



**AgEcon** SEARCH

RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

*No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.*

---

Hanf, E.: Prognose mittels Markov- Ketten. In: Reisch, E.: Quantitative Methoden in den Wirtschafts- und Sozialwissenschaften des Landbaues. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Band 4, Münster-Hiltrup: Landwirtschaftsverlag (1967), S. 279-296.

---



# Prognose mittels Markov-Ketten

Von E. HANF, Stuttgart-Hohenheim

Einleitung .....	279
1 Endliche Markov-Ketten mit stationären Transitionswahrscheinlichkeiten .....	280
1.1 Die Wahrscheinlichkeitsmatrix P .....	280
1.2 Die Potenzen von P .....	281
1.3 Die Konvergenz .....	282
1.4 Einige Definitionen .....	282
2 Ein numerisches Beispiel für eine reguläre Markov-Kette .....	283
3 Empirische Definition und Interpretation eines Markov-Prozesses .....	284
3.1 Die Erstellung der P-Matrix .....	285
3.2 Die Interpretation der P-Matrix und ihrer Potenzen .....	286
4 Die Erwartungswerte der Markov-Kette .....	287
4.1 Kennwerte der Klassen und Erwartungswerte .....	287
4.2 Erwartungswert der N-ten Klasse .....	292
5 Methodisch bedingte Voraussetzungen für die Anwendbarkeit .....	292
5.1 Die Beobachtungsvariable .....	292
5.2 Die Menge der Beobachtungselemente .....	293
5.3 Referenzperiode und Extrapolierbarkeit .....	293
5.4 Die Erwartungswerte .....	294
6 Vergleich mit anderen Schätzmethoden .....	295
6.1 Beschränkungen und Information .....	295
6.2 Mögliche Verbesserungen der beschriebenen Methode .....	296

## Einleitung

In der vorliegenden Studie soll berichtet werden, inwieweit Markov-Ketten als geeignete Hilfsmittel zur Schätzung der zukünftigen Entwicklung von Variablen erscheinen. Man unterscheidet nach der Art der Zustände des durch die Markov-Kette zu beschreibenden Systems ergodische, absorbierende, zyklische und reguläre Ketten. Untersucht wurden hier endliche, reguläre Markov-Ketten 1. Ordnung<sup>1)</sup>.

In den letzten Jahren ist diese Methode mehrfach zur Untersuchung dynamischer Strukturen, die sich durch die Belegung von Zuständen mit Elementen beschreiben lassen, verwendet worden.

---

<sup>1)</sup> In einer Kette n-ter Ordnung sind die Transitionswahrscheinlichkeiten Funktionen der n vorhergehenden Zustände.

Mehrere Anwendungen geben KEMÉNY und SNELL in ihrer ausgezeichneten Einführung in die Theorie der endlichen Markov-Ketten an. Beschrieben werden Prozesse aus verschiedenen Bereichen der Wissenschaften, so Diffusionsvorgänge durch eine permeable Membran, ein Kreuzungsversuch aus der Genetik und Veränderungen einer Gesellschaftsstruktur. JUDGE und SWANSON wenden die Methode auf die Untersuchung des Verhaltens einer Gruppe von Schweineproduzenten in Illinois an.

Im Rahmen der vorliegenden Arbeit soll gezeigt werden, daß unter gewissen Bedingungen und Einschränkungen Markov-Ketten zur Schätzung von Leistungskoeffizienten oder ähnlicher spezifischer Größen geeignet erscheinen.

## 1 Endliche Markov-Ketten mit stationären Transitionswahrscheinlichkeiten [14; 5; 3]

### 1.1 Die Wahrscheinlichkeitsmatrix $P$

Gegeben ist eine Menge von endlich vielen, namentlich  $N$ , wohlunterschiedenen, determinierten Zuständen  $S_1, S_2, \dots, S_N$ . Die Beschreibung dieser Zustände kann durch Eigenschaften (verbal und qualitativ), durch Zahlen bzw. Zahlengruppen (skalar oder vektoriell) erfolgen. Das System ist bezüglich der Zustände abgeschlossen.

Ferner existiert eine endliche Menge von Elementen, die sich in diesem System „bewegen“. Die Zustandsänderungen erfolgen zu bestimmten, nicht notwendig äquidistanten Zeitpunkten und dann sprunghaft. Diese Zeitpunkte und die damit verbundenen Veränderungen im System werden in der Folge als „Schritte“ bezeichnet.

Die Beschreibung der Schritte in einem solchen System kann auf verschiedene Arten erfolgen.

a) Man kann das System verbal durch logische Relationen aufgrund der kausalen Zusammenhänge beschreiben<sup>1)</sup>: Wenn sich ein Element zur Zeit  $t$  im Zustand  $S_i$  befindet, dann befindet es sich zur Zeit  $t+1$  im Zustand  $S_j$ .

b) Man beschreibt die „Bewegungen“ des Elementes im System durch eine Funktion  $f_i(i)$ ,  $i = 1, \dots, N$ , sie gibt das Verhalten des Elementes bei einem Sprung zur Zeit  $t$  in Abhängigkeit vom jeweiligen Ist-Zustand an.

c) Die Markov-Ketten beschreiben die Veränderung im System auf folgende Weise:

Zu jedem möglichen Anfangszustand  $i = 1, \dots, N$  zur Zeit  $t$  gibt es eine Folge von Wahrscheinlichkeiten  $p_{ik}$ ,  $k = 1, \dots, N$ ,  $0 \leq p_{ik} \leq 1$ ,  $\sum_{k=1}^N p_{ik} = 1$ . Der Wert  $p_{ik}$

gibt an, wie groß die Wahrscheinlichkeit ist, daß ein Element, das sich zur Zeit  $t$  im Zustand  $S_i$  befindet, zur Zeit  $t+1$  im Zustand  $S_k$  zu finden ist. Das Ergebnis der Transition hängt nur vom „Anfangszustand“ zur Zeit  $t$  ab;  $p_{ik}$  heißt Transitionswahrscheinlichkeit, die aus den  $p_{ik}$  gebildete Matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1N} \\ p_{21} & \dots & \dots & p_{2N} \\ \dots & \dots & \dots & \dots \\ p_{N1} & \dots & \dots & p_{NN} \end{pmatrix} \quad (1)$$

heißt Transitionsmatrix [14, S. 24 ff].

<sup>1)</sup> Diese Art der Beschreibung ist hypothetisch, sie verbietet sich durch ihren Umfang von selbst.

Die Matrix  $P$  beschreibt alle im abgeschlossenen System möglichen „Bewegungen“ der Elemente, die in einem Schritt vorgehen können, gewichtet durch ihre Wahrscheinlichkeiten.

Befindet sich ein Element zur Zeit  $t$  mit Sicherheit im Zustand  $S_j, j \leq N$ , ist ferner  $z$  ein Vektor mit den Komponenten  $z_i, i = 1, \dots, N, z_i = 0$  für  $i \neq j, z_j = 1$ , so ergibt

$$z \cdot P = p_j \quad (2)$$

einen Vektor, dessen  $N$  Komponenten die Elemente der  $j$ -ten Zeile der  $P$ -Matrix sind. Die  $p_{jk}, k = 1, \dots, N$ , geben an, mit welcher Wahrscheinlichkeit dieses Element nach einem Schritt in irgendeinem der  $N$  Zustände zu finden ist [11, S. 5 ff].

Kennt man für den Zustand des Elementes zur Zeit  $t$  nur einen Wahrscheinlichkeitsverteilungsvektor  $z$  mit den Komponenten  $z_k, k = 1, \dots, N, 0 \leq z_k \leq 1, \sum_{k=1}^N z_k = 1$ , so läßt sich

$$z P = p_j$$

wie oben interpretieren [11, S. 5 ff].

Bezeichnen die Komponenten  $z_k$  des Vektors  $z$  die Anzahl der Elemente des Systems, die sich zur Zeit  $t$  im Zustand  $S_k$  befinden, dann ergibt das Produkt

$$z P = n \text{ mit } n = (n_1, n_2, \dots, n_N) \quad (3)$$

die zu erwartende Verteilung der Elemente nach einem Schritt.

## 1.2 Die Potenzen von $P$

Die Bedeutung der Markov-Ketten zur Beschreibung zeitlich veränderlicher Systeme liegt in den Interpretationsmöglichkeiten der Potenzen der  $P$ -Matrix.

Zur Verdeutlichung betrachte man ein abgeschlossenes System mit 2 Zuständen  $S_1$  und  $S_2$ . Die Transitionsmatrix sei

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = P^{(1)} \quad (4)$$

Die Matrix  $P^{(2)} = P^2$  ergibt sich als

$$P^2 = P^{(2)} = \begin{pmatrix} p_{11}^{(2)} & p_{12}^{(2)} \\ p_{21}^{(2)} & p_{22}^{(2)} \end{pmatrix} = \begin{pmatrix} p_{11}^2 + p_{12} p_{21} & p_{11} p_{12} + p_{12} p_{22} \\ p_{21} p_{11} + p_{22} p_{21} & p_{21} p_{12} + p_{22}^2 \end{pmatrix} \quad (5)$$

mit  $0 \leq p_{ik}^{(2)} \leq 1, \sum_{k=1}^2 p_{ik} = 1$  für  $i = 1, 2$ .

Jeder Wert  $p_{ik}^{(2)}$  gibt an, wie groß die Wahrscheinlichkeit ist, daß ein Element, startend in  $S_i$ , nach zwei Schritten in  $S_k$  zu finden ist.

Zur Erläuterung diene  $p_{12}^{(2)}$ . Der Übergang von  $S_1$  nach  $S_2$  ist in diesem System in 2 Schritten auf 2 Arten möglich:

1.  $S_1 - (p_{11}) \rightarrow S_1 - (p_{12}) \rightarrow S_2$  und
2.  $S_1 - (p_{12}) \rightarrow S_2 - (p_{22}) \rightarrow S_2$

Die beiden Wege schließen sich gegenseitig aus, ihre Wahrscheinlichkeiten addieren sich, auf jedem Wege finden hintereinander zwei unabhängige Ereignisse statt, ihre Wahrscheinlichkeiten werden multipliziert; diese Überlegung führt zu

$$p_{12}^{(2)} = p_{11} p_{12} + p_{12} p_{22}$$

Für alle anderen Sprünge erhält man analog die  $p_{ik}^{(2)}$  von (5).

In einem System von  $N$  Zuständen beschreiben die Elemente der Matrix

$$P^n = P^{n-1} \cdot P = \left( p_{ik}^{(n)} = \sum_{j=1}^N p_{ij}^{(n-1)} \cdot p_{jk} \right) \quad (6)$$

den Übergang von  $S_i$  nach  $S_k$  in  $n$  Schritten.

Die Transitionswahrscheinlichkeiten  $p_{ik}$  sind während des Betrachtungszeitraumes stationär, also konstant.

Benützt man eine der oben angeführten Definitionen für  $z$ , so ergibt die Folge

$$zP, zP^2, zP^3, \dots, zP^n, \dots \quad (7)$$

eine Beschreibung des zeitlichen Verlaufs der „Bewegungen“ im System nach  $1, 2, \dots, n$  Schritten als Funktion der Anfangszustände zur Zeit  $t = 0$ .

### 1.3 Die Konvergenz

Die Folge der Potenzen von  $P$  konvergiert [14, S. 71] gegen eine Matrix  $A$ :

$$\lim P^n (n \rightarrow \infty) = A \quad (8)$$

$$A = e \cdot a \quad (9)$$

In (9) ist  $e$  ein Spaltenvektor, dessen  $N$  Komponenten den Wert 1 haben, während  $a$  ein Vektor ist, dessen Komponenten Wahrscheinlichkeiten sind, also gilt  $0 \leq a_i \leq 1$ ,

$$i = 1, \dots, N, \sum_{i=1}^N a_i = 1.$$

Die Matrix  $A$  besteht nach (9) aus lauter gleichen Zeilenvektoren  $a_i, i = 1, \dots, N$ . Nach einer hinreichend großen Anzahl von Schritten nähert sich also das System einem „Gleichgewichtszustand“. Die Wahrscheinlichkeit, nach einer hinreichend großen Zahl von Schritten in einem Zustand  $S_k$  zu sein, ist unabhängig vom Anfangszustand  $S_i$ . Es stellt sich eine Endverteilung ein.

Zur Berechnung von  $A$ , also der Endverteilung im Gleichgewichtszustand, bieten sich zwei Wege an [13, S. 4].

- Man berechnet  $P^n$  und gibt eine Größe  $d$  derart an, daß — falls nur das Maximum der  $d_k = p_{ik} - p_{jk}$  für alle Paare  $i$  und  $j$  und für alle  $k$  kleiner als  $d$  ist — die Zeilen der Matrix  $P_n$  als gleich zu betrachten sind. Einer der Zeilenvektoren ist dann die Endverteilung  $a$ .
- Aus der Grenzwerteigenschaft von  $A$  ergibt sich die Gleichung  $aP = a$  [14, S. 71]. Man findet also die Endverteilung als Lösung des Gleichungssystems

$$X(P - I) = 0, \quad \sum_{i=1}^N x_i = 1 \quad (10)$$

mit  $I$  als Einheitsmatrix und  $X$  als Vektor mit den Komponenten  $x_i$ .

Die Grenzmatrix  $A$  findet, abgesehen von ihrer Bedeutung als Charakterisierende des Gleichgewichtszustandes, Anwendung zur Berechnung einiger Folgerungen, die sich aus der mathematischen Struktur der Markov-Ketten ergeben [14, S. 207], so zur Berechnung der Fundamentalmatrix  $Z$ , der Matrix  $M$  der „mean first passage times“ und deren Streuungen.

### 1.4 Einige Definitionen

Zum allgemeineren Verständnis der Theorie seien hier noch einige aus [14] entnommene Definitionen angeführt.

- a) Eine endliche Markov-Kette ist ein stochastischer Prozeß, gekennzeichnet durch eine endliche Zahl von Zuständen. Die Wahrscheinlichkeit, in einen anderen Zustand überzugehen, hängt nur vom jeweiligen Ausgangszustand ab.
- b) Eine Markov-Kette heißt regulär, wenn ein Element von jedem Zustand als Anfangszustand, nicht notwendig in einem Schritt, in jeden beliebigen Zustand als Endzustand übergehen kann.  
 Etwas mathematischer formuliert, besagt diese wichtige Definition:  
 Eine Markov-Kette heißt regulär, wenn es eine Potenz  $P^n$  gibt, für die alle  $p_{ik}^{(n)}$  von Null verschieden sind.
- c) Ein Zustand, den ein Element nicht mehr verlassen kann, falls es ihn je erreicht, heißt absorbierend. (Ein Zustand  $S_i$  heißt absorbierend, wenn  $p_{ii} = 1$  ist).  
 Für Anwendungen sind die unter b) und c) definierten speziellen Markov-Ketten von Bedeutung.

## 2 Numerisches Beispiel für eine reguläre Markov-Kette

Im folgenden Beispiel wird ein System mit 5 Zuständen beschrieben. Man überzeugt sich an Hand der  $P$ -Matrix leicht, daß es mit maximal 3 Schritten möglich ist, aus jedem Zustand als Anfangszustand in jeden anderen als Endzustand überzugehen. Die Transitionswahrscheinlichkeit<sup>1)</sup>  $p_{34} = 0.18$  z. B. gibt die Wahrscheinlichkeit eines Übergangs von  $S_3$  nach  $S_4$  in einem Schritt an.

$P$	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	Zeilen- summe
$S_1$	0,67	0,31	0,02	0,00	0,00	1,00
$S_2$	0,02	0,73	0,24	0,01	0,00	1,00
$S_3$	0,00	0,04	0,78	0,18	0,00	1,00
$S_4$	0,00	0,00	0,15	0,80	0,05	1,00
$S_5$	0,00	0,00	0,02	0,17	0,81	1,00

Im folgenden werden einige der Potenzen  $P^n$  aufgeführt:

	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	
$P^3 = P^{(3)}$	$S_1$	0,30	0,46	0,19	0,03	0,02
	$S_2$	0,03	0,42	0,42	0,12	0,01
	$S_3$	0,00	0,07	0,53	0,35	0,05
	$S_4$	0,00	0,01	0,18	0,61	0,20
	$S_5$	0,00	0,00	0,08	0,35	0,57
$P^6 = P^{(6)}$	$S_1$	0,11	0,35	0,36	0,16	0,02
	$S_2$	0,02	0,23	0,43	0,27	0,05
	$S_3$	0,01	0,07	0,38	0,42	0,12
	$S_4$	0,00	0,03	0,22	0,50	0,25
	$S_5$	0,00	0,01	0,15	0,44	0,40
$P^{10} = P^{(10)}$	$S_1$	0,04	0,20	0,39	0,30	0,07
	$S_2$	0,01	0,12	0,36	0,38	0,13
	$S_3$	0,01	0,06	0,31	0,44	0,18
	$S_4$	0,00	0,04	0,24	0,47	0,25
	$S_5$	0,00	0,03	0,21	0,46	0,30

<sup>1)</sup> In einem mathematisch definiertem Markov-Prozeß sind die Transitionswahrscheinlichkeiten im allgemeinen a priori bekannt, was für Anwendungen in der Regel nicht gilt.



Die Grenzmatrix mit dem Vektor  $a$  der Endverteilung ergibt sich als  $A = e \cdot a^1$ ):

$$A = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} (0,01 \ 0,05 \ 0,25 \ 0,45 \ 0,24)^1$$

Der Einfachheit wegen sei angenommen, daß sich zur Zeit  $t = 0$  in jedem der Zustände 100 Elemente befinden, es ist also

$$n = (100, 100, 100, 100, 100).$$

Anzahl der Elemente nach  $m$  Schritten in den Klassen

	$m$	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	Summe
$n$	0	100	100	100	100	100	500
$n P$	1	69	108	121	116	86	500
$n P^2$	2	48	105	129	132	86	500
$n P^3$	3	35	97	140	145	83	500
$n P^4$	4	26	87	148	157	82	500
$n P^5$	5	19	18	152	168	82	500
$n P^6$	6	15	69	155	178	83	500
$n P^7$	7	12	62	155	186	85	500
$n P^8$	8	10	55	154	193	88	500
$n P^9$	9	8	50	153	199	90	500
$n P^{10}$	10	7	45	151	204	93	500
$n A$		5	25	125	225	120	500

Die Zeilen  $n P^i$ ,  $i = 1, \dots, 10$ , geben die zu erwartenden Verteilungen der Elemente nach  $i$  Schritten. Die Tabelle gibt interessante Einblicke in die Struktur des Prozesses. Bemerkenswert ist insbesondere das Verhalten der Belegung von  $S_5$ . Es erklärt sich natürlich aus den Matrizen  $P^i$ . Zunächst muß  $n_5$  sinken, da die 19% Abgänge die 5% Zugänge aus  $S_4$  überwiegen. Mit zunehmender Belegung von  $S_4$  und zunehmenden Transitionswahrscheinlichkeiten  $p_{i5}$  beginnt dann die Belegung von  $S_5$  zu steigen. Die letzte Zeile quantifiziert die zu erwartende Endverteilung, vorausgesetzt, daß sich der Prozeß ungestört hinreichend lange entwickeln kann<sup>2)</sup> [8; 18; 1; 21].

### 3 Empirische Definition und Interpretation eines Markov-Prozesses

Die Anwendung von Markov-Ketten verlangt als erstes das Auffinden der den Prozeß beschreibenden Matrix  $P$ . Die Anwendung impliziert einige Voraussetzungen und Fol-

<sup>1)</sup> Die Matrix  $A$  besteht also aus 5 gleichen Zeilen, die durch den Vektor  $a$  gegeben sind. Die Komponenten  $a_i$  sind notwendig von Null verschieden, da es sich um eine reguläre Markov-Kette handelt.

<sup>2)</sup> G. G. JUDGE und E. R. SWANSON [13] untersuchten z. B. eine Gruppe von 83 Schweineproduzenten in Illinois in den Jahren 1946-58. Als Zustände wurden die Klassen der Anzahl der Würfe pro Jahr gewählt. Sie untersuchten die Klassenfrequentierung zu Beginn und am Ende der Referenzperiode im Vergleich zur Gleichgewichtslösung, ferner mittels absorbierender Zustände einen möglichen Verlauf der zukünftigen Entwicklung.

gerungen über die weiter unten noch zu sprechen sein wird. Das Datenmaterial, aus denen  $P$  abgeleitet werden soll, kann in zwei Formen vorliegen.

- a) Über den zeitlichen Verlauf der zu beobachtenden Variablen liegen für jedes einzelne Element Informationen in Form einer Zeitreihe vor. Für jedes einzelne Element kann für jeden Zeitpunkt der Referenzperiode ausgesagt werden, welchem klassifizierten Zustand das Element angehört.
- b) Die Information beschränkt sich darauf, daß nur jeweils die Anzahl der Elemente bekannt ist, die zu einem bestimmten Zeitpunkt der Referenzperiode die verschiedenen Zustände belegen. Die Anzahl der Wechsel muß also aus der jeweiligen Ist-Verteilung extrahiert werden.

### 3.1 Die Erstellung der $P$ -Matrix

Der dieser Arbeit zugrunde liegende Beispielfall entspricht den Gegebenheiten unter a). Die zu untersuchende Variable ist die durchschnittliche Milchleistung pro Kuh und Jahr (im folgenden kurz als d. Ml. bezeichnet), die Beobachtungselemente sind hier die kleinsten verfügbaren Verwaltungseinheiten, für die die Variable ausgewiesen wird, die Kreise der Bundesrepublik<sup>1)</sup>.

An Stelle der a priori vorhandenen Transitionswahrscheinlichkeiten des mathematisch definierten Markov-Prozesses treten jetzt empirisch gewonnene relative Häufigkeiten. Die determinierten Zustände des Markov-Prozesses werden durch Intervalle der Beobachtungsvariablen  $y$  ersetzt. Soll der Prozeß durch  $N$  klassifizierte Zustände beschrieben werden, so erhält man folgende Klassen

$$\begin{array}{ccccccccc} K_1 & & K_2 & & K_3 & & \dots & & K_N & & K_{N+1} & (11) \\ y < g_1 & & g_1 \leq y < g_2 & & g_2 \leq y < g_3 & \dots & g_{N-1} \leq y < g_N & & g_N \leq y & & \end{array}$$

wobei die  $g_i$ ,  $i = 1, \dots, N$ , als „Grenz“-werte bezeichnet werden sollen. Die Äquidistanz der  $g_i$  ist nicht notwendig.

Durch die Klassifizierung der Beobachtungsvariablen werden wohlunterschiedene Zustände eingeführt. Beobachtet werden in der Referenzperiode (1951–64) die Sprünge der kreisweise ausgewiesenen Werte der d. Ml. von der Klasse  $K_i$  im Jahre  $t$  und der Klasse  $K_k$  im Jahr  $t+1$ . Man erhält so die Matrix

$$H = (h_{ik}), \quad i = 1, \dots, N, \quad k = 1, \dots, N$$

der Häufigkeit der Sprünge vom Anfangszustand  $i$  zum Endzustand  $k$  in Schritten von einem Jahr. Durch die Umformung

$$\bar{p}_{ik} = \frac{h_{ik}}{\sum_{k=1}^N h_{ik}} \quad (12)$$

erhält man eine Matrix  $\bar{P}$ . ANDERSON und GOODMAN [2] haben gezeigt, daß die Werte der „größten Mutmaßlichkeit“ (maximum likelihood) [16, S. 40 ff] für die stationären Transitionswahrscheinlichkeiten die relativen Häufigkeiten (12) sind. Wir betrachten nun die Matrix  $\bar{P}$  als die  $P$ -Matrix der Markov-Kette<sup>2)</sup>. Im vorliegenden Beispiel wurde die Zahl der Klassen  $N = 18$ , die unterste Grenze  $g_{\min} = g_1 = 2400$  l/Kuh und Jahr,

<sup>1)</sup> Unberücksichtigt blieben die 8 Kreise des Saarlandes und 99 kreisfreie Städte, sowie Berlin, da ihre Zeitreihen aus verschiedenen Gründen nicht vergleichbar bzw. unvollständig waren. Sie umfassen ca. 2,0% der Milchviehbestände der Bundesrepublik.

<sup>2)</sup> Für statistische Tests, auch bezüglich der Stationarität, an Markov-Ketten [2]

TABELLE 1 Klasseneinteilung der Variablen „durchschnittliche Milchleistung je Kuh und Jahr (DML)“ (in 1000 l/Kuh und Jahr)

i	Klasse	untere obere Grenze		Kennwerte der Klassen $q(i)$
		$g_{i-1}$ Intervall	$g_i$	
1	$K_1$		DML < 2,4	2,25
2	$K_2$	2,4 ≤	DML < 2,5	2,45
3	$K_3$	2,5 ≤	DML < 2,6	2,55
4	$K_4$	2,6 ≤	DML < 2,7	2,65
5	$K_5$	2,7 ≤	DML < 2,8	2,75
6	$K_6$	2,8 ≤	DML < 2,9	2,85
7	$K_7$	2,9 ≤	DML < 3,0	2,95
8	$K_8$	3,0 ≤	DML < 3,1	3,05
9	$K_9$	3,1 ≤	DML < 3,2	3,15
10	$K_{10}$	3,2 ≤	DML < 3,3	3,25
11	$K_{11}$	3,3 ≤	DML < 3,4	3,35
12	$K_{12}$	3,4 ≤	DML < 3,5	3,45
13	$K_{13}$	3,5 ≤	DML < 3,6	3,55
14	$K_{14}$	3,6 ≤	DML < 3,7	3,65
15	$K_{15}$	3,7 ≤	DML < 3,8	3,75
16	$K_{16}$	3,8 ≤	DML < 3,9	3,85
17	$K_{17}$	3,9 ≤	DML < 4,0	3,95
18	$K_{18}$	4,0 ≤	DML	4,15

die oberste Grenze  $g_{\max} = g_N = 4000$  l/Kuh und Jahr und die Klassenbreite (äquidistant)  $g_{n+1} - g_n = 100$  l/Kuh und Jahr gewählt (s. Tabellenteil Nr. 1).

Es ist offenbar, daß die Beschreibung eines Prozesses einer in einem Bereich stetigen Variablen desto detaillierter wird, je kleiner die Klassifizierungsbreite gewählt wird.

Andererseits wird bei kleinen Klassenbreiten die Zahl  $\sum_{k=1}^N h_{ik}$  der Ereignisse der Zugehörigkeit zur Klasse  $i$  klein und dadurch die Schätzung der  $p_{ik}$  schlecht. Ferner hängt aus dem gleichen Grund die Zahl der Klassen von der verfügbaren Menge an Einzelbeobachtungen ab.

Ungleich schwieriger und komplizierter wird die Erstellung der Matrix  $P$ , wenn der Fall b) vorliegt, die Information über den Fluß der Elemente in der Referenzperiode also lediglich aus den jeweiligen Frequentierungen der Zustände durch Elemente abzuleiten ist. In der amtlichen Statistik wird zum Beispiel nur die Anzahl der Betriebe je Betriebsgrößenklasse ausgewiesen, über den tatsächlich eingetretenen Fluß von Elementen, hier Betrieben, kann man unmittelbar daraus keine Information gewinnen. Aber auch aus diesen Zeitreihen der aggregierten Werte lassen sich Transitionswahrscheinlichkeiten ableiten. Ich verweise dazu auf die unten angeführte Literatur [17; 6; 15; 20, S. 270–292].

### 3.2 Die Interpretation der $P$ -Matrix und ihrer Potenzen

Versteht man die Matrix  $\bar{P}$  als  $P$ -Matrix eines Markov-Prozesses — über Einschränkungen wird weiter unten noch zu sprechen sein — dann ergeben sich gemäß der Gleichungen (2), (3) und (7) Interpretationsmöglichkeiten über das Verhalten von Elementen, das sind hier Kreise, in Abhängigkeit von Zeit und Anfangszustand. Zu beachten ist hierbei,

daß man sinngemäß keine Aussagen über ein bestimmtes Verhalten eines bestimmten Elementes machen kann, sondern lediglich stochastische Aussagen bezüglich des zu erwartenden Verhaltens einer Menge von Elementen (s. Tab. 2).

Sind die Klassen  $K_i$  im Sinne wachsender Werte der Beobachtungsvariablen angeordnet, dann bedeutet der Übergang eines Elementes in eine Klasse höherer Ordnung ein „Wachstum“ bezüglich der Variablen. Aus der  $P$ -Matrix lassen sich folgende Informationen in Bezug auf dieses Anwachsen gewinnen.

- a) Die Elemente der Hauptdiagonalen sind ein Maß für die Mobilität des Systems. Sie beschreiben die Übergänge von  $S_i$  nach  $S_i$  in einem Schritt (-also das Verbleiben in  $S_i$ ). Je größer die  $p_{ii}$  sind, desto geringer sind die einschrittigen (jährlichen) Veränderungen. Sind alle  $p_{ii} = 1$ , so finden in dem System überhaupt keine Veränderungen statt.
- b) Eine Aufwärtsentwicklung, d. h. ein Übergang von einem  $K_f$  nach einem  $K_g$  mit  $g > f$ , ist solange wahrscheinlicher als ein Rückgang, solange

$$\sum_{k=1}^{f-1} p_{ik} < \sum_{k=f+1}^N p_{ik} \quad (13)$$

gilt. Der Vergleich der Summen der  $p_{ik}$  rechts und links der Hauptdiagonalen gibt Aufschluß darüber, inwieweit vom Anfangszustand  $i$  aus noch ein Wachstum zu erwarten ist (s. Tab. 2).

Die Potenzen  $P^n$  und die Vektorfolgen  $z P^n$  beschreiben, wie in 1.1 beschrieben, die zu erwartenden Verteilungen der Elemente nach  $n$  Schritten. Die Matrix  $A$  liefert die Endverteilung bei ungestörter Entwicklung des Prozesses gemäß der in  $P$  postulierten stochastischen Gesetze (Tab. 3).

Beschreibt  $z$  einen gewissen Anfangszustand zur Zeit  $t = 0$ , also die Verteilung der Elemente in der Basisperiode, so wird der Prozeß unter Umständen über die Basisperiode mittels der Folge  $z P^n$  extrapolierbar und kann zur schätzenden Beschreibung der zukünftigen Entwicklung des Systems angewandt werden (Tab. 4).

## 4 Die Erwartungswerte der Markov-Kette

### 4.1 Kennwerte der Klassen und Erwartungswerte

Gemäß 3.2 werden Veränderungen bezüglich der Beobachtungsvariablen, im zugrundeliegenden Fall der d. MI., beschrieben durch die Veränderungen der Zugehörigkeit von Elementen, hier der Kreise, zu den einzelnen Klassen. Ein Anwachsen der d. MI. kommt dadurch zum Ausdruck, daß die Wahrscheinlichkeit einer Klasse höherer Leistung anzugehören, in Abhängigkeit vom Anfangszustand zur Zeit  $t = 0$  und von der Zeit  $t = n$  wächst.

Um Aussagen über den funktionalen Verlauf der Variablen zu machen, genügt es nicht, die dahinter verborgene Veränderung der Struktur zu betrachten. Dazu noch einige weitere Definitionen.

Jeder Klasse  $K_i$  wird ein Kennwert  $q_i$  zugeordnet. Seien zur Zeit  $t = 0$   $W_{ij}$ ,  $j = 1, \dots, M_i$ , die Werte, die die Variable innerhalb des Intervalls  $i$  annimmt, und  $H_{ij}$  die Anzahl der Elemente, die den Wert  $W_{ij}$  annehmen, dann sei der Kennwert  $q_i$  der Klasse  $i$

$$q_i = \frac{\sum_{j=1}^{M_i} H_{ij} W_{ij}}{\sum_{j=1}^M H_{ij}} \quad (14)$$

TABELLE 2 Die  $P$ -Matrix (456 Kreise der Bundesrepublik, 1951–1964)

$i$	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$	$K_6$	$K_7$	$K_8$	$K_9$	$K_{10}$	$K_{11}$	$K_{12}$	$K_{13}$	$K_{14}$	$K_{15}$	$K_{16}$	$K_{17}$	$K_{18}$
1	0,77	0,12	0,06	0,02	0,01	0,01												
2	0,10	0,28	0,28	0,23	0,07	0,02	0,01	0,00	0,01									
3	0,02	0,08	0,28	0,35	0,18	0,08	0,02											
4	0,01	0,03	0,06	0,23	0,37	0,19	0,08	0,01	0,01	0,01								
5	0,01	0,01	0,03	0,08	0,26	0,33	0,20	0,05	0,02	0,01								
6				0,03	0,08	0,29	0,32	0,16	0,06	0,04	0,01	0,01						
7					0,02	0,09	0,28	0,33	0,18	0,08	0,01	0,01						
8					0,01	0,07	0,14	0,26	0,25	0,17	0,07	0,02	0,01					
9						0,02	0,07	0,11	0,26	0,33	0,12	0,07	0,01	0,01				
10						0,01	0,03	0,05	0,13	0,29	0,24	0,16	0,06	0,02	0,01			
11							0,0	0,02	0,05	0,13	0,25	0,29	0,16	0,08	0,02			
12								0,01	0,02	0,03	0,15	0,26	0,27	0,14	0,09	0,03		
13									0,01	0,02	0,04	0,20	0,24	0,18	0,16	0,12	0,02	0,01
14										0,01	0,01	0,08	0,19	0,29	0,18	0,14	0,06	0,04
15											0,01	0,02	0,10	0,19	0,21	0,23	0,16	0,08
16												0,01	0,04	0,05	0,17	0,24	0,25	0,24
17													0,01	0,01	0,02	0,05	0,16	0,29
18															0,01	0,02	0,06	0,10

TABELLE 3 Wahrscheinlichkeiten der Übergänge von  $K_{10}$  nach  $K_i$ ,  $i = 1, \dots, 18$  (10. Zeile der Matrizen  $P^n$ )

$n$	Basis	1964	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$	$K_6$	$K_7$	$K_8$	$K_9$	$K_{10}$	$K_{11}$	$K_{12}$	$K_{13}$	$K_{14}$	$K_{15}$	$K_{16}$	$K_{17}$	$K_{18}$
1	1965							0,01	0,03	0,05	0,13	0,29	0,24	0,16	0,06	0,02		0,01		
2	66						0,01	0,02	0,04	0,06	0,10	0,17	0,17	0,18	0,12	0,07	0,03	0,02	0,01	0,01
3	67						0,01	0,02	0,04	0,06	0,08	0,13	0,13	0,16	0,13	0,09	0,06	0,04	0,02	0,02
4	68						0,01	0,02	0,03	0,05	0,07	0,10	0,11	0,14	0,13	0,10	0,08	0,06	0,04	0,05
5	69						0,01	0,02	0,03	0,05	0,06	0,09	0,09	0,13	0,12	0,10	0,08	0,08	0,05	0,09
6	70						0,01	0,02	0,03	0,04	0,05	0,07	0,08	0,11	0,12	0,10	0,09	0,09	0,07	0,13
7	71						0,01	0,02	0,04	0,04	0,05	0,07	0,07	0,10	0,11	0,10	0,09	0,09	0,08	0,17
8	72						0,01	0,02	0,03	0,03	0,04	0,06	0,06	0,09	0,10	0,09	0,09	0,09	0,08	0,21
9	73						0,01	0,01	0,02	0,03	0,04	0,05	0,06	0,08	0,09	0,09	0,08	0,10	0,09	0,25
10	74						0,01	0,01	0,02	0,03	0,03	0,05	0,05	0,08	0,09	0,08	0,08	0,10	0,09	0,28
11	75							0,01	0,02	0,02	0,04	0,04	0,05	0,07	0,08	0,08	0,08	0,10	0,10	0,31
88	—									0,01	0,01	0,01	0,02	0,03	0,05	0,05	0,07	0,11	0,13	0,52

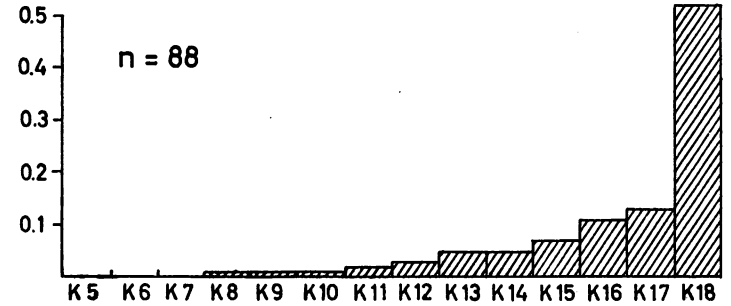
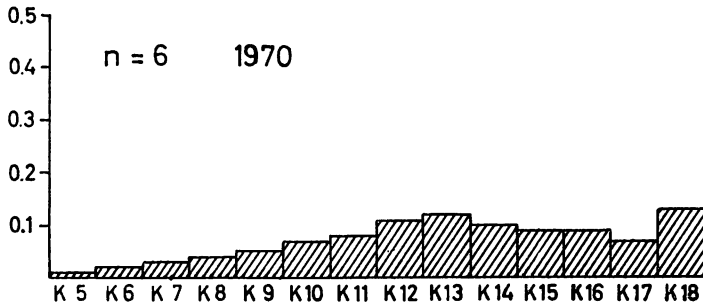
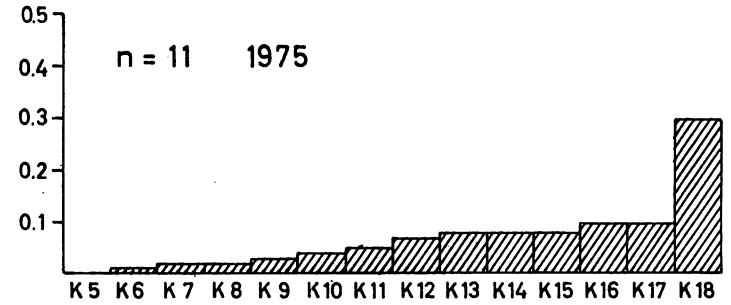
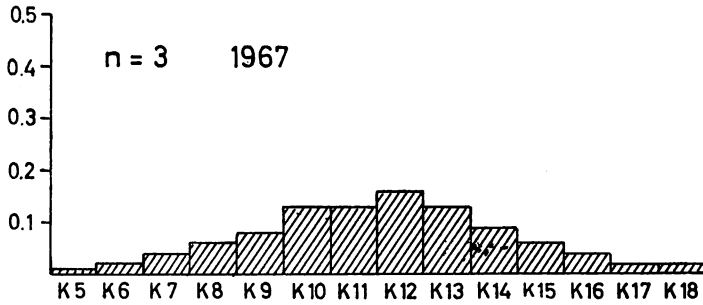
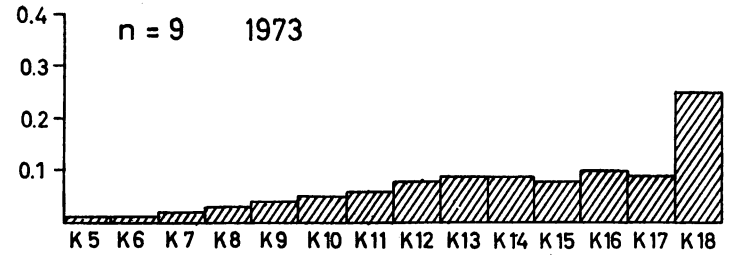
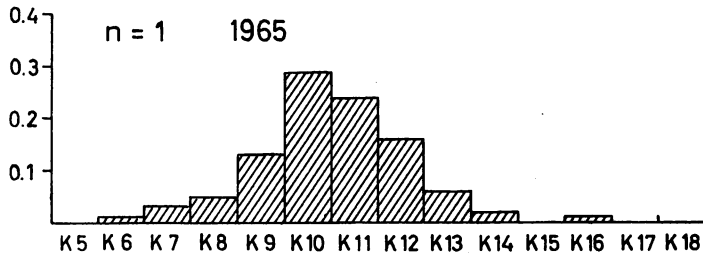


Abb. 1 Graphische Darstellung der  $p_{10,t}^{(n)}$

TABELLE 4 Prozentuale Verteilung der Kreise auf Klassen der DML (in v. H.)

Potenz von P	Jahr	bis 3000 1/Kuh u. Jahr	bis 3500	bis 4000	über 4000
	1951	69,4	13,6	15,4	1,6
	1957	59,0	17,1	20,5	3,4
	1964	22,9	33,2	19,2	24,7
1	1965	19,2	34,1	23,9	22,8
6	1970	8,2	27,2	36,8	27,8
11	1975	4,5	18,7	40,7	36,1
88 <sup>1)</sup>	—	0,7	7,0	40,5	51,8
Absorbierende Markov-Kette <sup>2)</sup>					
1	1965	19,2	34,0	19,4	27,4
6	1970	8,7	26,4	23,7	41,7
11	1975	4,4	16,8	21,0	57,8

<sup>1)</sup> Beschreibt die Verteilung bei hinreichend langer, ungestörter Entwicklung (Gleichgewichtsverteilung)

<sup>2)</sup>  $P_{18,18} = 1$

Handelt es sich bei der Klasse  $K_i$  um eine stärker belegte innere Klasse ( $i = 2, \dots, N-1$ ), so sind die Werte der Beobachtungsvariablen der  $\sum_j^{M_i} H_{ij}$  Elemente rein zufällig, und es besteht kein Grund eine andere als eine Gleichverteilung innerhalb des Intervalles anzunehmen. Anstelle des Wertes (14) tritt dann die Schätzung

$$q_i = \frac{g_i + g_{i+1}}{2} \quad (15)$$

wo  $g_i$  die in (11) definierten „Grenz“-werte der Klassen sind. Durch den Wert  $q_i$  wird die Klasse  $K_i$ ,  $g_i \leq y < g_{i+1}$ , jetzt durch einen einzigen Wert, einen idealisierten Meßwert, repräsentiert. Die Repräsentation ist desto genauer, je kleiner die Klassenbreite  $g_{i+1} - g_i$  gewählt wurde<sup>1)</sup>.

Da die Folge  $p_{ik}$ ,  $k = 1, \dots, N$ , die Wahrscheinlichkeit angibt, mit der ein Element, startend in der Klasse  $K_i$ , gekennzeichnet durch den Wert der Variablen  $q_i$ , sich nach einem Schritt in einem Zustand  $K_k$ , gekennzeichnet durch den Wert  $q_k$ , befindet, ist

$$EM(i) = \sum_{k=1}^N p_{ik} q_k \quad (16)$$

der Erwartungswert der Variablen  $y$  nach einem Schritt in Abhängigkeit vom Anfangszustand  $i^2)$ .

<sup>1)</sup> Bezüglich der Klassenbreite siehe auch 3.1

<sup>2)</sup> Für ein Element, dessen Wert  $y$  zur Zeit  $t = 0$  von  $q_i$  verschieden ist, ist dieser  $EM$ -Wert mit einem systematischen Fehler in der Größenordnung der halben Intervallbreite behaftet.

TABELLE 5 Die Erwartungswerte der DML<sup>1)</sup>  $E(i, n)$  (1000 l/Kuh u. Jahr)

Basis 1964 $n$	Jahr	$K_4$	$K_1$	$K_2$	$K_3$	$K_9$	$K_{10}$	$K_{11}$	$K_{16}$	$K_{17}$	$K_{18}$
1	1965		2,31	2,55	2,64	3,21	3,30	3,41	3,93	4,04	4,17
Variations- koeffizient (in %) c.a.			8,	10,	8,	6,	6,	6,	4,	4,	3,
6	1970		2,69	2,96	3,05	3,50	3,59	3,70	4,03	4,06	4,09
Variations koeffizient			34,	32,	29,	18,	16,	14,	4,	2,	5,
11	1975		3,05	3,29	3,37	3,74	3,81	3,88	4,05	4,07	4,08
Variations koeffizient			45,	39,	37,	20,	18,	15,	3,	1,	5,

<sup>1)</sup> Korrektur gegenüber  $EM(i, n)$ : Jährliches Anwachsen von  $q_N$  um 10 1, beginnend 1964 mit 4,229.

Man erhält mit der Definition (16) die Folge „der Erwartungswerte der Markov-Kette“

$$EM(i, n) = \sum_{k=1}^N p_{ik}^{(n)} q_k \quad (17)$$

für die Klassen  $i = 1, \dots, N$  und die Potenzen  $n = 1, 2, 3, \dots$ , das ergibt  $N$  Zeitreihen von Erwartungswerten als Funktion des Anfangszustands zur Zeit  $t = 0$  (s. Tab. 5). Die  $q_k$  sind per definitionem abhängig von der Belegung der Klasse  $k$  mit Elementen und dadurch variabel in der Zeit. Für stärker belegte innere Klassen  $K_i$  bleibt jedoch zu jeder Zeit die obige Überlegung bezüglich der Näherung durch (15) richtig. Anders verhalten sich die offenen Klassen  $K_1$  bzw.  $K_N$ . Ihr Verhalten hängt eng mit der Art des betrachteten Prozesses zusammen. Handelt es sich bei der Beobachtungsvariablen um eine mit der Zeit wachsende Variable, wie die d. MI, so heißt das, daß die Belegung der Klasse  $K_1$  abnimmt und die verbleibenden Elemente gegen  $g_1$  wachsende Werte der Variablen annehmen. Der Kennwert der Klasse,  $q_1$ , wächst also gegen  $g_1$ , gleichzeitig jedoch nehmen die Wahrscheinlichkeiten  $p_{11}$  ab. Bezeichnet  $q_1^{(n)}$  den zeitlichen Verlauf von  $q_1$ , dann gilt für den Fehler  $F_1$  von  $EM(i, n)$  folgende Einschränkung; der Fehler wird klein:

$$[p_{11}^{(n)} (g_1 - q_1) \geq F_1 = p_{11}^{(n)} q_1^{(n)} - p_{11}^{(n)} q_1] \rightarrow (0)^1 \quad (18)$$

Gleiches Verhalten zeigen alle Klassen, die im Laufe des Prozesses durch die „Entvölkerung“ darunterliegender Klassen zu „untersten“ Klassen werden.

Unter der Annahme, daß wesentliche Verschiebungen des Kennwertes  $q_k$  der kleinsten Klasse  $K_k$  erst eintreten, wenn sie wenig belegt sind, also ihre  $p_{ik}$  klein sind, bedeuten sie keine Komplikation für die Interpretation der Werte  $EM(i, n)$ .

Wesentlich anders verhält sich der Wert  $q_N$  der obersten und offenen Klasse. Im Laufe eines Prozesses mit wachsenden Werten der Variablen  $y$  wird 1. die obere offene Klasse im allgemeinen ständig stärker mit Elementen belegt, und 2. wachsen die Werte der

<sup>1)</sup> (0) bezeichnet hier einen kleinen, von Null verschiedenen Wert, da  $p_{11}$  in einer regulären Markov-Kette für große  $n$  von Null verschieden ist.



Variablen der bereits in der obersten Klasse vorhandenen Elemente mit der Zeit. Der Wert  $q_N$  wächst. Sei  $d_{q_N}(n)$  der Betrag, um den sich  $q_N$  vom Zeitpunkt  $t = 0$  bis zum Zeitpunkt  $t = n$  geändert hat, dann kann man dem Wert  $EM(i, n)$  folgenden korrigierten Erwartungswert gegenüberstellen

$$E(i, n) = \sum_{k=1}^{N-1} p_{ik}^{(n)} q_k + p_{iN}^{(n)} q_N + p_{iN}^{(n)} d_{q_N}(n) \quad (19)$$

#### 4.2 Erwartungswert der $N$ -ten Klasse

Für Anwendung und Interpretation wichtig ist die folgende Eigenschaft des Erwartungswertes der  $N$ -ten Klasse

$$EM(N, n) = \sum_{k=1}^{N-1} p_{Nk}^{(n)} q_k + p_{NN}^{(n)} q_N \quad (20)$$

Per definitionem sind alle  $q_k$  kleiner als  $q_N$ , es gilt also

$$q_k + d_k = q_N, d_k > 0 \text{ für alle } k = 1, \dots, N-1 \quad (21)$$

Aus

(20) und (21) folgt

$$EM(N, n) = \sum_{k=1}^N p_{Nk}^{(n)} q_N - \sum_{k=1}^{N-1} p_{Nk}^{(n)} d_k = q_N - D \quad (22)$$

mit

$$D = \sum_{k=1}^{N-1} p_{Nk}^{(n)} d_k \geq 0, \text{ daher also} \quad (23)$$

$$EM(N, n) \leq q_N. \quad (24)$$

In (23) und (24) gilt das Gleichheitszeichen genau dann, wenn alle  $p_{Nk}$ ,  $k = 1, \dots, N-1$ , Null sind. Der Erwartungswert der höchsten Klasse  $K_N$  ist also stets kleiner als ihr Kennwert  $q_N$ , es sei denn, es gelte  $p_{NN} = 1$ , die Kette wäre also nicht regulär, sondern absorbierend.

Stellt man den Erwartungswerten der Markov-Kette (17) die um die Veränderung von  $q_N$  bereinigten Werte (19) gegenüber, so ergibt sich die Differenz

$$dE(i, n) = p_{iN}^{(n)} d_{q_N}(n) \quad (25)$$

Für alle Fälle in denen  $dE(i, n)$  hinreichend klein gegen  $EM(i, n)$  ist, stellt also (17) eine Näherung für (19) dar.

## 5 Methodisch bedingte Voraussetzungen für die Anwendbarkeit

### 5.1 Die Beobachtungsvariablen

Für alle Elemente des Systems muß ein Sprung von  $K_i$  nach  $K_k$  eine vergleichbare Bedeutung haben. Die Variable muß also einen relativen oder, um einen Ausdruck aus der Physik zu verwenden, spezifischen Charakter haben. Durch die Bezugsgröße wird die Vergleichbarkeit sichergestellt<sup>1)</sup>. Größen solcher Art sind Leistungskoeffizienten wie die durchschnittliche Milchleistung pro Kuh und Jahr. Wählte man bei Kreisen als

<sup>1)</sup> Eine typische Größe dieser Art aus der Physik ist das „spezifische Gewicht“

Elementen die Anzahl der Milchkühe als Variable, könnte sie obige Forderung nicht erfüllen, da hierbei Veränderungen gleicher absoluter Größe vollkommen verschiedene Bedeutung haben können. Vergleichbare Größen wären jedoch<sup>1)</sup> „Anzahl der Milchkühe je ha LN“ bzw. „Anzahl je ha Ackerfutterfläche“.

### 5.2 Die Menge der Beobachtungselemente

Die Menge der Elemente muß bezüglich der Zustände in dem Sinn homogen sein, daß, zumindest theoretisch, jedes Element jeden Zustand erreichen kann. Sofern nicht ein Zustand absorbierend ist, sorgt die Markov-Kette dafür, daß nicht alle Elemente in einem Zustand enden.

Naturgemäß haben bei Erstellung der P-Matrix noch nicht alle Elemente die ganze Skala der Zustände durchlaufen; die Markov-Kette soll ja gerade in den zu erwartenden Fluß der Elemente Einblick geben. Vorausgesetzt wird, daß ein solches Durchlaufen aller Zustände unter entsprechender Entwicklung der äußeren Variablen möglich wäre. Gleich wie bei funktionalen Vorschätzungen u. a. der Verlauf der berechneten Funktion die Extrapolierbarkeit beschränkt, können bei Markov-Ketten die zwischenzeitlichen Verteilungen regulativ wirken<sup>2)</sup>.

### 5.3 Referenzperiode und Extrapolierbarkeit

Da die Methode auf der Anwendung der Theorie endlicher Markov-Ketten mit stationären Transitionswahrscheinlichkeiten beruht, sind also während der Untersuchungsperiode die Wahrscheinlichkeiten  $p_{ik}$  der einschrittigen Sprünge als konstant zu betrachten. Das schränkt sowohl die Referenzperiode als auch den Zeitraum für die denkbare Extrapolation erheblich ein. Gemäß der geschilderten Erhebungsmethode gehen nur zwei Variablen in den Prozeß ein, Beobachtungsvariable und Zeit. Die  $p_{ik}$  sind also Ausdruck gewisser äußerer Variablen, die Änderungen der Variablen verursachen, sofern man der betrachteten Variablen kein autonomes Wachstum zugesteht. Die Markov-Kette kann also den Prozeß nur solange beschreiben, wie die Entwicklung aller äußeren Faktoren die gleiche bleibt wie in der Referenzperiode, bzw. kann sie den Prozeß nur so beschreiben, als ob das der Fall wäre.

Die unter 5.1–3 gemachten Einschränkungen verstehen sich für alle mittels Markov-Ketten zu untersuchenden Probleme. Es sind ähnliche, wie sie bei Anwendung fast aller mathematischer Methoden entstehen und geprüft werden müssen. Inwiefern sie die Anwendbarkeit einer Methode einschränken, hängt oft vom Grad der Abstraktion und vom Zweck der Untersuchung ab. Abgesehen von diesen Punkten stellt die empirische Definition noch Anforderungen an das Datenmaterial. Selbstverständlich müssen hinreichend lange Zeitreihen von hinreichend vielen Elementen vorliegen, um einigermaßen gesicherte Schätzungen der Transitionswahrscheinlichkeiten zu erhalten. In der Wirkung bedeutet das insbesondere eine Einschränkung bezüglich der Anzahl der zu bildenden Klassen der Variablen und damit eine Beschränkung der Genauigkeit. Die noch unter 5.4 zu behandelnden Beschränkungen beziehen sich auf den Fall, daß man über die Information bezüglich der Struktur des Systems, der „Strömung“ der Elemente, hinaus eine Aussage über die zeitliche Abhängigkeit der Beobachtungsvariablen unter Benutzung der Erwartungswerte der Markov-Kette machen will.

<sup>1)</sup> (zumindest theoretisch)

<sup>2)</sup> Durchgängigkeit durch das System kann auch als Annahme konzipiert sein, um die notwendigen Aufwendungen zur Erreichung einer solchen Entwicklung zu beleuchten.

#### 5.4 Die Erwartungswerte

wurden gemäß

$$EM(i, n) = \sum_{k=1}^N p_{ik}^{(n)} q_k \quad (17)$$

definiert. Das sind  $N$  Zeitreihen für die Beobachtungsvariable in Abhängigkeit von den  $N$  möglichen Anfangszuständen zur Zeit  $t = 0$ . Demgegenüber steht der korrigierte Wert

$$E(i, n) = \sum_{k=1}^N p_{ik}^{(n)} q_k + p_{iN} dq_N(n) \quad (19)$$

für den Fall wachsender  $q_N$ . Die Frage ist nun, in welchen Fällen (17) eine brauchbare Näherung für (19) darstellt.

- a) Es werde angenommen, daß  $dq_N(n) = 0$  für alle  $n$  sei. Der Kennwert der obersten Klasse wird damit als konstant vorausgesetzt. Die Entwicklung der Beobachtungsvariablen stagniert in der Umgebung dieses Wertes. Dann sind die Gleichungen (17) und (19) identisch. Der Markov-Prozeß beschreibt dann einen Lern- bzw. Anpassungsprozeß der zur Zeit  $t = 0$  in einem Zustand  $S_k$ ,  $k \leq N-1$ , befindlichen Menge von Elementen. Die Potenzen  $P^n$  und die daraus abgeleiteten Werte  $EM(i, n)$  beschreiben den Grad der Anpassung zur Zeit  $t = n$ . Die Folge der  $P^n$  konvergiert gegen  $A$ , die Folgen der  $EM(i, n)$  konvergieren gegen den gemeinsamen Grenzwert

$$EM(i, \infty) = \sum_{k=1}^N a_{ik} q_k \quad (26)$$

Der Grenzwert  $EM$  und die Matrix  $A$  sind ein Maß für die bei gleichbleibender Entwicklung mögliche Anpassung. Das Ergebnis entspricht der Vorstellung, daß in einem solchen Anpassungsprozeß nicht notwendig alle Elemente auf Grund ihrer „physischen“ Voraussetzungen in der Lage sind, sich vollkommen anzupassen. Falls diese vollkommene Anpassung als Hypothese untersucht werden soll, definiert man  $K_N$  als absorbierenden Zustand mit  $p_{NN} = 1$ , die Folgen  $P^n$  und  $EM(i, n)$  beschreiben die Anpassung, die Grenzverteilung  $a$  hat die Form  $(0, 0, 0, \dots, 0, 1)$  und der Grenzwert  $EM(i, \infty)$  ist gleich  $q_N^1$ .

- b) Es werde angenommen, daß  $dq_N(n) > 0$ ,  $dq_N(n+1) > dq_N(n)$  gelte, dann beschreiben die Werte (19)  $E(i, n)$  die Abhängigkeit der Beobachtungsvariablen von der Zeit unter Berücksichtigung des Wachstums von  $q_N$ . Die Frage ist, unter welchen Voraussetzungen dennoch  $EM(i, n)$  als gute Näherung für  $E(i, n)$  dienen kann.

Es gilt (25)  $dE(i, n) = E(i, n) - EM(i, n) = p_{iN}^{(n)} dq_N(n)$  für alle  $i$ .  $EM(i, n)$  ist genau dann eine hinreichend gute Näherung für  $E(i, n)$ , wenn  $dE(i, n)$  hinreichend klein ist;  $dE$  ist seinerseits eine Funktion des Wachstums  $dq_N(n)$  des Kennwertes des letzten Zustands und der Wahrscheinlichkeit  $p_{iN}^{(n)}$  der Übergänge von  $S_i$  nach  $S_N$  in  $n$  Schritten. Im allgemeinen wächst in einem Wachstumsvorgang die Wahrscheinlichkeit der Übergänge in die letzte Klasse für kleinere Anfangszustände  $S_i$  mit der Zeit. Da gleichzeitig auch  $q_N$  wächst, wird die Abweichung  $dE(i, n)$  größer. Gibt man für die Abweichung  $dE(i, n)$  eine maximale Schranke  $K$  vor, dann gilt (17) solange als Schätzung für (19),

<sup>1)</sup> G. G. JUDGE und E. R. SWANSON, (12; 13) definierten in ihrem Schweineproduzenten-Beispiel  $K_1$  und  $K_N$  als absorbierende Zustände und untersuchten damit die Zahl der Restproduzenten im Gleichgewichtsfall.

solange nur  $dE(i, n) \leq K$  gilt. Da  $dE$  eine Funktion des Anfangszustandes  $i$  ist, wird die erlaubte Näherungszeit  $T$  ebenfalls vom Anfangszustand abhängen.  $T(i)$  ist bei Kenntnis von  $k, p_{iN}^{(n)}, dq_N(n)$  berechenbar. Ist, wie es allgemein der Fall sein wird,  $dq_N$  unbekannt, dann ist die Zeit der gültigen Schätzung  $T(i)$  zumindest mit  $dq_N$  als Parameter einschränkbar.

Zusammenfassend läßt sich sicher als Ergebnis formulieren: Der Wert  $EM(i, n)$  ist für kleine  $i$ , in unserem Beispiel also Klassen niedriger durchschnittlicher Milchleistung und/oder für kleine Werte von  $t$  („kurz“-fristig) eine hinreichend gute Näherung für den Wert  $E(i, n)$ .

Unter gewissen Voraussetzungen können also die Markov-Kette und ihre Erwartungswerte zur Beschreibung von Lern- bzw. Anpassungsprozessen und — lediglich unter Berücksichtigung des bisher gesagten — zur Ermittlung „kurz“-fristiger Prognosen dienen. Was im Einzelfall unter kurzfristig zu verstehen ist, hängt unter anderem von der Länge der betrachteten Zeitreihen, von der Homogenität der Kette bezüglich der Elemente, von der Struktur der P-Matrix, von der Veränderung der extremen Kenngrößen in der Zeit und von den Einflüssen äußerer Faktoren ab.

## 6 Vergleich mit anderen Schätzmethoden

### 6.1 Beschränkungen und Information

Ein Vergleich von Markov-Ketten als Mittel zur Schätzung gewisser spezifischer bzw. Leistungskoeffizienten mit anderen Methoden ist nur beschränkt möglich. Nach der oben beschriebenen mathematischen Struktur in dieser einfachen Form eignen sie sich u. U. für kurzfristige Prognosen. Bei Berücksichtigung der Art des Eingehens von äußeren Faktoren kann das nur bestätigt werden. Ein Vergleich kann also nur mit Methoden ähnlicher Informationsstruktur stattfinden.

Modelle, die weitere Mengen an Information verwenden und daher zur langfristigen [10] Projektion geeignet erscheinen, sind nicht vergleichbar. Solche Methoden wären die „synthetischen“ oder „naiven“ Modelle [19], die, ausgehend von einer Vorausschätzung der Nachfragefunktion nach landwirtschaftlichen Produkten, die Produktion den veränderten Nachfrageverhältnissen anpassen; ähnlich ungeeignet zum Vergleich ist auch das Input-Output-Modell von Heady und Carter [10; 9], das interregionale und intersektorale Warenströme berücksichtigt.

Vergleichbar dagegen sind Methoden, die auf der Schätzung von Trends beruhen. Sie sind das relativ einfachste Verfahren. Die Kritik [19, S. 44 ff, S. 52] richtet sich gegen die grundsätzlichen Annahmen, auf denen die Anwendung der Methode beruht, z. B. ein im Vergleich zur Basisperiode unverändertes Preisniveau und die gleichmäßigen Einwirkungen des technischen Fortschritts auf die Prozesse. Die gleichen Einwendungen lassen sich selbstverständlich auch gegen die Anwendungen von Markov-Ketten in der geschilderten einfachsten Form erheben. Damit wird jedoch der in 5.4 gesteckte Rahmen nicht berührt.

Markov-Ketten sind bezüglich der mindestens notwendigen Information ein vergleichsweise „einfaches“ Verfahren. Sie haben die gleichen Vorteile wie Trendschätzungen, nämlich (unter Vorbehalten) anwendbar zu sein, wenn über gewisse gesamtwirtschaftliche Aspekte des Problems unvollständige, mangelhafte oder gar keine Information vorliegt. Ferner teilen sie mit den Trendschätzungen den Anwendungsbereich. Nach G. SCHMIDT ist deren Anwendung am ehesten berechtigt auf Leistungskoeffizienten [19, S. 53].

TABELLE 6 Vergleich einiger Schätzungen

		(1)	(2)	(3)	(4)
Jahr Po- tenz von $P$	Variable	Statistisches Jahrbuch der BRD (Wj 64/65)	Reguläre Mar- kov-Kette EWDML kor- rigiert, lineare Zunahme von $q_{18}$ ; jähr- lich 10 1	Absorbie- rende Mar- kov-Kette $p_{18,18} = 1$	Absorbierende Markov-Kette $p_{18,18} = 1$ ; line- are Korrektur
1965 $n = 1$	Milchkühe in 1000 St.	5.826	5.712 <sup>1)</sup>	5.663 <sup>2)</sup>	5.712 <sup>1)</sup>
	Milchproduktion in 1000 l	21.020	20.407 <sup>3)</sup>	20.113	20.494
	dMl in 1000 l	3.608	3.57	3.55	3.59
1970 $n = 6$	Milchkühe in 1000 St.		5.712	5.660	5.712 <sup>1)</sup>
	Milchproduktion in 1000 l		21.351	21.180	21.798
	dMl in 1000 l		3.74	3.74	3.82
1975 $n = 11$	Milchkühe in 1000 St.		5.712	5.657	5.712 <sup>1)</sup>
	Milchproduktion in 1000 l		22.112	21.961	22.909
	dMl in 1000 l		3.87	3.88	4.01

<sup>1)</sup> Hierbei wurde unterstellt, daß der Viehbestand 1964 unverändert bleibt.

<sup>2)</sup> Hierbei wurden geringe Veränderungen der Bestände der Klassen in der Größenordnung bis 5% innerhalb von 10 Jahren angenommen.

<sup>3)</sup> Geschätzte Produktion = Summe über (Erwartungswert dMl der Klasse  $\times$  Anzahl Milchkühe i. d. Klasse)

Im Rahmen der Anwendbarkeit liefert eine Markov-Kette aber weitergehende Informationen:

- Die  $P$ -Matrix und ihre Potenzen geben Aufschluß über das zu erwartende Verhalten der der Beobachtung zugrunde gelegten Elementenmenge, und zwar abhängig von dem bereits zur Zeit  $t = 0$  erreichten Leistungsniveau.
- Das Modell liefert eine theoretische Gleichgewichtsverteilung bei gleichbleibender Entwicklung.
- Eine Interpretation der Erwartungswerte liefert Information über die zu erwartenden Werte der Beobachtungsvariablen.

Eine Markov-Kette kann demgemäß unter gleichermaßen „mageren“ Bedingungen zusätzliche Informationen über die Struktur des Prozesses geben.

## 6.2 Mögliche Verbesserungen der Methode

Eine Reihe von Einschränkungen der Wirksamkeit von Markov-Ketten als Methode der Schätzung beruhen darauf, daß die Wahrscheinlichkeiten  $p_{ik}$  stationär und der

Einfluß äußerer Faktoren nicht berücksichtigt wurde. Beide Einschränkungen sind nicht notwendig.

L.G. Telser [20, S. 286] schlägt eine Methode vor, die Transitionswahrscheinlichkeiten  $p_{ik}$  zeitlich variabel als Funktionen gewisser äußerer Variablen darzustellen

$$P_{ik}(t) = f_{ik}(Z_{1t}, Z_{2t}, \dots, Z_{nt}) \quad (27)$$

Jede der Transitionswahrscheinlichkeiten  $p_{ik}$  wird mittels durch Regression oder anderer Methoden gewonnener Funktionen  $f_{ik}$  dargestellt. Die  $p_{ik}$  sind also nicht während des gesamten Prozesses konstant, sondern sie ändern sich in Abhängigkeit von gewissen äußeren Variablen, die Einfluß auf die Transitionen haben. Insbesondere ist es dadurch möglich, die Transitionswahrscheinlichkeiten zeitlich variabel anzunehmen. Unter Beibehaltung der stationären  $p_{ik}$  und unter Berücksichtigung der Veränderungen von  $q_N$  würde sich der Bereich der Anwendbarkeit ebenfalls erweitern. Zu untersuchen wären folgende Möglichkeiten.

- a) Für den Fall, daß weitergehende Informationen über das Verhalten von Elementen bei extrem hohen Leistungsniveau vorliegen, könnten die Werte  $q_N(n)$  in einem gesonderten Markov-Prozeß untersucht werden, der zur Korrektur des ursprünglichen Prozesses angewandt wird. Der Anpassungsprozeß wird dadurch auf ein höheres Niveau verlagert.
- b) Die Größe  $dq_N(n)$  wird mittels anderer Untersuchungsmethoden geschätzt. Eine Untersuchung über die Größe des relativen Fehlers, der bei Nichtberücksichtigung dieses Wachstums von  $q_N$  entsteht, gibt Aufschluß über die Größenordnung des Fehlers bzw. der Korrektur.

$$F_R(i, n) = \frac{p_{iN}^{(n)} dq_N^{(n)}}{\sum_{k=1}^N p_{ik}^{(n)} q_k + p_{iN}^{(n)} dq_N(n)} = \frac{1}{1 + \frac{\sum_{k=1}^N p_{ik}^{(n)} q_k}{p_{iN}^{(n)} dq_N}}$$

$$F_R(i, n) \leq \frac{1}{\frac{\sum_{k=1}^N p_{ik}^{(n)} q_k}{p_{iN}^{(n)} dq_N}} = \frac{p_{iN}^{(n)} dq_N(n)}{\sum_{k=1}^N p_{ik}^{(n)} q_k} \quad (28)$$

Mittels dieser Korrektur läßt sich das Verfahren (im Rahmen anders verursachter Grenzen) auch zur Beschreibung anderer als Lern- bzw. Anpassungsprozesse verwenden.

### Literatur

1. ADELMAN, L. G.: A Stochastic Analysis of the Size Distribution of Firms. J. Amer. Statist. Assoc. 53, 893 (1958)
2. ANDERSON, T. W. und GOODMAN, L. A.: Statistical Inference about Markov-Chains. Annals of Mathem. Stat. 28, (1957)
3. BHARUCHA—REID, A. T.: Elements of the Theory of Markov-Processes and Their Applications. New York—Toronto—London 1960
4. BOSTWICK, D.: Yield Probabilities as a Markov-Process. Agr. Econ. Res. 14, Nr. 1, 49 (1962)

<sup>1)</sup> Beträgt z. B.  $\sum_{k=1}^N p_{ik}^{(10)} q_k = 3800$ ,  $p_{iN}^{(10)} = 0,25$  und  $dq_N(10) = 400$ , so ist  $F_R = (i, n) = 0,026 \approx \approx 2,6\%$ . (Die Werte wurden dem zugrundegelegten Beispiel entnommen).

5. CHUNG KAI LAI: Markov—Chains with Stationary Transition Probabilities. Berlin—Göttingen—Heidelberg 1960
6. GOODMAN, L. A.: A Further Note on Miller's Finite Markov-Processes in Psychology. *Psychometrika* 18, 245 (1953)
7. Ders.: Simplified Runs Test and Likelihood Ratio Tests for Markoff Chains. *Biometrika*. 45, 181 (1958)
8. HART, P. E. und S. J. PRAIS: The Analysis of Business Concentration, a Statistical Approach. *J. Royal Statist. Soc. Ser. A*, 119, Nr. 10, 150 (1956)
9. HEADY, I. O. und H. O. CARTER,: An Input-Output Analysis Emphasizing Regional and Comodity Sectors of Agriculture. *Iowa Agr. Exp. Stat. Bul.* 469, Ames, Iowa 1959
10. Dies.: Input-Output Models as Techniques of Analysis for Interregional Competition. *Farm Econ.* 41, 978 (1959)
11. HOWARD, R. A.: Dynamic Programming and Markov-Processes. Cambridge, Mass.
12. JUDGE, G. G. und E. R. SWANSON,: Markov-Chains: Basic Concepts and Suggested Uses in Agricultural Economics. *Austr. J. Agr. Econ.* 6, 49 (1962)
13. Dies.: Markov-Chains: Basic Concepts and Suggested Uses in Agricultural Economics. Department of Agric. Econ. Agric. Exp. Station, University of Illinois. College of Agr., Urbana, Ill.; Research Report AERR 49, Dez. 1961
14. KEMENY, J. G. und J. L. SNELL,: Finite Markov-Chains. Princeton—New Jersey—Toronto—London—New York 1960
15. LEE, T. C., JUDGE, G. G. und T. TAKAYAMA: On Estimating the Transition Probabilities of a Markov-Process. *J. Farm. Econ.* 47, No. 3, 742 (1965)
16. LINDER, A.: Statistische Methoden. Basel und Stuttgart 1960
17. MILLER, G. A.: Finite Markov-Processes in Psychology *Psychometrika* 17 149 (1952)
18. PRAIS, S. J.: Measuring Social Mobility. *J. Royal Statist. Soc. Ser. A*, 118, Nr. 7, 56 (1955)
19. SCHMITT, G.: Methoden und Möglichkeiten der langfristigen Vorausschätzungen der Agrarproduktion EWG-Studien, Reihe Landwirtschaft Nr. 3. Brüssel 1961
20. TELSER, L. G.: Least-Squares Estimates of Transition Probabilities. In: *Measurement in Economics*, hrsg. von C. F. CHRIST u. a. Stanford, Calif. 1963
21. Ders.: The Demand for Branded Goods as Estimated from Consumers Panel Data. *Rev. Econ. Stat.* 44 (1962)
22. WEINSCHENCK G. und G. SCHELLER: Voraussichtliche Entwicklung der Nahrungsmittelproduktion in Westdeutschland. Göttingen 1959 — Entwicklungstendenzen der Erzeugung und des Verbrauchs von Nahrungsmitteln in der EWG (1956 — 1965). EWG-Studien, Reihe Landwirtschaft Nr. 2. Brüssel 1960