Vom Aussterben der Geschlechter : einige Betrachtungen über Markoffsche Ketten mit Beispielen aus der Familienstatistik

Autor(en): **Ineichen, Robert**

Objekttyp: **Article**

Zeitschrift: **Mitteilungen der Naturforschenden Gesellschaft Luzern**

Band (Jahr): **21 (1967)**

PDF erstellt am: **13.09.2024**

Persistenter Link: https://doi.org/10.5169/seals-523385

Vom Aussterben der Geschlechter

Einige Betrachtungen über Markoffsche Ketten
mit Beispielen aus der Familienstatistik

von ROBERT INEICHEN, Luzern

1. Einleitung

> *Évidemment* tous les noms doivent s'éteindre... Un mathématicien pourrait calculer comment la réduction des noms ou titres aurait lieu, d'après la *probabilité* des naissances toutes féminines ou toutes masculines ou mélangées et la probabilité du défaut de naissances dans un couple quelconque.

So schreibt Alphonse de Candolle (1806—1893) in seiner «Histoire des sciences et des savants depuis deux siècles» [13] [1]; der Genfer Naturforscher – übrigens ein Sohn des den Botanikern wohlbekannten Augustin-Pyramus de Candolle – zitiert in diesem Zusammenhang den englischen Anthropologen Francis Galton, der sich intensiv mit der Frage beschäftigt hat, wieso einst verbreitete Geschlechter mit der Zeit nur noch wenige Vertreter haben oder sogar ganz aussterben. – Ist dies nun wirklich so «évident», wie de Candolle schreibt? Was läßt sich mit Hilfe der Wahrscheinlichkeitsrechnung über die Verminderung oder die Zunahme der Anzahl der Träger eines Familiennamens aussagen, welches ist insbesonders die Wahrscheinlichkeit dafür, daß ein Geschlecht ausstirbt?

Mit diesen Fragen wollen wir uns im folgenden etwas beschäftigen. Zwei Gründe haben uns veranlaßt, diesen Problemen nachzugehen: Einmal der, daß derartige Überlegungen auf Grund von Zahlenmaterial aus der Schweiz wohl noch kaum angestellt worden sind [2], und im weitern, weil – wie in den letzten Jahrzehnten erkannt worden ist – Fragestellung und Lösungsmethoden eine Übertragung oder Erweiterung in Bereiche gestatten, die von der Familienstatistik recht weit abliegen, so etwa in die Theorie chemischer oder physikalischer Kettenreaktionen, in die Unternehmungsforschung (Operations research) bei der Theorie der Warteschlangen oder schließlich auf gewisse biologische Probleme, vor allem in der Genetik. – Da vielleicht diese Zusammenhänge auch für einen größeren Leserkreis von einem gewissen Interesse sind, sollen sie im folgenden ebenfalls gestreift werden. Dazu ist es notwendig, etwas weiter auszuholen.

[1] Diese Bezeichnungen in eckiger Klammer verweisen auf die am Schlusse angeführte Literatur.

[2] Die Anregung zu diesen Überlegungen verdanke ich Herrn Professor Dr. E. Batschelet, Basel/Washington.

2. Vom Schema der unabhängigen Versuche zu den Markoffschen Ketten

In den Naturwissenschaften werden häufig zufällige Prozesse (stochastische Prozesse) beobachtet. Man denke an die Brownsche Bewegung, jene zufallsartige Bewegung kleiner Teile in einer Flüssigkeit, erzeugt durch die auftreffenden bewegten Moleküle der Flüssigkeit, oder an den Austritt von Elektronen aus der Glühkathode einer Röhre oder schließlich an den radioaktiven Zerfall.

Wir wollen zunächst einen wesentlich einfacheren zufälligen Prozeß betrachten. Wir denken uns eine Urne, die eine Anzahl Kugeln enthalte, welche zum Teil mit E_1, zum Teil mit E_2, ... usw. bis E_s bezeichnet sind; im übrigen seien die Kugeln völlig gleich gestaltet und die Urne gut durchmischt. Wir ziehen eine Kugel und achten auf die Aufschrift, dann legen wir die Kugel wieder zurück, mischen und führen einen weitern Zug aus, usw. Wir stellen fest: Die Wahrscheinlichkeit, daß in einem Versuch die Kugel E_i erscheint, $1 \leqq i \leqq s$, *hängt nicht von der Nummer des Versuches und nicht von den in früheren Versuchen eingetroffenen Ereignissen ab.* Diese Wahrscheinlichkeit werde mit p_i bezeichnet. Dieses *Schema einer Folge unabhängiger Versuche* ist eines der wichtigsten Schemata der Wahrscheinlichkeitsrechnung. – Die Wahrscheinlichkeit, daß in n derartigen unabhängigen Versuchen das Ereignis E_1 genau m_1-mal, das Ereignis E_2 genau m_2-mal, ... das Ereignis E_s genau m_s-mal eintritt, wobei natürlich $m_1 + m_2 + \ldots + m_s = n$, läßt sich elementar berechnen. Sie ist gegeben durch die Formel

$$W_n(m_1, m_2, \ldots m_s) = \frac{n!}{m_1!\, m_2! \ldots m_s!} \cdot p_1^{m_1}\, p_2^{m_2} \ldots p_s^{m_s} \tag{1}$$

Den wichtigsten Spezialfall von nur *zwei* möglichen Ereignissen, die sich gegenseitig ausschließen, hat bekanntlich als erster Jakob Bernoulli (1655—1705) behandelt; man spricht dann vom *Bernoullischen Schema;* die Bernoullische Formel stellt einen Sonderfall von Formel (1) dar.

Nun nehmen wir eine naheliegende Verallgemeinerung vor: Die Wahrscheinlichkeit, daß in einem Versuch E_i erscheint, *hänge davon ab, welches Ereignis im unmittelbar vorangegangenen Versuch eingetreten ist.* Wir denken uns nun s verschiedene Urnen, jede enthalte wieder eine Anzahl solcher Kugeln, die zum Teil mit E_1, mit E_2 ... mit E_s beschriftet seien. Die Zusammensetzung der Urnen mit diesen Kugeln sei aber verschieden. Wir bezeichnen die Wahrscheinlichkeit, aus der Urne Nummer i eine

Kugel mit der Aufschrift E_k zu ziehen, mit P_{ik}. Wir gehen nun so vor: Zu Beginn der Versuchsfolge sei die Wahrscheinlichkeitsverteilung der E_i gegeben. Wenn nun E_i eingetreten ist, so ziehen wir aus der Urne Nummer i eine Kugel. Mit der Wahrscheinlichkeit P_{i1} erscheint dann E_1, mit der Wahrscheinlichkeit P_{i2} erscheint $E_2, \ldots$ mit der Wahrscheinlichkeit P_{is} erscheint E_s. – Die P_{ik} sind also *bedingte* Wahrscheinlichkeiten, es ist

$$P_{ik} = W\,(E_k\,/\,E_i)$$

Ein solches Versuchsschema ist zuerst vom russischen Mathematiker A. A. Markoff (1856—1922) systematisch untersucht worden; man nennt es eine *Markoffsche Kette.*

Allgemein spricht man im folgenden Falle von einer *Markoffschen Kette*[3]: Es sei eine Folge von Versuchen gegeben; als Ergebnis eines Versuches möge genau eines der endlich oder abzählbar unendlich vielen Ereignisse $E_1, E_2, E_3 \ldots$ eintreten; diese Ereignisse sollen sich paarweise ausschließen. *Die Wahrscheinlichkeit für den Eintritt des Ereignisses* E_i *in irgendeinem Versuch hängt nun nur davon ab, welches Ereignis im vorangegangenen Versuch eingetreten ist und nicht von den in noch früheren Versuchen eingetretenen Ereignissen.*

In einer etwas andern Terminologie spricht man oft von einem *physikalischen System*, das sich in einem bestimmten Zeitpunkt in einem der *Zustände oder Phasen* E_i, mit $i = 1, 2, \ldots$ befinden kann und nur in den gegebenen Zeitpunkten $t_1, t_2, \ldots$ von einem Zustand in einen andern übergehen kann. Die bedingte Wahrscheinlichkeit $P_{ik} = W\,(E_k\,/\,E_i)$ heißt dann die *Übergangswahrscheinlichkeit*, also die Wahrscheinlichkeit des Überganges vom Zustand E_i in den Zustand E_k.

Diese Übergangswahrscheinlichkeiten lassen sich übersichtlich in der *Übergangsmatrix* anordnen:

$$(2) \qquad M_1 = \begin{Vmatrix} P_{11} & P_{12} & P_{13} & \ldots\ldots. \\ P_{21} & P_{22} & P_{23} & \ldots\ldots. \\ P_{31} & P_{32} & P_{33} & \ldots\ldots. \\ . & . & . & \ldots\ldots. \\ . & . & . & \ldots\ldots. \end{Vmatrix}$$

[3] Genauer: eine *einfache, homogene* Markoffsche Kette. *Einfach,* weil nur vom Ergebnis des unmittelbar vorangehenden Versuches abhängig und nicht wie in komplizierteren Fällen vom Ergebnis von $m > 1$ vorangehenden Versuchen; *homogen,* weil unabhängig von der Nummer des Versuches. Vgl. [2], [3].

Offenbar ist diese Matrix quadratisch; ihre Elemente sind nicht negativ und die Summe der Elemente einer *Zeile* ist 1.
Zusammen mit der Wahrscheinlichkeitsverteilung der Zustände E_1, $E_2, \ldots$ zu Beginn der Versuchsfolge definiert die Übergangsmatrix die Markoffsche Kette vollständig.

Ein einfaches *Beispiel:* Das Wetter eines Ortes, das während eines Tages herrscht, lasse sich durch zwei Zustände «schlecht» (E_1) und «gut» (E_2) charakterisieren. Die Übergangswahrscheinlichkeiten wären dann gegeben durch

$$M_1 = \begin{Vmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{Vmatrix} .$$

Hier gibt also zum Beispiel P_{12} die Wahrscheinlichkeit eines Überganges von einem Schlechtwettertag zu einem Gutwettertag. Allerdings setzen wir dabei voraus, daß die Wahrscheinlichkeit für gutes oder schlechtes Wetter an einem Tage nur davon abhängt, welcher Zustand am vorangehenden Tage geherrscht habe. Die Meteorologen sagen uns übrigens, daß $P_{11} > P_{12}$ und $P_{22} > P_{21}$. Hier kommt eine gewisse Beharrungstendenz der Wetterlagen zum Ausdruck.

Für das weitere Vorgehen bringen wir noch die folgenden Formeln aus der Theorie der Markoffschen Ketten:
Mit $P_{ij}(n)$ wollen wir die Übergangswahrscheinlichkeit bezeichnen, in n Versuchen vom Zustand E_i in den Zustand E_j zu gelangen.
Nun ist

$$(3) \qquad P_{ij}(2) = \sum_k P_{ik} P_{kj} \; ,$$

denn für ein festes k ist die Wahrscheinlichkeit des Überganges $E_i \rightarrow E_k \rightarrow E_j$ gegeben durch $P_{ik}P_{kj}$; anschließend hat man einfach über alle k zu summieren. Aus (3) erkennt man nun sofort, daß die Matrix der Übergangswahrscheinlichkeit in zwei Schritten gegeben ist durch

$$M_2 = M_1 \cdot M_1 = M_1^2.$$

Durch entsprechende Überlegungen findet man die *Markoffsche Gleichung*

$$(4) \qquad P_{ij}(n) = \sum_k P_{ik}(m)\, P_{kj}(n-m),$$

wobei m eine natürliche Zahl mit $1 \leqq m < n$ ist.

Mit Hilfe von (4) läßt sich für die Matrix der *Übergangswahrscheinlichkeiten in n Schritten* die Formel gewinnen

(5) $M_n = M_1{}^n$. – Man vergleiche [2] oder [3].

3. Der Galton-Watson-Prozeß

Wir wollen nun versuchen, das einleitend skizzierte Problem als eine Markoffsche Kette mit gewissen besondern Eigenschaften aufzufassen. Wir kommen dann auf einen stochastischen Prozeß, den wir im Anschluß an T. Harris in [5] als *Galton-Watson-Prozeß*[4] bezeichnen und wie folgt definieren:

a) Wir denken uns Objekte, die weitere Objekte *derselben Art*, ihre Nachkommen, erzeugen können, also zum Beispiel Männer und ihre männlichen Nachkommen.
b) Die am Anfang gegebene Menge solcher Objekte nennen wir die *nullte* Generation, ihre Nachkommen bilden die erste Generation, deren Nachkommen die zweite Generation usw.
c) Wir richten unser Augenmerk auf die *Anzahl* solcher Objekte jeder Generation und bezeichnen mit Z_0, Z_1, Z_2 ... die Anzahl dieser Objekte in der nullten, ersten, zweiten ... Generation. Diese Z_i, mit $i = 0, 1, 2, \ldots$, sollen *zufällige Variable* darstellen. Wir werden stets $Z_0 = 1$ voraussetzen; die Wahrscheinlichkeit, daß $Z_1 = k$, mit $k = 0, 1, 2, \ldots$, bezeichnen wir mit p_k:

$$W(Z_1 = k) = p_k\,, \qquad \sum_k p_k = 1$$

Für jedes Objekt in irgendeiner Generation sei nun die Wahrscheinlichkeit, im Laufe seines ganzen Lebens k Nachkommen zu haben (k = 0, 1, 2,...), wieder durch das obige p_k *gegeben.* Diese Wahrscheinlichkeit hänge also nicht etwa davon ab, wieviel Objekte schon vorhanden sind, sie variiere auch nicht mit der Zeit.

[4] Nach F. Galton, der sich — wie in der Einleitung erwähnt — mit dem Problem des Aussterbens der Geschlechter befaßt hat und H. W. Watson, der eine erste mathematische Lösung dieses Problems gegeben hat (1874). — Der Galton-Watson-Prozeß stellt einen besonders einfachen Typ eines *Verzweigungsprozesses (branching process)* dar; man vergleiche das Buch [5] von T. Harris.

d) $Z_0, Z_1, Z_2, \ldots$ sollen eine (einfache, homogene) *Markoffsche Kette* bilden. Die Wahrscheinlichkeit, daß in der (n+1)ten Generation eine Anzahl Objekte, zum Beispiel j, vorkommen werden, hängt also nur davon ab, wie groß die Anzahl i der Objekte der n-ten Generation war:

$$W(Z_{n+1} = j / Z_n = i) = P_{ij}, \quad i, j, n = 0, 1, 2 \ldots; Z_0 = 1 \tag{6}$$

Diese P_{ij} stellen also wieder die im vorangehenden Abschnitt definierten *Übergangswahrscheinlichkeiten* dar.

Diese Übergangswahrscheinlichkeiten sollen nicht mit der Zeit veränderlich sein. Ist $Z_n = 0$, so ist offenbar mit der Wahrscheinlichkeit 1 auch $Z_{n+1} = 0$:

$$W(Z_{n+1} = 0 / Z_n = 0) = P_{00} = 1 \tag{7}$$

Als bedingte Wahrscheinlichkeit ist die Übergangswahrscheinlichkeit (6) nicht definiert für jene i, für die $W(Z_n = i) = 0$.

In dieser allgemeinen Formulierung ist das durch den Galton-Watson-Prozeß gegebene Schema auf recht verschiedenartige Beispiele anwendbar.
Wir werden in den weitern Ausführungen vor allem das folgende Beispiel betrachten: Wir nehmen einen *männlichen* Neugeborenen als *Stammvater eines Geschlechtes*, als Begründer einer Linie (0. Generation). p_k stelle die Wahrscheinlichkeit dar, daß dieser Neugeborene im Laufe seines Lebens k männliche Nachkommen erzeuge, $k = 0, 1, 2 \ldots$ Seine Nachkommen bilden die 1. Generation. Unabhängig von Einflüssen der Vererbung, der Zeit oder der Umwelt bestehe nun für jeden männlichen Neugeborenen der 1. Generation wieder die Wahrscheinlichkeit p_k, genau k männliche Nachkommen zu haben, wiederum $k = 0, 1, 2, \ldots$ Die Menge aller männlichen Nachkommen der 1. Generation bilden die 2. Generation; für jeden männlichen Neugeborenen der 2. Generation bestehe wieder die Wahrscheinlichkeit p_k, k männliche Nachkommen zu haben usf. Wir fragen nach der Wahrscheinlichkeit, in einer Generation eine bestimmte Anzahl männlicher Nachkommen zu finden, vor allem nach der Wahrscheinlichkeit, daß in einer Generation 0 männliche Nachkommen sein werden, daß also die Linie erlischt, das *Geschlecht ausstirbt*.

Es mag überraschen, daß ein solches Schema auch auf gewisse *physikalische Vorgänge* paßt; A. Weinberg und E. Wigner schreiben in [11], «The physical Theory of Neutron Chain Reactors»: «*It is intresting to note that some of the basic mathematical methods developed by A. Lotka*[5] *and V. Volterra to deal with multiplication in biological populations are applicable to the study of fluctuations in neutron chain reactors.*» Bei einer solchen *Kettenreaktion* wird durch ein Neutron (0. Generation) geeigneter Energie ein schwerer Kern gespalten. Dabei entstehen einige neue Neutronen, beim Uran 235 zum Beispiel zwei bis drei (1. Generation). Besitzen diese wieder die geeignete Energie, so vermögen sie neue Kerne zu spalten, wobei wieder neue Neutronen entstehen (2. Generation), usf. Die Zahl der Neutronen kann lawinenartig anwachsen, falls nicht so viele Neutronen aus dem Prozeß herausgelöst werden können, daß sich dieser stabilisiert. Solche Betrachtungen hat zum Beispiel E. Schrödinger in [9] angestellt.

Wie N. Semenoff [10] gezeigt hat, können auch *chemische Kettenreaktionen* in gewissen Fällen durch einen Galton-Watson-Prozeß beschrieben werden.

Schließlich weisen wir noch auf folgende Situation aus einem ganz andern Bereich von Anwendungen hin: An einem Schalter, an dem immer nur ein Kunde abgefertigt werden kann, wird ein Kunde bedient (0. Generation). Alle Kunden, die während der Zeit eintreffen, da dieser Kunde abgefertigt wird, gelten als seine «Nachkommen» und bilden die erste Generation; es sei p_k die Wahrscheinlichkeit, daß während der Bedienungszeit des Kunden der 0. Generation k weitere Kunden ankommen und Schlange stehen müssen. Für jeden Kunden der ersten Generation bestehe wieder die Wahrscheinlichkeit p_k, daß während seiner Bedienungszeit k weitere Kunden eintreffen; $k = 0, 1, 2, \ldots$ Alle Kunden, die eintreffen, während ein Kunde bedient wird, der der 1. Generation angehört, bilden die 2. Generation usw. Natürlich stellt sich gerade hier die Frage, ob das Schema des Galton-Watson-Prozesses hier nicht zu vereinfachend sei: die p_k könnten beispielsweise zeitlich variieren. Wir kommen darauf unten noch zurück. Eine ähnliche Situation mag sich beim Anfall von Reparaturen einer bestimmten Warengattung stellen: Ein Stück wird repariert (0. Generation), während dieser Zeit kommen mit der Wahrscheinlichkeit p_k k weitere Stücke an (1. Generation) und müssen warten usw. Es stellt sich unter anderm die Frage nach der Wahrscheinlichkeit dafür, daß die Warteschlange abbricht. Die *Theorie der Warteschlangen*, ein bedeutsames Gebiet der *Unternehmensforschung*, untersucht derartige Probleme; man vergleiche etwa [1]. Wir müssen es uns versagen, hier noch weitere Beispiele anzuführen; Hinweise auf analoge Problemstellungen in der *Genetik* gibt zum Beispiel W. Feller in [2].

Wir skizzieren indessen noch kurz eine Situation, auf die unser Schema schlecht passen würde: Vor dem roten Licht einer Verkehrsampel warte ein Fahrzeug (0. Generation). Alle Fahrzeuge, die während des Wartens dieses Fahrzeuges eintreffen, bilden die erste Generation; p_k sei die Wahrscheinlichkeit, daß während dieses Wartens k Fahrzeuge eintreffen, $k = 0, 1, 2, \ldots$ Nun erscheint kurz grünes Licht, das Fahrzeug der 0. Generation und wohl auch etliche Fahrzeuge der 1. Generation können weiterfahren, einzelne kommen vielleicht nicht mehr durch. Alle Fahrzeuge, die jetzt eintreffen, während noch Fahrzeuge der 1. Generation weiter warten, bilden die zweite Generation usf. Hier wird im allgemeinen unser Schema nicht anwendbar sein: Die ersten Fahrzeuge der 1. Generation werden in der Regel beim Aufleuchten von Grün sofort durchkommen, für sie ist dann $p_0 = 1$, $p_k = 0$ für $k = 1, 2, \ldots$; für gewisse spätere Fahrzeuge der 1. Generation dürfte hingegen $p_0 \neq 1$ sein und damit auch $p_k \neq 0$ für $k = 1, 2, \ldots$ Damit sind die Voraussetzungen für den Galton-Watson-Prozeß nicht mehr gegeben.

[5] Vgl. zum Beispiel [6].

Wir wollen abschließend noch ausdrücklich auf die *vereinfachenden Annahmen* hinweisen, die dem Galton-Watson-Prozeß zugrunde liegen: Wir setzen die p_k als *zeitunabhängig* voraus. Wenn die p_k also etwa die Wahrscheinlichkeit eines männlichen Neugeborenen angeben, im Laufe seines Lebens k männliche Nachkommen zu haben, so soll diese Wahrscheinlichkeit im Laufe der Zeit nicht ändern. Bestimmen wir nun die relativen Häufigkeiten von männlichen Neugeborenen mit $k = 0, 1, 2, \ldots$ männlichen Nachkommen, um Schätzwerte für die p_k zu erhalten, so müssen wir indessen darauf hinweisen, daß diese relativen Häufigkeiten im Laufe der Jahre ziemlichen *Schwankungen* unterworfen sind, wie zum Beispiel A. Burckhardt in der Arbeit [12] sehr schön zeigt.
Diese p_k sollen weiter auch nicht davon abhängen, *wie viele Objekte* der betreffenden Art *schon vorhanden* sind. Die einzelnen Objekte erzeugen also unabhängig voneinander weitere Objekte derselben Art. Diese Voraussetzung kann sicher bei der Beschreibung gewisser biologischer Populationen nicht gemacht werden, wo in verschiedenen Fällen die Gesamtzahl der schon vorhandenen Objekte auf die Erzeugung von Nachkommen einen Einfluß haben kann.
Schließlich setzen wir voraus, daß die p_k auch nicht davon abhängen sollen, ob das Objekt, das Nachkommen erzeugt, aus einer *Familie mit vielen oder mit wenigen Nachkommen* stammt. Diese Voraussetzung dürfte für unser Problem des Aussterbens der Geschlechter nur bedingt zutreffen; es ist durchaus denkbar, daß ein Nachkomme aus einer Familie mit vielen Brüdern die Tendenz haben könnte, auch eine größere Anzahl von Nachkommen zu haben als ein anderer aus einer kleineren Familie.

4. *Erzeugende Funktionen*

Wir betrachten die Folge $\{Z_0, Z_1, Z_2, \ldots\}$ von Zufallsvariablen, wobei wie im vorangehenden Abschnitt ausgeführt, Z_i die Anzahl Objekte der i-ten Generation angibt. Wiederum sei ferner $Z_0 = 1$ und es sei die Wahrscheinlichkeitsverteilung für Z_1 gegeben durch $\{p_0, p_1, p_2, \ldots p_\omega\}$, wenn ω die größtmögliche Anzahl von Nachkommen eines Objektes ist. Man führt nun zweckmäßig eine *erzeugende Funktion* ein (vgl. zum Beispiel Feller in [2]):

$$(8)\qquad f(s) = p_0 + p_1 s + p_2 s^2 + \ldots + p_\omega s^\omega$$

Wir bemerken sofort:

$$(9)\qquad f'(1) = p_1 + 2p_2 + 3p_3 + \ldots + \omega\, p_\omega = \sum_k k p_k = m$$

stellt den *Erwartungswert* der Zahl der direkten Nachkommen eines Objektes dar; ferner ist

$$(10)\qquad f(1) = \sum_k p_k = 1 \quad \text{und} \quad f(0) = p_0.$$

Nun suchen wir die erzeugenden Funktionen für $Z_2, Z_3, \ldots$; wir stellen fest:

a) Nach den bekannten Sätzen über Multiplikation und Addition von Wahrscheinlichkeiten und unter Verwendung von Formel (1) gibt

$$(11)\qquad w = \sum \frac{r!}{m_0!\, m_1! \ldots m_\omega!}\, p_0^{m_0}\, p_1^{m_1} \ldots p_\omega^{m_\omega}$$

die Wahrscheinlichkeit w dafür, daß r Objekte genau t Nachkommen im Laufe ihres Lebens haben, wenn sich die Summation über alle m_k, $k = 0, 1, 2 \ldots \omega$, erstreckt mit

$$m_0 + m_1 + m_2 + \ldots + m_\omega = r \quad \text{und}$$
$$0 \cdot m_0 + 1 \cdot m_1 + 2 \cdot m_2 + \ldots + \omega \cdot m_\omega = t.$$

Nach dem polynomischen Lehrsatz (E. Netto [7]) ist nun aber die Summe (11) mit den beiden obigen Summationsbedingungen gerade der Koeffizient von s^t in $[f(s)]^r$.

b) Die Wahrscheinlichkeit, daß das Objekt der 0. Generation (der «Stammvater») gerade r Nachkommen in der 1. Generation («Söhne») und t Nachkommen in der 2. Generation («Enkel») habe, läßt sich aus (11) durch $p_r \cdot w$ berechnen, das heißt durch den Koeffizienten von s^t in $p_r[f(s)]^r$.

c) Somit ergibt sich die Wahrscheinlichkeit dafür, daß der Begründer des Geschlechtes, der Stammvater, in der 2. Generation t Nachkommen hat, welches auch immer die Zahl seiner Nachkommen in der 1. Generation sei, also die Wahrscheinlichkeit $W(Z_2 = t)$ als Koeffizient von s^t in der Summe

$$p_0 + p_1 f(s) + p_2 [f(s)]^2 + \ldots + p_\omega [f(s)]^\omega = f[f(s)] = f_2(s)$$

Das heißt doch: Die erzeugende Funktion für Z_2 ist die zusammengesetzte Funktion

$$f_2(s) = f\,[f(s)].$$

d) Durch vollständige Induktion finden wir, daß die erzeugende Funktion für Z_{n+1}, also für die Zahl der Objekte der $(n+1)$ten Generation rekursiv gegeben ist durch

$$(12) \qquad f_1(s) = f(s) \quad \text{und} \quad f_{n+1}(s) = f\,[f_n(s)],\, n = 1, 2, \ldots$$

Wir bemerken noch, daß sich diese Ergebnisse auch durch die Benützung allgemeiner Sätze über erzeugende Funktionen von Summen gleich verteilter Zufallsvariablen beweisen lassen; diese finden sich zum Beispiel bei W. Feller in [2].

Beispiel: Es sei $p_0 = 0{,}5$; $p_1 = 0{,}25$; $p_2 = p_3 = 0{,}125$; $p_k = 0$ für $k > 3$; also $m = 0{,}875$.
Dann ist $f(s) = 0{,}5 + 0{,}25s + 0{,}125\, s^2 + 0{,}125\, s^3$; $f_2(s) = f\,[f(s)]$ ergibt dann die erzeugende Funktion für Z_2, aus der wir die Wahrscheinlichkeiten $W(Z_2 = k)$, $k = 0, 1 \ldots 9$ finden können. Die Rechnung ergibt zum Beispiel
$W(Z_2 = 0) = 6{,}7 \cdot 10^{-1}$; $W(Z_2 = 1) = 1{,}2 \cdot 10^{-1}$; $W(Z_2 = 2) = 7{,}8 \cdot 10^{-2}$, usf. schließlich $W(Z_2 = 9) = 2{,}4 \cdot 10^{-4}$.
Analog ergibt $f_3(s) = f\,[f_2(s)]$ die Wahrscheinlichkeiten $W(Z_3 = k)$, $k = 0, 1, \ldots 27$, etwa $W(Z_3 = 0) = 7{,}6 \cdot 10^{-1}$; $W(Z_3 = 1) = 6{,}9 \cdot 10^{-2}$, usf., schließlich $W(Z_3 = 27 =)\, 1{,}8 \cdot 10^{-12}$.

5. Die Wahrscheinlichkeit des Aussterbens der Linie

Wir gehen wieder aus von der Folge der Zufallsvariablen $\{Z_0, Z_1, Z_2, \ldots\}$, wobei also Z_i die Anzahl Glieder der i-ten Generation darstellt und $Z_0 = 1$ ist. Wenn ein Glied dieser Folge, etwa Z_n, den Wert 0 annimmt, dann nehmen wegen Gleichung (7)

$$W(Z_{n+1} = 0 / Z_n = 0) = P_{00} = 1$$

auch alle folgenden mit Sicherheit ebenfalls den Wert 0 an: *Die Linie stirbt aus, das Geschlecht erlischt.*

Dieses Erlöschen des Geschlechtes ist offenbar nur möglich, wenn $p_0 \neq 0$; wir setzen deshalb im folgenden $0 < p_0 < 1$ voraus.

Nun suchen wir die Wahrscheinlichkeit q_n, daß das Geschlecht in oder vor der n-ten Generation erlischt:

$$q_n = W(Z_n = 0).$$

Nach unsern Ausführungen über erzeugenden Funktionen im vorigen Abschnitt ist

$$q_n = W(Z_n = 0) = f_n(0).$$

Da $f(s) = \sum_{k=0}^{\omega} p_k s^k$ im Intervall $0 < s < 1$ mit zunehmendem s ebenfalls zunimmt, folgt (vgl. W. Feller in [2]) aus

$$q_1 = f(0) = p_0, \text{ daß } q_2 = f[f(0)] = f(q_1) > f(0) = q_1,$$

und durch vollständige Induktion

$$q_{n+1} = f(q_n) > f(q_{n-1}) = q_n.$$

Das heißt aber: Die Folge der q_n strebt monoton gegen ein q, das der Gleichung

$$q = f(q) = p_0 + p_1 q + p_2 q^2 + \ldots + p_\omega q^\omega \text{ genügt.} \tag{13}$$

Dieses der Gleichung (13) genügende q stellt somit die *Wahrscheinlichkeit dafür dar, daß das Geschlecht im Laufe der Zeit ausstirbt.*

Übrigens läßt sich Gleichung (13) im Anschluß an J. Steffensen [8] sehr anschaulich deuten:

Die Wahrscheinlichkeit q des Aussterbens des Geschlechtes ist eine zusammengesetzte Wahrscheinlichkeit aus

- der Wahrscheinlichkeit p_0, daß in der 1. Generation 0 männliche Nachkommen vorhanden sind,
- der Wahrscheinlichkeit $p_1 q$, daß in der 1. Generation 1 männlicher Nachkomme vorhanden ist und die dadurch begründete Linie ausstirbt,
- der Wahrscheinlichkeit $p_2 q^2$, daß in der 1. Generation 2 männliche Nachkommen vorhanden sind und beide dadurch begründete Linien aussterben usf., schließlich aus der Wahrscheinlichkeit $p_\omega q^\omega$, daß in der 1. Generation ω männliche Nachkommen vorhanden sind und alle dadurch begründeten Linien aussterben.

Um dieses q zu finden, haben wir somit die im Intervall $0 \leqq s \leqq 1$ liegenden Wurzeln der Gleichung

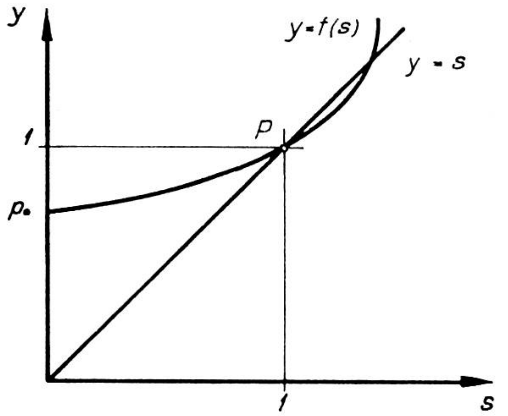
(14) $$s = f(s) = p_0 + p_1 s + p_2 s^2 + \ldots + p_\omega s^\omega$$

zu suchen. Sicher ist (wegen $\sum_{k=0}^{\omega} p_k = 1$) $s = 1$ eine Lösung.

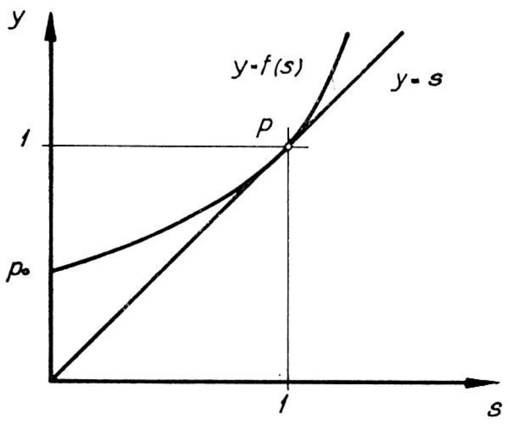
Nun gilt der folgende *grundlegende Satz* (vgl. zum Beispiel T. Harris [5]):

Wenn $m = \sum_{k=0}^{\omega} kp_k \leqq 1$, dann ist die Wahrscheinlichkeit für das Erlöschen des Geschlechtes 1. – Wenn $m > 1$, dann ist diese Wahrscheinlichkeit q die einzige nicht-negative Lösung, die kleiner als 1 ist, die die Gleichung (14) besitzt.

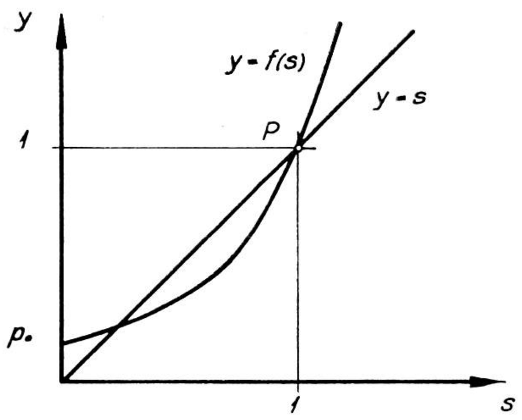
Auch dieser Satz läßt sich anschaulich deuten: Wir betrachten dazu die Graphen der beiden Funktionen $y = s$ und $y = f(s)$, die sich wegen $f(1) = 1$ im Punkte (1/1) schneiden. Der Graph von $y = f(s)$ ist konvex, somit sind nur die drei folgenden Fälle denkbar:



a) Zweiter Schnittpunkt außerhalb $0 \leqq s \leqq 1$



b) Berührung in P(1/1)



c) Zweiter Schnittpunkt innerhalb $0 \leqq s \leqq 1$

Betrachten wir nun die Steigung von $y = f(s)$ im Punkte $P(1/1)$, so erkennen wir, daß die beiden Fälle a) und b) durch $f'(1) = \sum_{k=0}^{\omega} kp_k = m \leqq 1$ und der dritte Fall c) durch $f'(1) = \sum_{k=0}^{\omega} kp_k = m > 1$ charakterisiert sind; mit andern Worten: Die Wahrscheinlichkeit q, daß ein Geschlecht ausstirbt, ist dann und nur dann kleiner als 1, wenn der *Erwartungswert der direkten Nachkommen eines Objektes größer als 1 ist,* wenn also die «mittlere Anzahl von Nachkommen» größer als 1 ist.

6. *Numerische Beispiele für das Aussterben der Geschlechter*

a) Wir wollen nun die im Intervall $0 \leqq s \leqq 1$ liegenden Wurzeln der Gleichung

(14) $$s = f(s) = p_0 + p_1 s + p_2 s^2 + \ldots + p_\omega s^\omega$$

suchen, wobei also die p_k, $k = 0, 1, 2, \ldots \omega$, die Wahrscheinlichkeiten dafür darstellen, daß ein Vertreter einer Generation im Laufe seines ganzen Lebens genau k männliche Nachkommen hat; ferner erinnern wir an die im dritten Abschnitt aufgeführten vereinfachenden Annahmen.

Um nun numerische Beispiele betrachten zu können, müssen wir für die p_k geeignete Zahlen einsetzen. Um Schätzwerte für diese p_k zu finden, nehmen wir eine passend ausgewählte Menge von N männlichen Neugebornen. Wir stellen fest, daß N_k davon im Laufe ihres

Lebens k männliche Nachkommen haben, $k = 0, 1, 2, \ldots \omega$, und berechnen die relativen Häufigkeiten $f_k = N_k/N$. Mit diesen f_k, oder wenn notwendig mit ausgeglichenen Werten bauen wir dann ein *Modell für die oben geschilderte vereinfachte Situation.*

b) Obwohl die Fragestellung für diese $f_k = N_k/N$ außerordentlich einfach ist, ist es nicht ohne weiteres möglich, aus vorhandenen Unterlagen eine Antwort zu gewinnen. Stichproben auf den Zivilstandsämtern sind infolge der getrennten Führung der Geburten- und Familienregister mit zahlreichen Schwierigkeiten verbunden. Aus den einläßlichen bevölkerungsstatistischen Publikationen des Eidgenössischen Statistischen Amtes läßt sich für unsere Frage auch nichts entnehmen. Das statistische Material von Versicherungskassen führt in der Regel nicht alle Kinder auf, sondern nur jene unterhalb einer gewissen Altersgrenze; zudem wäre die Anzahl männlicher Versicherter durch zusätzliche Annahmen zu jener oben genannten Menge N von Neugeborenen zu erweitern. Zahlreiche Publikationen familienstatistischer Art schließlich, so die sehr reichhaltige Studie von A. Moser [19] und die Arbeiten von A. Burckhardt [12], F. Kaufmann [16], A. Miller [17] und L. Solari [21], haben andere Zielsetzungen und ergeben deshalb zum Teil wenig für unsere Frage.
Immerhin zeigen gerade die zitierten Arbeiten von A. Moser und A. Burckhardt, wie unsere f_k Änderungen unterworfen sind, so etwa deshalb, weil infolge des Rückganges der Sterblichkeit immer mehr Neugeborene das heiratsfähige Alter erreichen und weil sich deutlich eine säkulare Steigerung der Heiratsfähigkeit feststellen läßt; man vergleiche auch die Arbeit [20] von A. Moser. – Wir weisen deshalb nochmals darauf hin, daß unsere Beispiele *Modelle* für stark vereinfachte Situationen darstellen. Es könnte das Ziel einer besondern Arbeit sein, entsprechende Überlegungen durchzuführen für p_k, die zum Beispiel mit der Zeit variieren.

c) Infolge der besondern Art der Registerführung der *Korporationsverwaltung Luzern*[6] ließen sich für *Luzerner Korporationsbürger* die $f_k = N_k/N$ einfach erheben. Wir ließen anfangs 1964 für die $N = 225$ in den Jahren 1880 ... 89 geborenen Korporationsbürger die Anzahl männlicher Nachkommen feststellen und erhielten folgende Tabelle:

(15)

k	0	1	2	3	4	5
$f_k = N_k/N$	0,630	0,196	0,102	0,058	0,009	0,005
$\sum_{k=0}^{5} f_k = 1{,}000$						

Setzen wir nun diese Werte für die p_k in die Gleichung (14) ein, so stellen wir sofort fest, daß die Gleichung

$$s = 0{,}630 + 0{,}196s + 0{,}102s^2 + \ldots + 0{,}005s^5$$

[6] Die Korporationsgemeinde entspricht der «Burgergemeinde» in andern Städten.

im Intervall $0 \leqq s \leqq 1$ nach dem im fünften Abschnitt mitgeteilten Satz nur die eine Wurzel $s = 1$ hat, da $m = 0{,}635 < 1$, das heißt Geschlechter, deren p_k die durch Tabelle (15) gegebenen Werte aufweisen, *sterben unter den von uns gemachten Voraussetzungen mit Sicherheit aus.*

Aufzeichnen auf einfach logarithmischem Papier legt den Versuch nahe, durch die Punkte (k/f_k) für $k \geqq 1$ eine Gerade nach der Methode der kleinsten Fehlerquadratsumme zu legen. Die Rechnung ergibt

$$(16) \qquad p_k = 0{,}654 \cdot e^{-0{,}977\,k} = p_1 \cdot 0{,}377^{\,k-1}, \quad k \geqq 1.$$

Dieser Ausgleich bewirkt natürlich, daß im so veränderten Modell auch für $k > 5$ eine nicht verschwindende Wahrscheinlichkeit vorhanden ist. Für p_0 ergibt sich dann aus $p_0 = 1 - \sum_{k=1}^{\infty} p_k$ der Wert $p_0 = 0{,}605$. – Die p_k stellen also in diesem Falle für $k \geqq 1$ eine geometrische Folge mit dem Quotienten $Q = 0{,}377$ dar.

An Stelle von Gleichung (14) haben wir dann die Gleichung

$$(17) \qquad s = p_0 + p_1 s + p_2 s^2 + \ldots \text{ oder mit Berücksichtigung von (16)}$$

$$s = p_0 + p_1 s + p_1\,0{,}377 s^2 + \ldots \quad \text{zu untersuchen.}$$

Wegen $\sum_{k=0}^{\infty} p_k = 1$ ist die rechts stehende Reihe in unserm Intervall $0 \leqq s \leqq 1$ konvergent.

Setzen wir für p_0 und p_1 die eben angegebenen Werte ein und summieren wir die geometrische Reihe, so erhalten wir schließlich die quadratische Gleichung $0{,}377 s^2 - 0{,}982 s + 0{,}605 = 0$, die wiederum in $0 \leqq s \leqq 1$ nur die Lösung $s = 1$ hat. Auch $m = \sum_{k=0}^{\infty} k p_k = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot 0{,}377\, p_1 + 3 \cdot 0{,}377^2\, p_1 + \ldots$ läßt sich einfach berechnen: Für $x = 0{,}377$ wird $\lim\limits_{n \to \infty} S_n$ berechnet, wobei $S_n = \sum_{k=1}^{n} k x^{k-1}$

$$= \frac{1 - x^n}{(1-x)^2} - \frac{n x^n}{1-x}, \text{ weil } (1-x)\,S_n = -n x^n + \sum_{k=0}^{n-1} x^k.$$

Man findet wieder $m = 0{,}635$.

d) Um ein anderes Modell untersuchen zu können, haben wir 1961 in den Familienregistern des *Zivilstandskreises Römerswil* die 449 Familien der vor 1896 geborenen Männer in bezug auf die Anzahl der männlichen Nachkommen untersucht; die große Mehrzahl der dabei untersuchten Männer ist in der zweiten Hälfte des 19. Jahrhunderts geboren. Wir haben die folgenden Zahlen erhalten:

Anzahl männliche Nachkom.	0	1	2	3	4	5	6	7	8	9	10
Anzahl Familien	102	111	107	62	42	14	8	1	1	0	1
relative Häufigkeit	0,227	0,247	0,238	0,138	0,094	0,031	0,018	0,002	0,002	0	0,002

Aus den Angaben von A. Moser in [20] dürfen wir schließen, daß in diesem Zeitraum vielleicht etwa 55 Prozent der Geborenen im Laufe ihres Lebens eine Familie gegründet haben. Somit läßt sich ungefähr der Prozentsatz der Geborenen berechnen, die im Laufe ihres Lebens $k = 0, 1, 2, \ldots 10$ männliche Nachkommen erhalten haben. Aufzeichnen auf einfach logarithmischem Papier legt es hier nahe, die Werte für $k \geqq 2$ nach der Methode der kleinsten Fehlerquadratsumme linear auszugleichen. Dann können wir unserm Modell die folgenden Zahlen für die p_k zu Grunde legen:

$$p_k = 0{,}53 \cdot e^{-0{,}70k} = p_2 \cdot 0{,}497^{k-2},\ k \geqq 2. \tag{18}$$

Dann wird $p_0 + p_1 = 1 - \sum\limits_{k=2}^{\infty} p_k = 0{,}74$; für p_0 und p_1 nehmen wir dementsprechend und gemäß den obigen Berechnungen $p_0 = 0{,}60$ und $p_1 = 0{,}14$.

Setzen wir diese Zahlen für p_k in Gleichung (17) ein, so führen analoge Betrachtungen wie im vorangehenden Beispiel auf die quadratische Gleichung

$$0{,}56s^2 - 1{,}16s + 0{,}60 = 0,$$

die im Intervall $0 \leqq s \leqq 1$ wiederum nur die Lösung $s = 1$ hat. Das heißt, daß auch in diesem Modell die Linien mit Sicherheit erlöschen.

e) In den letzten Jahren sind auch Untersuchungen über die «*ideale Familiengröße*» durchgeführt worden; A. Miller berichtet in [17] und [18] darüber, D. Hanhart in [15]. A. Miller zitiert in [17] die in der nachstehenden Tabelle aufgeführten Ergebnisse einer deutschen Umfrage aus dem Jahre 1958; er weist darauf hin, daß wir annehmen dürfen, eine Umfrage in der Schweiz würde nicht wesentlich verschiedene Resultate liefern.

Als *ideale Familiengröße* wird betrachtet:
von 0,6 Prozent aller Befragten die kinderlose Ehe;
von 3,5 Prozent aller Befragten die Ehe mit 1 Kind;
von 46,4 Prozent aller Befragten die Ehe mit 2 Kindern;
von 37,9 Prozent aller Befragten die Ehe mit 3 Kindern;
von 10,1 Prozent aller Befragten die Ehe mit 4 Kindern.

Da A. Miller ferner die *enge Beziehung* zwischen der «idealen» oder also der «gewünschten» Familiengröße und der tatsächlichen Kinderzahl feststellt, ist es gegeben, auch die obigen Zahlen zur Konstruktion eines Modells zu benützen. Nach A. Moser [20] dürfen wir annehmen, daß im Zeitraum, in dem die Befragten geboren worden sind, etwa 70 Prozent im Laufe ihres Lebens heiraten. Das gibt uns die Möglichkeit, den Prozentsatz der Geborenen zu berechnen, die im Laufe ihres Lebens 0, 1, 2, ... Kinder (männliche und weibliche Nachkommen!) haben. Wir verwenden in unserm Modell die so erhaltenen Zahlen für die Wahrscheinlichkeit π_k eines Neugeborenen, im Laufe seines Lebens k Kinder zu haben; wir bekommen so

$$\pi_0 = 0{,}30; \pi_1 = 0{,}02; \pi_2 = 0{,}33; \pi_3 = 0{,}27; \pi_4 = 0{,}07;$$

$$\text{wir setzen} \quad \pi_5 = 1 - \sum_{k=0}^{4} \pi_k = 0{,}01.$$

Die gesuchten Werte für die p_k finden wir nun, wenn wir mit Hilfe der π_k die Wahrscheinlichkeit berechnen, daß ein Neugeborener im Laufe seines Lebens k = 0, 1, 2, ... *männliche* Nachkommen erhält. Es genügt für diese Untersuchung, anzunehmen, die Wahrscheinlichkeit, daß eine Geburt eine Knabengeburt sei, betrage $^1/_2$.
Dann ist nach den elementaren Sätzen der Wahrscheinlichkeitsrechnung

$$p_0 = \pi_0 + {}^1/_2 \cdot \pi_1 + {}^1/_4 \cdot \pi_2 + {}^1/_8 \cdot \pi_3 + {}^1/_{16} \cdot \pi_4 + {}^1/_{32} \cdot \pi_5 = 0{,}43$$
$$p_1 = {}^1/_2 \cdot \pi_1 + 2 \cdot {}^1/_4 \cdot \pi_2 + 3 \cdot {}^1/_8 \cdot \pi_3 + 4 \cdot {}^1/_{16} \cdot \pi_4 + 5 \cdot {}^1/_{32} \cdot \pi_5 = 0{,}30$$
$$p_2 = {}^1/_4 \cdot \pi_2 + 3 \cdot {}^1/_8 \cdot \pi_3 + 6 \cdot {}^1/_{16} \cdot \pi_4 + 10 \cdot {}^1/_{32} \cdot \pi_5 = 0{,}21$$

analog $p_3 = 0{,}053$, $p_4 = 0{,}005$, $p_5 = 0{,}002$.

Bilden wir nun Gleichung (14), so ist im Intervall $0 \leqq s \leqq 1$ wiederum nur die Wurzel $s=1$ vorhanden, da $m = \sum_{k=0}^{5} kp_k = 0{,}91 < 1$; auch solche Linien erlöschen mit Sicherheit. – Zu ähnlichen Überlegungen könnten Zahlen herangezogen werden, die D. Hanhart in [15] gibt.

f) Man kann noch einen grundsätzlich andern Weg einschlagen, um zu Werten für die p_k zu kommen: Wir untersuchen die *Nachkommenschaft eines Stammvaters* und bestimmen in jeder Generation die Anzahl männlicher Nachkommen, die im Laufe ihres Lebens $k = 0, 1, 2, \ldots \omega$ männliche Nachkommen erhalten haben. Wir bezeichnen mit z_{mk} die beobachtete Anzahl von Vertretern der m-ten Generation, die in der $(m+1)$ten Generation genau k männliche Nachkommen gehabt haben. Nach T. Harris [4] gilt dann der Satz, daß die *Maximum-Likelihood-Schätzung* für die p_k gegeben ist durch

$$p_k = \sum_{m=0}^{n} z_{mk}/N. \tag{19}$$

Dabei haben die z_{mk} die oben angegebene Bedeutung; $(n+1)$ ist die Anzahl der beobachteten Generationen und N die Gesamtzahl, der in diesen Generationen beobachteten männlichen Glieder, also $N = 1 + z_1 + z_2 + \ldots z_n$, wenn z_i die in der i-ten Generation beobachtete Anzahl von männlichen Gliedern ist.

Die z_{mk} können dann aus genealogischen Handbüchern gewonnen werden. So verfolgen etwa die Bände des «Genealogischen Handbuches zur Schweizergeschichte» [14] unter anderm das Ziel, in ihren Stammtafeln eine vollständige Deszendenz der männlichen Generationen der angeführten Familien zu bringen.

Wir haben zum Beispiel die Stammtafel der *Familie Segesser von Brunegg* ausgezählt vom Ende des 13. Jahrhunderts bis zur Generation, deren letzte Vertreter um 1900 gestorben sind. Nach den Ausführungen in [14] besitzen übrigens nur wenige schweizerische Familien eine so vollständige Genealogie wie die Familie der Segesser von Brunegg. Die Ergebnisse sind die folgenden:

$p_0 = 0{,}631$, $p_1 = 0{,}102$, $p_2 = 0{,}057$, $p_3 = 0{,}080$, $p_4 = 0{,}063$, $p_5 = 0{,}029$, $p_6 = 0{,}023$; es ist $p_7 + p_8 + \ldots + p_\omega = 0{,}015$; die Zahl 7 für die männlichen Nachkommen tritt nicht auf, 8 und 9 treten je einmal auf, schließlich tritt als größte Zahl männlicher Nachkommen 15 einmal auf; N beträgt 176.

Natürlich müssen wir hier wieder an die Problematik dieser Schätzung erinnern; die Maximum-Likelihood-Funktion von T. Harris ist unter den einschränkenden Voraussetzungen aufgestellt worden, die wir oben für die p_k als erfüllt angenommen haben.
In einem Modell, das auf diesen Schätzwerten für die p_k aufgebaut würde, wäre übrigens $m > 1$; nach dem im fünften Abschnitt mitgeteilten grundlegenden Satz, wäre die Wahrscheinlichkeit q des Erlöschens der Linie also kleiner als 1. Mit andern Worten: Mit diesen Werten als p_k ergibt sich eine gewisse von Null verschiedene Wahrscheinlichkeit $1-q$, daß *die Linie nicht erlischt.*
So gewonnene Schätzwerte können sehr verschieden sein, je nach der Stammtafel, die ausgezählt wird. Für die Familie der *Meyer von Knonau* zum Beispiel finden wir in [14] für den Zeitraum vom 14. Jahrhundert bis zur Generation, deren letzte Vertreter um 1900 gestorben sind:

$$p_0 = 0{,}436, p_1 = 0{,}179, p_2 = 0{,}308, p_3 = 0{,}077;$$
$$\sum_{k=0}^{3} p_k = 1; N = 39.$$

g) Zu *Vergleichszwecken* erwähnen wir noch die von A. Lotka in [6] mitgeteilten Ergebnisse aus Untersuchungen an Hand der *Statistik der USA;* er findet für unsere p_k:

$$p_0 = 0{,}4825; p_1 = 0{,}2126; p_k = p_1 \cdot 0{,}5893^{k-1}, k \geqq 1;$$
$$m = 1{,}260.$$

Diese Resultate basieren auf der Bevölkerungsstatistik der Jahre 1920 und 1930, die weiße Rasse betreffend.

7. Zur Übergangsmatrix dieses Prozesses

Wir haben durch Gleichung (6) die Übergangswahrscheinlichkeit

$$P_{ij} = W(Z_{n+1} = j / Z_n = i)$$

definiert und bereits im zweiten Abschnitt darauf hingewiesen, wie sich die Übergangswahrscheinlichkeiten in einer Matrix darstellen lassen.

Es sei nun

$$M_1 = \left\| \begin{matrix} P_{00} & P_{01} & P_{02} & \cdots\cdots \\ P_{10} & P_{11} & P_{12} & \cdots\cdots \\ P_{20} & P_{21} & P_{22} & \cdots\cdots \\ . & . & . & \cdots\cdots \\ . & . & . & \cdots\cdots \\ . & . & . & \cdots\cdots \end{matrix} \right\| .$$

Die P_{ij} lassen sich nach (11) berechnen; es ist also

$$P_{ij} = \Sigma \frac{i!}{m_0!\ m_1!\ \ldots\ m_\omega!} \cdot p_0^{m_0}\, p_1^{m_1} \ldots p_\omega^{m_\omega}, \text{ dabei ist}$$

zu summieren über alle m_k, $k = 0, 1, 2, \ldots$, mit $m_0 + m_1 + \ldots m_\omega = i$ und $0 \cdot m_0 + 1 \cdot m_1 + \ldots + \omega \cdot m_\omega = j$. P_{ij} finden wir somit als Koeffizienten von s^j in der i-ten Potenz der erzeugenden Funktion f(s), also in $[f(s)]^i$.

Einige Zeilen und Spalten von M_1 lassen sich ohne große Rechnung durch die p_k ausdrücken, zum Beispiel

$P_{0j} = 1$ für $j = 0$ und $P_{0j} = 0$ für $j \neq 0$,
$P_{1j} = p_j$ für $0 \leqq j \leqq \omega$, und $P_{1j} = 0$ für $j > \omega$,
$P_{i0} = p_0^i$ und $P_{i1} = i p_0^{i-1} \cdot p_1$ für $i = 0, 1, \ldots$; weiter ist etwa
$P_{22} = 2\, p_0 p_2 + p_1^2$, $P_{32} = 3 p_0^2 p_2 + 3\, p_0 p_1^2$ usf.

Somit hat M_1 für unsern Prozeß folgendes Aussehen

$$M_1 = \left\| \begin{matrix} 1 & 0 & 0 & 0 \ \cdots\cdots\cdots\cdots\cdots\cdots \\ p_0 & p_1 & p_2 & p_3 \ \cdots\cdots\ p_\omega \cdots\cdots \\ p_0^2 & 2p_0p_1 & (2p_0p_2 + p_1^2) & \cdots\cdots\cdots\cdots \\ p_0^3 & 3p_0^2p_1 & (3p_0^2p_2 + 3p_0p_1^2) & \cdots\cdots\cdots \\ p_0^4 & 4p_0^3p_1 & \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \end{matrix} \right\|$$

In der Terminologie der Markoffschen Ketten heißt der Zustand $Z_n = 0$ ein *absorbierender* Zustand, da er offenbar nicht mehr verlassen werden kann.

Mit Hilfe von Formel (5) läßt sich nun weiter für jedes natürliche $g > 1$ die Wahrscheinlichkeit

$$W(Z_{n+g} = j / Z_n = i) = P_{ij}(g)$$

als Element der Matrix $M_g = M_1{}^g$ finden.

Wir bemerken noch, daß $\lim\limits_{n \to \infty} W(Z_n = k) = 0$ für $k = 1, 2, \ldots$, wenn nur der Erwartungswert von Z_1, also m, endlich ist. *Ferner geht* $Z_n \to \infty$ *mit der Wahrscheinlichkeit* $1-q$ *und* $Z_n \to 0$ *mit der Wahrscheinlichkeit* q. In diesen Aussagen kommt die *Instabilität* der Folge der Z_n zum Ausdruck; für den Beweis vgl. T. Harris [5].

LITERATUR

Wahrscheinlichkeitsrechnung und verwandte Gebiete

[1] C. Churchman, R. Ackoff, E. Arnoff, Operations Research, eine Einführung in die Unternehmensforschung, Wien/München 1961

[2] W. Feller, An Introduction to Probability Theory and its Applications, 2. Auflage, New York/London, 1957

[3] M. Fisz, Probability Theory and Mathematical Statistics, 3. Auflage, New York/London, 1963

[4] T. Harris, Branching Processes, The Annals of Mathematical Statistics, Vol. XIX., 1948

[5] T. Harris, The Theory of Branching Processes (Grundlehren der mathematischen Wissenschaften Band 119), Berlin/Göttingen/Heidelberg, 1963

[6] A. Lotka, Théorie analytique des associations biologiques, II^e partie (Actualités scientifiques et industrielles No. 780), Paris, 1939

[7] E. Netto, Lehrbuch der Combinatorik, 2. Auflage, New York, ohne Jahrgang

[8] J. Steffensen, Deux problèmes du calcul des probabilités, Annales de l'Institut Henri Poincaré, Vol. III, 1932

Physik und Chemie

[9] E. Schrödinger, Probability Problems in Nuclear Chemistry, Proceedings of the Royal Irish Academy, Vol. LI, 1945—48

[10] N. Semenoff, Chemical Kinetics and Chain Reactions, Oxford, 1935

[11] A. Weinberg, E. Wigner, The physical Theory of Neutron Chain Reactors, Chicago, 1958

Bevölkerungsstatistik, Familienstatistik, Verschiedenes

[12] A. Burckhardt, Über Kinderzahl und jugendliche Sterblichkeit in früheren Zeiten, Zeitschrift für Schweizerische Statistik, 1907

[13] A. de Candolle, Histoire des sciences et des savants, 2. Auflage, Genf/Basel, 1885

[14] Genealogisches Handbuch zur Schweizer Geschichte, herausgegeben von der Schweizerischen heraldischen Gesellschaft, Band III, Zürich 1908—16

[15] D. Hanhart, Der Zürcher Arbeiter und sein Leitbild von der idealen Familiengröße, Schweizerische Zeitschrift für Volkswirtschaft und Statistik, 1963

[16] F. Kaufmann, Geburtenkontrolle und Bevölkerungsentwicklung im Rahmen der industriellen Gesellschaft, Zeitschrift für Präventivmedizin, 1962

[17] A. Miller, Die Fruchtbarkeit der schweizerischen Bevölkerung, Schweizerische Zeitschrift für Volkswirtschaft und Statistik, 1961

[18] A. Miller, Psychische und soziale Faktoren bei der Planung der Familiengröße, Zeitschrift für Präventivmedizin, 1962

[19] A. Moser, Familienstatistik und Bevölkerungsvermehrung, Mitteilungen des statistischen Büros des Kantons Bern, Nr. 45, 1962

[20] A. Moser, Heiratshäufigkeit und Bevölkerungsreproduktion, Schweizerische Zeitschrift für Volkswirtschaft und Statistik, 1963

[21] L. Solari, Evolution récente de la fécondité en Suisse, Schweizerische Zeitschrift für Volkswirtschaft und Statistik, 1956