane di cui alla lettera *c*) e le corrispondenti percentuali degli altri Stati.

TABELLA I. — *Capacità contributiva per testa e gravame fiscale in alcuni Stati.*

STATI	Reddito nazionale nel 1925 al netto del minimo di esistenza		Gettito delle imposte erariali e locali in miliardi di dollari	Percentuali del reddito netto assorbite dalle imposte	Rapporto delle percentuali italiane alle percentuali degli altri Stati
	Totale miliardi di dollari *a*	Dollari per testa	*b*	100 *b* : *a*	
	1	2	3	4	5
Italia . . .	2.13	53	0.87	38.10	1.0
Francia . . .	5.37	136	1.57	29.20	1.3
Belgio . . .	1.28	164	0.175	13.30	2.9
Regno Unito .	15.58	344	3.91	27.30	1.4
Stati Uniti .	58.60	514	7.77	13.30	2,9

stanza che la sua memoria era destinata ad essere presentata alla Commissione americana per i debiti, in occasione delle trattative dello scorso novembre con la Delegazione italiana. Per la stessa ragione abbiamo ritenuto di attenerci, senza variazione, ai criteri da lui adottati.

I dati sul gettito delle imposte in Italia risultano da informazioni ufficiali ricevute dalla Direzione Generale delle Imposte Dirette del Ministero delle Finanze e concordano con quelli contenuti nella citata memoria dello SHIRRAS. Per la

Le cifre sono così eloquenti che non richiedono commenti troppo diffusi. Il cittadino italiano, al quale è stato attribuito un minimo di esistenza che sta appena tra il 48 e l'80 $^0/_0$ di quello attribuito ai cittadini degli altri Stati ex alleati, ha, tuttavia, una capacità contributiva pari a circa 1/3 di quella dei francesi e dei belgi, inferiore a 1/6 di quella degli inglesi e poco maggiore di 1/10 della corrispondente capacità contributiva degli americani (col. 2). E, ciò non ostante, è in Italia che il fisco preleva la massima percentuale del reddito nazionale al netto delle sussistenze (col. 4 e 5). Adottando tale percentuale come indice del gravame fiscale sopportato dagli italiani, possiamo dire che tale gravame è pari a 1 e 1/3 del corrispondente gravame fiscale della Francia, 1 e 2/5 di quello del Regno Unito e quasi 3 volte di quello del Belgio e degli Stati Uniti.

Queste grandi differenze fra l'indice del gravame fiscale dell'Italia e l'indice corrispondente degli altri Stati dipendono anzitutto dal forte ammontare totale dei tributi pagati nel nostro Paese; e, in secondo luogo, dal fatto che poco meno della metà del basso reddito nazionale italiano è assorbita dal minimo indispensabile alla sussistenza della popolazione. Perciò, i forti tributi vengono prelevati su un reddito netto che in Italia rappresenta una percentuale del reddito totale più bassa che altrove (1).

3. — Per quanto siano abbastanza istruttivi i risultati precedenti, essi sono, comunque, lungi dal dare una idea comparativa sufficien-

---

Francia, si veda: *Progression des impôts de 1913 à 1925 en France et aux divers pays.* « Bulletin de la Statistique Générale de la France », tome XIV, Juillet, 1925. Per il Regno Unito: *The Statesman Year Book, 1925.* Per il Belgio: *The Inter-Ally Debts.* « National Industrial Conference Board Inc. », New York, 1925, p. 99. Nel limite del possibile, anche queste cifre sono state controllate su fonti ufficiali.

(1) Tutto ciò è fatto manifesto dalle cifre seguenti:

	Percentuale del reddito totale lordo assorbito dalle imposte	dalle sussistenze
Italia	21.4	47.5
Francia	20.3	30.6
Belgio	10.0	26.8
Regno Unito	20.6	18.0
Stati Uniti	11.1	16.3

Che in Italia l'ammontare dei tributi sia elevatissimo è provato dal fatto che tale ammontare rappresenta già una percentuale massima del reddito totale lordo per rispetto agli altri Stati. In Italia è pure massima la percentuale di quella parte del reddito che non sopporta gravame tributario perchè rappresenta il consumo minimo indispensabile per mantenere in efficienza la popolazione.

temente corretta del gravame fiscale sopportato dai cittadini dei cinque Stati in esame.

Anche se in due o più Stati il reddito netto individuale è approssimativamente lo stesso ed è press'a poco uguale la percentuale di esso prelevata dal fisco, ciò non significa ancora che il gravame fiscale sopportato dai cittadini di tali Stati sia uniforme.

Un fattore che influisce certamente sul gravame fiscale e che deve essere preso anzitutto in considerazione è la composizione e la stabilità del reddito, che variano notevolmente da uno Stato all'altro. Il reddito agricolo è molto soggetto a variare di anno in anno per l'influenza di circostanze meteorologiche incontrollabili e perchè i prodotti agricoli, i quali si conservano meno bene e meno a lungo dei prodotti industriali propriamente detti, debbono essere commerciati in breve intervallo ai prezzi imposti dalle condizioni momentanee della domanda e dell'offerta (1). Perciò, nel confrontare il gravame fiscale dei diversi Stati non si può trascurare di tener conto, in generale, della diversa variabilità del reddito (2) e, in particolare, della proporzione che sul reddito totale rappresenta, in ciascuno di essi, il reddito agricolo. Orbene, fra gli Stati che si considerano nel presente studio, la percentuale della ricchezza fondiaria rispetto alla ricchezza totale e la percentuale del reddito agricolo rispetto al reddito complessivo, a giudicare dalle cifre seguenti, per quanto un po' frammentarie, sembrano essere massime in Italia (3).

Tabella II. — *Ricchezza fondiaria e reddito agricolo.*

STATI	Percentuale della ricchezza fondiaria alla ricchezza totale	Percentuale del reddito agricolo al reddito totale
Italia	38 % (1925)	37 % (1914) 32 % (1925)
Stati Uniti.	21 % (1920)	—
Francia.	13 % (1925)	30 % (1914)
Belgio	—	15 % (1914)

(1) C. Gini, *A comparison etc.*, cit., p. 46.

(2) Un paese come l'India, scrive lo Shirras (art. cit.), soggetto ai monsoni, che costituiscono la vena giugulare della prosperità indiana, non può essere tassato nella stessa misura di un altro, nel quale la carestia non è da temere.

(3) Riferite da C. Gini, *A comparison, etc.*, cit., pp. 45-46.

Anche la composizione del reddito va tenuta presente, quando si confrontino la capacità contributiva e il gravame fiscale degli Stati. Infatti, a parità di altre condizioni, il gravame fiscale è più grande negli Stati in cui è maggiore la proporzione del reddito non guadagnato (reddito da capitale). Anche senza riferire delle cifre, che, comunque, dovrebbero essere desunte per mezzo di coefficienti dai dati sul reddito, è verisimile ritenere che l'Italia, che un grandissimo poeta chiamò « la grande Proletaria », ritragga una rilevante percentuale del suo reddito dal lavoro dei cittadini.

4. — Un altro fattore che influisce sul gravame fiscale degli Stati è la concentrazione dei redditi.

È evidente che due Stati, i quali abbiano uno stesso ammontare di reddito imponibile (reddito netto), saranno diversamente gravati da una imposta che prelevi $x\ ^0/_0$ di tale ammontare, se in essi è diversa la concentrazione dei redditi. In quello dei due Stati nel quale, rispetto all'altro, hanno una relativa prevalenza i redditi piccoli, il fisco dovrà gravare la mano su questi e prelevare forti aliquote sulla parte di tali redditi che supera il minimo di esistenza dei rispettivi possessori, per garantirsi il desiderato gettito di imposta. Al contrario, nello Stato in cui, rispetto all'altro, hanno una relativa prevalenza i redditi elevati, il fisco potrà, mediante forti falcidie sulla parte disponibile di tali redditi alti, ottenere più facilmente e con minore sacrifizio dei contribuenti il desiderato gettito di imposta (1).

Poichè le differenze nella concentrazione dei redditi fra Stato e Stato sono veramente cospicue, c'è da immaginare che l'importanza di questo fattore sul gravame fiscale relativo di più Stati possa essere anche considerevolissima. Le indagini che seguono tendono a precisare statisticamente siffatta importanza, e a permettere un passo ulteriore nella valutazione comparativa del gravame fiscale sopportato dai cittadini dei cinque Stati elencati nella tabella I.

---

(1) Queste ed altre osservazioni circa l'importanza della concentrazione dei redditi rispetto al gravame fiscale degli Stati possono leggersi, fra l'altro, in: Sir JOSIAH STAMP, *The Taxable Capacity of Ireland*. « Economic Journal », 1921, p. 337. Vedasi anche: Sir JOSIAH STAMP, *Wealth and Taxable Capacity*. London, 1922, p. 118; *Inventaire de la situation Financière de la France au debut de la treizième législature, presenté par* M. CLÉMENTEL, *Ministre des Finances*. Paris 1924, p. 125; e, per qualche cenno: V. TANGORRA, *La pressione tributaria in Italia*. « Rivista internazionale », 1922. Parecchi anni prima il GINI aveva indicato il metodo, che applicheremo in questo articolo, per dare una misura sintetica di tale importanza.

5. — Poichè in tutti e cinque gli Stati esiste una imposta sul reddito globale, si conosce per ciascuno il numero dei redditieri e per quattro Stati l'ammontare dei redditi accertati ai fini dell'imposta, distinti secondo date classi di reddito (1).

Sarebbe facile, previa correzione ed integrazione dei dati originari, per eliminare l'influenza esercitata sulle serie dalle esenzioni e dagli abbattimenti consentiti dalla legislazione fiscale di ciascuno Stato, di comparare sinteticamente la distribuzione dei redditieri e dei redditi in base ai noti esponenti α (del PARETO) e β e δ (del GINI). Ma, non si raggiungerebbe, per questa via, lo scopo di valutare l'influenza della diversa concentrazione dei redditi nazionali sul gravame fiscale dei vari Stati.

È invece necessario seguire la via del confronto diretto tra le seriazioni dei redditieri e dei redditi preventivamente corrette ed integrate per le dette ragioni, uniformandone la classificazione per tutti gli Stati.

Il criterio da seguire è semplice, per quanto ne sia molto faticosa l'attuazione. In primo luogo, è necessario esprimere le classi di reddito dei diversi Stati in un'unica moneta, ad esempio, come faremo, in dollari. Ma, poichè le classificazioni dei redditi dei vari Stati si riferiscono ad anni diversi, nel fare le trasformazioni è necessario tener conto delle variazioni che ha subito il potere di acquisto del dollaro dal periodo per il quale si conosce la classificazione dei redditi americani al periodo per il quale si conosce la classificazione dei redditi degli altri Stati (2).

---

(1) I dati per l'Italia sono stati ufficialmente raccolti in occasione dell'applicazione dell'imposta complementare sul reddito e comunicati dalla Direzione Generale delle Imposte Dirette del Ministero delle Finanze. I dati per la Francia sono pubblicati in: « Bulletin de Statistique et de législation comparée », février 1925, pp. 226-227; quelli per il Belgio sono tratti da uno studio inedito del MAHAIM: *La fortune de la Belgique*, che dovrà uscire in una pubblicazione dell'Istituto Solvay, e che ho potuto consultare perchè è stato gentilmente comunicato dall'autore al prof. GINI. In tale scritto però è contenuta solo la classificazione dei redditieri e non dei redditi accertati. La classificazione dei redditieri e dei redditi del Regno Unito trovasi in: *Sixty-fourth Report of the Commissioners of his Majesty's Inland Revenue for the year ended 31st March, 1921.* London, 1922, e quelli degli Stati Uniti nel volume: *Treasury Department, United States Internal Revenue. Statistics of Income from Returns of net Income for 1920*, Washington, 1922.

(2) Per ridurre le lire del 1924 in dollari del 1920 basta, come è ovvio, adottare come cambio il prodotto del corso medio del dollaro in Italia nel 1924 per il numero indice dei prezzi agli Stati Uniti nel 1924 diviso per il numero indice dei prezzi agli Stati Uniti nel 1920.

Naturalmente noi non possiamo escludere che dall'uno all'altro dei paesi posti

In questo modo, i redditi espressi nelle singole monete nazionali risulteranno distinti secondo le classi di reddito espresse in dollari di un dato anno, o, in altri termini, in dollari che hanno uno stesso potere di acquisto.

Ciò fatto, si tratta di uniformare le classificazioni dei redditieri e dei redditi, adottando per tutti gli Stati che si considerano una stessa classificazione arbitraria, o anche la classificazione dei redditieri e dei redditi di uno degli Stati che si considerano: ad esempio, come faremo, quella degli Stati Uniti.

Naturalmente, questo scopo si può raggiungere solo in maniera approssimata, valendosi degli esponenti α e β.

Dopo avere controllato che tutte le serie dei redditieri e dei redditi di cui disponevamo seguono assai davvicino la legge di PARETO, mediante interpolazioni parziali abbiamo determinato il numero presuntivo di coloro che, in ciascun paese, posseggono i redditi corrispondenti alle classi di reddito adottate nelle statistiche americane e l'ammontare dei redditi rispettivi (1).

Le statistiche finanziarie del Belgio indicano bensì il numero dei redditieri accertati ai fini della *Surtaxe*, classificati secondo il loro reddito, ma non forniscono invece le cifre dell'ammontare dei redditi percepiti. Per non rinunciare a compiere, rispetto al Belgio, indagini analoghe a quelle degli altri Stati, abbiamo determinato l'ammontare approssimato dei redditi belgi, facendo i calcoli in base alla relazione approssimata studiata dal GINI: $\beta = \alpha - 1$ (2).

I risultati dei calcoli figurano nell'annessa tabella III, nella quale, però, per ragioni di brevità, le classi di reddito originariamente stabilite sono state alquanto raggruppate.

---

a confronto varii la rigorosità della tassazione, in modo che le distribuzioni dei redditi accertati rispecchino con diverso grado di fedeltà le distribuzioni dei redditi reali. Una rigorosa comparabilità, da tale punto di vista, non si può garantire in questo come in ogni altro confronto internazionale che si basa sulle statistiche finanziarie.

(1) Non ostante le critiche rivolte alla legge di PARETO, i calcoli accennati nel testo possono abbastanza bene servire nel caso presente. Vale la pena di ricordare che nel 1906 si fece un impiego ufficiale del metodo di PARETO in Inghilterra. Ricorda lo STAMP che un *British Select Committee* fu incaricato di indagare sulla possibilità di graduare l'*Income Tax*. Il BOWLEY fece uso della legge di PARETO applicandola ai dati esistenti per stimare il numero e l'ammontare dei redditi sopra 5.000 sterline. Vedasi: *Report by the Select Committee*, 1906, cit. da J. STAMP, *British Income and Properties*, London, 1922, pp. 330 e segg.

(2) C. GINI, *Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza*. «Giornale degli economisti», 1909. C. GINI, *Indici di concentrazione e di dipendenza* «Biblioteca dell'economista», serie V, Torino 1912.

Le cifre hanno bisogno di qualche chiarimento. Esse non dicono già quante persone percepiscono nei singoli paesi redditi compresi fra i limiti indicati in ciascuna classe e l'ammontare dei redditi rispettivi, ma solo il numero dei redditieri che posseggono redditi compresi entro certi limiti e l'ammontare dei redditi rispettivi, che si possono calcolare per i diversi paesi sulla base delle statistiche finanziarie. Questo non ci impedirà, per brevità di espressione, di parlare, in seguito, di numero di redditieri e di ammontare di redditi dei vari Stati, senza ripetere, di volta in volta, l'avvertenza qui fatta e senza richiamare le limitazioni insite nelle cifre, parendoci, dopo tutto, sufficiente quanto se ne è detto.

La prima classe di reddito considerata è diversa per i vari Stati. Le statistiche dei redditi dei Stati Uniti cominciano a considerare solo i redditi superiori a 1.000 dollari, superiori, cioè, esattamente, a una somma decupla del minimo di esistenza da noi assegnato agli americani. Abbiamo voluto, analogamente, per ragioni che esporremo in seguito, conoscere anche per gli altri Stati il numero dei redditieri che posseggono redditi almeno decupli del minimo di esistenza assegnato ai rispettivi cittadini, e l'ammontare dei redditi corrispondenti. Cosicchè, la prima classe di reddito indicata nella tabella III, comprende redditi di 480-1000 dollari per l'Italia; redditi di 600-1000 dollari per la Francia ed il Belgio; redditi di 750-1000 dollari per il Regno Unito.

Nella tabella IV che segue, le cifre della tabella III sono ridotte, facendo uguale a 1.000 per i vari Stati il numero totale dei redditieri con redditi superiori a 1.000 dollari e l'ammontare dei redditi rispettivi. Nella prima linea della tabella IV le cifre in parentesi sono ottenute dividendo il numero dei redditieri che posseggono redditi compresi fra il minimo adottato per i vari Stati e 1.000 dollari e l'ammontare dei redditi rispettivi, per il numero dei redditieri con redditi superiori a 1.000 dollari e per l'ammontare dei redditi superiori a 1.000 dollari. Cosicchè, i totali di ciascuna colonna, che sono uguali a 1.000, non comprendono le cifre della prima linea, scritte tra parentesi.

Le differenze fra la concentrazione dei redditi dei diversi Stati risultano veramente notevoli. Su 1.000 redditieri con più di 1.000 dollari di reddito, in Italia, coloro che percepiscono redditi compresi tra 1.000 e 2.000 dollari rappresentano una frazione del totale dei redditieri assai più elevata che negli altri Stati, e l'ammontare dei redditi rispettivi in Italia costituisce da solo più della metà dell'ammontare totale dei redditi superiori a 1.000 dollari. Viceversa, in

Tabella III. — *Redditieri e redditi rilevati in cinque Stati ai fini dell'imposta sul reddito e classificati secondo classi uniformi di reddito.*

Classi di reddito migliaia di dollari	ITALIA		FRANCIA		BELGIO		REGNO UNITO		STATI UNITI	
	Numero dei redditieri	Ammontare dei redditi in lire migliaia	Numero dei redditieri	Ammontare dei redditi in franchi migliaia	Numero dei redditieri	Ammontare dei redditi in franchi migliaia	Numero dei redditieri	Ammontare dei redditi in L.st migliaia	Numero dei redditieri	Ammontare dei redditi in $ migliaia
dal min. a 1	799.062	8.978.970	1.369.280	9.656.500	121.099	1.264.400	1.501.450	246.770	—	—
1–2	165.821	3.541.110	717 018	8.824.200	61.129	1.124.150	1·445.120	389.740	14.086.006	17.819.437
2–3	28.760	1.164 300	167.892	3.701.500	13.658	442.250	277.370	136.160	2.569.316	6.184.544
3–4	9.708	520.150	65.330	2.082.930	5.348	246.200	132.540	93.440	894.559	3.067.086
4–5	4.762	335.540	32.606	4.335.140	2.681	161.060	83.030	75.710	442.557	1.972.521
5-6	2.694	208.930	20.990	957.570	1 555	113.580	48.930	55.070	177.147	969.505
6–8	1.797	228.150	21.579	1.324.050	1.687	154.450	56.870	79.410	186.955	1.283.466
8–10	838	115.040	11.094	853.230	876	104.650	31.730	57.440	91.340	815.361
10–15	832	151.640	12.375	1.248.310	926	149.400	36.300	91.220	103.570	1.253.338
15–20	297	77.650	5.001	688.380	345	74.540	15.100	53.550	44.531	765.354
20–40	268	109.108	6.358	1.502.500	365	126.710	21.010	119.530	54.008	1.468.819
40-60	54	38.916	1.234	497.620	78	49.660	6.100	60.390	13.054	629.619
60–100	20	22.618	642	404.290	44	45.000	4.126	64.160	7.308	548.953
100–150	4	7.521	178	193.110	15	25.590	1.576	39.230	2.191	265.512
150–300	3	8.158	95	189.660	10	29.571	1.209	50.374	1.063	215.138
300–1000	—	—	25	114.450	4	45.389	403	38.828	391	216.434
oltre 1000	—	—	2	35.160	—	—	36	15.868	4	29.920
Totale	1.014.920	15.507.801	2.431.699	33.608.600	209.820	4.156.600	3.662.900	1.666.820	18.674.000	37.504.003

TABELLA IV. — *Ripartizione dei redditieri fatto = 1000 il loro numero e dei redditi fatto = 1000 il loro totale secondo classi di reddito uniformi in cinque Stati*

Classi di reddito migliaia di dollari	ITALIA		FRANCIA		BELGIO		REGNO UNITO		STATI UNITI	
	Redditieri	Reddito	Redditieri	Reddito	Redditieri	Reddito	Redditieri	Reddito	Redditieri	Reddito
dal min. a 1	(3701.79)	(1344.64)	(1288.83)	(403.15)	(1364.94)	(437.17)	(694.64)	(173.76)	—	—
1-2	768.19	542.39	674.90	368.50	689.00	388.69	668.58	274.44	754.31	475.11
2-3	133.24	178.33	158.03	154.44	153.94	152.91	128.33	95.88	107.59	164.90
3-4	44.97	79.67	61.49	86.96	60.28	85.13	61.32	65.80	47.90	81.78
4-5	22.06	51.39	30.69	55.74	30.22	55 69	38.41	53.31	23.70	52 59
5-6	12.48	32.00	19.76	39.98	17.53	39.27	22.64	38.78	9.49	25.85
6-8	8.32	34.95	20.31	55.28	19.01	53.40	26.31	55.92	10.01	34.22
8-10	3.88	17.62	10.44	35.62	9.87	36.18	14.68	40.45	4.89	21.74
10-15	3.85	23.23	11.65	52.12	10.44	51.66	16.79	64 23	5.55	33.42
15-20	1.38	11.89	4.71	28.74	3.89	25.77	6.99	37.71	2.38	20.41
20-40	1.24	16.71	5.98	62.73	4.11	43.81	9.72	84.17	2 89	39.16
40-60	0.25	5.96	1.16	20.78	0.88	17.17	2.82	42.52	0.70	16.79
60-100	0.11	3.46	0.60	16.88	0.50	15.56	1.91	45 18	0.39	14.64
100-150	0.02	1.15	0.17	8.06	0.17	8.85	0.73	27.62	0.12	7.08
150-300	0.01	1.25	0.09	7.92	0.11	10.22	0.56	35.47	0.06	5.74
300-1000	—	—	0.02	4.78	0.05	15.69	0.19	27.34	0.02	4.51
oltre 1000	—	—	0.00	1.47	—	—	0.02	11.18	0.00	2.06
	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00

Italia sono poco numerosi coloro che posseggono redditi elevati e i redditi rispettivi rappresentano una frazione del totale dei redditi superiore a 1.000 dollari più piccola che negli altri Stati.

Ma, la bassissima concentrazione dei redditi italiani diviene ancor più manifesta se si tien conto dei redditi compresi fra il decuplo del minimo di esistenza e 1.000 dollari. Il numero dei redditieri accertati con redditi compresi fra 480 e 1.000 dollari è poco meno che quadruplo del numero di coloro che posseggono redditi superiori a 1.000 dollari e l'ammontare dei rispettivi redditi è 1,3 volte maggiore. Naturalmente, non si possono istituire confronti con gli altri Stati, perchè, dato che il minimo di esistenza è diverso dall'uno all'altro, la classe di redditi compresa fra il decuplo del minimo e 1.000 è diversamente ampia. Sarebbe possibile confrontare fra loro solamente la Francia ed il Belgio, ma le differenze fra i due Stati non sono — almeno per ciò che si riferisce ai redditi non troppo alti — notevolmente sensibili.

6. — Un metodo sintetico che serve per valutare l'importanza della diversa concentrazione dei redditi sul gravame fiscale di due o più Stati è il così detto *metodo dei coefficienti tipo* (1).

Il gravame fiscale sopportato da due o più Stati è certamente lo stesso quando, in ciascuno di essi, i redditi di uguale entità sono colpiti con le stesse aliquote di imposta.

Orbene, supponiamo che uno stesso sistema di aliquote progressive di imposta — o aliquote tipo — colpisca i redditi di più Stati. Il prodotto totale presunto dell'imposta di ciascuno Stato, diviso per il totale dei redditi che si suppongono soggetti all'imposta stessa, darà un quoziente diverso da Stato a Stato; e ciò perchè, quantunque l'imposta sia immaginata in modo da rappresentare per tutti gli Stati lo stesso gravame, influisce, sul suo gettito, la diversa concentrazione dei redditi colpiti. La percentuale del reddito colpito, assorbita dall'imposta, sarà minore negli Stati nei quali, prevalendo, più che altrove, i redditi bassi, hanno maggior giuoco le aliquote di imposta minime, e sarà proporzionalmente maggiore là dove prevalgono relativamente i redditi elevati, ed hanno così maggior giuoco le aliquote elevate.

---

(1) Così chiamato dal GINI ed applicato alla distribuzione della ricchezza già in: *Appunti di statistica raccolti dalle lezioni del prof.* GINI *dallo studente* ETTORE BITTARELLO, Padova, 1913-14.

Per una esposizione sistematica del metodo dei coefficienti tipo, vedasi: C. GINI, *Quelques considérations au sujet de la construction des nombres indices des prix et des questions analogues.* « Metron », Vol. IV, 1924, p. 28 segg.

Cosicchè, mediante il procedimento delle aliquote tipo, applicate alle distribuzioni dei redditi di alcuni Stati, si mette in evidenza, in maniera sintetica:

a) il diverso rendimento, nei vari Stati, di una imposta ugualmente gravosa, dovuto alla diversa concentrazione dei redditi in ciascuno di essi;

b) il diverso gravame rappresentato nei vari Stati in cui i redditi hanno una differente concentrazione, da una imposta che assorba una uguale percentuale dei redditi tassati;

c) la necessità di applicare in uno Stato *A*, nel quale la concentrazione dei redditi è relativamente bassa, aliquote più gravose che in uno Stato *B*, nel quale la concentrazione dei redditi è più elevata, quando si voglia prelevare in entrambi la stessa percentuale del reddito imponibile;

d) che uno Stato *A*, il quale preleva una percentuale del reddito tassabile *n* volte più elevata di quanto non faccia uno Stato *B*, in realtà sottopone i contribuenti a uno sforzo fiscale più che *n* volte maggiore se i redditi sono meno concentrati in *A* che in *B*, e meno che *n* volte maggiore se i redditi sono più concentrati in *A* che in *B*.

Abbiamo supposto che, nei cinque Stati ex-alleati, i redditi accertati siano stati sottoposti alle stesse percentuali di imposta che colpivano i redditi americani nel 1920.

Si otterrebbe allora un certo gettito, il quale può essere logicamente ragguagliato, tanto all'ammontare dei redditi colpiti, cioè superiori a 1.000 dollari, quanto all'ammontare dei redditi accertati superiori al decuplo del minimo di esistenza, gli uni e gli altri considerati al lordo oppure al netto del minimo di esistenza dei redditieri.

Il rapporto del gettito dell'imposta all'ammontare dei redditi superiori a 1.000 dollari al lordo del minimo di esistenza dice la reale falcidia subita con l'imposta dai redditi tassati, e l'analogo rapporto all'ammontare di detti redditi al netto dice quale porzione della parte di reddito che misura la capacità contributiva dei redditieri è prelevata dal fisco.

Più importanti, sotto il punto di vista dei confronti internazionali, sono i due rapporti fra il gettito dell'imposta e l'ammontare lordo e netto dei redditi che superano il decuplo del minimo di esistenza. Il primo dice quale falcidia, mediante l'imposta, opera lo Stato sui redditi di coloro che, tenuto conto delle diverse condizioni e del diverso tenore di vita fra Stato e Stato, hanno un reddito superiore a un minimo che, nei vari Stati in cui il reddito è percepito, può soddisfare gli stessi o analoghi bisogni, e il secondo dice quale porzione della parte disponibile di tali redditi — la quale misura la capacità contributiva dei loro possessori — è prelevata dal fisco.

Facendo i rapporti in conformità di questo secondo criterio, interessa anche di conoscere, per i vari Stati, quale frazione dei possessori di redditi superiori al decuplo del minimo di esistenza risulterebbe esonerata dall'imposta, se in ciascuno venissero adottate le percentuali americane, e quale frazione i redditi rispettivi rappresentano sui redditi totali accertati.

I risultati dei calcoli eseguiti in conformità dei precedenti criteri figurano nella tabella V.

Qualunque sia il criterio adottato nel calcolo della percentuale del reddito prelevata mediante l'imposta, il risultato mette sempre in luce una cifra minima per l'Italia, una cifra massima per il Regno Unito e cifre intermedie per gli altri Stati.

Applicando una imposta progressiva ugualmente gravosa nei cinque Stati, se ne otterrebbe un gettito minimo in Italia, appunto perchè, essendovi i redditi meno concentrati, hanno maggior giuoco le aliquote più basse, che colpiscono i redditi minimi. Per apprezzare meglio al giusto valore i risultati conseguiti dalle ricerche effettuate, limitiamoci a considerare le cifre della colonna 10 della tabella V, che, a quanto abbiamo detto, devono essere considerate come le più significative, per i confronti internazionali, se esse indicano, in realtà, quale porzione della parte disponibile dei redditi superiori a un minimo, che nei paesi in cui i redditi sono percepiti dà analoghe soddisfazioni, è prelevata dal fisco.

7. — Nella tabella VI, le cifre della colonna 10 della tabella V sono ridotte in numeri indici, adottando come base, di volta in volta, la percentuale di rendimento in ciascuno Stato di una imposta progressiva sul reddito ugualmente gravosa in tutti gli Stati.

Il significato delle cifre, del resto evidente, può essere facilmente spiegato mediante un esempio. Prendiamo la seconda cifra della quarta colonna. Essa dice che un'imposta sul reddito ugualmente gravosa nel Regno Unito e in Francia, rende nel primo stato 2,8 volte di più che nel secondo, e il suo reciproco, cioè la quarta cifra della seconda colonna, dice che per prelevare la stessa percentuale del reddito imponibile nei due Stati, basta nel Regno Unito sottoporre il reddito imponibile ad un gravame che è appena 0.36 volte quello corrispondente della Francia. Queste differenze dipendono

Tabella V. — *Risultati di una imposta sul reddito ugualmente gravosa in cinque Stati e basata sulle percentuali di reddito prelevate negli Stati Uniti nel 1920 con la « Income Tax » e la « Surtax ».*

STATI	Minimo di esistenza in dollari	Percentuale dei redditieri accertati con redditi compresi fra il decuplo del minimo di esistenza e 1000 dollari.	Percentuale dei redditi accertati compresi fra il decuplo del minimo di esistenza e 1000 dollari.	Rapporto percentuale del gettito della imposta ai redditi accertati superiori a 1000 dollari — al lordo del minimo di esistenza		Rapporto percentuale del gettito della imposta ai redditi accertati superiori a 1000 dollari — al netto del minimo di esistenza		Rapporto percentuale del gettito della imposta ai redditi accertati che superano il decuplo del minimo di esistenza — al lordo del minimo di esistenza		Rapporto percentuale del gettito della imposta ai redditi accertati che superano il decuplo del minimo di esistenza — al netto del minimo di esistenza	
				%	Percentuali degli S. U. = 100	%	Percentuali degli S. U. = 100	%	Percentuali degli S. U. = 100	%	Percentuali degli S. U. = 100
	1	2	3	4	5	6	7	8	9	10	11
Italia.	48	78.7	57.3	1.36	47.4	1.39	46.0	0.58	20.2	0.61	20.2
Francia	60	56.3	28.7	3.61	125.8	3.70	122.4	2.57	89.5	2.68	88.7
Belgio	60	57.7	30.4	3.99	138.0	4.10	135.8	2.78	96.9	2.90	96.0
Regno Unito	75	41.0	14.8	8.53	297.2	8.74	289.4	7.27	253.3	7.53	249.3
Stati Uniti	100	—	—	2.87	100.—	3.02	100.0	2.87	100.—	3.02	100.—

TABELLA VI. — *Rapporti percentuali fra le percentuali d'imposta, rispetto al reddito netto dei vari Stati superiore al decuplo del minimo di esistenza, degli Stati indicati in testa delle colonne per le corrispondenti percentuali degli Stati indicati nelle righe.*

	Italia	Francia	Belgio	Regno Unito	Stati Uniti
Italia	100	493.3	475.4	1234.4	495.1
Francia	22.8	100	108.2	280.9	112.7
Belgio	21.0	92.4	100	259.1	104.1
Regno Unito	8.1	35.6	38.5	100	40.1
Stati Uniti	20.2	88.7	96.0	249.3	100

appunto dalla diversa concentrazione dei redditi nei due Stati. Se i redditi fossero quasi ugualmente concentrati, infatti, i detti rapporti si aggirerebbero intorno all'unità. È quanto avviene, approssimativamente, per i tre Stati Francia, Belgio e Stati Uniti. I rapporti percentuali fra le percentuali di imposta dei tre Stati si aggirano infatti intorno a 100. Poichè, tenendo conto del differente minimo di esistenza, i redditi non sono troppo diversamente distribuiti in Francia, Belgio e Stati Uniti, un'imposta ugualmente gravosa dà in questi tre paesi press'a poco lo stesso gettito percentuale, e per prelevare in ciascuno una eguale percentuale del reddito imponibile basta sottoporre i redditi rispettivi press'a poco allo stesso gravame fiscale.

Ben altrimenti avviene quando si instituisca il confronto tra Francia, Belgio e Stati Uniti da un lato e Italia dall'altro; fra gli stessi tre Stati e il Regno Unito; fra l'Italia e il Regno Unito. Il fatto principale che risalta da siffatti confronti è la posizione di sfavore in cui si trova l'Italia a causa della bassissima concentrazione dei redditi italiani. *Una imposta progressiva sul reddito ugualmente gravosa nei vari Stati, rende in Italia poco più di* $\frac{1}{5}$ *di quello che rende in Francia, in Belgio e negli Stati Uniti e rende meno di* $\frac{1}{12}$ *di quello che rende nel Regno Unito; per prelevare in Italia, mediante una imposta progressiva, la stessa percentuale del reddito imponibile, occorre sottoporre i redditi accertati ad uno sforzo quasi cinque volte maggiore che in Francia, Belgio e Stati Uniti, e oltre 12 volte maggiore che nel Regno Unito.*

Questa conclusione è particolarmente importante per la retta interpretazione delle cifre contenute nella penultima e nell'ultima colonna della tabella I. Se il fisco preleva in Italia, sotto tutte le forme, una parte del reddito nazionale al netto del minimo di sussistenza della popolazione 1.3 — 1.4 volte maggiore che in Francia e che nel Regno Unito; 2.9 volte maggiore che nel Belgio e negli Stati Uniti, anche prescindendo dalla minore stabilità o dalla più sfavorevole composizione del reddito nazionale italiano, non si potrebbe concludere che il gravame fiscale è in Italia 1.3 — 1.4 volte più forte che in Francia e negli Stati Uniti, o 2.9 volte più forte che nel Belgio. *Il gravame fiscale sopportato dall'Italia è più forte di quanto non appaia da dette cifre, perchè la bassa concentrazione dei redditi degli italiani richiede uno sforzo più che proporzionale che negli altri Stati per prelevare mediante le imposte una stessa percentuale del reddito nazionale.*

8. — Le conseguenze di questa situazione di fatto si fanno notevolmente risentire sulla organizzazione fiscale dei vari Stati. Se in Italia bisogna prelevare, mediante imposte, percentuali del reddito più elevate che negli altri paesi; e, se la concentrazione dei redditi tende a deprimere, rispetto agli altri Stati, il rendimento delle imposte, in Italia è necessario escogitare tutti i mezzi, anche vessatori, per vincere siffatta influenza e per assicurare allo Stato e agli Enti locali i mezzi necessari per la vita di un grande paese civile.

Esamineremo partitamente i principali di detti mezzi e i loro effetti sul gravame fiscale sopportato dai cittadini, i quali tendono tutti ad assicurare allo Stato e agli Enti locali una parte del reddito persino dei contribuenti più poveri.

Cominciamo col confrontare, desumendoli da fonti ufficiali, i minimi di esenzione dalle principali imposte dirette in vigore nei diversi Stati.

Tabella VII. — *Minimi di esenzione delle imposte dirette per un contribuente ammogliato con tre figli* (*dollari*).

Denominazione delle imposte	Italia	Francia	Belgio	Regno Unito	Stati Uniti
Imposte sui terreni e sulle case . .	zero	zero	zero	—	—
Redditi di capitale	zero	zero	zero	1515	
Redditi industriali e commerciali .	22	83			
Redditi professionali, salari e pensioni	26	310-360	120-359	1683	3.500
Salari degli impiegati di Stato .	zero				
Salari di impiegati pubblici . . .	32				
Supertassa . . .	244	826	120-359	9620	

Quasi tutti i minimi imponibili sono più bassi in Italia che negli altri Stati e le differenze sono talora veramente fortissime.

Ma il confronto diretto delle cifre può dare un'idea fallace, se non si tien conto di alcuni elementi che ne debbono aiutare l'interpretazione. Li schematizziamo assai rapidamente.

a) In Italia, le imposte, salvo la imposta complementare, non ammettono minimi imponibili diversi, a seconda dei carichi di famiglia del contribuente; cosicchè, se per i contribuenti singoli le differenze fra Stato e Stato sono inferiori a quelle che risultano dalla tabella, esse diventano più grandi quando si faccia l'ipotesi di contribuenti con famiglia numerosa.

b) In Italia, a differenza degli altri paesi, il contribuente, che possegga diverse categorie di reddito mobiliare, paga l'imposta anche sulla categoria del suo reddito che non raggiunge il minimo imponibile.

c) Generalmente, il minimo imponibile viene dedotto dal reddito dei contribuenti per ricavare il reddito tassabile (abbattimento alla base). Fra gli Stati considerati fa eccezione a questa regola solo l'Italia, dove il contribuente con un reddito che sorpassa appena il minimo imponibile paga l'imposta sul reddito intero. In altri termini, in Italia non esiste il minimo di esenzione dalle imposte per i contribuenti che posseggono redditi superiori a quelli indicati nella tabella VII.

d). La maggior parte delle imposte dei vari Stati prevedono detrazioni dal reddito imponibile per ricavarne il reddito tassabile giustificate da ragioni di famiglia o altre. Siffatto genere di detrazioni, nel caso dell'Italia, è attuato solo per la imposta complementare.

Queste considerazioni bastano a convincere che realmente l'Italia colpisce con le imposte dirette redditi molto più bassi degli altri Stati, e che se, nei comuni più piccoli del Belgio, il limite di esenzione della supertassa è inferiore a quello dell'Italia, tale inferiorità, quando si tenga conto dell'insieme delle imposte e del loro meccanismo di applicazione, è più apparente che reale.

9. — Se le imposte dirette giungono, in Italia, a colpire anche redditi molto bassi, le aliquote di tassazione sono nel nostro paese particolarmente elevate.

Qualche esempio concreto giova per convincersene meglio di una lunga discussione.

Sapponiamo alcune categorie di redditi commerciali, o professionali, e di salari e vediamo le percentuali di imposta che li colpiscono in Italia (Ricchezza mobile e imposta complementare) e nel Regno Unito (*Income Tax* e *Super Tax*) (1). Le cifre dello specchio della pagina seguente sono calcolate supponendo 1 sterlina come equivalente a 120 lire e considerando per i due paesi dei contribuenti che non fruiscono di detrazioni per carichi famigliari.

Le percentuali italiane, per le categorie di reddito previste negli esempi, sono costantemente superiori alle corrispondenti percentuali britanniche. È bensì vero che la Supertassa inglese è progressiva e raggiunge percentuali elevatissime, ma essa colpisce solamente i redditi sopra 2000 sterline, cioè sopra 240 000 lire italiane, e non raggiunge le aliquote massime che sui redditi di 30 000 sterline, pari a

---

(1) L'esempio è tratto dal *Memorandum drafted by the General Confederation of Italian Industries, about Italy's Burden of Taxation, with special Reference to Industry and Trade as an Answer to the Federation of British Industries*, annesso alla memoria di G. Borgatta. *The fiscal Burden upon the Italian Joint Stock Companies*. Roma, Libreria dello Stato, 1925, p. 58. I calcoli sono però stati completamente rifatti, tenendo conto del corso della sterlina nell'autunno del 1925 e rettificando le aliquote dell'imposta complementare italiana.

TABELLA VIII. — *Imposte su alcuni redditi guadagnati tipo in Italia e nel Regno Unito.*

	Imposte sul reddito		
	commerciale	professionale	salario, assegno, pensione
Lire Sterline 200 = Lire 24.000			
Italia	19.4 %	17.8 %	13.9 %
Regno Unito	zero	zero	zero
Lire Sterline 500 = Lire 60.000			
Italia	20.5 %	18.5 %	14.8 %
Regno Unito	7.3 %	7.3 %	7.3 %
Lire Sterline 1000 = Lire 120.000			
Italia	21.2 %	19.3 %	15.6 %
Regno Unito	13.8 %	13.8 %	13.8 %

3.600.000 lire italiane. In Italia, l'imposta complementare è graduata dall'uno al 10 %, ma comincia a colpire i redditi da 3000 lire in su (25 sterline) e raggiunge il massimo prima del milione di lire (circa 8.300 sterline).

10. — Solo perseguendo fin i più piccoli redditièri e colpendoli con aliquote d'imposta molto elevate il fisco italiano riesce a prelevare le elevatissime percentuali del reddito privato di cui ha bisogno per fronteggiare le già immense e ognor crescenti spese pubbliche.

Cercheremo in concreto di renderci conto dei diversi criteri con cui vengono colpiti i redditi in Italia, nel Regno Unito e negli Stati Uniti e quale è il risultato fiscale delle tassazioni. Supponendo una distribuzione dei redditi uniforme nei tre paesi ed applicando a tali redditi le percentuali d'imposta che realmente li colpirebbero in ciascuno si ottiene un rendimento fiscale il quale, ridotto in percentuale del reddito imponibile totale, dà un idea sintetica delle differenze fra le serie dei coefficienti.

Questo procedimento corrisponde al metodo classico della popolazione tipo, che il GINI ha generalizzato sotto la denominazione di *metodo della composizione tipo* (1).

Come tipo, assumeremo la distribuzione reale dei redditi italiani di categoria C_1 di cinque province rappresentative, classificate me.

(1) C. GINI, *Quelques considérations, etc.* cit., pp. 28-37.

diante apposito spoglio dell'elenco ufficiale dei contribuenti privati possessori di redditi incerti e variabili eseguito dalla Direzione generale della Statistica.

Per ciò che concerne l'Italia e il Regno Unito, sono state utilizzate le percentuali d'imposta calcolate dal prof. EINAUDI per alcuni redditi tipo (1) applicando ai redditi di ciascuna classe la media delle percentuali che colpiscono il reddito minimo e il reddito massimo della classe stessa. Per gli Stati Uniti, ho calcolato le percentuali in un modo analogo, tenendo conto dell'*Income Tax* e della *Surtax* (2). In ogni caso, si è immaginato un contribuente capo di una famiglia normale, composta cioè dei genitori e di tre figli (3).

Il confronto fra i risultati conseguiti con le imposte italiane e quelli che si sarebbero ottenuti con le corrispondenti imposte inglesi e americane è dei più contrastanti. Il fisco italiano preleva già sui redditi minimi percentuali d'imposta uguali a quelle che il fisco inglese preleva su redditi quasi cento volte più elevati e che il fisco americano preleva su redditi quasi duecento volte più grandi. I redditi fra 60.000 e 100.000 lire sono colpiti in Italia con imposte 2,6 volte più forti che nel Regno Unito e 9 volte più forti che negli Stati Uniti. I risultati di queste differenze si compendiano in poche cifre. In Italia, i redditi accertati pagano in media più del 15 per cento, mentre pagherebbero, col sistema inglese, 0,33 per cento e col sistema Americano, 0.13 per cento. In Inghilterra e in America, su 8742 contribuenti sarebbero sottoposte all'imposta solo 20 persone, e appena il 4,4 per cento del reddito totale accertato verrebbe colpito dal fisco. La percentuale media di tassazione italiana è rispettivamente 46 volte e 117 volte più elevata che quella ottenuta sugli stessi redditi nel Regno Unito e negli Stati Uniti. Se supponiamo nei tre Stati esente dall'imposta la parte dei redditi professionali che rappresenta il minimo di esistenza personale dei contribuenti, le percentuali effettive o ipotetiche della tassazione italiana, inglese e americana diventano rispettivamente 18.32 % 0.39 % e 0.16 % e i rapporti fra la prima e le altre due diventano 47 e 115.

(1) L. EINAUDI, *Il bilancio inglese. Osservazioni e confronti italiani.* « Corriere della Sera », 6 maggio 1925.

(2) L'*Income Tax* colpisce col 4 % i primi 4000 dollari di reddito al disopra del limite di esenzione e con l'8 % i successivi. La *Surtax* è invece progressiva, colpisce solo i redditi sopra 6000 dollari e si calcola con una apposita tabella. Vedasi: R. H. MONTGOMERY, *Income Tax Procedure, 1924.* New York, 1924, p. 192. Se nella tabella i redditi da 60.000-100.000 lire figurano colpiti in America solo col 2 % si è perchè in realtà i redditi di 60.000 lire non pagano nulla e quelli di 100.000 lire pagano il 4 %: in media, dunque, calcoliamo il 2 %.

(3) Il minimo imponibile per una famiglia normale è stimato a $ 3500 secondo: *Tax Burden and Public Expenditure.* « National Industrial Conference Board », New York, 1925, p. 28.

Tabella IX. — *Redditi di ricchezza mobile della categoria* C_1 *in cinque provincie italiane. Gettito reale delle imposte che li colpiscono in Italia e gettito presunto se i redditi fossero colpiti con le corrispondenti aliquote inglesi e americane.*

Categorie di redditi L. it.	Numero dei redditieri	Ammontare dei redditi L. it.	Imposta di R. M. e complementare		Imposte presunte secondo le aliquote inglesi (Income Tax e Supertax)		Imposte presunte secondo le aliquote degli Stati Uniti (Income Tax e Surtax)	
			%	L. it.	%	L. it.	%	L. it.
640-750	693	473.742	6.79	32.167	—	—	—	—
750-1.000	913	828.588	8.69	71.917	—	—	—	—
1.000-2.000	2.273	3.586.012	11.20	401.633	—	—	—	—
2.000-2.001	—	—	—	—	—	—	—	—
2.001-3.000	1.354	3.711.640	14.40	534.476	—	—	—	—
3.000-4.000	804	3.024 417	14.40	435.516	—	—	—	—
4.000-5.000	672	3.232.744	14.40	465.515	—	—	—	—
5.000-7.500	768	4.798.391	16.30	782.138	—	—	—	—
7.500-10.000	608	5.418.118	16.30	883.153	—	—	—	—
10.000-15.000	346	4.427.775	16.66	737.667	—	—	—	—
15.000-20.000	135	2.508.953	16.82	422.006	—	—	—	—
20.000-25.000	56	1.324.750	17.00	225.208	—	—	—	—
25.000-30.000	44	1.288 180	17.00	218 991	—	—	—	—
30.000-50.000	43	1.747.907	17.25	301.519	—	—	—	—
50.000-60.000	13	755.250	17.25	130 281	—	—	—	—
60.000-100.000	16	1.232.532	17.76	218.898	6.87	84.675	2.00	24.650
100.000-120.000	3	350.000	17.76	62.160	6.87	24.045	5.03	17.605
120.000-240.000	1	150.000	18.44	27 660	12.08	18.120	7.05	10.575
Totale (A) . . .	8.742	38.858.029						
Minimo esistenza		6.373.967						
Differenza (B)		32.484.062						
Totale imposta				5.950.905		126 840		52.830
Imposta $^0/_0$ rispetto ad A. .			15.31		0.33		0.13	
Imposta $^0/_0$ rispetto B. . .			18.32		0.39		0.16	

Queste cifre, è forse superfluo avvertirlo, non significano già che il fisco inglese e americano prelevano percentuali insignificanti dei redditi professionali, ma dicono solo che ne preleverebbero percentuali insignificanti se i redditi professionali fossero distribuiti nei rispettivi paesi così come lo sono in Italia.

La realtà è che nel Regno Unito e negli Stati Uniti è possibile esentare dall'imposta o colpire moderatamente i bassi redditi su cui il fisco italiano grava fortemente la mano, perchè in quei paesi i redditi molto elevati, che mancano in Italia, rappresentano una frazione cospicua del totale dei redditi e tollerano forti falcidie le quali compensano largamente la perdita fiscale causata dalle larghe esenzioni accordate ai contribuenti meno ricchi.

11. — Altre considerazioni gioveranno a porre sempre meglio in luce l'entità del gravame fiscale oggi sopportato dal contribuente italiano.

Secondo una distinzione generalmente accettata, le imposte si dividono in dirette e indirette, a seconda che abbiano carattere costante, durevole o almeno continuo, oppure che colpiscano, non qualità o possessi, ma circostanze, fatti particolari o atti permanenti. Una larga categoria di imposte indirette è costituita dalle imposte sui consumi. Negli ordinamenti fiscali odierni, per le imposte indirette si può dire che non esistono minimi di esenzione: esse gravano anche su coloro che percepiscono redditi piccolissimi, per il solo fatto che li consumano o li impiegano in determinati modi. È fatto riconosciuto che la scarsa concentrazione e la scarsa stabilità dei redditi sono due importanti elementi che determinano gli Stati a dare un grande sviluppo alle tassazioni indirette (1). D'altra parte, è pure evidente che, quanto più il largo sviluppo dato da uno Stato alle imposte indirette dipende dal fatto che fra i suoi cittadini prevalgono i redditi piccoli e poco stabili, tanto più elevato risulta il gravame fiscale da essi sopportato.

Queste considerazioni bastano a chiarire il significato delle cifre contenute nella tabella X. Esse indicano il rapporto fra le imposte indirette e il totale delle imposte erariali e locali dei vari Stati. Quantunque non pretendano ad una esattezza rigorosa, date le difficoltà che si incontrano nel classificare e nel rendere comparabili le imposte dei vari Stati, pure le differenze fra Stato e Stato sono abbastanza rilevanti perchè si possano ritenere significative.

(1) « Se in Prussia, scrive il NITTI, 95 81 % di tutta la popolazione ha redditi inferiori a 3000 marchi e 66.85 % ha redditi inferiori a 900 marchi, ed è in posizione non stabile, solo le imposte indirette possono costituire la base delle entrate ». V. *Principi di Scienza delle finanze*, Napoli, 1912, p. 312.

TABELLA X. — *Imposte indirette e imposte sui consumi in rapporto al totale delle imposte, (compresi i prodotti netti dei monopoli e delle imprese pubbliche) nei vari Stati.*

STATI	Percentuale sulle imposte erariali e locali in totale delle	
	Imposte indirette	Imposte sui consumi
Italia (1924-25)	58.3	46.6
Francia (1924)	61.7	28.0
Belgio (1924)	60.8	23.7
Regno Unito (1924-25) . . .	40.0	30.0

Il grande sviluppo dato in Italia alle imposte sui consumi da una chiara idea della notevole partecipazione che sono chiamate a dare alle spese pubbliche anche quelle classi di cittadini i quali percepiscono redditi minimi, e che probabilmente li spendono, interamente o quasi, per provvedere alla sussistenza propria e delle rispettive famiglie.

12. — Per apprezzare al giusto valore le ripercussioni delle imposte sui consumi sulle classi meno abbienti, il TIVARONI ha calcolato il gravame di imposte che colpisce in Italia la spesa settimanale di una famiglia operaia. In un memorandum preparato dell'onorevole CABRINI, del quale ho avuto comunicazione, i calcoli del TIVARONI sono stati rettificati e aggiornati in modo approssimativo al 1925, cercando di tener conto, fra l'altro, dell'effetto del dazio doganale sul grano di recente introduzione e di un piccolo consumo di bevande alcooliche.

TABELLA XI. — *Bilancio settimanale di una famiglia operaia (Torino). (Operaio ammogliato con tre figli).*

	Spesa	Ammontare delle imposte	%
Alimenti	142.86	28.42	19.9
Vestiario	27.12	4.29	15.8
Abitazione.	7.10	1.01	15.5
Calore e luce.	9.62	0.45	4.7
Tabacco e varie.	32.86	6.55	19.9
Imposte dirette (imposta complementare)	3.00	3.00	100.0
Totale . .	222.56	43.81	19.7

Sulla modesta spesa settimanale per il mantenimento di 223 lire, pari a 9 dollari, la famiglia operaia tipo italiana paga, almeno, dollari 1,8 d'imposte, di cui appena 12 centesimi di dollaro sotto forma di tributi diretti. Le cifre riferite rappresentano solo un minimo, perchè alcune imposte di importanza secondaria sono state trascurate nel computo.

13. — Possiamo ormai rapidamente raccogliere e sistemare le conclusioni suggerite dalle indagini precedenti. In Italia, la percentuale che le somme prelevate sotto qualsiasi forma di imposta erariale e locale rappresentano rispetto al reddito nazionale al netto del minimo di sussistenza della popolazione risulta: 1.3 volte più elevata che in Francia; 1.4 volte più elevata che nel Regno Unito; 2.9 volte più elevata che nel Belgio e negli Stati Uniti.

Queste cifre, tuttavia, non danno un'idea adeguata della pressione fiscale comparativa dell'Italia rispetto agli altri Stati. La pressione fiscale italiana è più grave di quanto quelle cifre lascerebbero credere:

I) perchè il reddito nazionale dell'Italia comprende una frazione più elevata che negli altri Stati di redditi agricoli, i quali sono molto soggetti a variare di anno in anno, e di redditi da puro lavoro, che sono più aleatori dei redditi da capitale. Le imposte riescono infatti più pesanti per i contribuenti che non possono contare sulla stabilità e sulla sicurezza dei propri redditi.

II) perchè i redditi italiani sono molto meno concentrati che i redditi degli altri Stati. Perciò, una imposta sul reddito, che si proponesse nei vari Stati di prelevare una eguale percentuale dei redditi tassabili, dovrebbe essere basata in Italia su aliquote in media 5 volte più gravose che non negli Stati-Uniti e 12 volte più gravose che non nel Regno Unito.

Cosicchè se con una data imposta si preleva in Italia, ad esempio, una percentuale del reddito tripla rispetto agli Stati-Uniti, la pressione fiscale che costituisce in Italia quella imposta non è solo tripla, ma è molto più grande.

Il fatto che in Italia i redditi presentano una media bassa ed una distribuzione relativamente uniforme fa sì che il fisco sia costretto, per assicurare il gettito necessario alla vita dello Stato, a non risparmiare alcuni redditi che in altri Stati appaiono infimi.

Perciò, in Italia, i limiti minimi di esenzione dalle imposte dirette sono quasi sempre più bassi che negli altri Stati; non si esonera dal pagamento dell'imposta il reddito inferiore al minimo imponibile che sia percepito insieme a redditi mobiliari di altre categorie ; i minimi imponibili non costituiscono una detrazione sui redditi tassabili come avviene altrove; ed esistono solo per l'imposta complementare le detrazioni per carico di famiglia così frequenti negli altri paesi.

Perciò, le aliquote iniziali delle imposte dirette italiane, che colpiscono generalmente redditi molto piccoli, sono più forti delle aliquote iniziali di altri Stati, che colpiscono redditi grandi, e rappresentano quindi un gravame fortissimo per il contribuente.

Perciò il sistema fiscale italiano dà una forte preponderanza alle imposte indirette e alle imposte sui consumi, col risultato di colpire anche i redditi inferiori ai minimi di esenzione dalle imposte dirette.

G. H. KNIBBS.

The Growth of human populations, and the laws of their increase.

The rapid rate at which the world's population has increased during the last century and a quarter, and the various difficulties of maintaining such a rate till the end even of the present century, have given an impulse to the study of the fondamental laws of population-increase. A very elementary supposition long ago indicated by VERHULST, and recently independently developed by Prof. PEARL, and examined bu Mr. UDNY YULE and others, has been advanced as substantially embodying the totality effect of the true law of increase. This paper proposes to shew that such a view of population growth is not justified, and it is believed that it would lead to an unsatisfactory estimate of the world's population limits.

In a restricted region, a group of organisms increasing in number under the impulse to reproduce themselves, the impulse acting with constant intensity, but its effectiveness at any moment being reduced in the ratio, that the further number still possible at that moment bears to the total number which can possibly live in the region in question, grows less and less rapidly, relatively. This diminution in relative growth occurs in such a way that the co-ordinates expressing its numbers, with time as the variable, conform to what has been called a « logistic » curve. Initially this curve is concave upwards in its middle region it is sensibly linear ; after that it is convex upwards. Going back to past time it approaches its lowest limit asymptotically: going forward it approaches its highest asymptotically; and it is symmetrical on either side of its middle point.

Attempts have been made recently to show that the growing population of the world substantially conform to this curve. That such a curve should characterise the mode of increase, is deduced from the fundamental assumption and can easily be derived by expressing the relation differentially and integrating the equation. It can possess the characteristics mentioned only when the inherent reproductive impulse is constant, excepting in so far as it is limited by opportunity of effectually acting, this effectualness being dependent upon the limitation of the possibility of future population.

The nature of the organism, and that of its environment, are supposed to be constant except that the restriction of the latter limits the result. This assumption, therefore, raises at once the question as to whether, taken in the mass, human beings can be regarded as organisms of essentially constant qualities, living in an environment which also is sensibly constant. Obviously merely minute changes in the nature and extent of oscillations about the average essential character of human beings, and minute oscillations in the nature of their environment, could of course, quite properly, be disregarded. In part, they would tend to cancel one another in any large summations, and the « logistic » curve would represent the average condition of things, it being understood that at any moment the actual numbers, etc. would only approximately conform thereto.

To anyone who considers for a moment the enormous length of time that man has been upon the earth, the fact that, with his powers of reproduction, he has not even yet attained in number to 2,000,000,000 shows that the fondamental assumption, set out above, cannot be even approximately true. In order to adequately appreciate the significance of the very low average rate of man's power to increase his numbers let us assume that he has already attained the number of 2,000,000,000, and assuming different periods for the development of such a number, estimate the rate of increase necessary to attain it, in differing periods of time. We then, obtain the following rather startling results:

For a single pair of human beings to attain to a population of 2,000,000,000 in the periods of time shewn hereunder, the percentages of increase per annum would have to be as follows:

In	2.000	years,	per cent.	per an.	1.04165
»	10.000	»	»	»	0.20745
»	100.000	»	»	»	0.02072
»	1.000.000	»	»	»	0.00207
»	10.000.000	»	»	»	0.000207

In other words if uhman beings had existed on the earth for only as much as 1,000,000 years, their rate of increase over all was slightly less than 21 persons per milion per annum. Such calculations show at what an extremely slow rate increase would have taken had it been uniform, and in view of the fact that, if the rate of increase of last century had been maintained for less than 2,000 years, the whole population of the world could have sprung from a single pair, one sees that the fundamental assumption of the « logistic » theory is unquestionably invalid, even if it happened to nearly represent more recent growths of population, which however it does not.

The traces still existing of past civilisation, suggest that cause profoundly affecting the rate of reproduction have existed, but of these is no adequate historical record. Either the earth has been subject to tremendous vicissitudes of which we know practically nothing, but of which there is perhaps geological evidence, or else the powers of self-destruction of the human being himself can be so evoked as to practically annihilate civilisations. Nor is it wholly impossible that there are cosmic disasters of which no historical or astronomical account remain, but which can be imagined from phenomena that have come to hand through surveys of the solar system and stellar universe now being made.

It may be remarked that the study of changes on the solar surface has not yet covered any great range of time. It is becoming increasingly clear, however, that climatological changes upon the earth are due very largely indeed to phisical alterations, changes of energy, etc., of the solar surface. Sufficient is not yet known to be able to say whether these cosmic factors have caused profound changes in man's power to increase in number or not, nevertheless it is quite possible that the rate of his increase has been very dependente upon great meteorological factors, the like of which are only beginning to be sistematically studied. Even so far as man's physical environment is concerned, there is every reason to believe that the very naive assumption underlying the development of the « logistic » curve is invalid. It probably is also invalid to suppose that his intrinsic nature is fundamentally unchanged. There is no doubt for example, that the accessions of knowledge, particularly during the 19 th Century, which came to man, were such as to enable him far more efficiently to exploit, to his advantage, the material resources of his environment. The ability to do this easily made it possible for a population expansion perhaps larger than that which has existed in historical times, for the civilisations of China, India, Babylon, Egypt, Peru and Mexico, without doubt did not attain to the enormous developments that have characterised more modern times.

One sees therefore that the vicissitudes of Nature, with which we are familiar through meteorological studies, may profoundly check human increase; the developments of agriculture, the powers of transportation which enable man to transfer such commodities as he needs to where they are most serviceable, and those accessions of knowledge which have made the earth more productive for him (as for example, through schemes of irrigation and fertilisation) have opened up possibilities for supporting a much larger number of human beings.

On the other hand it has to be borne in mind that since human desires have no intrinsic limit, we tend to surround ourselves with something more than is necessary for actually maintaining life, namely with luxuries, expressed not only in our foods, but also in our clothing and housing. Even compared with say two or three hundred years ago, the elaborations of civilisation are very striking. Thus some measure of human effort goes into the development of « luxury », which possibly could have gone into the development of increased amounts of food and the maintenance of off-spring.

Man also has greatly lengthened his life by better national and individual hygiene: this affects his economic power, efficiency, and his possible reproductive efficiency.

When all these things are borne in mind the inadequacy, for the study of human increase, of the assumptions leading to the derivation of the « logistic » curve are clearly recognised. When closely studied populations are found not to increase in the way supposed, but it is found that (a) by making the population-unit the number of the greatest possible population; (b) by making the unit of time different for the different series of results, and (c) by selecting a region of the whole curve which with these modifications most closely fits the actual numbers, the latter *seem* in fairly good agreement. This agreement is however fortuitous, not a rational conseguence of correct assumptions, and is not therefore an indication of the soundness of the hypothesis reviewed.

The question under review is not merely an academic one of no practical moment; it is already influencing the national policy of peoples whose governing classes are being influenced by the future outlook. Were territories can no longer provide the food-supplies for their inhabitants by means of home-production, they are dependent upon exchanges of commodities necessary to them, and therefore are at once influenced by the developments and attitudes of peoples in other territories and by their production. And in view of the fact that with the world's growing population food-production difficulties are already in evidence, it powerfully influences the question of migration, inasmuch as national, racial and linguistic differences and diverse social ideals operate strongly to prevent free movement.

We shall first consider briefly the evidence that the « logistic » type of development does not apply to human populations. For the United States the census-data from 1790 to 1920, give the position of the upper asymptote of the curve at somewhat less than 200 millions. If we accept this figure as the maximum possible population, and take the successive population expressed as a ratio of this subtracted from unity, it will give the factor to be multiplied into r, the assumed constant inherent impulse to increase, which however has been modified by the diminution of the furter possible increase, viz., from Y, originally, to $Y - y$. For the « logistic » theory is that

$$(dy/dt)/y = r\,(1 - y/Y)$$

y being the population at any time in question, Y the greatest possible population, and r the inherent impulse to increase. This last in the United States was about 36 per cent. per decennium or say 3.1226 per cent. per annum, that is 3.074847 per cent. at every moment; in other words $r = 0.03074847$. Calculating the instantaneous rate for each decennium, and referring these to the middle years thereof, then comparing them with the means of the istantaneous rates (based on the above value for r modified by the above formula), at the beginning and end of each decennium we get:

Istantaneous rates per 10.000 persons. U. S. A.

Year	1755	1805	1815	1825	1835	1845	1855
Actual rates	300	310	286	289	283	307	304
« Logistic » rates	300	298	294	290	284	277	265
Year	1855	1865	1875	1885	1895	1905	1915
Actual rates	304	204	263	227	188	191	139
« Logistic » rates	265	254	239	220	201	178	156

The steady diminution shewn in the supposititious « logistic » rates is not exhibited by the actual rates, notwithstanding that the 200 million was so taken as to fit them as well as possible. For example for 1845 the rate actually had increased from the original 300 to 307 instead of dropping to 277: it was as high as 304 instead of being 265 for 1855, and as high as 263 instead of being 239 for 1875.

These considerable irregularities in the rate of increase were exhibited in the successive values for the population of Australia. Again referring them to the middle year, and expressing them as rations of the population at the commencement of each decennium, they were as follow:

Decennial Ratios of Increase for Australia 1790-1920.

Year	1795	1805	1815	1825	1835	1845	1855
Ratios of increase . . .	2.48	2.23	2.89	2.09	2.71	2.13	2.83
Year	1855	1865	1875	1885	1895	1905	1915
Ratios of increase . . .	2.83	1.44	1.35	1.41	1.20	1.18	1.22

From 1870 to 1920 the quinquennial rates of increase can be very exactly found. They are:

Quinquennial ratios of Increase for Australia 1870-1920.

Year	1872 1/2	77 1/2	82 1/2	87 1/2	92 1/2	97 1/2	1902 1/2	7 1/2	12 1/2	17 1/2
Ratio of increase	1.15	1.18	1.21	1.17	1.11	1.08	1.07	1.10	1.12	1.09

The arithmetical average is 1,128 and the quinquennial ratio over all 1.12626.

The population numbers for England and Wales from 1801 to 1911 are also not well fitted by the « logistic » curve. From 1811 to 1871 they are *very closely* given by the parabola $10.14 + 1.745\ T_1 + 0.055\ T_1^2$ and then for the next four decades by the straight line $22{,}59 + 3.34\ T_2$, the unit of the times T_1 and T_2 being a decennium.

If we take the populations of France for the years shewn below, excluding Alsace-Lorraine from 1876 onward, the differences (in millions) do not at all support the « logistic » hypothesis.

Population differences. France 1801-1911.

Year.	Diff.	Year.	Diff.	Year.	Diff.
1801		1841		1876	
	1.76		1.12		1.80
1806		1846		1881	
	0.24		0.46		0.64
1811		1851		1886	
	0.65		0.26		0.17
1816		1856		1891	
	2.42		1.34		0.17
1821		1861		1896	
	—		0.05		0.42
1836		1866		1901	
	0.72		—		0.30
1841		—		1906	
	—		—		0.36
—		—		1911	

(The population have been obtained by small linear interpolations for the periods in question). Although a « logistic » curve can be drawn close to the curve of the successive population-figures, it does not well represent them, and to fit it to them is without meaning.

In this connection it is well to turn to an examination of the characteristics of population-growth in the United States. The view of F. A. Walker, the American statistician and economist, who had studied the question, was that the remarkable change that operated to limit the rate of growth of the people of the United States, had its origin in social and industrial changes. In 1790 there were only about 600,000 white families, few either very rich or very poor. At that time food was abundant, domestic tranquillity prevailed, and and both social traditions and religious beliefs encouraged fecundity. The land was but partially settled, and — it may be added — the people occupied a virtually unrestricted area. Up to 1850, particularly in the North, instead of mechanical labour being employed, the farm labourers consisted largely of young men, who worked for a few years to acquire sufficient ready money to marry. These conditions allowed the population to expand quite freely. But between 1840 and 1850 a change came. There was a rise in the standard of living, artificial « necessities » were multiplied, domestic service was extended, women were introduced into factory labour. Then in 1861 to 1865 came the war of secession, causing a defect — it was estimated — of 1,765,000 persons at the 1870 enumeration. Although there had been an enormous immigration, particularly from Ireland and Germany, in the successive decennia, viz.:

Immigration in the U. S. America 1820-1910.

Decennium ending	1830	40	50	60	70	80	90	1900	10
Millions of immigrants	0.143	0.599	1.713	2.593	2.314	2.812	5.246	3.844	7.754 ,

the old rate of growth was never re-established. What these facts reveal is that social and economic factors operate as very potent cause of variation in the rate of growth of populations. The percentage of urban population grew more slowly from 1790 to about 1831 than from 1831 to say 1880. Thus in the two periods the urban populations grew as follow:

$$1790 - 1831 \qquad 3.40 + 0.70t_1$$
$$1831 - 1880 \qquad 6.34 + 0.28t_2$$

in which t_1 are the years after 1790 and t_2 those after 1831. These were due to the change of the poeple in their industrial character.

Before considering the matter further it is well to consider the nature of the reproductive element in a population, viz., the women of fertile age and the things which ensure or hinder fertility. Obviously the potential reproductivity of a population depends upon the following elements, viz.:

(i) Its constitution according to sex and age;

(ii) The health of its individuals;

(iii) The possibilities of their response to the reproductive impulse;

(iv) The possible effectiveness of this response in securing a normal costitution according to sex and age.

For married women of all ages from 13 to the end of life the numbers per 1000 who bore from 0 to 20 or more children were found in Australia in 1911 to be as follows (1)

Frequency of wives bearing 0 to 20 children and over. — Australia, 1911.

No. of children	0	1	2	3	4	5	6	7	8	9	10
No. of wives	143	150	145	123	101	80	64	51	41	33	26

No. of children	11	12	13	14	15	16	17	18	19	20	Over 20
No. of wives	17	12	7	3	1.6	.8	.32	.15	.06	.02	.05

(1) G. H. Knibbs, *Math Theory of Population,* p. 336.

These results gave about the number 3.76 per wife, which is evidently less than the physiologically possible number for healthy wives. Thus the question of the measure of the response, (iii) above is an important element in the law of growth of a population. Moreover all social and economic factors which affect the average interval between the age of husband and that of wife affect the fecundity (1). Thus one sees that health, contraceptive knowledge and practice, affect the numbers of the issue, and these again are in turn affected by social traditions and ideals, despite the easy circumstances of a territory.

That changes of this last kind take place and that they have some relation to the productivity of a country is well revealed in the history of Australia from say 1860 to the present time. The following are quinquennial results:

Marriage, birth, and increase rates, etc. and production per head. Australia 1860-1923.

Periods	1860-4	1865-9	1870-4	1875-9	1880-4	1885-9	1890-4
Population (1)	1.233	1.474	1.722	1.995	2.347	2.835	?
Marriage (2)	8.56	7.67	7.12	7.25	7.90	7.86	3 273
Births (3)	42.54	40.28	37.59	35.59	35.08	35.30	6.82
Births p. marriage	4.97	5.25	5.28	4.91	4.44	4.49	33.34
Increase (4)	38.52	37.42	30.40	31.74	38.04	32.84	4.89
Production (5) £	23.4	28.5	30.7	31.4	30.4	29.9	24.50

Periods	1890-4	1895-9	1900-4	1905-9	1910-4	1915-9	1920-3
Population (1)	3.273	3.585	3.855	4.158	4.651	4.905	5.575
Marriage (2)	6.82	6.64	7.10	7.65	8.72	7 68	8.52
Births (3)	33.34	28.29	26.58	26.57	27.72	25.69	24.71
Births p. marriage	4.89	4.26	3.74	3.47	3.18	3.35	2 90
Increase (4)	24.50	16.84	13.54	17.04	28.34	13.14	20.4
Production (5) £	29.9	29.9	31.6	37.9	43.5	57.5	65.6

(1) Millions. (2) Per 1000 mean population. (3) Per 1000 mean population. (4) Per 1000 mean population. (5) Pounds stirling per head of population. The mean of the populations as at 31 Dec. 1861 and 31 Dec. 1862 are taken as the population at the middle of the 5-year periods, and similarly for each following quinquennium.

(1) See, *Diisogeny in Australia*, G. H. Knibbs, *Ibid.*, pp. 356-361 and the results of Körösi, *Budapest.* « Phil. Trans » Vol. 186, pt. II, pp. 781-875.

The period 1915 to 1919 is the war period and its affect is obvious.

The period 1920-3 is a four-year period only.

The ratios of the totals of the birth-rates to the totals of the marriage-rates shew a steady fall with the lapse of time, the general trend of which may be roughly represented by

$$5.47 - 0.21\ T$$

in which T denotes the number of quinquennia after 1860-4, or the number of years June 1862 divided by 5.

These results shew that the response in Australia to the marriage and reproductive impulses have diminished during the last 60 years, despite the increasing value of the production per head of population, and the easy going circumstances of life therein. The production had increased from 1860-4 to 1870-4; it then averaged about £ 30.1 for 30 years, when it again progressed at the rate of about $1.6T^2$, in which T denotes a 5-year period. We thus see that a population's economic production may be accompanied by influences which can rob it of its appropriate effect on the measure of increase. The development of secondary industries reinforced the urban populations, the percentage divisions in Australia on 4 th April, 1921. being as follows:
Metropolitan 43.01: Provincial 19.09; Rural 37.35; Migratory 0.55; Total 100. This large aggregation in the cities probably has had much to do with the limited rate of increase.

A study of the relative productive activity of Australia from 1871 to the present time (the figure being corrected by wholesale price-indexes up to 1907 and by production price-indexes from 1908 onwards), and a comparison with the increase, marriage, and birth-rates, and vith the ratios of the birth-rates to marriage rates, shews that very often fluctuations in the relative productive activity have been followed slightly afterwards by somewhat similar fluctuations in the rates of increase. On the whole the increase may be expected to become greater in the year following the increase of the relative productivity. This is shewn in the following figures, starting with 1891 for the productive activity (P. A.) and 1892 for the increase (Inc.) per 10,000.

Relative Productive Activity and Population-increase. Australia 1892-923.

P. A.	762	761	793	686	906	746	654	695	839	899	921	949	942
Inc.	200	193	176	140	158	132	107	147	148	145	172	170	217
P. A.	1062	1070	1000	984	1026	817	958	908	897	868	828	925	967
Inc.	234	336	378	310	159	— 5	— 104	130	198	438	203	181	225

The births and marriages per 100.000 ratios of the births to marriages from 1891 onwards were as follows:

Births and Marriages per 100,000 and their ratios, Australia, 1891-922.

B.	3447	3365	3085	2843	2727	2716	2671	2529	2641	2623	2657	2676	2659
M.	747	674	608	655	703	732	726	667	702	725	749	787	776
B÷M	4.60	5.00	5.07	4.34	3.88	3.71	3.68	3.80	3.76	3.62	3.55	3.40	3.43
B.	2669	2673	2720	2860	2815	2790	2706	2657	2627	2500	2353	2545	2495
M.	790	837	879	906	863	876	907	814	681	659	780	962	859
B÷M	3.38	3.19	3.09	3.16	3.26	3.19	2.99	3.26	3.86	3.79	3.02	2.65	2.90

The above results indicate that the economic factor acts definitely. In general favourable circumstances lead to an increase of marriage and birth-rates, but the development of « easy-going » tendencies may militate quite appreciably against the increase which otherwise would have taken place.

A factor which should be noted is that an improved hygiene has operated to preserve life during the years of higher economic efficiency. Tuberculosis is yielding thereto with most valuable economic results, since life is not terminated before economic efficiency is reached. It is true that cancer is increasing, but cancer is an old-age disease, and its baleful effects do not ordinarily appear till the age of highest economic efficiency has been passed.

It is necessary to bear in mind that in different localities and different peoples the reactions to advancing therapeutics and hygiene are very diverse. This is seen in the very different « average ages » attained by all born in a particular country (the expectation of life at birth). It is also seen in the great differences in the mortality of infants during the first 12 months of life (infantile mortality) and also in the death-rates of the populations. The next three Tables shew these.

It may be mentioned that the expectation of life at date of birth has been greatly increased during the last four decades, but as stated is very different in different countries. In the following Tables, it is given for males (M) and for females (F) in years:

Expectation of Life at Birth.

Country	Period	M. years	F.	Country	Period	M. years	F.
Australia	1920–2	59.16	63.29	U. S. America	1901–10	48.34	51.92
Denmark	1906–10	54.90	57.90	Belgium	1901–10	45.39	48.84
Norway	1901–10	54.82	57.70	Germany	1901–10	44.82	48.33
Sweden	1901–10	54.53	56.98	Italy	1901–10	43.60	44.40
Netherlands	1900–9	51.00	53.40	Finland	1891–1900	42.90	45.60
England & Wales	1901–10	48.53	52.38	Austria	1900–1901	37.77	39.87
France	1908-13	48.50	52.42	—	—	—	—

Results such as the above are of course reflected in the rates of infantile mortality, per 1000 births recorded, and also in the crude death-rates, per 1000 of the mean population. These are as follow, the number following the country being the year of the present century.

Infantile Mortality per 1000 births.

Country	Year	Rate	Country	Year	Rate	Country	Year	Rate
New Zealand	24	40	Scotland	23	79	Austria	20	157
Australia	»	57	Canada (1)	24	79	Italy	17	158
Norway	21	63	S. Africa (b) W.	23	82	Japan	22	166
Sweden	»	65	Canada (2)	24	86	Jamaica	»	177
Irish Free State	23	66	Finland	23	92	Ceylon	»	188
Netherlands	22	67	France	»	96	Hungary	21	197
Great Britain	23	70	Bulgaria	19	109	Roumania	20	231
Switzerland	21	74	Belgium	21	115	Russia in E.	09	248
England & Wales	24	75	Canada (3)	11	128	Chile		283
U. S. America (a)	21	76	Germany	»	134	Canada (1) excl. Quebec. (2) is for Ontario. (3) Quebec.		
Denmark	»	77	Spain	22	141			

(a) The results for the U. S. America are for the registration area only.
(b) The results for South Africa are for whites only.

Crude Death-rates per 1000 mean Population.

Country	Year	Rate	Country	Year	Rate	Country	Year	Rate
New Zeland	24	8.3	Great Britain	24	12.6	Czeco-Slovakia	22	17.8
Australia	»	9.5	Switzerland	22	12.9	Bulgaria	19	19.9
S. Africa W.	23	9.7	Finland	23	13.8	Spain	22	20.5
Canada (1)	24	9.8	Belgium	22	13.9	Hungary	»	20.8
Canada (2)	»	10.8	Canada (3)	21	14.2	Japan	»	22.3
Netherlands	22	11.4	Scotland	24	14.4	Jamaica	»	22.9
Sweden	23	11.4	Irish F. State	»	14.8	Ceylon	»	27.8
U. S. America	22	11.8	North Ireland	»	15.8	Russia in E.	09	28.9
Norway	»	11.9	France	23	17.0	Chile	23	32.8
Denmark	»	11.9	Austria	22	17.1	—		
England & Wales	24	12.2	Italy	21	17.5	—		

The effect of war on the population growth of the United States of America has already been noticed. Recently the birth rates were considerably affected in European countries. Thus the pre-war, war and post-war rates were about as follow :

	Prussia	Bavaria	Austria	Belgium	France	Britain	Switzerland
Pre-War	21.1	21.3	18.5	16.5	11.6	23.0	19.5
War	15.1	15.8	14.3	12.1	10.7	20.8	18.8
Post-War	24.2	26.5	23.0	21.3	20.5	22.8	20.9

Australia affords a striking example of the effect of a great economic disaster. From 1891 to 1902 the number of sheep was reduced from over 106 millions to 53.7 millions and, from 1894 to 1902 of cattle from over 12 millions to 7.1 millions. The population increase per 1000 per annum changed as follows :

From 31 December of	1883-1891	1891-1898	1898-1903
Increase per 1000 p. a.	32.7	17.7	13.9

Although not *wholly* attributable to economic losses it was mainly so.

A careful review of all the factors affecting population increase indicate the following as the main determining elements:

1. The reproductive potential of, and therefore the increase of, any population ultimately depends upon:

a) Its constitution in respect of sex ad age;

b) Its health at all ages, and particularly at the child-bearing ages of its females;

c) The natural resources available to it;

d) The efficiency of its knowledge of the means of exploiting such resources;

e) Its constant social traditions.

2. The effectiveness of the reproductive potential is uniform only so long as *a*) to *e*) remain unchanged, and for such time the law of increase is $P_t = P_0 e^{rt}$, r being the constant inherent impulse to increase expressed as an instantaneous rate.

3. In a new country with relatively very great resources, the increase of which arises from within itself, such increase is the measure of its effective reproductive impulse, all factors affecting it remaining constant.

4. Variations of social tradition and changes in economic facts may operate to increase or diminish the annual rate of increase.

5. Studied more minutely, the rate exhibits an annual periodicity (but this can ordinarily be ignored) because births and deaths have such periodicity.

6. The rate of increase can become greater in many ways; among others *a*) by the discovery of resources previously unknown or ignored; *b*) by the more skilful exploitation of the known resources; *c*) by social changes reducing the standard of living, so long as these do not hinder the reproductive efficiency of the population or the impulse to increase; *d*) by the maintenance of desire for offspring; *e*) by the improvement of the general health of the population or the normalising of their activities.

7. From the standpoint of incréase the elaboration of a standard of living is advantageous only so long as it increases normally the productive efficiency without impairing the reproductive impulse.

8. All unnecessary luxury lowers the possibilities of increase by deflecting human activities towards their production instead of leaving such activities available for food and other supplies under normal conditions.

9. Since thrifty and efficient living demands the exercise of intelligence and will, and since also these are essential to secure optimum conditions for reproduction, ethical factors have a far-reaching influence upon the rates of increase, and of this there is the clearest evidence.

10. The highest efficiency of the human units of a population presuppose the application of the results of a knowledge of heredity, in the reproduction of the human being himself, as regards *a*) his intellectual and mental powers generally; *b*) the normal development of his social characteristics; and *c*) the genius of his social-economic world-system.

11. Under any system which admits of an increasing world-population, there inevitably comes a time when the efforts to satisfy human needs will exhibit the law of diminishing returns.

12. So many factors of adjustment co-operate in the result, that the optimum scheme of meeting such needs cannot be stated *a priori*, nevertheless certain elements of the optimum conditions are discernible.

13. These optimum conditions are: *a*) the danger of hostile collisions shall be reduced to a minimum; *b*) constructive birth-control shall become a world-ideal and all populations shall submit to its discipline; *c*) such migrations shall occur as eliminate all unnecessary movement, and minimise the effort of response to human needs; *d*) every territory shall become an integral element in the world-effort; *e*) the disposition to dominate or take advantage by financial or military power shall give place to that of helpfulness in the whole issue.

When the real factors of population increase are considered, it is self-evident that the « logistic » theory cannot be regarded as satisfactory. The present decrease of the rate of advance in most countries, and perhaps particularly in the United States of America and perhaps also in Russia and the fact that Australia has not yet exhibited the rate of increase of 3 per cent, per annum, these things are due to factors which are *not* expressed by the law

$$r_t = r_0 (1 - y/Y).$$

This is not even empirically an expression which really represents the facts. It can fit them in a general way, and then only wery imperfectly, wen the units of t can be varied arbitrarily, the distance between the asymptotes, and the range of the curve may also vary arbitrarily. It then lends a false support to the notion that the logistic theory is confirmed.

Given a different distribution of things *essential* to the maintenance of human life and the satisfaction of human needs, many more people could live in the world than it now maintains: needless luxury, collisions of interest, and wanton destruction limit the numbers greatly. One has only to survey the facts, patent everywhere in the world, to see that humanity has not yet a definite answer to the question whether it is better that very large numbers live humbly, or a smaller number with a few living very luxuriously and many with needless luxury at the expense of great numbers.

It is however beyond doubt that the standards of living prevailing in some countries are such that they limit the possibilities of their population. For this and similar reasons any empirical fitting of the variations of populations with time do not represent the intrinsic laws of population-development, but merely reflect the vicissitudes of the natural resources, and of the ethical, social, and economic facts prevailing at particular periods in the variable time.

Direttore responsabile: CORRADO GINI.

Città di Castello - Società Anonima Tipografica «Leonardo da Vinci».

The International Review of Statistics METRON is published four times a year, the four numbers making a volume of 700 to 800 pages in all.

It accepts original articles on statistical methods and on the applications of statistics to the different spheres of activity, and reviews or discussions of results obtained by statistical method in various fields of science, or such material as may be of interest to the statistician. A bibliography is annexed of all works or Reviews presented or received in exchange.

Articles and reviews may be written in English, Italian, French or German. Manuscripts in English, French or German should be typewritten. Contributors wil receive free of charge 25 copies of their publications issued.

Manuscripts submitted for publication should be adressed to *Prof. Corrado Gini, Istituto di Statistica e Politica Economica, R. Università di Roma (Italy)*, or to the member of the Editorial Committee who represents the writers's country. Contributors are requested to retain one copy of each manuscript sent, as, in case of non acceptance, the Editors will not be responsible for the safe return of the original.

Proposals for exchange made by Reviews or other periodicals, and all publications sent in exchange, or as complimentary copies, should be addressed to Prof. Corrado Gini.

All applications of subscribers, as well as the sums for the subscriptions, are to be made payable to *Amministrazione del Metron, Istituto di Statistica. R. Università di Padova, Italy.*

The subscription rate for each volume is **20 sh.** (draft) in Europe and **5 dollars** (draft) in other parts of the world, post paid; singles copies **6 sh.** and **1 $^1/_2$ dollars** respectively, each post paid. For Italy and countries with less favorable exchange, the subscription rate is **100 it. lire** and for single copies **30 it. lire**, each post paid.

Die Internationale Statistische Zeitschrift METRON erscheint jährlich in 4 Heften im Gesamtumfang von 700-800 Seiten.

Die Zeitschrift veröffentlicht Originalaufsätze über die Methode der Statistik und die Anwendung der Statistik auf die verschiedenen Zweige der Wissenschaften, sowie Uebersichten und Erörterungen über die Ergebnisse der statistischen Methode auf den verschiedenen Wissenschaftsgebieten, soweit die für den Statistiker von Interesse sind. Sie enthält ferner ein Verzeichnis aller unentgeltlich oder im Austauschverkehr eingehenden Bücher und Zeitschriften.

Die zur Veröffentlichung eingesandten Aufsätze und Mitteilungen können in deutscher, italienischer, französischer und englischer Sprache verfasst sein. Deutsche, französische und englische Manuskripte müssen mit der Maschine geschrieben sein. Beiträge werden nicht honoriert. Jeder Verfasser erhält unentgeltlich 25 Sonderabdrücke seiner Arbeit.

Die Manuskripte, deren Veröffentlichung gewünscht wird, sind an Herrn *Prof. Corrado Gini, Istituto di Statistica e Politica Economica, R. Università di Roma (Italien)* oder an das Mitglied des Direktion-Komitees, das den Staat des Mitarbeiters vertritt, zu richten.

Die Verfasser werden gebeten, eine Abschrift des eingesandten Manuskriptes zurückzubehalten, da die Schriftleitung für den Fall, dass die eingesandte Arbeit nicht veröffentlicht wird, keine Gewähr für deren Rücksendung übernimmt.

Austauschanträge für andere Zeitschriften und alle Veröffentlichungen, die unentgeltlich oder im Austausch zur Verfügung gestellt werden, sind an Herrn Prof. Corrado Gini zu richten.

Die neuen Abonnements-Anfragen, sowie die Zahlungen für die Abonamentes, sind an *Amministrazione del Metron, Istituto di Statistica, R. Università di Padova (Italien)* zu richten.

Der postfreie Bezugspreis für jeden Band ist **20 sh.** (chèque) in europäischen Ländern, und **5 dollars** (chèque) in extra-europäischen Ländern, für das einzelne Heft **6 sh.** beziehugsweise **1 $^1/_2$ dollars.** Für Italien und die Länder mit schwächerer Valuta, **100 it. lire**, und **30 it. lire** für das einzelne Heft.

BIBLIOTECA DEL "METRON„ - "METRON„ LIBRARY
BIBLIOTHÈQUE DU "METRON„ - "METRON„ 'S BIBLIOTHEK

SERIE **A** — Problemi di attualità - Problèmes d'actualité Gegenwärtige Fragen

SERIES **A** — Problems of the moment

1. - A. ANDRÉADÈS - *La population anglaise avant, pendat et apres la grande guerre*

10 lires pour l'Italie et les pays ayant un change plus défavorable **5 Frs. suisses** pour les autres pays

SERIE **B** — Memorie scientifiche - Mémoires scientifiques Wissenchaftliche Arbeiten

SERIES **B** — Scientific Memoirs

1. - F. SCHINDLER - *Das Volksvermögen Voralbergs*

25 lires pour l'Italie **70.000 couronnes** pour l'Autriche

8 Frs. suisses pour la Suisse et les autres pays

2. - F. SAVORGNAN - *La scelta matrimoniale - Studi statistici*

12 lires pour l'Italie et les pays ayant un change plus défavorable

6 Frs. suisses pour les autres pays

3. - F. v. FELLNER - *Die Verteilung des Volksvermögens und Volkseinkommens der Länder der Ungarischen Heiligen Krone zwischen dem heutigen Ungarn und den Successions-Staaten*

10 lires pour l'Italie et les pays ayant un change plus défavorable

5 Frs. suisses pour les autres pays

4. - MARIO BALESTRIERI - *I costumi alimentari della popolazione italiana dal 1910 al 1921* con prefazione del Prof. CORRADO GINI

15 lires.

Gli abbonati del *Metron* che domandano *direttamente* all'Amministrazione le opere pubblicate nella *Biblioteca del «Metron»* ricevono uno sconto, sul prezzo di copertina, del 30%. Le spese di porto restano a carico dell'acquirente.

Les abonnés du *Metron*, qui commandent directement à l'Administration les ouvrages publiés par la *Bibliothèque du «Metron»* reçoivent un rabais de 30% sur le prix indiqués. Les frais de port restent à la charge de l'acheteur.

Those subscribers to the *Metron* who obtain directly from the Administration works published in the *«Metron» Library*, receive a discount, on the marked price, of 30%. The cost of carriage must be borne by the buyer.

Den Abonnenten der Zeitschrift *Metron* welche die von der *«Metron» 's Bibliothek* veröffentichen Werke daselbst beziehen, kommt ein Bonus von 30% des angeschlagen en Preises zugute.