Werk

Titel: Zu der von R. v. Mises gegebenen Begründung der Wahrscheinlichkeitsrechnung. 2. M...

Autor: DÖRGE, K.

Ort: Berlin

Jahr: 1936

PURL: https://resolver.sub.uni-goettingen.de/purl?266833020_0040|log13

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Zu der von R. v. Mises gegebenen Begründung der Wahrscheinlichkeitsrechnung.

Zweite Mitteilung:

Allgemeine Wahrscheinlichkeitstheorie.

Von

Karl Dörge in Köln.

Das Ziel meiner unter diesem Titel veröffentlichten Arbeiten[1]) ist es, unter möglichst weitgehender Aufrechterhaltung der beiden v. Misesschen Ansätze, das ist der Definition der Wahrscheinlichkeit als limes der relativen Häufigkeiten und des Prinzips des ausgeschlossenen Spielsystems, die Wahrscheinlichkeitsrechnung zu axiomatisieren.

Meine Idee ist die: Ich axiomatisiere nicht das einzelne Kollektiv, sondern den Bereich aller „Grundkollektivs" und die „zulässigen" Auswahlen durch Eigenschaften gegeneinander und leite aus den Grundkollektivs durch Operationen alle betrachteten Kollektivs ab. Insbesondere erreiche ich damit, daß ich definieren kann, welche Kollektivs unabhängig sind im Bereich aller Kollektivs und daß ich dann die Multiplikationsregel beweisen kann, ohne dies bei der Definition vorauszusetzen.

In der ersten als Theorie des Glückspiels betitelten Arbeit habe ich vorausgesetzt, daß jedes Grundkollektiv ohne Nachwirkung ist, d. h. daß der Ausfall an irgendeiner Stelle des Kollektivs von dem Ausfall an anderen Stellen unabhängig ist, und daß zwischen je zwei Grundkollektivs eine Abhängigkeit nicht besteht, d. h. daß der Ausfall in einem Kollektiv nicht den Ausfall in den anderen beeinflußt.

In der vorliegenden Arbeit zeige ich, daß sich meine Axiomatisierung auch dann durchführen läßt, wenn die Kollektivs eine Nachwirkung haben und wenn zwischen ihnen eine Abhängigkeit besteht. Viel Neues ist zu den Ansätzen der ersten Arbeit nicht hinzugekommen, vielmehr bildet diese das ganze Gerüst auch der allgemeinen Theorie.

Eine Kenntnis der Wahrscheinlichkeitsrechnung und der ersten Mitteilung ist für das Folgende nicht erforderlich.

Der Einwand, daß für die Theorie in meiner ersten Note ein nur triviales Beispiel angegeben wird, trifft auch für diese zweite Note zu. Ob die Theorie also jemals innerhalb der Mathematik angewandt werden

[1]) Die erste Mitteilung steht in Bd. 32 dieser Zeitschrift.

wird, ist nicht gewiß. Es besteht aber kein Zweifel daran, daß die Physiker von ihren — zu unendlichen idealisierten — Experimentalreihen statt meiner Grundkollektivs das Erfülltsein meiner Axiome annehmen, also sie dauernd anwenden.

Hiermit ist die von mir beabsichtigte Formalisierung der Wahrscheinlichkeitsrechnung mittels unendlicher Folgen abgeschlossen.

Ich hoffe, diese Theorie später zu finitisieren, da selbstverständlich jede Wahrscheinlichkeitstheorie mit unendlichen Versuchsreihen als Elementen nur als eine vereinfachte Theorie der in Wahrheit endlichen Folgen zu gelten hat.

Die vorliegenden Arbeiten behandeln die „Häufigkeitstheorie", das ist die Wahrscheinlichkeitstheorie, welche bei statistischen Untersuchungen gebraucht wird. Ich bin nicht der Meinung, daß die Wahrscheinlichkeitstheorie damit erschöpft ist, glaube vielmehr, daß für physikalische Probleme noch andere Ansätze benötigt werden. Diese finden sich in allen älteren Lehrbüchern. Ob man sie noch als mathematisch bezeichnen will, ist belanglos.

Was fehlt, ist der Übergang von den physikalischen Ansätzen zur Häufigkeitstheorie. Hier besteht das Anwendungsproblem. Ich werde eine teilweise Antwort auf dieses und auf das Induktionsproblem, d. i. die Frage: Warum induzieren wir mit größerem Recht aus einer größeren Anzahl von Beobachtungen als aus einer kleineren? mitteilen.

Um unwesentliche Weitläufigkeiten zu vermeiden, beschränke ich mich auf Kollektivs mit endlich vielen Merkmalen.

§ 1.

Merkmalfolge, Auswahl, Teilung.

Wenn man das Roulette einmal dreht, so ergibt sich als Resultat des Versuches eine Zahl aus der Reihe 0, 1, 2, ..., 36. Denkt man sich nun den Versuch unendlich oft ausgeführt, so hat man eine unendliche Folge von Elementen erhalten, von denen jedes eine Zahl aus der Reihe 0, 1, ..., 36 ist. Eine solche Folge werden wir als eine Merkmalfolge bezeichnen, die Ziffern (eindimensionalen Punkte) 0, 1, ..., 36 als die Merkmale, die Menge $\{0, 1, \ldots, 36\}$ als die Merkmalmenge. Nun ist die Annahme, daß unendlich oft gespielt sei, eine Idealisation, welcher im Konkreten, genau genommen, nichts entspricht. In einer mathematischen Theorie, wie sie hier entwickelt werden soll, tritt an Stelle der unendlich häufigen Wiederholung ein die unendliche Folge erzeugendes Gesetz. Wir definieren daher so: Definition: $\mathfrak{F}$ sei eine — gesetzlich bestimmte — unendliche Folge von Elementen: $a_1, a_2, \ldots$. Jedes sei ein Punkt aus der

endlichen Punktmenge $M = \{P_1, \ldots, P_l\}$. Dann heißt $\mathfrak{F}$ eine *Merkmalfolge*, jedes P_λ ein *Merkmal* und die Menge M die *Merkmalmenge* von $\mathfrak{F}$.

Die Wahrscheinlichkeitsrechnung beschäftigt sich nun mit der relativen Häufigkeit des Auftretens der Ziffern 0, 1, ..., 36 in der Folge der Roulettedrehungen, aber auch z. B. in der Folge, die entsteht, wenn man nur auf jedes 10^{te} Element achtet, also $a_{10}, a_{20}, \ldots$. Wir werden sagen, daß diese Folge entsteht durch Anwendung der Auswahl 10, 20, ... Demgemäß definieren wir: Definition: Unter einer *Auswahl* $\mathsf{A} = \alpha_1, \alpha_2, \ldots$ verstehen wir eine unendliche Folge wachsender ganzer positiver Zahlen: $\alpha_1 < \alpha_2 < \alpha_3 \ldots$.

Die Dichte der Auswahl A definieren wir so: $z(\nu)$ sei die Anzahl der α unterhalb ν. $\overline{\lim} \frac{z(\nu)}{\nu}$ ist dann die Dichte von A. A heißt dicht, wenn die Dichte von A größer als 0 ist, d. h. wenn es eine kleine positive Zahl η gibt, so daß immer wieder einmal $z(\nu) > \nu \cdot \eta$ ist.

Aus Gründen formaler Vereinfachung werden wir uns später, und zwar von einer Stelle an, wo dies noch ausdrücklich angegeben wird, auf Auswahlen beschränken, welche dicht sind.

Zwei Auswahlen $\mathsf{A} = \alpha_1, \alpha_2, \ldots$ und $\mathsf{B} = \beta_1, \beta_2, \ldots$ heißen *gleich*, wenn sie völlig übereinstimmen, d. h. $\alpha_\nu = \beta_\nu$ für alle ν. Weil es sich in der Wahrscheinlichkeitsrechnung aber immer nur um Grenzwerte von relativen Häufigkeiten handeln wird, wo endlich viele Elemente in den Merkmalfolgen keine Rolle spielen, ist der Begriff der Gleichheit von Auswahlen von geringer Wichtigkeit. Vielmehr heißen A, B *fast gleich*, wenn $\alpha_\nu = \beta_\nu$ für fast alle ν, A, B *fast total verschieden*, wenn $\alpha_\nu \neq \beta_\nu$ für fast alle ν und schließlich A, B *total verschieden*, wenn $\alpha_\nu \neq \beta_\nu$ für alle ν. Endlich heißen in dem einzigen noch übrigen Falle, das ist, wenn $\alpha_\nu = \beta_\nu$ für unendlich viele ν und gleichzeitig $\alpha_\nu \neq \beta_\nu$ für unendlich viele ν, die Auswahlen A und B *halb gleich* oder auch *halb verschieden*.

Man treffe nun aus der Folge der Roulettedrehungen nacheinander erst die Auswahl $\mathsf{A} = \alpha_1, \alpha_2, \ldots$ und dann aus dem Resultat die Auswahl $\mathsf{B} = \beta_1, \beta_2, \ldots$. Offenbar hat man dann im ganzen aus der Folge unendlich viele Stellen ausgewählt und zwar die Stellen mit den Indizes $\alpha_{\beta_1}, \alpha_{\beta_2}, \ldots$. Wir definieren daher: Definition: Unter dem *Produkt der beiden Auswahlen* $\mathsf{A} = \alpha_1, \alpha_2, \ldots$ $\mathsf{B} = \beta_1, \beta_2, \ldots$ verstehen wir die Auswahl $\alpha_{\beta_1}, \alpha_{\beta_2}, \ldots$. Diese bezeichen wir mit $\mathsf{A}\,\mathsf{B}$. Sukzessive ergibt sich dann als Operation, welche entsteht, wenn man beliebig viele Auswahlen $\mathsf{A}_1, \mathsf{A}_2, \ldots, \mathsf{A}_k$ nacheinander ausführt, wieder eine Auswahl, nämlich ihr Produkt, welches wir mit $\mathsf{A}_1\,\mathsf{A}_2 \ldots \mathsf{A}_k$ bezeichnen.

Bei dieser Multiplikation von Auswahlen gilt offenbar das assoziative Gesetz. Dagegen gilt nicht das commutative Gesetz. Es habeu z. B.,

wenn $\mathsf{A} = 2, 2^2, \ldots$, $\mathsf{B} = 3, 3^2, 3^3, \ldots$ ist, AB und BA sogar kein einziges gemeinsames Element, weil AB nur Ziffern aus A, BA nur solche aus B enthält.

Die Auswahl $\mathsf{E} = 1, 2, 3, \ldots$, die Identität, welche aus jeder Merkmalfolge die ganze Folge auswählt, spielt bei der Auswahlmultiplikation die Rolle der 1, weil offenbar für jedes A ist: $\mathsf{AE} = \mathsf{EA} = \mathsf{A}$.

Wir haben den Begriff Auswahl $\mathsf{A} = \alpha_1, \alpha_2, \ldots$ eingeführt, weil wir neben $\mathfrak{F} = a_1, a_2, \ldots$ auch die Teilfolgen $a_{\alpha_1}, a_{\alpha_2}, \ldots$ nötig haben. Wir bezeichnen diese Folge nun mit $\mathfrak{F}\mathsf{A} = a_{\alpha_1}, a_{\alpha_2}, \ldots$ und nennen $\mathfrak{F}\mathsf{A}$ die Merkmalfolge, welche *durch Anwendung der Auswahl* A *auf die Merkmalfolge* $\mathfrak{F}$ *oder durch Auswahl* A *aus* $\mathfrak{F}$ *entsteht.*

Man denke sich in einem Roulettesaal zwei Roulettes betrieben. Das erste liefert — idealisiert — die Merkmalfolge $\mathfrak{F}_1 = a_1, a_2, \ldots$. das zweite $\mathfrak{F}_2 = b_1, b_2, \ldots$. Die Wahrscheinlichkeitsrechnung hat sich dann u. a. auch dafür zu interessieren, wann z. B. gleichzeitig, d. h. an gleicher Stelle ν bei $\mathfrak{F}_1$ eine gerade Ziffer und bei $\mathfrak{F}_2$ irgendein vorgegebenes Resultat erscheint. Es handelt sich hier um die Betrachtung der Teilfolge, welche aus $\mathfrak{F}_2$ entsteht, indem man die Elemente herausgreift an den Stellen, an welchen in $\mathfrak{F}_1$ eine gerade Ziffer herauskommt. Das geschieht durch eine Auswahl aus $\mathfrak{F}_2$ und zwar eine solche, die durch eine andere Merkmalfolge definiert ist, physikalisch könnte man sagen, eine Auswahl, welche nicht durch ein arithmetisches Gesetz, sondern durch Naturereignisse bestimmt ist. Gleichwohl muß zur Vermeidung von Mißverständnissen betont werden, daß in unserer mathematischen Theorie andere als durch ein mathematisches Gesetz bestimmte Auswahlen nicht möglich sind, insbesondere ist auch die betrachtete Auswahl, da die Merkmalfolgen, insbesondere also $\mathfrak{F}_1$, durch ein mathematisches Gesetz bestimmt sind, durch eine mathematische Vorschrift charakterisiert. Etwas allgemeiner ist die entsprechende Betrachtung von $\mathfrak{F}_1\mathsf{A}$ und $\mathfrak{F}_2$, wo also zuerst aus $\mathfrak{F}_1$ eine Teilfolge herausgegriffen wird und diese Teilfolge dann benutzt wird, um aus $\mathfrak{F}_2$ auszuwählen. Die Auswahlen dieses betrachteten Typus sind so bestimmt: Definition: $\mathfrak{F} = a_1, a_2, \ldots$ sei eine Merkmalfolge, M ihre Merkmalmenge. $N < M$ eine Teilmenge von M. A sei eine Auswahl. Aus $\mathfrak{F}\mathsf{A} = a_{\alpha_1}, a_{\alpha_2}, \ldots$ greife man alle Elemente heraus, die zu N gehören. Sie mögen, nach den Indizes geordnet, sein: $a_{\alpha_{\gamma_1}}, a_{\alpha_{\gamma_2}}, \ldots$. Wir nehmen an, diese Folge möge nicht abbrechen. Dann ist $\mathsf{T} = \gamma_1, \gamma_2, \ldots$ eine Auswahl. Wir sagen, diese sei die *Teilung von* $\mathfrak{F}\mathsf{A}$ *auf* N und bezeichnen sie als $\mathsf{T}(\mathfrak{F}\mathsf{A}, N)$. Wir benutzen auch die Bezeichnung $\mathsf{T}(\mathsf{A}, \mathfrak{F}, N)$ und sprechen von der Teilung von A an $\mathfrak{F}$ auf N.

$\mathsf{T}(\mathsf{A}, \mathfrak{F}, N)$ besteht also aus den Stellen, an welchen $\mathfrak{F}\mathsf{A}$ ein Element aus der Teilmenge N von der Merkmalmenge M von $\mathfrak{F}$ hat.

Es ist nun unbequem, daß $\mathsf{T}(\mathfrak{F}\mathsf{A}, N)$ nicht für alle A, N definiert ist, sondern nur dann, wenn unbegrenzt viele Stellen $\gamma_1, \gamma_2, \gamma_3, \ldots$ herauskommen. Um diesen Ausnahmefall formal nicht ausschalten zu müssen, führen wir daher noch eine uneigentliche Auswahl Λ ein, indem wir festsetzen:

Definition: Wenn $\mathfrak{F}\mathsf{A}$ entweder gar keine oder nur endlich viele Elemente hat, welche zu N gehören, so ist bisher $\mathsf{T}(\mathfrak{F}\mathsf{A}, N)$ nicht definiert. Wir setzen $\mathsf{T}(\mathfrak{F}\mathsf{A}, N)$ dann gleich Λ und nennen Λ auch eine Auswahl, *die uneigentliche Auswahl.*

Die Definition des Produktes von Auswahlen ergänzen wir jetzt nach Einführung von Λ noch durch die Festsetzung: Produkte von Auswahlen, in denen der Faktor Λ vorkommt, sind gleich Λ zu setzen.

Sind $\mathfrak{F}_1$, $\mathfrak{F}_2$ zwei Merkmalfolgen, und bildet man $\mathfrak{F}_1' = \mathfrak{F}_1\mathsf{T}(\mathfrak{F}_2, N)$, so entsteht dies durch eine Auswahl aus $\mathfrak{F}_1$ und zwar sagen wir, $\mathfrak{F}_1'$ *entsteht durch Teilung an* $\mathfrak{F}_2$. $\mathfrak{F}_1\mathsf{T}(\mathfrak{F}_1, N)$ ist die Merkmalfolge, welche entsteht, wenn man in $\mathfrak{F}_1$ alle Glieder fortläßt bis auf die, welche zur Teilmenge N gehören. Wenn man also nur auf Elemente aus N achtet. Definition: Wir bezeichnen $\mathfrak{F}_1\mathsf{T}(\mathfrak{F}_1, N)$ auch mit $\mathfrak{F}_{1N}^*$ oder, wenn N bekannt ist, mit $\mathfrak{F}_1^*$ und sagen, dies entstehe durch — *eigentliche oder echte — Teilung* von $\mathfrak{F}_1$ auf N.

§ 2.

Wahrscheinlichkeit, Additionsregel, Divisionsregel.

$\mathfrak{F} = a_1, a_2, \ldots$ sei eine Merkmalfolge, $M = \{P_1, \ldots, P_l\}$ ihre Merkmalmenge. Man greife P_λ heraus. Die Anzahl, wie oft unter $a_1, \ldots, a_\nu$ das P_λ vorkommt, werde mit $z(\nu)$ bezeichnet. $\frac{z(\nu)}{\nu}$ ist dann die relative Häufigkeit von P_λ in $\mathfrak{F}$.

Definition: Wenn $\lim\frac{z(\nu)}{\nu}$ existiert, so sagen wir, dieser limes sei die *Wahrscheinlichkeit von* P_λ *in* $\mathfrak{F}$ und setzen $\lim\frac{z(\nu)}{\nu} = w(P_\lambda)$. Offenbar ist $w(P_\lambda)$ eine Zahl zwischen 0 und 1.

Ist $w(P_\lambda) = 0$, so heißt dies, daß P_λ schließlich selten in $\mathfrak{F}$ auftritt, es bedeutet aber nicht, daß das Auftreten von P_λ in $\mathfrak{F}$ unmöglich ist. Ebenso bedeutet $w(P_\lambda) = 1$, daß P_λ sehr oft, jedoch nicht, daß es mit Sicherheit immer in $\mathfrak{F}$ auftritt.

Wenn jedes Merkmal $P_1, \ldots, P_l$ von $\mathfrak{F}$ eine Wahrscheinlichkeit $w(P_1), \ldots, w(P_l)$ hat, so nennen wir das System dieser Zahlen die *Verteilung von* $\mathfrak{F}$ und kürzen dies ab mit $\mathfrak{B}$. Die Wahrscheinlichkeitsrechnung hat es fast ausschließlich nur mit Merkmalfolgen zu tun, zu welchen Verteilungen gehören.

$\mathfrak{F}$ habe jetzt eine Verteilung und es sei N Teilmenge von M, nämlich $N = \{P_{\alpha_1} \dots P_{\alpha_k}\}$[2]). $z_N(\nu)$ gebe an, wie oft unter den $a_1, \dots a_\nu$ ein Element aus N vorkommt. Dann ist offenbar $z_N(\nu) = z_{P_{\alpha_1}}(\nu) + \dots + z_{P_{\alpha_k}}(\nu)$, also ist $\lim \frac{z_N(\nu)}{\nu}$ vorhanden und zwar gleich $w(P_{\alpha_1}) + \dots + w(P_{\alpha_k})$. Wir setzen $\lim \frac{z_N(\nu)}{\nu} = w(N)$ und nennen es die *Wahrscheinlichkeit von N in $\mathfrak{F}$*. Offenbar ist dies wieder eine Zahl zwischen 0 und 1. Sind N und L zwei elementfremde Teilmengen von M, $N = \{P_{\alpha_1}, \dots P_{\alpha_k}\}$, $L = \{P_{\beta_1}, \dots P_{\beta_l}\}$, so ist offenbar $w(N+L) = \Sigma w(P_{\alpha_\varkappa}) + \Sigma w(P_{\beta_\lambda})$, also $= w(N) + w(L)$. Offenbar ergibt sich entsprechend für endlich viele, je zwei elementfremde Teilmengen von $M: N_1, N_2, \dots N_s: w(M) = \Sigma w(N_\sigma)$, das ist die *Additions- oder Mischungsregel.*

Schließlich gilt auch schon jetzt, also bevor wir über die Merkmalfolgen irgendwelche Axiome gefordert haben: Satz: *$\mathfrak{F}$ habe die Merkmalmenge M. Es sei $N < M$, $w(N) \neq 0$. Dann gilt: Wenn zu $\mathfrak{F}$ eine Verteilung $\mathfrak{B}$ gehört, so gehört auch zu $\mathfrak{F}_N^*$ eine Verteilung. Man bezeichne dann die Wahrscheinlichkeiten in $\mathfrak{F}$ mit w, in $\mathfrak{F}^*$ mit w^*. Dann gilt für jede Teilmenge A von N:*

$$w^*(A) = w(A) \cdot \frac{1}{w(N)}.$$

Beweis: Es genügt offenbar wegen der Additionsregel, den Satz für die speziellen Teilmengen von N zu beweisen, welche nur aus einem Punkte bestehen. Es sei also P_λ ein Punkt von N. Ferner sei $\mathfrak{F}_N^* = a_{\alpha_1}, a_{\alpha_2} \dots$. Wir suchen nun unter diesen diejenigen auf, welche P_λ sind, diese bezeichne man fortlaufend mit $a_{\alpha_{\beta_1}}, a_{\alpha_{\beta_2}}, \dots$. $z(\varrho)$ sei die Anzahl der P_λ aus $a_{\alpha_1}, a_{\alpha_2}, \dots a_{\alpha_\varrho}$. Wir haben zu untersuchen $\frac{z(\varrho)}{\varrho}$. Wenn nun die Folge der $\beta_1, \beta_2, \dots$ nur aus endlich vielen Zahlen besteht, so ist offenbar $w^*(P_\lambda) = 0$ und $w(P_\lambda) = 0$ und dann stimmt offenbar die Behauptung. Andernfalls kann man die ϱ von der Stelle β_1 ab einschließen in ein Intervall $\beta_\nu, \beta_{\nu+1}$, so daß $\beta_\nu \leqq \varrho < \beta_{\nu+1}$. $z(\varrho)$ ist dann offenbar ν und die relative Häufigkeit $\frac{z(\varrho)}{\varrho}$ liegt zwischen $\frac{\nu}{\beta_\nu}$ und $\frac{\nu}{\beta_{\nu+1}}$. Mit wachsendem ϱ wächst ν nun über alle Grenzen und es ist unter Berücksichtigung von

$$w(N) \neq 0 \text{ der } \lim \frac{\nu}{\beta_\nu} = \lim \frac{\nu}{\alpha_{\beta_\nu}} \cdot \frac{\alpha_{\beta_\nu}}{\beta_\nu} = w(P_\lambda) \cdot \frac{1}{w(N)}.$$

---

[2]) Wir werden später auch sagen: N ist ein neues Merkmal, welches aus $P_{\alpha_1}, \dots, P_{\alpha_k}$ durch Mischung entsteht.

Ferner ist

$$\frac{\nu}{\beta_{\nu+1}} = \frac{\nu+1}{\beta_{\nu+1}} - \frac{1}{\beta_{\nu+1}}$$

und daher auch

$$\lim \frac{\nu}{\beta_{\nu+1}} = w(P_\lambda) \cdot \frac{1}{w(N)}.$$

Mithin ist wirklich $\lim \frac{z(\varrho)}{\varrho}$ vorhanden und ist gleich $w(P_\lambda) \frac{1}{w(N)}$, w. z. b. w. Offenbar sind nach diesem Satz die Wahrscheinlichkeiten der Punkte aus N in $\mathfrak{F}^*$ auch so bestimmt: Die Verhältnisse der Wahrscheinlichkeiten sind die gleichen wie in $\mathfrak{F}$. Da die Summe der Wahrscheinlichkeiten 1 sein muß, sind sie wirklich dadurch bestimmt.

§ 3.

Die Axiome.

Die fortgesetzten Würfe mit einem Würfel geben, idealisiert, eine Merkmalfolge mit den Merkmalen 1, 2, ..., 6. Ist der Würfel symmetrisch, so sind die Wahrscheinlichkeiten, wie wir annehmen, jede $\frac{1}{6}$. Ist nun an einer Stelle die Augenzahl 5 gefallen, so wird dadurch, wie wir annehmen, das Resultat des folgenden Wurfes in keiner Weise beeinflußt oder, was dasselbe nach unserer Auffassung präziser beschreibt: Greift man aus der Folge nur die Würfe heraus, die unmittelbar hinter einem Wurf mit dem Resultat 5 kommen, so erhält man eine neue ausgewählte Merkmalfolge, in der aber die Wahrscheinlichkeiten sich keineswegs geändert haben. Allgemeiner: Trifft man aus der Folge zwei fast total verschiedene Auswahlen, so erhält man zwei Merkmalfolgen, zwischen deren Wahrscheinlichkeiten eine Abhängigkeit nicht besteht. Diese Eigenschaft der Würfelfolge ist es, die wir durch die Bezeichnung „*ohne Nachwirkung*" bezeichnen und welche im folgenden durch die Axiome präzisiert wird. Nicht alle Merkmalfolgen sind ohne Nachwirkung. Verfolgt man die Todesursachen der Einwohner einer Stadt, indem man feststellt, ob die Personen an einer ansteckenden Krankheit gestorben sind oder nicht, so erhält man — idealisiert — eine Merkmalfolge, welche jetzt jedoch — wie ich annehme — die Eigentümlichkeit hat: Hinter der Todesursache ansteckende Krankheit kommt häufiger die gleiche Todesursache als der Wahrscheinlichkeit in der Gesamtfolge entspricht, mit anderen Worten, greift man die Todesfälle heraus unmittelbar hinter einem durch ansteckende Krankheit, so erhält man eine neue Folge, und hierin hat die Ursache: ansteckende Krankheit eine größere Wahrscheinlichkeit als in der ursprünglichen Folge. In diesem Todesregister gibt es also eine Nachwirkung.

Hat man die Geburtenregister zweier benachbarter Städte, in denen das Geschlecht der Neugeborenen eingetragen ist, so können wir aus der Kenntnis des einen Registers nichts über das andere schließen. Hat man dagegen die Todesregister der Städte, enthaltend die Angaben, ob die Todesursache eine ansteckende Krankheit war oder nicht, so übt die Ursache ansteckende Krankheit auf die gleiche beim anderen Register, weil die beiden Städte benachbart sein sollten, eine Anziehung aus. Die Merkmalfolgen, die die Idealisierungen der Geburtenregister sind, nennen wir daher *fremd*, die der Todesregister nennen wir *verwandt*. Die Verwandtschaftsbeziehung wird im folgenden durch die Axiome präzisiert werden. Alle verwandten Merkmalfolgen werden wir zusammenfassen zu „Familien".

Als die Hauptelemente der aufzustellenden Wahrscheinlichkeitstheorie sind uns vorgelegt endlich oder unendlich viele Merkmalfolgen, die eindeutig zu Familien zusammengefaßt sind mit je endlich oder unendlich vielen Familiengliedern: $\mathfrak{G}_{11}, \mathfrak{G}_{12}, \ldots$; $\mathfrak{G}_{21}, \mathfrak{G}_{22}, \ldots$; $\ldots$ $\mathfrak{G}_{k1}, \mathfrak{G}_{k2}, \ldots$; $\ldots$. Diese Merkmalfolgen heißen die *Grundkollektivs*. Als erstes Axiom fordern wir:

Axiom I: *Jedes Grundkollektiv hat eine Verteilung.*

Die $\mathfrak{G}$ mit gleichem ersten Index bilden eine *Familie*. Zwei $\mathfrak{G}$ derselben Familie heißen *verwandt*, zwei $\mathfrak{G}$ mit verschiedenen ersten Indizes heißen *fremd*. Manche der $\mathfrak{G}$ heißen *ohne Nachwirkung*, manche *mit Nachwirkung*.

Dann gilt: Sind zwei Grundkollektivs einem dritten verwandt, dann sind sie untereinander verwandt.

Die Bezeichnung der $\mathfrak{G}$ mit Doppelindizes ist gewählt worden, um zu Anfang die Einteilung in Familien, die über das Verhalten der Grundkollektivs zueinander alles entscheiden wird, anschaulich zu betonen. Im folgenden sollen jedoch der Einfachheit halber die Grundkollektivs auch bezeichnet werden als $\mathfrak{G}_1, \mathfrak{G}_2, \ldots$ oder $\mathfrak{G}, \mathfrak{G}', \ldots$ oder ähnlich. Die Schreibweisen $\mathfrak{G}\, v\, \mathfrak{G}'$ und $\mathfrak{G}\, f\, \mathfrak{G}'$ bedeuten dann: $\mathfrak{G}$ verwandt $\mathfrak{G}'$ und $\mathfrak{G}$ fremd $\mathfrak{G}'$. Die Grundkollektivs werden nun außer dem gänzlich unproblematischen ersten Axiom weiteren Forderungen unterworfen werden. Als charakteristische Forderung seiner Theorie hat von Mises folgendes verlangt: Trifft man aus einem Grundkollektiv irgendeine Auswahl, ohne dabei die Merkmale der auszuwählenden Elemente zu benutzen, so erhält man eine neue Merkmalfolge mit ungeänderter Verteilung. Das ist die Forderung des ausgeschlossenen Spielsystems. Die Folge von Roulettedrehungen wird, idealisiert, aufgefaßt als ein Grundkollektiv und dann wird gesagt: Wenn man irgendwie auswählt — nur darf man nicht gerade z. B. sagen: Ich wähle die Drehungen, wo ein vorgegebenes Merkmal herauskommt, allgemeiner, man darf nicht die Merkmale der auszu-

wählenden Stellen schon benutzen — so ist es unmöglich, die ursprünglichen Grenzwerte der relativen Häufigkeiten der Ziffern 0, 1, ... 36 zu verändern. Meiner Ansicht nach wird hiermit zu viel verlangt. Ich lege daher, und es ergibt sich, daß dies ausreicht, der Theorie irgendeinen fest umrissenen Bereich von „Stellenauswahlen" zugrunde durch die *Definition:* Neben den Grundkollektivs $\mathfrak{G}$ habe man noch eine Gruppe[3]) von Auswahlen fest vorgegeben, bestimmt durch mathematische Gesetze. Diese Auswahlen bezeichne man mit Σ, Σ_1, ... und nenne sie die *„Stellenauswahlen"*. Die Gruppe bezeichne man mit $\mathfrak{S}$. Wir haben jetzt die Grundkollektivs und die Stellenauswahlen Σ. *Diese sind die (einzigen) Elemente, und deren Gesamtheit wird nicht mehr erweitert werden. Außerdem ist uns noch allein vorgegeben bei den verwandten Grundkollektivs ihre Verwandtschaft, d. h. wie wir später präzis formulieren werden, wir wissen, wie die Wahrscheinlichkeiten sich abändern, wenn man an verwandten Grundkollektivs teilt.*

Wendet man nun auf die Grundkollektivs jedesmal alle Stellenauswahlen an, so erhält man neue Merkmalfolgen, ich nenne sie von *erster Art*, und wir werden nunmehr verlangen, daß bei dieser Auswahl die Verteilungen der Grundkollektivs ungeändert bleiben. Hat man im Bereich dieser abgeleiteten Merkmalfolgen erster Art zwei, $\mathfrak{F}_1$, $\mathfrak{F}_2$, $\mathfrak{F}_1 = \mathfrak{G}_1 \Sigma_1$, $\mathfrak{F}_2 = \mathfrak{G}_2 \Sigma_2$, so werden wir auch die Frage zu beantworten haben: Was geschieht, wenn man in $\mathfrak{F}_1$ nur auf die Stellen achtet, bei denen in $\mathfrak{F}_2$ ein vorgegebenes Resultat auftritt? Das ist eine Frage an die Merkmalfolge, die aus $\mathfrak{F}_1$ entsteht durch eine Teilung an $\mathfrak{F}_2$: $\mathfrak{F}_1 \mathsf{T} (\mathfrak{F}_2, \ldots) = \mathfrak{G}_1 \Sigma_1 \mathsf{T} (\Sigma_2, \mathfrak{G}_2, \ldots)$. Man wird also durch die Gruppe der Stellenauswahlen dazu geführt, daneben die Teilungen der Σ an den $\mathfrak{G}$ zu betrachten, diese nenne man die *Teilungen erster Art*, und dann die Produkte der Stellenauswahlen mit den Teilungen erster Art auf die $\mathfrak{G}$ anzuwenden. Diese Produkte nenne man die *Auswahlen zweiter Art.* Wendet man eine solche auf ein $\mathfrak{G}$ an, so erhält man eine abgeleitete Merkmalfolge zweiter Art. Die Teilungen einer Auswahl zweiter Art an einem $\mathfrak{G}$ nenne man Teilungen zweiter Art. Die Produkte von Auswahlen höchstens zweiter Art mit Teilungen zweiter Art nenne man Auswahlen dritter Art. Die Merkmalfolgen, welche durch solche Auswahlen, angewandt auf die $\mathfrak{G}$, entstehen, heißen abgeleitete Merkmalfolgen dritter Art. Man sieht: Nachdem man, so fortschreitend, die Auswahlen ν-ter Art und die abgeleiteten Merkmalfolgen ν-ter Art erhalten hat, ergeben sich Aufgaben, welche dazu führen, noch die Teilungen der Auswahlen ν-ter Art an den $\mathfrak{G}$, die Teilungen

[3]) Unter einer Gruppe von Auswahlen verstehe ich eine Gesamtheit, welche neben je zwei auch deren Produkt enthält.

ν-ter Art, und die Produkte von Auswahlen höchstens ν-ter Art mit Teilungen ν-ter Art einzuführen, *die (zusammengesetzten) Auswahlen* $(\nu+1)$*-ter Art.* Diese erzeugen aus den Grundkollektivs die *abgeleiteten Merkmalfolgen* $(\nu+1)$*-ter Art.*

Definition: Zu den Stellenauswahlen, welche als die Auswahlen erster Art definiert sind, nehme man hinzu die Auswahlen ν-ter Art (für alle $\nu = 1, 2, \ldots$). Diese Gesamtheit bildet offenbar eine Gruppe. *Man nenne sie die Gruppe* $\hat{\mathfrak{S}}$. Durch die Aufgaben der Wahrscheinlichkeitsrechnung ist somit die ursprüngliche Gruppe $\mathfrak{S}$ der Stellenauswahlen vergrößert worden zu der umfassenderen Gruppe $\hat{\mathfrak{S}}$, welche nicht mehr nur aus Stellenauswahlen besteht, sondern welche zwar alle Stellenauswahlen enthält, außerdem aber, wenn sie die Auswahl A enthält, auch daneben jede Teilung $\mathsf{T}(\mathsf{A}, \mathfrak{G}, \ldots)$ enthält, ferner neben irgendwelchen Auswahlen auch deren Produkt.

Die Gruppe $\hat{\mathfrak{S}}$ kann mithin offenbar auch so charakterisiert werden:

1. Sie ist eine Menge, welche jede Stellenauswahl enthält.
2. Wenn A zu $\hat{\mathfrak{S}}$ gehört, so auch $\mathsf{T}(\mathsf{A}, \mathfrak{G}, \ldots)$ für jedes $\mathfrak{G}$.
3. $\hat{\mathfrak{S}}$ ist eine Gruppe.
4. Unnötig, d. h. außer zur Genüge der Forderungen 1, 2, 3, werden nicht Elemente in $\hat{\mathfrak{S}}$ aufgenommen, d. h. $\hat{\mathfrak{S}}$ ist die kleinste Gruppe (der Durchschnitt aller Gruppen) mit den Eigenschaften 1, 2, 3.

Über die Elemente von $\hat{\mathfrak{S}}$ gilt folgender Satz:

Definition: Unter einem *Normalprodukt* verstehe man ein Produkt von Stellenauswahlen und von Teilungen von Auswahlen aus $\hat{\mathfrak{S}}$ an irgendwelchen Grundkollektivs $\mathfrak{G}$. Dann gilt: Jedes Element aus $\hat{\mathfrak{S}}$ läßt sich als Normalprodukt darstellen.

Beweis: $\hat{\mathfrak{S}}$ besteht per definitionem aus allen Auswahlen ν-ter Art für $\nu = 1, 2, \ldots$. Wir schließen dann von ν auf $\nu+1$. Die Auswahlen erster Art sind selbst Stellenauswahlen, für sie gilt also die Behauptung. Die Auswahlen $(\nu+1)$-ter Art sind definiert als die Produkte von Auswahlen höchstens ν-ter Art mit Teilungen von Auswahlen ν-ter Art. Diese Teilungen sind selbstverständlich Teilungen von Auswahlen aus $\hat{\mathfrak{S}}$ und für die Auswahlen höchstens ν-ter Art gilt der Satz bereits nach unserer Annahme. Daraus folgt die Behauptung.

Wir wissen damit: Jede Auswahl aus $\hat{\mathfrak{S}}$ kann geschrieben werden in der Form $\mathsf{P}_1 \ldots \mathsf{P}_k$ (wo jedes $\mathsf{P}_\varkappa$ entweder ein Σ ist oder ein $\mathsf{T}(\mathsf{A}, \mathfrak{G}, \ldots)$), mit einem zu $\hat{\mathfrak{S}}$ gehörenden A. Es ist hingegen nicht behauptet worden, daß etwa diese Darstellung eindeutig ist.

Unsere Axiome der Wahrscheinlichkeitsrechnung bestehen darin, daß wir, nachdem wir zunächst verlangt haben, daß die Grundkollektivs eine Verteilung haben, nun fordern, was mit dieser Verteilung geschieht, wenn

wir aus den Grundkollektivs durch Auswahl neue Merkmalfolgen ableiten. Welche Auswahlen hierbei betrachtet werden, ist zunächst durch Konstruktion von $\hat{\mathfrak{S}}$ eingeschränkt worden. In Wirklichkeit war diese Konstruktion noch ein wenig zu grob, und wir werden jetzt für die Elemente von $\hat{\mathfrak{S}}$ bestimmen, welche wir außer Acht lassen — wir werden sie irregulär nennen — und welchen wir ihr Verhalten gegenüber den Grundkollektivs vorschreiben. Diese Vorschriften sind unsere Forderungen der Wahrscheinlichkeitsrechnung.

Wegen des letzten Satzes haben wir für die zu betrachtenden Auswahlen aus $\hat{\mathfrak{S}}$ die bequeme Darstellung als Normalprodukte. Wir werden uns in den Normalprodukten die Faktoren der Reihe nach einzeln ansehen, das sind die Stellenauswahlen und die Teilungen und werden dann von den Faktoren an ihrer Stelle im Normalprodukt einzeln verlangen, wie sie auf die Grundkollektivs einwirken sollen.

Nun machen wir vorerst diese Vorbemerkung: Das Normalprodukt $\Pi = \mathsf{P}_1 \ldots \mathsf{P}_k$ möge an der Stelle P_k eine Teilung $\mathsf{T}(\mathsf{A}, \mathfrak{G}', N)$ enthalten. Es soll nun Π angewandt werden auf $\mathfrak{G}$. Man kann das auch so machen. Man bilde erst $\mathfrak{G} \cdot (\mathsf{P}_1 \ldots \mathsf{P}_{k-1}) = \mathfrak{F}$ und dann $\mathfrak{F} \cdot \mathsf{T}(\mathsf{A}, \mathfrak{G}', N)$. Nehmen wir an, wir wüßten bereits irgendwoher, daß die Teilfolge $\mathfrak{F}$ von $\mathfrak{G}$ eine Verteilung hat. $\mathsf{P}_1 \ldots \mathsf{P}_{k-1}$ setze man gleich Γ. Auf $\mathfrak{G}\Gamma$ wird dann angewandt $\mathsf{T}(\mathfrak{G}'\mathsf{A}, N)$. Man erinnere sich nun an die in § 2 besprochene Operation der echten Teilung. Wenn nämlich $\mathfrak{G} = \mathfrak{G}'$ ist und $\mathsf{A} = \Gamma$, so haben wir genau diesen Fall vor uns. Wenn dann noch die Wahrscheinlichkeit von N in $\mathfrak{F} = \mathfrak{G}\Gamma$ nicht gerade verschwindet, dann wissen wir, daß bei der Teilung $\mathsf{T}(\mathfrak{G}\mathsf{A}, N)$ die Verhältnisse der Wahrscheinlichkeiten der Merkmale von N ungeändert bleiben. Beachten wir also, daß in diesem Falle nichts zu verlangen ist, daß wir also hier gebunden sind durch den Satz aus § 2.

Wie aber, wenn die Wahrscheinlichkeit von N in $\mathfrak{F}$ verschwindet? Dieser Fall ist von geringer Wichtigkeit und zur formalen Vereinfachung der Theorie soll er verboten werden. In diesem Falle ist die Auswahl $\mathsf{T}(\mathfrak{G}\mathsf{A}, N)$ offenbar nicht dicht. Erst recht ist dann offenbar das Normalprodukt $\mathsf{P}_1 \ldots \mathsf{P}_k$ nicht dicht. Ich werde nicht dichte Normalprodukte später nicht mehr in Betracht ziehen, indem ich sie in die Klasse der irregulären Normalprodukte tue, in welcher alles Unangenehme oder zu Schwierige abgelegt wird. Dazu setze ich jetzt nachträglich fest: *Jede Stellenauswahl aus* $\mathfrak{S}$ *ist dicht.*

Definition: Eine *nicht dichte Teilung* in einem Normalprodukte heißt bezüglich jedes Grundkollektivs $\mathfrak{G}$ irregulär.

Folgerung: *Mithin sind jetzt alle Normalprodukte ohne eine irreguläre Teilung (bezüglich irgendeines* $\mathfrak{G}$*) dicht.*

Nach dieser, den Ausschluß der nicht dichten Auswahlen erläuternden Vorbemerkung nehme man irgendein Normalprodukt $\Pi = \mathsf{P}_1 \ldots \mathsf{P}_k$ und irgendein Grundkollektiv $\mathfrak{G}$. $\mathsf{P}_1 \ldots \mathsf{P}_{k-1}$ heißt der *Abschnitt von* P_k *innerhalb* Π. Er sei $\mathsf{P}_1 \ldots \mathsf{P}_{k-1} = \Gamma$. Nehmen wir einmal an, wir wüßten bereits, daß $\mathfrak{G}\,\Gamma$ wieder eine Verteilung hat. Was wird nun vermutlich geschehen, wenn wir P_k anwenden?

I. P_k sei eine Stellenauswahl. Dann werden wir verlangen (in Übereinstimmung mit der Erfahrungstatsache, daß z. B. beim Glücksspiel bei hinreichend langem Spiel noch immer der Unternehmer gewonnen hat), daß $(\mathfrak{G}\,\Gamma)\,\mathsf{P}_k$ wieder eine Verteilung hat, und zwar wieder die gleiche wie $\mathfrak{G}\,\Gamma$.

II. P_k sei eine Teilung: $\mathsf{T}(\mathsf{A}, \mathfrak{G}', N)$. Diese Auswahl T sei dicht.

1.) Und es sei $\mathfrak{G} \neq \mathfrak{G}'$. $\mathfrak{G}\,\Gamma$ ist jetzt eine Teilfolge von $\mathfrak{G}$. Aus dieser wird ausgewählt mit Hilfe des Ausfalls von $\mathfrak{G}'\,\mathsf{A}$. Nun sind $\mathfrak{G}$ und $\mathfrak{G}'$ fremd, daher besteht, wie wir annehmen, zwischen $\mathfrak{G}\,\Gamma$ und $\mathfrak{G}'\,\mathsf{A}$ kein Zusammenhang. Wir werden daher fordern, daß $(\mathfrak{G}\,\Gamma)\,\mathsf{P}_k$ wieder eine Verteilung hat und wieder dieselbe wie $\mathfrak{G}\,\Gamma$. Im Falle II_1, der bei weitem der wichtigste ist, werden wir die Teilung T „*in* Π *einfach regulär bezüglich* $\mathfrak{G}$“ nennen.

Von den Fällen $\mathfrak{G} \vee \mathfrak{G}'$ heben wir hervor:

2.) $\mathfrak{G} = \mathfrak{G}'$, $\mathfrak{G}$ ohne Nachwirkung, A fast total verschieden zu Γ. Dann sind $\mathfrak{G}\,\Gamma$ und $\mathfrak{G}\mathsf{A}$ zwar demselben Grundkollektiv entnommen, es sind aber höchstens endlich oft dieselben Stellen ausgewählt. Ferner ist dieses $\mathfrak{G}$ ohne Nachwirkung. Zwischen $\mathfrak{G}\,\Gamma$ und $\mathfrak{G}\,\mathsf{A}$ besteht daher wiederum keine Abhängigkeit. Wir werden daher wieder fordern, daß $(\mathfrak{G}\,\Gamma)\,\mathsf{P}_k$ die gleiche Verteilung hat wie $\mathfrak{G}\,\Gamma$. Im Falle II_2 heißt T „*im* Π *bezüglich* $\mathfrak{G}$ *kritisch regulär*“.

3.) $\mathfrak{G} = \mathfrak{G}'$, $\Gamma = \mathsf{A}$. Das ist der vorweggenommene Fall, wo $\mathfrak{G}\,\Gamma$ einer echten Teilung unterworfen wird und hier ist durch den Satz in § 2 bereits bestimmt, was geschieht. In diesem Falle heißt T „*in* Π *bezüglich* $\mathfrak{G}$ (*kritisch*) *singulär*“.

Es bleiben noch übrig die Fälle: $\mathfrak{G} = \mathfrak{G}'$ und $\mathfrak{G}$ mit Nachwirkung; $\mathfrak{G} = \mathfrak{G}'$, $\mathfrak{G}$ ohne Nachwirkung, Γ, A halb gleich; $\mathfrak{G} \neq \mathfrak{G}'$ aber $\mathfrak{G} \vee \mathfrak{G}'$, im Ganzen bleiben also übrig $\mathfrak{G} \vee \mathfrak{G}'$ bis auf II_2. II_3. Jetzt wird also $\mathfrak{G}\,\Gamma$ geteilt an $\mathfrak{G}'\,\mathsf{A}$ und es sind $\mathfrak{G}, \mathfrak{G}'$ verwandt. Die beiden vorgenommenen Fälle II_2, II_3 sind offenbar die einzigen, in welchen ohne Rücksicht auf die spezielle Verwandtschaft zwischen $\mathfrak{G}$ und $\mathfrak{G}'$ eine Aussage erwartet werden kann. Liegt jedoch eine konkrete Theorie vor, so wird, wenn eben überhaupt die Verwandschaften der Grundkollektivs ins Gewicht fallen, in einigen weiteren Fällen bekannt sein, wie hier P_k die Verteilung von $\mathfrak{G}\,\Gamma$ abändert. Es kommt nunmehr darauf an, was wir

über die Verwandtschaft der Grundkollektivs wissen. Da hier alles möglich ist, können wir hier eine spezielle Annahme nicht allgemein machen. Wir nehmen aber an: Uns ist von vornherein bekannt bei einigen Paaren von Grundkollektivs $\mathfrak{G}$, $\mathfrak{G}'$ und einigen Paaren von Auswahlen Γ, A — es spielen nur solche aus $\hat{\mathfrak{S}}$ eine Rolle —, wobei $\mathfrak{G}\Gamma$ wieder eine Verteilung hat, wie sich diese abändert bei der Teilung $\mathsf{T}(\mathfrak{G}', \mathsf{A}, N)$, wobei N eine Teilmenge der Merkmalmenge von $\mathfrak{G}'$ ist. Wir wollen dann die Teilung T hinter Γ für $\mathfrak{G}$ wesentlich kritisch nennen. Ist nun im Normalprodukt Π der Faktor $\mathsf{P}_k = \mathsf{T}(\mathfrak{G}'\mathsf{A}, N)$, $\mathsf{P}_1 \ldots \mathsf{P}_{k-1} = \Gamma$ und T wesentlich kritisch bezüglich $\mathfrak{G}$ hinter Γ, so heiße T *in* Π *bezüglich* $\mathfrak{G}$ *wesentlich kritisch.* Wenn wir dann also die Verteilung von $\mathfrak{G}\Gamma$ noch kennen, so ist uns mithin durch unsere vorgegebene Bekanntschaft mit der Verwandtschaft von $\mathfrak{G}$ und $\mathfrak{G}'$ auch die Verteilung von $\mathsf{P}_1 \ldots \mathsf{P}_k$ bekannt.

Wissen wir hingegen nichts über die Einwirkung von T auf $\mathfrak{G}\Gamma$, ist uns also nicht genug über die Verwandtschaft von $\mathfrak{G}$ und $\mathfrak{G}'$ bekannt, so heiße T *in* Π *bezüglich* $\mathfrak{G}$ *irregulär.*

Schließlich heißt die Teilung T *in* Π *bezüglich jedes* $\mathfrak{G}$ *irregulär*, wenn sie nicht dicht ist.

Von den Teilungen eines Normalproduktes haben somit je nach ihrer Stellung im Normalprodukt bezüglich eines jeden Grundkollektivs $\mathfrak{G}$ einige die Bezeichnung irregulär erhalten. Von ihnen wissen wir nicht, wie sie bei Anwendung von Π auf $\mathfrak{G}$ wirken, die übrigen aber, und das sind wegen II_1 die wichtigsten und sozusagen die meisten, sind uns in ihrer Einwirkung an ihrer Stelle im Normalprodukt auf $\mathfrak{G}$ bekannt.

Wir wissen also jetzt von allen Normalprodukten bis auf die, welche bezüglich $\mathfrak{G}$ irreguläre Teilungen enthalten, was wir als deren Wirkung auf $\mathfrak{G}$ zu fordern haben.

Um dies kurz ausdrücken zu können, führen wir noch eine Bezeichnung ein. Die zu einer Merkmalfolge gehörende Verteilung ist ein System von l Zahlen $\mathfrak{V}: w(P_1), \ldots, w(P_l)$. Die Abänderung der Verteilung wollen wir jetzt immer dadurch bewerkstelligen, daß wir jede Zahl aus $\mathfrak{V}$ mit einer geeigneten Zahl multiplizieren. Wir werden mithin unter einem *Multiplikator von* $\mathfrak{V}$ ein System von l nicht negativen reellen Zahlen verstehen und mit K, L oder ähnlichen Buchstaben abkürzen. Wir wollen diese Multiplikatoren, weil sie l Zahlen enthalten, l stellig nennen. Ist dann $K = [\tau_1, \ldots, \tau_l]$, so wollen wir unter $\mathfrak{V}K$ die Verteilung verstehen $w^* P(\mathrm{P}_1), \ldots, w^*(P_l)$, worin

$$w^*(P_1) = w(P_1)\tau_1, \ \ldots, \ w^*(P_l) = w(P_l)\cdot\tau_l$$

ist. Als Produkt von Multiplikatoren $K = [\tau_1, \ldots, \tau_l]$, $L = [\sigma_1, \ldots, \sigma_l]$ definieren wir den Multiplikator $K \cdot L = [\tau_1\sigma_1, \ldots, \tau_l\sigma_l]$. Entsprechend

wird das Produkt von beliebig vielen Multiplikatoren definiert. Offenbar ist dann $(\mathfrak{B}K)L = \mathfrak{B}(KL)$. Wir schreiben dafür $\mathfrak{B}KL$ und entsprechend für beliebig viele Multiplikatoren. Die Erhaltung einer Verteilung ist dann gleichbedeutend mit der Multiplikation mit dem Einheitsmultiplikator $E = [1, 1, \ldots, 1]$.

Den Multiplikator für eigentliche Teilung können wir angeben. Hat $\mathfrak{F}$ die Merkmalmenge M und bildet man $\mathfrak{F}^*_N$, wobei etwa $M = (P_1, \ldots, P_l)$, $N = (P_{\alpha_1}, \ldots, P_{\alpha_r})$ sei, so ist die Verteilung $\mathfrak{B}^*$ von $\mathfrak{F}^*$ offenbar

$$\mathfrak{B}\left[0, 0, \ldots, \frac{1}{w(N)}\ 0, \ldots, \frac{1}{w(N)}, \ldots, \frac{1}{w(N)}, \ldots, 0\right],$$

wobei an den Stellen $\alpha_1, \ldots, \alpha_r$ die Zahl $\frac{1}{w(N)}$ steht und sonst Nullen. Wir wollen diesen l stelligen Multiplikator mit $\left[\frac{\mathfrak{B}}{N}\right]$ bezeichnen. Hat man dann aus einem Grundkollektiv mittels $\mathsf{P}_1 \ldots \mathsf{P}_{k-1}$ abgeleitet, und weiß man bereits, daß dabei die Verteilung $\mathfrak{B}K_1 \ldots K_{k-1}$ herausgekommen ist, so wird danach offenbar die Operation P_k, wenn sie die singuläre Teilung $\mathsf{T}(\mathsf{A}, \mathfrak{G}, N)$ ist, die Verteilung multiplizieren mit $\left[\frac{\mathfrak{B}K_1 \ldots K_{k-1}}{N}\right]$. Was wir von der Einwirkung der Normalprodukte auf die Grundkollektivs fordern wollen, läßt sich jetzt so ausdrücken. Das Normalprodukt sei $\Pi = \mathsf{P}_1 \ldots \mathsf{P}_k$. Es werde angewandt auf $\mathfrak{G}$. Die Verteilung von $\mathfrak{G}$ sei l stellig. Wir benutzen daher jetzt l stellige Multiplikatoren. Wir ordnen nun den Stellenauswahlen, den bezüglich $\mathfrak{G}$ einfach regulären und den bezüglich $\mathfrak{G}$ kritisch regulären Teilungen von Π den Multiplikator E zu. Den wesentlich kritischen Teilungen gehört ein Multiplikator zu, den wir in der konkreten Theorie haben, hier in der allgemeinen Theorie offen lassen. Ist P_k eine kritisch singuläre Teilung auf die Teilmenge N und haben die $\mathsf{P}_1, \ldots, \mathsf{P}_{k-1}$ bereits die Multiplikatoren $K_1, \ldots, K_{k-1}$ erhalten, so ordne man P_k zu den Multiplikator

$$\left[\frac{\mathfrak{B}K_1 \ldots K_{k-1}}{N}\right].$$

Auf diese Weise erhält sukzessive jede singuläre Teilung einen Multiplikator. Etwaigen irregulären Teilungen ordnen wir nichts zu. Enthält dann Π keine bezüglich $\mathfrak{G}$ irregulären Teilungen, so gehört zu jedem Faktor von Π bezüglich $\mathfrak{G}$ ein Multiplikator und man erhält $K_1, \ldots, K_k$. Und wir verlangen: $\mathfrak{G}\Pi$ hat eine Verteilung und zwar gerade $\mathfrak{B}K_1 \ldots K_k$. Wir wollen unsere Forderung präzis und knapp, aber vollständig wiederholen und zum Teil etwas ergänzen.

Zusammenfassung: Ein Produkt von Stellenauswahlen mit Teilungen von Auswahlen aus $\dot{\mathfrak{S}}$ an den $\mathfrak{G}$ heißt ein Normalprodukt.

Mit Rücksicht auf folgendes soll noch definiert werden:

A, B seien zwei Auswahlen aus $\hat{\mathfrak{S}}$, also als Normalprodukte darstellbar. Dann heißt B hinter A bezüglich $\mathfrak{G}$ regulär, oder singulär, oder kritisch, wenn es Normalproduktdarstellungen $\Pi(\mathsf{A})$, $\Pi(\mathsf{B})$ von A und B gibt, so daß die Normalprodukte $\Pi(\mathsf{A})\,\Pi(\mathsf{B})$ hinter $\Pi(\mathsf{A})$ regulär oder singulär oder kritisch sind, d. h. hinter $\Pi(\mathsf{A})$ bezüglich $\mathfrak{G}$ nur reguläre Teilungen oder nur reguläre und singuläre und mindestens eine singuläre Teilung oder nur nicht irreguläre und mindestens eine kritische Teilung aufweisen.

Trifft keins zu, so heißt B hinter A bezüglich $\mathfrak{G}$ irregulär.

§ 4.

Abgeleitete Kollektivs erster Art.

Man wende auf jedes Grundkollektiv $\mathfrak{G}$ alle für $\mathfrak{G}$ nicht irregulären Auswahlen aus $\hat{\mathfrak{S}}$ an. Es entstehen neue Merkmalfolgen $\mathfrak{G}\,\mathsf{A}$. Diese heißen die *abgeleiteten Kollektivs erster Art.* Weil zu $\mathfrak{S}$, also auch zu $\hat{\mathfrak{S}}$ die identische Auswahl E gehörte, so umfaßt der Bereich der abgeleiteten Kollektivs erster Art auch die Grundkollektivs $\mathfrak{G}$. Dieser Bereich ist erzeugt durch die vorgegebenen Grundkollektivs und durch die vorgegebenen Stellenauswahlen — außerdem hängt er davon ab, welche Abhängigkeiten bei den verwandten $\mathfrak{G}$ wir vorgegeben haben — und die Wahrscheinlichkeitsrechnung handelt von diesen Kollektivs.

Wir wollen statt abgeleitetes Kollektiv erster Art kurz Kollektiv sagen. Jedes solche ist in der Form $\mathfrak{G}\,\mathsf{A}$ vorgegeben, wo A eine bezüglich $\mathfrak{G}$ reguläre oder singuläre Auswahl aus $\hat{\mathfrak{S}}$ ist. Von diesem Kollektiv sagen wir, es entstamme aus $\mathfrak{G}$. Nun kann es sein, daß das gleiche Kollektiv aus demselben Grundkollektiv durch verschiedene nicht irreguläre Auswahlen hervorgeht. Und es kann auch sein, daß das gleiche Kollektiv aus verschiedenen Grundkollektivs zugleich entstammt. Beides können wir nicht widerlegen. Wir können jedoch zeigen, und das ist von großer Wichtigkeit, weil es hierauf beruht, daß wir für den Bereich aller Kollektivs den Begriff der Unabhängigkeit definieren können, das gleiche Kollektiv kann nicht zugleich aus zwei verschiedenen Familien entstammen. Wir müssen nur eine triviale Ausnahme machen.

Definition: Ein Kollektiv heißt konstant, wenn es unter den Merkmalen nur eins mit nicht verschwindender Wahrscheinlichkeit — diese ist dann offenbar 1 — gibt.

Dann gilt der Satz:

Es ist unmöglich, daß ein nicht konstantes Kollektiv aus zwei Grundkollektivs verschiedener Familien entstammt.

Beweis: Andernfalls wäre $\mathfrak{R} = \mathfrak{G}\,\mathsf{A} = \mathfrak{G}'\,\mathsf{B}$, $\mathfrak{G} \neq \mathfrak{G}'$, A und B aus $\hat{\mathfrak{S}}$. Die Merkmalmenge von $\mathfrak{R}$ enthielte wenigstens zwei Merkmale mit nicht verschwindender Wahrscheinlichkeit, etwa P und Q. Die Operation $\mathsf{A}\,\mathsf{T}(\mathfrak{G}\,\mathsf{A}, \{P\})$ ist dann offenbar, indem man A als ein für $\mathfrak{G}$ nicht irreguläres Normalprodukt Π schreibt, im ganzen eine Auswahl aus $\hat{\mathfrak{S}}$ und zwar ist sie für $\mathfrak{G}$ nicht irregulär. Der Multiplikator von $\mathsf{T}(\mathfrak{G}\,\mathsf{A}, \{P\})$ ist ferner wegen der Bestimmung von P von E verschieden, weil das Merkmal Q von $\mathfrak{G}\,\mathsf{A}$ fortfällt. Andererseits ist die Operation $\mathsf{T}(\mathfrak{G}'\,\mathsf{B}, \{P\})$ einerseits gleich $\mathsf{T}(\mathfrak{G}\,\mathsf{A}, \{P\})$ und andererseits, da $\mathfrak{G} \neq \mathfrak{G}'$, in dem Normalprodukt Π das $\mathsf{T}(\mathfrak{G}'\,\mathsf{B}, \{P\})$ bezüglich $\mathfrak{G}$ regulär und daraus folgt, daß zu T der Multiplikator E gehört und daß die Verteilung von $\mathfrak{G}\,\mathsf{A}$ und $\mathfrak{G}\,\mathsf{A}\,\mathsf{T}(\mathfrak{G}\,\mathsf{A}, \{P\})$ dieselbe ist. Das ist ein Widerspruch, womit der Satz indirekt bewiesen ist.

Zu jedem Kollektiv $\mathfrak{R}$, welches nicht konstant ist, gibt es also eine eindeutig bestimmte Familie — man nenne sie $\Phi(\mathfrak{R})$ — der es entstammt, wo hingegen ein eindeutiges Stammgrundkollektiv nur angegeben werden kann in dem nicht allgemeinen, wenngleich in den Anwendungen wichtigsten Falle, wenn die Familie nur ein Grundkollektiv enthält.

Zur Vereinheitlichung der Ausdrucksweise setze man als Stammfamilie jedes konstanten Kollektivs irgendeine beliebige der Familien, aus denen es entstammt, eindeutig fest.

$\Phi(\mathfrak{R})$ ist nun eine eindeutige Funktion für den Bereich aller Kollektivs.

Wir nennen dann zwei Kollektivs *verwandt*, wenn sie zur gleichen Familie gehören, andernfalls *fremd*. Offenbar sind dann immer zwei Kollektivs, wenn sie zu einem dritten verwandt sind, untereinander verwandt.

Entscheidend für das Verhalten der Kollektivs zueinander in der praktisch wichtigsten Beziehung zueinander — nämlich darin, ob bei der Verbindung, wie wir später sagen werden, die Multiplikationsregel gilt — ist ein Begriff, der ein klein wenig weniger besagt als ihre Fremdheit. Nämlich: Definition: *Zwei nichtkonstante Kollektivs heißen unabhängig:*

1. *wenn sie fremd sind,*

2. *wenn sie demselben Grundkollektiv ohne Nachwirkung entstammen und aus diesem durch fast total verschiedene nicht irreguläre Auswahlen ableitbar sind. Ein konstantes Kollektiv und ein anderes heißen immer unabhängig.*

Der Zusatz über konstante Kollektivs ist vielleicht unerwartet. Er beruht einfach darauf, daß, wie wir sehen werden, die Multiplikationsregel, d. i. der Satz, dem zuliebe der Begriff der Unabhängigkeit eingeführt ist, wenn eins der beiden Kollektivs konstant ist, trivial sein wird.

Für je zwei Kollektivs steht jetzt fest, ob sie unabhängig sind oder abhängig.

Es ist jetzt aus formalen Gründen zweckmäßig, den Auswahlen aus $\hat{\mathfrak{S}}$ auch bezüglich den abgeleiteten Merkmalfolgen einen Regularitätscharakter zu geben.

Es sei $\mathfrak{F}$ ein abgeleitetes Kollektiv. Es kann dann vielleicht auf verschiedene Weise in der Form $\mathfrak{G}\mathsf{A}$ geschrieben werden, worin $\mathfrak{G}$ ein Grundkollektiv, A eine bezüglich $\mathfrak{G}$ nicht irreguläre Auswahl aus $\hat{\mathfrak{S}}$ ist. Man denke sich alle Schreibweisen von $\mathfrak{F}$ aufgestellt: $\mathfrak{F} = \mathfrak{G}_1\mathsf{A}_1 = \mathfrak{G}_2\mathsf{A}_2 = \ldots$. Dann nenne man B *bezüglich* $\mathfrak{F}$ *regulär*, wenn einmal B hinter A_ν bezüglich $\mathfrak{G}_\nu$ regulär ist, *singulär*, wenn einmal B hinter A_ν bezüglich $\mathfrak{G}_\nu$ singulär und *kritisch*, wenn B einmal hinter A_ν bezüglich $\mathfrak{G}_\nu$ kritisch ist.

Dagegen heiße B *irregulär* bezüglich $\mathfrak{F}$, wenn B für jedes ν hinter A_ν bezüglich $\mathfrak{G}_\nu$ irregulär ist.

Dann kann zwar gleichzeitig B für $\mathfrak{F}$ regulär, singulär und kritisch sein, aber es ist, wenn eins davon zutrifft, nicht gleichzeitig irregulär.

Offenbar gilt dann: *Reguläre Auswahlen erzeugen aus Kollektivs wieder Kollektivs mit unveränderter Verteilung. Singuläre Auswahlen erzeugen aus Kollektivs wieder Kollektivs und die Verhältnisse der Wahrscheinlichkeiten der wirklich auftretenden Merkmale bleiben unverändert. Kritische Auswahlen lassen ebenfalls die Kollektiveigenschaft unangetastet.* Wie aber die Verteilung sich abändert, ist nicht allgemein bekannt und hängt von der Art der Verwandtschaften ab.

Zur Formulierung der folgenden Hauptsätze führen wir die Bezeichnung ein: Eine nichtdünne Teilung an einem Grundkollektiv oder abgeleiteten Kollektiv heiße *zulässige Teilung*. Dann gelten:

Hauptsatz I: *Wendet man auf ein Kollektiv eine Stellenauswahl an, so erhält man wieder ein Kollektiv, und zwar mit ungeänderter Verteilung.*

Beweis: Eine Stellenauswahl ist offenbar für jedes Kollektiv eine reguläre Auswahl.

Hauptsatz II: *Wendet man auf ein Kollektiv eine zulässige Teilung durch ein unabhängiges Kollektiv an, so erhält man wieder ein Kollektiv, und zwar mit unveränderter Verteilung.*

Beweis: Auf $\mathfrak{K}$ soll $\mathsf{T}(\mathfrak{K}_1, N)$ angewandt werden.

Da $\mathfrak{K}_1$ zu $\mathfrak{K}$ unabhängig ist, ist entweder $\Phi(\mathfrak{K}_1) \neq \Phi(\mathfrak{K})$, also $\mathfrak{K}$, $\mathfrak{K}_1$ fremd oder $\mathfrak{K}$ und $\mathfrak{K}_1$ entstammen demselben $\mathfrak{G}$ ohne Nachwirkung mittels fast total verschiedener Auswahlen. In jedem Falle ist dann die Teilung an $\mathfrak{K}_1$ eine reguläre Operation für $\mathfrak{K}$, woraus die Behauptung folgt.

Hauptsatz III: *Wendet man auf ein Kollektiv $\mathfrak{K}$ mit der Verteilung $\mathfrak{B}$ eine zulässige eigentliche Teilung auf die Teilmenge N von der*

Merkmalmenge von $\mathfrak{K}$ an, so entsteht wieder ein Kollektiv mit der Verteilung $\left[\frac{\mathfrak{B}}{N}\right]$.

Die Verhältnisse der Wahrscheinlichkeiten der nicht herausfallenden Merkmale bleiben mithin unverändert.

Beweis: Offenbar ist die eigentliche Teilung für jedes Kollektiv eine singuläre Operation.

Schließlich gilt, wie noch zur Vollständigkeit formuliert werde, wenn man definiert:

Definition: Zwei abhängige Kollektivs $\mathfrak{K}_1, \mathfrak{K}_2, \mathfrak{K}_1 = \mathfrak{G}_1 \mathsf{A}_1, \mathfrak{K}_2 = \mathfrak{G}_2 \mathsf{A}_2$ heißen *kritisch abhängig zueinander*, wenn alle nichtdünnen Teilungen durch jedes für das andere kritisch sind.

Offenbar gilt dann:

Hauptsatz IV: *Jede zulässige Teilung eines Kollektivs durch ein kritisch abhängiges Kollektiv erzeugt wieder ein Kollektiv.*

Über die neue Verteilung können wir eine Aussage nicht machen, es ist aber bei Festlegung der Verwandtschaftsbeziehung von $\mathfrak{G}$ und $\mathfrak{G}'$ vorgeschrieben, mit welchen Multiplikatoren sich die Verteilungen multiplizieren.

Aus den Hauptsätzen I, II, III, IV folgt: Geht man von den durch die Grundkollektivs erzeugten Kollektivs $\mathfrak{K}$ aus und wendet nacheinander in beliebiger Reihenfolge die Operationen an: Stellenauswahl, zulässige Teilung durch ein unabhängiges Kollektiv, eigentliche Teilung und Teilung durch ein kritisches Kollektiv, so kommt man aus dem Bereich aller Kollektivs nicht heraus. Insbesondere bleibt, wenn man nur Stellenauswahl und Teilung durch unabhängige Kollektivs anwendet, die Verteilung invariant, während bei weiterer Hinzunahme der eigentlichen Teilung einzelne Merkmale fortfallen, die Verhältnisse der Wahrscheinlichkeiten der übrigbleibenden aber unverändert bleiben. Kommen außerdem Teilungen durch kritisch abhängige Kollektivs hinzu, so kann aber auch dies nicht mehr behauptet werden.

§ 5.

Verbindung.

Die wichtigste Operation der Wahrscheinlichkeitsrechnung ist die Verbindung. Dieselbe soll jetzt besprochen werden. Es seien $\mathfrak{A}$ und $\mathfrak{B}$ Merkmalfolgen $\mathfrak{A} = a_1, a_2, \ldots, \mathfrak{B} = b_1, b_2, \ldots$. Die Merkmale von $\mathfrak{A}$ seien die ϱ-dimensionalen Punkte $P_1, \ldots, P_k$, sie bilden die Merkmalmenge M. Die Merkmale von $\mathfrak{B}$ seien die σ-dimensionalen Punkte $Q_1, \ldots, Q_l$, sie bilden die Merkmalmenge U.

Hat $P_\varkappa$ die Koordinaten $p_1, \ldots, p_\varrho$ und Q_λ die Koordinaten $q_1, \ldots, q_\sigma$, *so verstehen wir unter* $P_\varkappa \times Q_\lambda$ *den* $(\varrho + \sigma)$*-dimensionalen Punkt* $r_1, \ldots, r_{\varrho+\sigma}$, *wobei* $r_1 = p_1, \ldots, r_\varrho = p_\varrho, r_{\varrho+1} = q_1, \ldots, r_{\varrho+\sigma} = q_\sigma$ *ist.* Unter der *Punktmenge* $M \times U$ verstehen wir die Gesamtheit der kl Punkte $P_\varkappa \times Q_\lambda$. Unter der *Merkmalfolge* $\mathfrak{A}\mathfrak{B}$ verstehen wir die folgende $c_1, c_2, \ldots$, wobei c_ν gleich dem Produkte der Punkte ist, die hier mit a_ν, b_ν bezeichnet sind, also $c_\nu = a_\nu \times b_\nu$. Die Merkmalmenge von $\mathfrak{A}\mathfrak{B}$ ist also $\varrho + \sigma$ dimensional. Sie ist gewiß in $M \times U$ enthalten. Entsprechend definieren wir *das Produkt von mehr als zwei Merkmalfolgen.* Bei dieser Multiplikation gilt das assoziative, aber nicht das kommutative Gesetz. Statt Produkt sagen wir auch Verbindung.

Wir betrachten jetzt Produkte von Kollektivs erster Art. Die Sätze, auf welche wir unsere Überlegungen stützen, sind diese: Teilt man ein Kollektiv $\mathfrak{A}$ erster Art zulässig durch ein zweites unabhängiges Kollektiv erster Art, so erhält man aus $\mathfrak{A}$ ein Kollektiv erster Art mit der gleichen Verteilung wie $\mathfrak{A}$. Ferner: Erzeugt die gleiche für $\mathfrak{A}$ und $\mathfrak{B}$ nicht irreguläre Auswahl A aus $\mathfrak{A}$ und $\mathfrak{B}$ neue Kollektivs erster Art, so sind mit $\mathfrak{A}$ und $\mathfrak{B}$ die neuen ausgewählten Kollektivs wieder unabhängig.

Wir beweisen nun

Satz: *Seien* $\mathfrak{A}$, $\mathfrak{B}$ *unabhängige Kollektivs erster Art. Man bilde* $\mathfrak{A}\mathfrak{B}$. *Dann gehört zu* $\mathfrak{A}\mathfrak{B}$ *wieder eine Verteilung. Nennt man die Wahrscheinlichkeiten von* $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{A}\mathfrak{B}$ *der Reihe nach* w_1, w_2, w_3, *so ist*

$$w_3(P_\varkappa \times Q_\lambda) = w_1(P_\varkappa) \cdot w_2(Q_\lambda).$$

Beweis: Man teile $\mathfrak{B}$ auf die Punktmenge, welche allein aus dem Punkt Q_λ besteht. Es sei zunächst $w_2(Q_\lambda) \neq 0$. Man erhält die Auswahl $\mathsf{A}: \alpha_1, \alpha_2, \ldots$. In $\mathfrak{A}\mathsf{A}: a_{\alpha_1}, a_{\alpha_2}, \ldots$ greife man die Elemente heraus, welche $P_\varkappa$ sind, es seien $a_{\alpha_{\beta_1}}, a_{\alpha_{\beta_2}}, \ldots$. Dann ist, weil $\mathfrak{A}\mathsf{A}$ dieselbe Verteilung wie $\mathfrak{A}$ hat, $\lim \frac{\nu}{\beta_\nu} = w_1(P_\varkappa)$ und $\lim \frac{\beta_\nu}{\alpha_{\beta_\nu}} = w_2(Q_\lambda)$. Daraus folgt $\lim \frac{\nu}{\alpha_{\beta_\nu}} = \lim \frac{\nu}{\beta_\nu} \cdot \frac{\beta_\nu}{\alpha_{\beta_\nu}}$ existiert und ist gleich $w_1(P_\varkappa)\, w_2(Q_\lambda)$. Andererseits ist aber $\lim \frac{\nu}{\alpha_{\beta_\nu}} = w_3(P_\varkappa \times Q_\lambda)$, also $w_3(P_\varkappa \times Q_\lambda) = w_1(P_\varkappa)\, w_2(Q_\lambda)$. Ist hingegen $w_2(Q_\lambda) = 0$, so ist offenbar die Formel selbstverständlich.

Entsprechend ergibt sich als Resultat der Verbindung von mehr als zwei unabhängigen Kollektivs $\mathfrak{A}_1, \ldots, \mathfrak{A}_n$ mit den Wahrscheinlichkeiten $w_1, \ldots, w_n$ und den Merkmalmengen $M_1, \ldots, M_n$ von den Dimensionen $\varkappa_1, \ldots, \varkappa_n$ eine neue Merkmalfolge von der Dimension $\varkappa_1 + \varkappa_2 + \ldots + \varkappa_n$ mit der Merkmalmenge $M_1 \times M_2 \times \ldots \times M_n$. Zu der neuen Merkmalfolge

$\mathfrak{A}_1 \ldots \mathfrak{A}_n$ gehört wieder eine Verteilung, und es ist die neue Wahrscheinlichkeit

$$w(P^1 \times P^2 \times \ldots \times P^n) = w_1(P^1) \cdot w_2(P^2) \ldots \cdot w_n(P^n),$$

wobei P^ν ein Merkmal von $\mathfrak{A}_\nu$ bezeichnet.

Definition: Jedes Produkt unabhängiger Kollektivs erster Art nennen wir ein *Kollektiv zweiter Art.*

Wir denken uns also ein Kollektiv zweiter Art nicht nur als Merkmalfolge gegeben, sondern als ein Produkt von mindestens zwei Faktoren, welche Kollektivs erster Art sind.

Eine Merkmalfolge kann dann durchaus als Merkmalfolge gleich einem Kollektiv zweiter Art und gleichzeitig gleich einem Kollektiv erster Art sein. Diese sind uns dann eben nur formal verschieden in ihrer Darstellung.

Definition: Ist $\mathfrak{A}$ ein Kollektiv zweiter Art $\mathfrak{A} = \mathfrak{A}_1 \ldots \mathfrak{A}_n$ ($\mathfrak{A}_\nu$ sind Kollektivs erster Art), $\mathfrak{B}$ ein Kollektiv erster Art, so heißen $\mathfrak{A}$ und $\mathfrak{B}$ *unabhängig*, wenn jedes $\mathfrak{A}$ zu $\mathfrak{B}$ unabhängig ist.

Sind $\mathfrak{A} = \mathfrak{A}_1 \ldots \mathfrak{A}_n$ und $\mathfrak{B} = \mathfrak{B}_1 \ldots \mathfrak{B}_m$ zwei Kollektivs zweiter Art, so heißen sie *unabhängig*, wenn jedes $\mathfrak{A}_\nu$ zu jedem $\mathfrak{B}_\mu$ unabhängig ist.

Damit ist für je zwei Kollektivs erster oder zweiter Art bestimmt, wann sie unabhängig sind.

Definition: Sind $M_1, \ldots, M_n$ die Merkmalmengen von $\mathfrak{A}_1, \ldots, \mathfrak{A}_n$, so sollen die Teilmengen von $M_1 \times \ldots \times M_n$, welche sich als Produkt $N_1 \times \ldots \times N_n$ darstellen lassen, wo N_ϱ eine Teilmenge von M_ϱ mit nicht verschwindender Wahrscheinlichkeit ist, als *zulässige Teilmengen* der Merkmalmenge von $\mathfrak{A}$, Teilungen von $\mathfrak{A}_1 \ldots \mathfrak{A}_n$ auf zulässige Teilmengen sollen als *zulässige Teilungen* bezeichnet werden. Teilmengen mit nicht verschwindender Wahrscheinlichkeit von Kollektivs erster Art sollen sämtlich als zulässig bezeichnet werden.

Jedes Kollektiv erster oder zweiter Art werde als ein Kollektiv bezeichnet. Der Bereich dieser Kollektivs wird nicht mehr wesentlich erweitert werden.

Offenbar ergeben sich dann folgende Sätze:

Hauptsatz 1: *Eine Stellenauswahl an einem Kollektiv erzeugt wieder ein Kollektiv mit unveränderter Verteilung.*

Beweis: Für Kollektivs erster Art ist der Satz bewiesen. Ist $\mathfrak{A} = \mathfrak{A}_1 \ldots \mathfrak{A}_n$ ein Kollektiv zweiter Art und Σ die Stellenauswahl, so ist $\mathfrak{A}\Sigma = \mathfrak{A}_1\Sigma \ldots \mathfrak{A}_n\Sigma$. Die $\mathfrak{A}_\nu\Sigma$ sind jetzt wieder Kollektivs erster Art mit der gleichen Verteilung wie $\mathfrak{A}_\nu$, und zwar sind je zwei wieder unabhängig.

Hauptsatz 2: $\mathfrak{K}_1$, $\mathfrak{K}_2$ *seien zwei unabhängige Kollektivs. Ist* T *eine zulässige Teilung von* $\mathfrak{K}_2$, *so ist* $\mathfrak{K}_1\,\mathsf{T}$ *wieder ein Kollektiv und wieder mit derselben Verteilung wie* $\mathfrak{K}_1$.

Beweis: Sind $\mathfrak{K}_1$, $\mathfrak{K}_2$ beide erster Art, so ist der Satz bewiesen. Sei also $\mathfrak{K}_1$ von zweiter Art, $\mathfrak{K}_1 = \mathfrak{A}_1 \ldots \mathfrak{A}_n$, $\mathfrak{K}_2$ von erster Art. Dann ist $\mathfrak{K}_1\,\mathsf{T} = \mathfrak{A}_1\,\mathsf{T} \ldots \mathfrak{A}_n\,\mathsf{T}$ und die $\mathfrak{A}_\nu\,\mathsf{T}$ sind unabhängig und haben dieselbe Verteilung wie $\mathfrak{A}_\nu$.

Ist $\mathfrak{K}_1$ von erster oder zweiter Art, $\mathfrak{K}_2 = \mathfrak{B}_1 \ldots \mathfrak{B}_m$ von zweiter Art, so schließe man durch den Schluß von m auf $m+1$.

Hauptsatz 3. *Macht man mit einem Kollektiv eine zulässige Teilung, so erhält man wieder ein Kollektiv* (die Verteilung folgt bereits aus dem Satz in § 2).

Beweis: Für Kollektivs erster Art ist der Satz bewiesen. Sei also $\mathfrak{K}$ von zweiter Art, $\mathfrak{K} = \mathfrak{A}_1 \ldots \mathfrak{A}_n$. Die Teilung gehe auf die Teilmenge $N_1 \times \ldots \times N_n$. Man wende nun erst auf $\mathfrak{K}$ die Teilung $\mathsf{T}(\mathfrak{A}_1, N_1)$ an. Diese Auswahl erzeugt aus $\mathfrak{A}_1 \ldots \mathfrak{A}_n$ die neuen Merkmalfolgen $\mathfrak{A}_1^*, \ldots \mathfrak{A}_n^*$. Diese sind nach Hauptsatz 2 und 3 aus § 4 wieder Kollektivs erster Art, und zwar sind sie wieder unabhängig. Dann mache man die Auswahl $\mathsf{T}(\mathfrak{A}_2^*, N_2)$ und schließe analog weiter.

Zusatz: Macht man mit einem Kollektiv eine beliebige (also auch nicht zulässige), jedoch nicht dünne Teilung, so kann man selbstverständlich den Satz aus § 2 anwenden. Es gibt also wieder eine Verteilung und man kann dieselbe angeben. Jedoch ist die neue Merkmalfolge nicht notwendig ein Kollektiv, und man kann z. B. nicht schließen, daß eine Stellenauswahl die Verteilung weiter unverändert läßt.

Hauptsatz 4: *Verbindet man unabhängige Kollektivs* $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ *mit den Merkmalmengen* $M_1, \ldots, M_n$ *und den Wahrscheinlichkeiten* $w_1, \ldots, w_n$, *so erhält man ein Kollektiv mit der Merkmalmenge* $M_1 \times \ldots \times M_n$. *Die neue Wahrscheinlichkeit ist*

$$w(P^1 \times \ldots \times P^n) = w_1(P^1) \ldots w_n(P^n),$$

wobei P^ν irgendein Merkmal von $\mathfrak{K}_\nu$ bedeutet.

Beweis: Ist $\mathfrak{K}_1 = \mathfrak{A}_1 \ldots \mathfrak{A}_l$, $\mathfrak{K}_2 = \mathfrak{B}_1 \ldots \mathfrak{B}_m, \ldots, \mathfrak{K}_n = \mathfrak{Z}_1, \ldots, \mathfrak{Z}_r$, so ist wegen der Gültigkeit des assoziativen Gesetzes

$$\mathfrak{K}_1 \ldots \mathfrak{K}_n = \mathfrak{A}_1 \ldots \mathfrak{A}_l\,\mathfrak{B}_1 \ldots \mathfrak{B}_m \ldots \mathfrak{Z}_1 \ldots \mathfrak{Z}_r.$$

Wegen der vorausgesetzten Unabhängigkeit von $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ sind jetzt je zwei der $\mathfrak{A}, \mathfrak{B}, \ldots, \mathfrak{Z}$ unabhängige Kollektivs erster Art. Der Satz über die Wahrscheinlichkeiten folgt jetzt aus dem entsprechenden Satze für Kollektivs erster Art.

Zu beachten ist, daß zulässige Teilmengen der Merkmalmenge von $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ jetzt auf Grund der früheren Definition nur solche Punkt-

mengen sind, welche sich als Produkte der Teilmengen der Merkmalmengen von $\mathfrak{A}_1, \mathfrak{A}_2, \ldots, \mathfrak{A}_r$ mit nicht verschwindenden Wahrscheinlichkeiten oder, was dasselbe ist, welche sich als Produkte zulässiger Teilmengen von $\mathfrak{R}_1, \ldots, \mathfrak{R}_n$ darstellen lassen.

Schließlich gilt der fast selbstverständliche Satz: Sind zwei untereinander unabhängige Kollektivs zu einem dritten unabhängig, so ist es auch ihr Produkt.

§ 6.

Verbindung abhängiger Kollektivs.

Bei der Verbindung unabhängiger Kollektivs ergaben sich Kollektivs, deren Wahrscheinlichkeiten durch die Multiplikationsregel bestimmt sind, und es gelten die elementaren Hauptsätze der Wahrscheinlichkeitsrechnung. Wir wollen jetzt untersuchen, ob bei abhängigen Kollektivs ähnliches gilt.

Wir betrachten Grundkollektivs und abgeleitete Kollektivs erster Art. Es seien $\mathfrak{R}$ und $\mathfrak{R}'$ kritisch abhängig. $\mathfrak{R}$ habe die Verteilung $w(P_1), \ldots, w(P_r)$ und $\mathfrak{R}'$ die Verteilung $w'(Q_1), \ldots, w'(Q_s)$. Dann ist also $\mathsf{T}(\mathfrak{R}_1 \{P_\varrho\})$ für $\mathfrak{R}'$ kritisch und zu diesem T gehöre der s-stellige Multiplikator

$$K^\varrho = [k_1^\varrho, \ldots, k_s^\varrho].$$

Wie man leicht bestätigt, hat dann in $\mathfrak{R} \cdot \mathfrak{R}'$ das Merkmal $P_\varrho \times Q_\sigma$ die Wahrscheinlichkeit

$$w(P_\varrho)\, w'(Q_\sigma)\, k_\sigma^\varrho \frac{1}{\sum\limits_\varrho k_\sigma^\varrho\, w'(Q_\varrho)}$$

oder, was dasselbe sein muß,

$$w(P_\varrho)\, w'(Q_\sigma)\, k_\sigma^\varrho \frac{1}{\sum\limits_\sigma k_\sigma^\varrho\, w'(P_\sigma)}.$$

Es steht also fest, daß zu der Verbindung zweier kritisch abhängigen Kollektivs eine Verteilung gehört.

Wir können aber schon jetzt nicht mehr behaupten, daß eine Stellenauswahl Σ aus $\mathfrak{R} \times \mathfrak{R}'$ etwas erzeugt, was wieder als Verbindung kritisch abhängiger Kollektivs aufgefaßt werden kann. Vielmehr sind selbstverständlich $\mathfrak{R}\Sigma$ und $\mathfrak{R}'\Sigma$ abgeleitete Kollektivs und sie sind abhängig, sie brauchen aber nicht kritisch abhängig zu sein. Ferner haben $\mathfrak{R}\Sigma$ und $\mathfrak{R}'\Sigma$ natürlich die gleichen Verteilungen wie $\mathfrak{R}$ und $\mathfrak{R}'$, aber selbst, wenn sie kritisch abhängig sein sollten, kann es sein, daß zu $\mathfrak{R}\Sigma$, $\mathfrak{R}'\Sigma$ andere Multiplikatoren gehören als zu $\mathfrak{R}$ und $\mathfrak{R}'$. Zur Aufrechterhaltung der Hauptsätze werden wir also bei der Verbindung von kritisch abhängigen Kollektivs weitere Einschränkungen machen müssen.

Bei mehr als zwei Kollektivs macht schon die Frage nach der Existenz der Verteilung Schwierigkeiten. Es genügt dazu nämlich nicht,

wenn man verlangt, daß je zwei Faktoren kritisch abhängig sind. Vielmehr sind wir genötigt, umständlicher vorzugehen.

Es seien $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ Kollektivs. Man gebe aus der Merkmalmenge von $\mathfrak{K}_\nu$ eine Teilmenge N_ν vor und betrachte die Auswahl der Stellen, an denen gleichzeitig in $\mathfrak{K}_1$ ein Merkmal von N_1 in $\mathfrak{K}_2$ ein Merkmal von N_2 usw., schließlich in $\mathfrak{K}_n$ ein Merkmal aus N_n steht. Diese Auswahl nenne man die (*gleichzeitige*) *Teilung an* $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ *auf die Teilmengen* $N_1, \ldots, N_n$[4]).

Diese Auswahl ist offenbar nichts anderes als das Produkt

$$\mathsf{T}_1(\mathfrak{K}_1, N_1)\,\mathsf{T}_2(\mathfrak{K}_2\,\mathsf{T}_1, N_2)\ldots\mathsf{T}_n(\mathfrak{K}_n\,\mathsf{T}_1\ldots\mathsf{T}_{n-1}, N_n).$$

Offenbar ist dies ein Normalprodukt, also eine Auswahl aus $\mathfrak{S}$. Es läßt sich vielleicht, wegen der verschiedenen Darstellungsmöglichkeiten der $\mathfrak{K}_\nu$ durch die Grundkollektivs, in verschiedener Weise als Normalprodukt schreiben.

Ist $\mathfrak{K}_{n+1}$ ein weiteres Kollektiv, so heißen:

Definition: $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ *vor* $\mathfrak{K}_{n+1}$ *schaltbar*, wenn die gleichzeitigen Teilungen an $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ für $\mathfrak{K}_{n+1}$ sämtlich nicht irregulär sind.

Ferner heißen $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ *verbindungsfähig*, wenn für $\nu = 1, \ldots, n-1$ die $\mathfrak{K}_1, \ldots, \mathfrak{K}_\nu$ vor $\mathfrak{K}_{\nu+1}$ schaltbar sind.

Z. B. sind unabhängige Kollektivs immer verbindungsfähig, weil die in Frage kommenden Teilungen dann sämtlich regulär sind.

Ferner gilt bei zwei Faktoren offenbar: Wenn sie kritisch abhängig sind, sind sie auch verbindungsfähig.

Die verbindungsfähigen Kollektivs sind nun gerade die, deren Verbindung eine Verteilung hat. Nämlich es gilt:

Satz: *Sind* $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ (*in dieser Reihenfolge*) *verbindungsfähig, so gehört zur Verbindung* $\mathfrak{K}_1 \ldots \mathfrak{K}_n$ *eine Verteilung.*

Beweis: Ist $n = 2$, so ist der Satz bewiesen. Wir nehmen daher an, wir hätten den Satz für $n-1$. Nun sind $\mathfrak{K}_1, \ldots, \mathfrak{K}_{n-1}$ verbindungsfähig, also hat $\mathfrak{K}_1 \ldots \mathfrak{K}_{n-1}$ eine Verteilung. Nimmt man also von $\mathfrak{K}_\nu$ ein Merkmal P^ν, so hat $P^1 \times \ldots \times P^{n-1}$ in $\mathfrak{K}_1 \ldots \mathfrak{K}_{n-1}$ eine Wahrscheinlichkeit und weil $\mathfrak{K}_1, \ldots, \mathfrak{K}_{n-1}$ vor $\mathfrak{K}_n$ schaltbar sind, so hat in der Teilfolge, die durch gleichzeitige Teilung von $\mathfrak{K}_1, \ldots, \mathfrak{K}_{n-1}$ auf $P_1 \times \ldots \times P^{n-1}$ aus $\mathfrak{K}_n$ entsteht, weiter P^n eine Wahrscheinlichkeit. Daraus folgt: Zu $P^1 \times \ldots \times P^n$ gehört in $\mathfrak{K}_1 \ldots \mathfrak{K}_n$ eine Wahrscheinlichkeit, nämlich das Produkt der beiden angegebenen Wahrscheinlichkeiten.

Die Produkte verbindungsfähiger Kollektivs haben also eine Verteilung, im übrigen haben sie im allgemeinen nicht alle Kollektiveigenschaften, indem sie nicht die Hauptsätze erfüllen. Wir wollen die Produkte solcher

[4]) Dies ist die Teilung von $\mathfrak{K}_1 \ldots \mathfrak{K}_n$ auf $N_1 \times \ldots \times N_n$.

verbindungsfähiger Kollektivs, die nicht zu je zwei unabhängig sind, daher als *Semikollektivs* bezeichnen. Die Geltung der Hauptsätze wissen wir nur von den Verbindungen unabhängiger Kollektivs. Zur Erweiterung dieses Bereiches versuchen wir nun, Spezialeigenschaften von den Semikollektivs, genauer von den in ihnen vereinigten Kollektivs erster Art, zu fordern, so daß nach Hinzunahme dieser speziellen Semikollektivs zu den Kollektivs zweiter Art ein Bereich entsteht, in welchem die Hauptsätze gelten.

Macht man eine gleichzeitige Teilung aller $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$, so erhalte man $\mathfrak{K}_1^*, \ldots, \mathfrak{K}_n^*$. Dann ist uns selbstverständlich wegen der Verbindungsfähigkeit, also der Kenntnis der Multiplikatoren innerhalb $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ die Verteilung der Verbindung jeden Teilsystems der $\mathfrak{K}^*$ bekannt. Es ist mithin — zwar nicht sicher, daß die Teilungen $\mathsf{T}(\mathfrak{K}_1^*, \ldots)$ für $\mathfrak{K}_2^*$ nicht irregulär sind, aber — aus den Multiplikatoren innerhalb $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ ein Multiplikator berechenbar, welcher die Verteilung von $\mathfrak{K}_2^* \mathsf{T}(\mathfrak{K}_1^*, N_1^*)$ bestimmt. Ebenso ist allgemein zu jedem System von Teilmengen $N_1^*, \ldots, N_{\nu-1}^*$ zu $\mathfrak{K}_1^*, \ldots, \mathfrak{K}_{\nu-1}^*$ ein Multiplikator errechenbar, welcher die Verteilung von $\mathfrak{K}_\nu^* \mathsf{T}_1(\mathfrak{K}_1^*, N_1^*) \mathsf{T}_2(\mathfrak{K}_2^* \mathsf{T}_1, N_2^*) \ldots \mathsf{T}_{\nu-1}(\mathfrak{K}_{\nu-1}^* \mathsf{T}_1 \ldots \mathsf{T}_{\nu-2}, N_{\nu-1}^*)$ durch die Verteilung von $\mathfrak{K}_\nu^*$ bestimmt für jedes $\nu = 2, \ldots, n$. Es soll also nochmals hervorgehoben werden: Durch die Annahme der Hintereinanderschaltbarkeit von $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ sind für die $\mathfrak{K}_1^*, \ldots, \mathfrak{K}_n^*$, die aus den $\mathfrak{K}$ durch gleichzeitige Teilung entstehen, Multiplikatoren bestimmt, welche sukzessive die Verteilung von $\mathfrak{K}_1^* \mathfrak{K}_2^*$, $\mathfrak{K}_1^* \mathfrak{K}_2^* \mathfrak{K}_3^*$, ... und von $\mathfrak{K}_1^* \mathfrak{K}_2^*, \ldots, \mathfrak{K}_n^*$ bestimmen.

Definition: $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ heißen *verbindbar*, wenn gilt

1. $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ sind verbindungsfähig.

2. Für jedes Normalprodukt Π, das zu $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ gleichzeitig regulär ist, gilt: $\mathfrak{K}_1 \Pi, \ldots, \mathfrak{K}_n \Pi$ sind wieder verbindungsfähig und die Multiplikatoren sind die gleichen wie zu $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$.

3. Ist Π Normalprodukt, so daß jeder Faktor entweder für jedes $\mathfrak{K}_\nu$ regulär ist oder für (mindestens) ein $\mathfrak{K}_\nu$ singulär, so mögen die singulären Teilungen von Π bezüglich $\mathfrak{K}_\nu$ gehen auf die Teilmengen $T_1^\nu, \ldots, T_{r_\nu}^\nu$ der Merkmalmengen von $\mathfrak{K}_\nu$, und es sei $D_\nu = T_1^\nu \ldots T_{r_\nu}^\nu$. Ferner möge $\mathfrak{K}_\nu^*$ aus $\mathfrak{K}_\nu$ durch eigentliche Teilung auf die Teilmenge D_ν entstehen. Dann gilt: Mit $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ sind auch die $\mathfrak{K}_1 \Pi, \ldots, \mathfrak{K}_n \Pi$ verbindungsfähig und die Multiplikatoren sind die gleichen wie die von $\mathfrak{K}_1^*, \ldots, \mathfrak{K}_n^*$.

Die Forderungen 2 und die allgemeinere 3 besagen offenbar: Die Multiplikatoren, die die Abhängigkeit der $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ bestimmen, sollen von Stellenauswahlen und Teilungen an unabhängigen Kollektivs unabhängig sein. Sie sind also nur abhängig von singulären Teilungen. Die Abhängigkeit hiervon ist offenbar von vornherein bestimmt, ohne daß darüber noch etwas ge-

fordert werden müßte. Es ist ungefähr so: Kollektivs sind verbindbar, wenn die *Art ihrer Verwandtschaft durch Stellenauswahlen und Teilungen an unabhängigen Kollektivs nicht berührt wird.*

Offenbar sind dann unabhängige Kollektivs verbindbar.

Die Verbindung verbindbarer Kollektivs, von denen nicht je zwei unabhängig sind, nenne man *Kollektivs dritter Art.*

Offenbar gilt: Ist Π für $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ regulär oder singulär, so sind mit $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ auch $\mathfrak{K}_1 \Pi, \ldots, \mathfrak{K}_n \Pi$ verbindbar.

In diesem Bereich aller Kollektivs erster, zweiter oder dritter Art nenne man schließlich zwei Kollektivs — solche erster Art werden als Produkt mit einem Faktor aufgefaßt — *unabhängig*, wenn jeder Faktor des einen Produkts von iedem Faktor des anderen Produkts unabhängig ist. Dann gilt in diesem Bereich:

Hauptsatz 1: *Ist $\mathfrak{K}$ ein Kollektiv und Σ eine Stellenauswahl, so ist $\mathfrak{K}\,\Sigma$ ein Kollektiv und hat die gleiche Verteilung wie $\mathfrak{K}$.*

Beweis: Ist $\mathfrak{K}$ Kollektiv dritter Art, so sei $\mathfrak{K} = \mathfrak{K}_1 \ldots \mathfrak{K}_n$ und jedes $\mathfrak{K}_r$ Kollektiv erster Art. Dann ist $\mathfrak{K}\,\Sigma = (\mathfrak{K}_1\,\Sigma)\,(\mathfrak{K}_2\,\Sigma), \ldots, (\mathfrak{K}_n\,\Sigma)$ und diese sind verbindbar und wegen Forderung 2 hat ihre Verbindung die gleiche Verteilung wie $\mathfrak{K}_1 \ldots \mathfrak{K}_n$.

Hauptsatz 2: *Es seien $\mathfrak{K}_1$, $\mathfrak{K}_2$ unabhängige Kollektivs. $\mathsf{T}(\mathfrak{K}_2, \ldots)$ sei zulässige Teilung von $\mathfrak{K}_2$. Dann ist $\mathfrak{K}_1\,\mathsf{T}$ ein Kollektiv mit der gleichen Verteilung wie $\mathfrak{K}_1$.*

Beweis: Offenbar ist $\mathsf{T}(\mathfrak{K}_2, \ldots)$ eine reguläre Operation für jeden Faktor von $\mathfrak{K}_1$.

Hauptsatz 3: *Jede zulässige Teilung an $\mathfrak{K}$ erzeugt aus $\mathfrak{K}$ wieder ein Kollektiv.*

Das folgt aus Bedingung 3.

Hauptsatz 4: *Bei der Verbindung unabhängiger Kollektivs gilt die Multiplikationsregel.*

Beweis: Der Satz folgt offenbar aus Hauptsatz 2.

§ 7.

Hinzunahme gemischter Kollektivs.

Ein Kollektiv ist für uns bisher entweder ein abgeleitetes Kollektiv erster Art oder ein Kollektiv zweiter Art. Jetzt soll ein letztes Mal, und zwar nicht mehr sehr wesentlich, der Begriff des Kollektivs erweitert werden, indem wir alle diejenigen Merkmalfolgen, die aus einem Kollektiv durch Mischung entstehen, ebenfalls noch als Kollektiv auffassen.

Definition: Unter einem *gemischten Kollektiv* verstehen wir eine Merkmalfolge, welche aus einem Kollektiv erster oder zweiter Art durch Mischung entsteht. Wir denken uns aber ein gemischtes Kollektiv gegeben

(nicht nur durch eine Merkmalfolge, sondern) durch ein Kollektiv erster oder zweiter oder dritter Art, das Originalkollektiv, und eine Merkmaltransformation[5]), wobei einfach Systeme von Merkmalen zu neuen Merkmalen zusammengefaßt werden, wobei die Additionsregel gilt, wie in § 2 besprochen wurde. Wir nennen ein gemischtes Kollektiv von erster, zweiter oder dritter Art, je nachdem das Originalkollektiv von erster oder zweiter oder dritter Art ist.

Nach § 1 gehört zu jedem gemischten Kollektiv eine Verteilung.

Definition: Eine Teilmenge der Merkmalmenge eines gemischten Kollektivs heiße *zulässig*, wenn es eine zulässige Teilmenge der Merkmalmenge des Originalkollektivs gibt, aus der sie durch die Mischungstransformation entsteht. Eine Teilung an einem gemischten Kollektiv heiße zulässig, wenn es eine Teilung auf eine zulässige Teilmenge ist. Offenbar ist dann bei gemischten Kollektivs erster Art jede Teilmenge nicht verschwindender Wahrscheinlichkeit und jede dichte Teilung zulässig,

Definition: Zwei gemischte Kollektivs heißen *unabhängig*, wenn die Originalkollektivs unabhängig sind.

Definition: Gemischte Kollektivs heißen *verbindbar*, wenn die Originalkollektivs verbindbar sind.

Wir verstehen nunmehr endgültig unter einem Kollektiv ein Originalkollektiv oder ein gemischtes Kollektiv. Je zwei von ihnen sind dann entweder unabhängig oder nicht. Zwei Produkte, wo jeder Faktor des einen von jedem Faktor des anderen unabhängig ist, sind wieder unabhängig.

Es gelten dann, wie jetzt leicht folgt, *die Hauptsätze* 1 *bis* 4 *mit wörtlich unverändertem Wortlaut im Bereich aller Kollektivs. Durch die Operationen: Stellenauswahl, zulässige Teilung durch unabhängige Kollektivs, zulässige Teilung durch das Kollektiv selbst, Verbindung von verbindbaren Kollektivs und (jetzt auch) Mischung kommt man also, von Kollektivs ausgehend, aus dem Bereiche aller Kollektivs nicht heraus.*

§ 8.

Theorie des Glückspiels.

In der ersten Mitteilung ist ein Spezialfall besprochen worden. Es wurde angenommen, daß sämtliche Grundkollektivs ohne Nachwirkung sind und daß jede Familie nur ein Glied enthält, daß also zwei verschiedene Grundkollektivs immer fremd sind.

---

[5]) Man vergleiche, wenn nötig, den Begriff Mischung und Merkmaltransformation bei R. v. Mises oder in einer meiner Noten: Math. Zeitschr. Bd. 32 oder Jahresbericht der D. M. V. Bd. 43.

Teilungen an einem anderen Grundkollektiv waren also immer regulär, und nie kritisch. Ferner wurden Teilungen am gleichen Grundkollektiv, wenn sie nicht regulär oder singulär waren, immer als irregulär bezeichnet, so daß wesentlich kritische Teilungen überhaupt nicht auftraten. Der Name wurde so gewählt, weil ich mir vorstellte, daß beim Glücksspiel als Haupttypus, etwa in einem Roulettesaal, einerseits die einzelnen Versuchsfolgen ohne Nachwirkung sind und andererseits zwischen den verschiedenen Spielen eine Abhängigkeit nicht besteht.

§ 9.

Beispiel einer Glücksspieltheorie.

Es sei $\mathfrak{G}_1 = 1, 1, 1, \ldots$, $\mathfrak{G}_2 = 1, 0, 1, 0, \ldots$, $\mathfrak{S}$ bestehe allein aus der Einheit E. Wir behaupten: Diese Grundkollektivs und diese Gruppe von Stellenauswahlen genügen den Axiomen I, II. Daß das Axiom I erfüllt ist, ist klar. Wir stellen nun die Gruppe $\hat{\mathfrak{S}}$ auf. Ist A: 1, 3, 5, ..., B: 2, 4, 6, ..., welche durch Teilung von $\mathfrak{G}$ entstehen, so ergibt sich leicht $\hat{\mathfrak{S}}$ als die Gesamtheit aller Produkte von Elementen A und B: $\mathsf{A}^{\alpha_1}\,\mathsf{B}^{\alpha_2}\,\mathsf{A}^{\alpha_3}\ldots$ $[\mathsf{A}^0 = \mathsf{B}^0 = \mathsf{E}]$. Wendet man eine solche Auswahl auf $\mathfrak{G}_2$ (für $\mathfrak{G}_1$ ist das selbstverständlich) an, so bekommt man ein konstantes Kollektiv. Daraus folgt, daß trivialerweise das Axiom II erfüllt ist.

§ 10.

Durch Zusatzaxiom über die Stellenauswahl spezialisierte Theorie und Gesetze der großen Zahlen.

Die hier angegebene Theorie ist insofern abstrakt, als darüber, welche Auswahlen denn nun Stellenauswahlen sind, nichts vorausgesetzt wurde. Nun hat die Wahrscheinlichkeitsrechnung bisher immer solche Fragen beantwortet, wie z. B. diese: Wie wahrscheinlich ist es, daß beim Würfeln mit einem Würfel auf zwei nacheinander fallende Fünfen eine 1 folgt? In unserer abstrakten Theorie lassen sich diese und ähnliche, wie überhaupt jede konkrete Frage nur durch einen hypothetischen Satz beantworten: „Wenn gewisse Auswahlen für gewisse Merkmalfolgen als Grundkollektivs Stellenauswahlen sind, ...“ Will man solche hypothetischen Sätze formal vermeiden, so kann das geschehen, indem man weitere Axiome postuliert, darüber, welche Auswahlen Stellenauswahlen sind. Eine solche durch ein Zusatzaxiom spezialisierte Theorie, in welcher die oben angedeuteten Fragen beantwortet werden, soll jetzt besprochen werden.

Die Auswahl $\varkappa, \varkappa + 1, \ldots$ bezeichne man mit $\mathsf{P}_\varkappa$ die Auswahl $1, 1+\lambda, 1+2\lambda, \ldots$ mit Δ_λ. Dann wird z. B. $\mathsf{P}_\varkappa \Delta_\lambda$ die Auswahl $\varkappa, \varkappa+\lambda, \varkappa+2\lambda, \ldots$. Neben den Axiomen I, II, fordere man jetzt noch folgendes

Axiom III: $\mathsf{P}_\varkappa$ $(\varkappa = 1, 2, \ldots)$ *und* Δ_λ $(\lambda = 1, 2, \ldots)$ *sind Stellenauswahlen.*

Dann sagen wir, es liege eine spezielle Theorie vor. Dann sind auch die Produkte von $\mathsf{P}'s^-$ und $\Delta' s$ Stellenauswahlen. Daraus folgt: Jede arithmetische Progression erzeugt eine Stellenauswahl.

Ist also $\mathfrak{A}$: $a_1, a_2, \ldots$ ein Kollektiv ohne Nachwirkung in dieser speziellen Theorie, so folgt, daß neben $\mathfrak{A} = \mathfrak{A}_1$ auch

$$\mathfrak{A}_2 = a_2, a_3, \ldots$$
$$\mathfrak{A}_3 = a_3, a_4, \ldots$$

Kollektivs und zwar unabhängige Kollektivs sind, von denen jedes die gleiche Verteilung hat wie $\mathfrak{A}$. Die Betrachtung der Verbindung $\mathfrak{A}_1 \cdot \mathfrak{A}_2 \cdot \mathfrak{A}_{2+\nu}$ liefert dann z. B. die Beantwortung der Frage nach der Wahrscheinlichkeit dafür, daß in $\mathfrak{A}$ die Merkmale $P_\varkappa$, P_λ aufeinander folgen und dann im Abstand ν das Merkmal P_ϱ folgt. Diese Wahrscheinlichkeit ist $w(P_\varkappa)\ w(P_\lambda)\ w(P_\varrho)$, unter w die Wahrscheinlichkeit in $\mathfrak{A}$ verstanden. Ferner ergibt sich z. B. die Wahrscheinlichkeit für P_ϱ, wenn man schon weiß, daß das $(\nu+1)$-te vorhergehende (es könnte auch das folgende sein) und das ν-te vorhergehende Element $P_\varkappa$, P_λ sind, als $w(P_\varrho)$.

Etwas ungenauer ausgedrückt, ergibt sich, wie dieses Beispiel statt des allgemeinen Falles zur Genüge zeigt: Wählt man aus einem Kollektiv eine neue Merkmalfolge aus, indem man zur Auswahl nur die Merkmale der dem auszuwählenden Element vorhergehenden oder folgenden Elemente benutzt — diese Benutzung muß aber geschehen durch Verknüpfung der P und Δ —, so ergibt sich ein neues Kollektiv mit unveränderter Verteilung.

§ 11.

Anwendung.

R. v. Mises hat erkannt, daß mittels der fünf Grundoperationen: Stellenauswahl, Teilung an sich selbst, Teilung an unabhängigen Kollektivs, Verbindung und Mischung alle wichtigen Probleme der Wahrscheinlichkeitsrechnung sich erledigen lassen. Um zwei der wichtigsten Beispiele zu behandeln, wollen wir jetzt nach dem Vorbilde von R. v. Mises die Gesetze der großen Zahlen — wenigstens, nur der kürzeren Ausdrucksweise wegen, in zwei konkreten Spezialfällen — behandeln. Für das erste Gesetz der großen Zahlen beginnen wir mit einigen vorbereitenden Begriffen: $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ seien unabhängige Kollektivs. Ihre Merkmalräume mögen dieselbe Dimension haben. P^ν bedeute ein Merkmal der Merkmalmenge M_ν von $\mathfrak{K}_\nu$. Durch Verbindung entsteht das neue Kollektiv $\mathfrak{K} = \mathfrak{K}_1 \ldots \mathfrak{K}_n$. Die Merkmale von $\mathfrak{K}$ sind die Punkte $P^1 \times P^2 \times \ldots \times P^n$ und die Merkmalmenge ist $M_1 \times M_2 \ldots \times M_n$. Mit $\mathfrak{K}$ mache man jetzt eine Mischung,

indem man $P^1 \times P^2 \times \ldots \times P^n$ ersetzt durch $P^1 + P^2 + \ldots + P^n$ [6]). Aus $\mathfrak{K}$ entsteht so das neue Kollektiv $\overline{\mathfrak{K}}$. Wir sagen, $\overline{\mathfrak{K}}$ sei entstanden durch Addition der unabhängigen Kollektivs $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ und schreiben kurz $\overline{\mathfrak{K}} = \mathfrak{K}_1 + \ldots + \mathfrak{K}_n$.

Man transformiere $\mathfrak{K}$ noch einmal, indem man jedes Merkmal P von $\mathfrak{K}$ durch $\frac{P}{n}$ ersetzt. Aus $\overline{\mathfrak{K}}$ entsteht so $\overline{\overline{\mathfrak{K}}}$. Da es durch eine Mischung entsteht, ist es wieder ein Kollektiv, wir sagen $\overline{\overline{\mathfrak{K}}}$ sei aus $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ durch Durchschnittsbildung entstanden und schreiben $\overline{\overline{\mathfrak{K}}} = \frac{\mathfrak{K}_1 + \ldots + \mathfrak{K}_n}{n}$. Gehen wir nun von einem einzigen Kollektiv $\mathfrak{A}$ aus, so sind in unserer speziellen Theorie die Folgen $\mathfrak{A}_1 = \mathfrak{A} = a_1, a_2, \ldots$; $\mathfrak{A}_2 = \mathfrak{A}\,\mathsf{P}_2 = a_2, a_3, \ldots$, $\mathfrak{A}_n = \mathfrak{A}\,\mathsf{P}_n = a_n, a_{n+1} \ldots$ wieder Kollektivs, und zwar sind sie voneinander unabhängig.

Dann sind $\overline{\mathfrak{A}}_n = \mathfrak{A}_1 + \ldots + \mathfrak{A}_n$ und $\overline{\overline{\mathfrak{A}}}_n = \frac{\mathfrak{A}_1 + \ldots + \mathfrak{A}_n}{n}$ wieder Kollektivs. Wir sagen: $\overline{\mathfrak{A}}_n$ entstehe aus $\mathfrak{A}$ durch n-fache Summenbildung, $\overline{\overline{\mathfrak{A}}}_n$ durch n-fache Durchschnittsbildung.

Ist $\mathfrak{K}$ ein Kollektiv mit den ϱ-dimensionalen Merkmalen $P_1, \ldots, P_l$ und den Wahrscheinlichkeiten w, so nennen wir den ϱ-dimensionalen Punkt $\Sigma P_\lambda w(P_\lambda) = T(\mathfrak{K})$ den Mittelwert von $\mathfrak{K}$. Aus dieser Bedeutung von T folgt sofort $\sum_\lambda (P_\lambda - T) w(P_\lambda) = 0$.

Neben dem Begriff des Mittelwertes brauchen wir noch den Begriff der Streuung: Der ϱ-dimensionale Punkt $\sqrt{\sum_\lambda (P_\lambda - T)^2 w(P_\lambda)} = \mathfrak{S}(\mathfrak{K})$ heißt die Streuung von $\mathfrak{K}$.

Sind dann $\mathfrak{K}_1, \ldots, \mathfrak{K}_n$ irgendwelche unabhängige Kollektivs gleicher Dimension mit den Mittelwerten $T^1, \ldots, T^n$, so ist,

$$T(\mathfrak{K}_1 + \ldots + \mathfrak{K}_n) = \sum_{P^1 \ldots P^n} (P^1 + \ldots + P^n)\, w_1(P^1) \ldots w_n(P^n),$$

summiert über alle Punktsysteme $P^1, \ldots, P^n$. Dies ist aber gleich

$$\sum P^1 w_1(P^1) [\sum w_2(P^2) \ldots w_n(P^n)] + \ldots$$
$$+ \sum_{P_n} P^n w_n(P^n) [\sum w_1(P^1) \ldots w_{n-1}(P^{n-1})] = T^1 + \ldots + T^n,$$

weil

$$\sum_{P^2 \ldots P^n} w_2(P^2) w_3(P^3) \ldots w_n(P^n) = \sum_{P^2} w_2(P^2) \ldots \sum_{P^n} w_n(P^n) = 1$$

ist, ebenso die übrigen eingeklammerten Faktoren. Ferner ist $T\left(\frac{\mathfrak{K}_1 + \ldots + \mathfrak{K}_n}{n}\right)$ $= \frac{1}{n} T(\mathfrak{K}_1 + \ldots + \mathfrak{K}_n)$, also gleich $\frac{T^1 + \ldots + T^n}{n}$.

---

[6]) Wir addieren Punkte, indem wir das koordinatenweise machen, ebenso werden wir quadrieren und Wurzelziehen von Punkten koordinatenweise machen.

Wir beweisen nun den Satz, daß das Streuungsquadrat eines Kollektivs, welches durch Addition von unabhängigen Kollektivs entsteht, gleich der Summe der Streuungsquadrate ist, also

$$(\mathfrak{S}(\mathfrak{R}_1 + \ldots + \mathfrak{R}_n))^2 = (\mathfrak{S}(\mathfrak{R}_1))^2 + \ldots + (\mathfrak{S}(\mathfrak{R}_n))^2.$$

Denn

$$(\mathfrak{S}(\mathfrak{R}_1 + \ldots + \mathfrak{R}_n))^2 = \sum_{P^1 \ldots P^n} [(P^1 + \ldots + P^n)] - [T^1 + \ldots + T^n]^2 w_1(P^1) \ldots w_n(P^n)$$

$$= \sum_{P'} (P^1 - T^1)^2 w_1(P') \Big[\sum_{P^2 \ldots P^n} w_2(P^2) \ldots w_n(P^n) \Big] + \ldots$$

$$+ \sum (P^n - T^n)^2 w_n(P^n) \Big[\sum_{P^1 \ldots P^{n-1}} w_1(P^1) \ldots w_{n-1}(P^{n-1}) \Big]$$

$$+ 2 \sum_{P' P^2} (P' - T')(P^2 - T^2) w_1(P) w_2(P^2) \Big[\sum_{P^3 \ldots P^n} w_3(P^3) \ldots w_n(P^n) \Big] + \ldots.$$

Hier folgen dem letzten entsprechende Glieder, in denen das Produkt von zwei Differenzen $(P^\varkappa - T^\varkappa)(P^\lambda - T^\lambda)$ auftritt. Solche Glieder sind aber 0, denn es ist z. B. das letzte ausgeschriebene Glied gleich

$$\Sigma (P^1 - T^1) w_1(P_1) \cdot \Sigma (P^2 - T^2) w_2(P^2) \cdot 1 = 0,$$

da die beiden ersten Faktoren verschwinden, also erhält man, da rechts alle eckigen Klammern den Wert 1 haben: $(\mathfrak{S}(\mathfrak{R}_1))^2 + \ldots + (\mathfrak{S}(\mathfrak{R}_n))^2$, was zu beweisen war.

Offenbar folgt daraus für die Durchschnittsbildung

$$\Big(\mathfrak{S}\Big(\frac{\mathfrak{R}_1 + \ldots + \mathfrak{R}_n}{n}\Big)\Big)^2 = \frac{(\mathfrak{S}(\mathfrak{R}_1))^2 + \ldots + (\mathfrak{S}(\mathfrak{R}_n))^2}{n^2}.$$

Wir wenden diesen Satz an auf die Durchschnittsbildung von unabhängigen Kollektivs mit gleicher Streuung. Wir wollen speziell als Kollektiv wählen $\mathfrak{A}_1, \ldots, \mathfrak{A}_n$, die wir oben aus $\mathfrak{A}$ erzeugt haben. Dann ist $\mathfrak{A} = a_1, a_2, \ldots$ und $\bar{\bar{\mathfrak{A}}}_n = \frac{a_1 + \ldots + a_n}{n}, \frac{a_2 + \ldots + a_{n+1}}{n} \ldots$. Weil jetzt $\mathfrak{A}_1, \ldots, \mathfrak{A}_n$ die gleiche Verteilung, also erst recht die gleiche Streuung $\mathfrak{S}(\mathfrak{A})$ haben, folgt $\mathfrak{S}(\bar{\bar{\mathfrak{A}}}_n) = \mathfrak{S}\frac{1}{\sqrt{n}}$. Bildet man mithin den n-ten Durchschnitt mit wachsendem n, so ergibt sich dabei die Streuung als ein ϱ-dimensionaler Punkt, welcher nach 0 konvergiert.

Aus diesem Ergebnis folgt jetzt, wie allgemein bekannt, das Gesetz der großen Zahlen so: Um den Mittelwert $T(\mathfrak{A})$ lege man die feste Kugel mit dem Radius R. Die Wahrscheinlichkeit dafür, daß ein Merkmal von $\mathfrak{A}$ in das Innere dieser Kugel fällt, sei w_{ni} die, daß es in das Äußere fällt, w_{na}. Wir berücksichtigen nun, daß $T(\mathfrak{A})$ gleichzeitig der Mittelwert von $\bar{\bar{\mathfrak{A}}}_n$ ist für jedes n. Dann folgt aus der Definition der Streuung, daß die Streuung des Kollektivs $\bar{\bar{\mathfrak{A}}}$ vom Nullpunkt mindestens die Entfernung $\sqrt{w_{na} R^2}$ hat. Weil aber der Radius R fest angenommen war, folgt dann, daß bei wachsendem n die Wahrscheinlichkeit w_{na} nach 0,

also $w_{n\,i}$ nach 1 konvergiert. Das ist das *erste Gesetz der großen Zahlen*. Man kann es ungefähr so aussprechen: Bei n-facher Durchschnittsbildung geht die Wahrscheinlichkeit dafür, daß ein Merkmal in das Innere einer unabhängig von n vorgelegten Kugel um den Mittelwert des Kollektivs fällt, mit wachsendem n nach 1, wird also mit wachsendem n fast zur Sicherheit.

Wir wollen noch kurz einen speziellen Fall des *zweiten Gesetzes der großen Zahlen* behandeln. Auf einem Tische liegen drei Würfel A, B, C. Wir nehmen an, daß bei jedem Würfel jeder Augenzahl eine Wahrscheinlichkeit zukommt, welche erperimentell festgestellt ist und angibt, welche Wahrscheinlichkeiten sich ergeben, wenn man mit diesem Würfel unendlich oft würfelt. Bei dem ersten Würfel seien die Wahrscheinlichkeiten der Ziffern 1 bis 6: $\alpha_1, \ldots, \alpha_6$ beim zweiten: $\beta_1, \ldots, \beta_6$, beim dritten: $\gamma_1, \ldots, \gamma_6$. Jemand greift nun einen Würfel nach Belieben heraus und wirft n-mal nacheinander. Es erscheine etwa n-mal die 5. Wir fragen: Wie groß ist die Wahrscheinlichkeit, daß mit dem ersten Würfel A gewürfelt wurde?

Damit wir eine Frage in unserer Theorie erhalten, müssen wir dieser Fragestellung erst in unserer Theorie ein neues Problem zuordnen, das wir für äquivalent halten. Dazu müssen wir uns jedenfalls an Stelle des einzelnen Vorganges eine (unendliche) Merkmalfolge konstruieren. Denken wir uns also den Vorgang: Ergreifen eines der drei Würfel und danach n-faches Werfen mit ihm und Ablesen der Augen unendlich oft wiederholt. Wir erhalten dann eine unendliche Merkmalfolge $\mathfrak{F}$. Die Merkmale sind $n+1$-dimensional. Die erste Komponente ist das Merkmal A, B, C, die zweite bis $(n+1)$-te Komponente sind Zahlen $a_1, \ldots, a_n$ der Reihe 1 bis 6. Wir machen jetzt folgende Hypothesen: 1. Achtet man in der Folge F nur auf die ersten Komponenten der Merkmale, so ergibt sich für A, B, C die gleiche Wahrscheinlichkeit, also $\frac{1}{3}$. 2. Achtet man nur auf die Merkmale, in denen die erste Komponente A bzw. B bzw. C ist, und hierin nur auf die $(\nu+1)$-te Komponente, also a_ν, so ergibt sich eine Merkmalfolge der Merkmale 1 bis 6 mit den Wahrscheinlichkeiten $\alpha_1, \ldots, \alpha_6$ bzw. $\beta_1, \ldots, \beta_6$ bzw. $\gamma_1, \ldots, \gamma_6$. 3. Die n Merkmalfolgen, die man erhält, wenn man nur alle die Glieder nimmt, in denen die erste Komponente A bzw. B bzw. C ist, und dort die $(\nu+1)$-te Komponente a_ν betrachtet, genügen, angesehen als Grundkollektivs, wenn man das Element E als Gruppe $\mathfrak{S}$ annimmt, den Axiomen I, II (hier braucht man Axiom III nicht zu fordern). Dann ergeben sich für die Folge folgende Wahrscheinlichkeiten: Für die $n+1$-dimensionalen Merkmale P: $A, a_1, \ldots, a_n$ bzw. $B, a_1, \ldots, a_n$ bzw. $C, a_1, \ldots, a_n$ der Reihe nach $\frac{1}{3}, \alpha_{a_1}, \ldots, \alpha_{a_n}$ bzw. $\frac{1}{3}, \beta_{a_1}, \ldots, \beta_{a_n}$ bzw. $\frac{1}{3}, \gamma_{a_1}, \ldots, \gamma_{a_n}$.

Die Frage: Weiß man, daß ein vorgelegtes System $\bar{a}_1, \ldots, \bar{a}_n$ herauskommt, wie groß ist dann die Wahrscheinlichkeit, daß die erste Komponente A ist, daß also mit dem Würfel A gewürfelt wurde, wird dann so beantwortet: Man teile $\mathfrak{F}$ auf die Teilmenge der Merkmalmenge, in welcher die zweite bis $(n+1)$-te Kompenente gerade $\bar{a}_1, \ldots, \bar{a}_n$ sind. Die Wahrscheinlichkeit für das Auftreten von A ist dann

$$\tfrac{1}{3}\alpha_{\bar{a}_1} \ldots \alpha_{\bar{a}_n} \cdot \frac{1}{\tfrac{1}{3}\alpha_{\bar{a}_1} \ldots \alpha_{\bar{a}_n} + \tfrac{1}{3}\beta_{\bar{a}_1} \ldots \beta_{\bar{a}_n} + \tfrac{1}{3}\gamma_{\bar{a}_1} \ldots \gamma_{\bar{a}_n}}.$$

Wir wollen nun zur Anfangsfrage zurückkehren und annehmen, es sei n-mal 5 gefallen. Ferner nehmen wir jetzt an, es sei $\alpha_5 > \beta_5$ und $\alpha_5 > \gamma_5$. Offenbar wächst dann der Quotient $\frac{\alpha_5^n}{\beta_5^n + \gamma^n}$ mit wachsendem n über alle Grenzen. Daher konvergiert $\frac{1}{3}\alpha_5^n \frac{1}{\frac{1}{3}\alpha_5^n + \beta_5^n + \gamma_5^n}$ mit wachsendem n nach 1, d. h. also: Die Wahrscheinlichkeit dafür, daß eine n-fache Serie von Fünfen mit dem Würfel geworfen wurde, für welchen 5 eine größere Wahrscheinlichkeit hat als für jeden anderen Würfel, geht mit wachsendem n nach 1, wird also mit wachsendem n fast zur Sicherheit.

Endlich wollen wir in Ergänzung einer früheren Bemerkung nochmals nachweisen, daß in einem Kollektiv ohne Nachwirkung der Zufall „ohne Gedächtnis" ist. Sei $\mathfrak{K} = a_1, a_2 \ldots$ ein solches Kollektiv in einer speziellen Theorie.

Es seien

$$\begin{aligned} \mathfrak{K}_2 &= a_2, a_3, \ldots \\ &\cdots\cdots\cdots \\ \mathfrak{K}_{100} &= a_{100}, a_{101}, \ldots \\ \mathfrak{K}_{101} &= a_{101}, a_{102}, \ldots \end{aligned}$$

$\mathfrak{K}^1, \ldots, \mathfrak{K}^{100}$ sind dann unabhängig. Ferner ist die Verbindung $\mathfrak{K}^1 \ldots \mathfrak{K}^{100}$ unabhängig zu $\mathfrak{K}^{101}$. Jede zulässige Teilung an $\mathfrak{K}^1 \ldots \mathfrak{K}^{100}$ ist mithin eine reguläre Operation an $\mathfrak{K}^{101}$ und ändert die Verteilung von $\mathfrak{K}^{101}$ nicht ab. Daraus folgt, daß der Ausfall in $\mathfrak{K}$ von dem Ausfall der hundert vorangegangenen Versuche nicht abhängt. Wenn das Roulette unsere Axiome erfüllt, so kann mithin aus der Kenntnis von 100 aufeinanderfolgenden Ergebnissen über das Folgende nichts ausgesagt werden, es kann weder vorausgesagt noch auch mit einer anderen Sicherheit als ohne die Kenntnis der 100 Resultate erwartet werden.

(Eingegangen am 8. Mai 1934.)