

SIEGFRIED GEISENBERGER/WOLFGANG MÄLICH

Informationstheoretische Messung regionaler Konzentrationserscheinungen

Dargestellt an der BIP-Konzentration in Baden-Württemberg

I. Einführung in die Problemstellung

In Heft 5/6 (1969) dieser Zeitschrift behandelte *Storbeck* Methodik und Problematik von Maßstäben der regionalen Konzentration. In der Tat gehören Konzentrationsmessungen zu den wichtigsten Aufgaben der empirisch ausgerichteten ökonomischen Forschung. Das Erkennen regionaler Verdichtungen, sei es zu einem Zeitpunkt als Querschnitt, sei es über die Zeit hinweg als Längsschnittanalyse, ist ein bedeutsames Diagnoseinstrument der Regionalpolitik.

Es ist nicht möglich, auf dem hier zur Verfügung stehenden Raum auf alle Aspekte räumlicher Verdichtungen einzugehen. Vielmehr soll nur ein besonders wichtiges und – soweit wir sehen – bisher vernachlässigtes Problem, und zwar die Konzentration lediglich klassifikatorisch ausgeprägter Merkmale, genauer analysiert werden.

Zur Verdeutlichung der Fragestellung seien zunächst einige maßtheoretische Bemerkungen vorangestellt. Für die Bestimmung jeder Maßzahl – und daher auch beim Konzentrationskoeffizienten – ist von entscheidender Bedeutung, von welcher Skala die vorhandenen Werte stammen¹. Will man z. B. eine Aussage über die zentrale Tendenz einer bestimmten Verteilung machen, so bestehen folgende Möglichkeiten:

1. Es liegen quantitative Daten vor, die einen kardinalen Vergleich der Abstände erlauben. Der Messung liegt demnach eine metrische Skala zugrunde, und das angemessene Maß mit minimalem Informationsverlust ist das arithmetische Mittel.
2. Man geht von komparativen Messungen aus. Hier ist nur noch ein Vergleich der Rangfolge möglich. Diese Daten beruhen somit auf einer topologischen oder ordinalen Skala. Als Maßzahl der zentralen Tendenz kann hier nur noch der Median verwendet werden, d. h. die Größe, der genauso viele Daten vorstehen, wie ihr nachfolgen, womit natürlich weniger Information vermittelt wird als durch das arithmetische Mittel unter 1.
3. Schließlich können klassifikatorische Werte vor-

liegen. Hier ist nur ein Vergleich der Unterschiedlichkeit der jeweiligen Werte erlaubt, da solchen Messungen eine Nominalskala zugrundeliegt. Als Ausdruck der zentralen Tendenz kann nur noch der Modus (d. h. der dichteste oder häufigste Wert) angegeben werden.

In entsprechender Weise sind maßtheoretische Überlegungen bei Erfassung der Streuung oder Konzentration vorzunehmen. Als Maßgrößen verwendet man hier in der theoretischen Statistik bei quantitativen Daten die Varianz (bzw. Standardabweichung), bei topologischen Größen die Quantilsabstände, und bei klassifikatorischen Merkmalen klafft hier eine Lücke. Dem Auffüllen dieser Lücke für regionale Probleme dienen die nachstehenden Ausführungen.

II. Relevanz des neuen Maßes in der Regionalforschung

Prüfen wir zunächst die Frage, inwieweit Daten mit klassifikatorischer Ausprägung in der Regionalforschung relevant sind. Als erstes Beispiel ist auf Warenströme zwischen Regionen hinzuweisen. Nimmt man als Ausgangsregion Baden-Württemberg, so können die Exporte nach anderen Bundesländern und den verschiedenen ausländischen Importländern aufgeschlüsselt werden. Berechnet man die jeweiligen Prozentanteile, so ergibt sich eine Verteilung, die über Länder definiert ist, was nur als klassifikatorische Skala interpretiert werden kann. Eine Charakterisierung dieser Verteilung mit den üblichen statistischen Maßen ist nur durch den häufigsten Wert möglich. Ferner wird versucht, durch graphische Darstellung der relativen Häufigkeiten ein anschauliches Bild der Verteilung zu vermitteln. Damit gelingt es jedoch nicht, die Struktur der Verteilung in eine Maßzahl zu verdichten, ohne die ein regionaler oder zeitlicher Vergleich von Unterschieden unmöglich ist.

Ein weiteres Beispiel für klassifikatorische Daten sind die Strukturen von Wanderungsströmen. Vom Standpunkt der statistischen Messung aus handelt es sich bei der Erfassung von Wanderungen um dasselbe Problem wie bei den Warenströmen. Alle Maße, die für letztere entwickelt wurden, sind unmittelbar auf Wanderungen anwendbar.

¹ Vgl. *Pfanzagl, J.: Die axiomatischen Grundlagen einer allgemeinen Theorie des Messens*. 2. Aufl. – Würzburg 1962. S. 11.

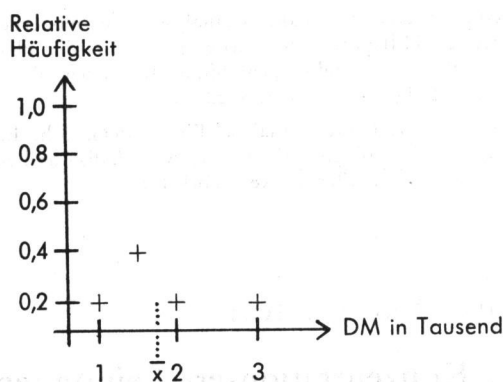


Abb. 1

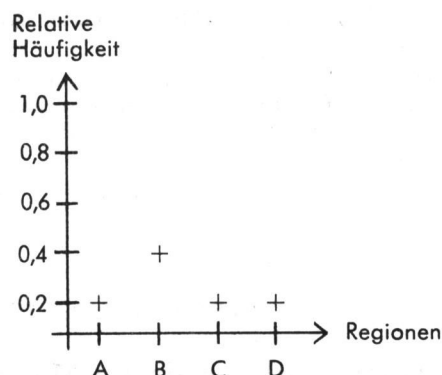


Abb. 2

Die Definition einer Verteilung auf der Nominalskala liegt schließlich auch dann vor, wenn BIP-Werte für Regionen gemessen wurden. Gemäß den obigen maßtheoretischen Überlegungen ist es daher statistisch nicht sinnvoll, z. B. vom arithmetischen Mittel der BIP-Werte für die 72 Stadt- und Landkreise von Baden-Württemberg zu sprechen. Dieser wichtige Grundgedanke sei an einem vereinfachten Beispiel illustriert:

Abb. 1 zeigt eine Verteilung, die auf metrischer Skala definiert wurde.

Die Berechnung des arithmetischen Mittels kann hier auf die übliche Weise erfolgen:

$$\bar{x} = 0,2 \cdot 1000 + 0,4 \cdot 1500 + 0,2 \cdot 2000 + 0,2 \cdot 3000 = 1800.$$

Dieser Wert wurde in Abb. 1 angedeutet. Zur Abb. 2 gelangt man, indem die BIP-Werte je Region in ihre Anteile umgerechnet werden. Die graphische Darstellung zeigt, daß in diesem Fall die Zuordnung eines Mittelwertes auf der Regionen-Achse nicht möglich ist, woran man besonders deutlich die Eigenart einer Nominalskala erkennen kann. Es ist daher nicht sinnvoll, vom mittleren BIP-Wert der 72 Kreise zu sprechen. Selbstverständlich können die BIP-Werte addiert und durch die entsprechende Anzahl dividiert werden, doch damit erhält man keinen Mittelwert im Sinne der theoretischen Statistik, sondern eine Aussage für das gesamte Land, die sachlich nicht interpretiert werden kann. Die einzige Aussage bezüglich der zentralen Tendenz aus Abb. 2 ist, daß Region B den höchsten BIP-Anteil aufweist (also dort der Modus vorliegt).

Entsprechende Maßprobleme ergeben sich, wenn eine Aussage über die Konzentration in Abb. 2 gewünscht wird. Bevor auf die spezifisch regionale Anwendung eingegangen wird, sollen in allgemeiner Form die statistischen Eigenschaften des dazu benötigten neuartigen Maßes in aller Kürze und ohne Anspruch auf mathematische Exaktheit erläutert werden.

III. Die Grundformel der Entropie-Messung und einige Verfeinerungen

Die Entropie als Maß der Ungewißheit oder des Informationsgehalts stammt aus der mathematischen Infor-

mationstheorie. Im folgenden soll unter bewußtem Verzicht auf mathematische Exaktheit heuristisch erklärt werden, was man sich unter der Entropie vorzustellen hat und welcher Art die verschiedenen formelmäßigen Ausdrücke sind.

Ausgangsüberlegung ist der plausible Gedanke, daß der Informationsgehalt einer Nachricht eine monoton fallende Funktion der Wahrscheinlichkeit p des Ereignisses ist. Hierbei ist zu beachten, daß „Wahrscheinlichkeit“ auch als empirische Häufigkeit verstanden werden kann. Unter Berücksichtigung einer Reihe weiterer Eigenschaften der gesuchten Funktion (z. B. Additivität bei unabhängigen Ereignissen, Informationsgehalt gleich Null beim sicheren Ereignis) folgt als explizite Form²:

$$h(p) = \log(1/p) = -\log p \quad (1)$$

wobei p die Wahrscheinlichkeit des Ereignisses darstellt.

Hat man nun eine Menge von Ereignissen $E_1, E_2, \dots, E_N$ mit zugeordneten Wahrscheinlichkeiten $p_1, p_2, \dots, p_N$, und ist ferner gesichert, daß eines mit Sicherheit eintreten wird, aber keine zwei gleichzeitig eintreten, so läßt sich der Grad der Unbestimmtheit des gesamten Systems numerisch bestimmen. Hierzu wird ein einzelnes Ereignis E_i mit Wahrscheinlichkeit p_i herausgegriffen. Gemäß obiger Gleichung (1) ist sein Informationsgehalt $-\log p_i$. Dieselbe Überlegung gilt jedoch auch für alle anderen Ereignisse. Da nun jedes Ereignis mit der Wahrscheinlichkeit p_i erwartet wird, gilt für den erwartenden Informationsgehalt (wie für jeden Erwartungswert der theoretischen Statistik) des gesamten Systems, und das ist die Entropie:

$$H(P) = -\sum_{i=1}^N p_i \log p_i \quad (2)$$

Diesen Ausdruck kann man als die Grundformel aller Entropie-Messungen ansehen. Es läßt sich zeigen, daß H nicht kleiner als Null und nicht größer als $\log N$ (bei einem System aus N Einzelereignissen) werden kann:

² Vgl. die ausführliche Darstellung bei Theil, H.: Economics and information theory. – Amsterdam 1967. Ferner Bohn, W. R.: Die informationstheoretische Messung von Struktursystemen des internationalen Handels. – Freiburg 1967. S. 48 ff., und Schultze, E.: Einführung in die mathematischen Grundlagen der Informationstheorie. – Berlin, Heidelberg, New York 1969.

$$0 \leq H(P) \leq \log N \quad (3)$$

Die Bezugsgröße der Logarithmen ist grundsätzlich beliebig, meist wird als Basis 2 oder die Eulersche Zahl verwendet. Die Maßeinheit der Entropie ist dann das bit (binary digit) oder nit. Bevor auf Ausweitungen der Gleichung eingegangen wird, sei am einfachen Beispiel die Aussage der Gleichung (2) kurz erläutert. Ein drastisch vereinfachtes System liegt beim Werfen einer Münze vor. Den beiden Ereignissen Kopf und Zahl sind bei der „fairen“ Münze die Wahrscheinlichkeiten $p_K = 0,5$ und $p_Z = 0,5$ zugeordnet. Folglich errechnet sich die Entropie gemäß (2) als: $-0,5 \cdot \log 0,5 - 0,5 \cdot \log 0,5 = 1$ bit. Weiß man jedoch aus früheren Versuchsreihen, daß tatsächlich (z. B. durch Verbiegung der Münze) $p_K = 0,7$ und $p_Z = 0,3$, dann ändert sich die Entropie zu: $+0,7 \cdot \log 0,7 + 0,3 \cdot \log 0,3 = 0,88$ bit. Die Entropie, die Unbestimmtheit, hat demnach abgenommen, ein durchaus plausibles Resultat; denn man kann von vornherein damit rechnen, daß eher „Kopf“ als „Zahl“ aufgrund der unterschiedlichen Wahrscheinlichkeiten auftreten wird.

An diesem Beispiel kann man sich auch den Zusammenhang zwischen Unbestimmtheit und Informationsgehalt verdeutlichen. Vor Kenntnis über den Ausgang des Versuchs herrscht Unbestimmtheit. Sie ist bei der fairen Münze größer. Nach Durchführung des Versuchs hat man eine Information erhalten, die bei der fairen Münze ebenfalls größer ist. Beides deshalb, weil die faire Münze den Ereignissen „Kopf“ und „Zahl“ die gleichen Eintrittswahrscheinlichkeiten gibt.

Die Verwendung dieses Maßes für regionale Probleme sei folgendermaßen illustriert. Gegeben sind die BIP-Werte für $N = 72$ Stadt- und Landkreise von Baden-Württemberg. Jeder Kreis hat einen bestimmten Anteil p_i an dem Land. Diese empirische Häufigkeit entspricht den obigen Wahrscheinlichkeiten. Im Land liegt maximale Gleichheit der BIP-Werte in den einzelnen Kreisen vor, wenn jeder Kreis den gleichen Anteil aufweist, d. h. $p_i = 1/N$ für alle Kreise. Umgekehrt hat man die größte Ungleichheit der BIP-Verteilung, wenn – hypothetischerweise – ein Kreis den gesamten BIP-Wert des Landes auf sich vereinigt, d. h. ein p_i ist 1, alle anderen sind Null. Berechnet man für diese beiden Extremsituationen die obige Entropiegleichung (2), so erhält man für die maximale Gleichheit $H = \log N$, für die größte Ungleichheit der BIP-Verteilung $H = 0$. Folglich ist die Entropie nichts anderes als ein Maß für die BIP-Gleichheit unter den Kreisen im gesamten Land. Die Aussage kann auch als Information über die Ungleichheit der BIP-Verteilung formuliert werden, wenn man den tatsächlichen Wert von der maximalen Entropie subtrahiert:

$$\log N - H(P) = \sum_{i=1}^N p_i \log N p_i = U(p_i) \quad (4)$$

Dieses Maß der Ungleichheit oder Konzentration einer Verteilung wird nur von den relativen Häufigkeiten (p_i) und der Anzahl der Elemente beeinflusst, nicht jedoch von der Art der Merkmalsausprägung. Genau aus diesem Grunde ist die Form (4) der Entropie aber für die Konzentrationsmessung klassifikatorischer Daten geeignet. Der nume-

rische Wert des Ausdrucks (4) schwankt zwischen 0 (es liegt überhaupt keine Ungleichheit vor) und $\log N$ bei totaler Ungleichheit. In bezug auf die größtmögliche Ungleichheit ist somit das Maß nicht normiert. Dies ist jedoch durchaus sinnvoll; was an einem Beispiel illustriert sei. Wenn eine Einkommensverteilung vorliegt dergestalt, daß ein Individuum das gesamte Einkommen, ein zweites nichts bezieht, so ist $U(p_i) = \log 2$. Sind jedoch insgesamt 999 Individuen mit Null-Einkommen vorhanden, dann beträgt $U(p_i) = \log 1000$, also einen weit größeren Wert. Die zweite Situation wird demnach als extremere Ungleichheit ausgewiesen.

Die Definition der Entropie in (4) läßt ohne weiteres die positive Eigenschaft erkennen, daß das Konzentrationsmaß unempfindlich gegenüber einem proportionalen Ansteigen oder Abfallen der Messungen bei allen Merkmalsausprägungen ist.

Schließlich sei auf die Ähnlichkeit des Entropie-Maßes mit der Varianz bei metrischen Skalen hingewiesen. Definiert man x_i als Merkmalsgrößen und p_i als entsprechende Wahrscheinlichkeiten, so ist die Varianz im diskreten Fall bestimmt durch:

$$s^2 = \sum_{i=1}^N \left[x_i - \left(\sum_{i=1}^N x_i p_i \right) \right]^2 p_i \quad (5)$$

Insbesondere im Vergleich mit Ausdruck (2) erkennt man die Ähnlichkeit der Maße, vor allem, daß bei der Varianz sich nicht nur die Wahrscheinlichkeiten, sondern auch die Merkmalswerte in (5) auswirken. Daß letzteres bei der Entropie nicht der Fall ist, rechtfertigt ihre Interpretation als Streuungsmaß für Messungen auf einer Nominalskala.

Berechnet man diesen Wert für mehrere Jahre, so kann in einer Größe die Entwicklung der Ungleichheit verfolgt werden. Will man für ein bestimmtes Jahr feststellen, ob die Ungleichheit im Bundesland A oder im Bundesland B größer ist, so genügt ein direkter Vergleich des Ausdrucks (4) nicht, denn dieses Maß ändert sich – wie erwähnt – mit der Anzahl der im Land vorhandenen Kreise. Ausdruck (4) ist daher zu normieren, indem jeweils die tatsächliche Ungleichheit in Prozent der maximal möglichen angegeben wird.

Die einfache Entropiegleichung in den Formen (2) und (4) kann in vielfacher Weise erweitert werden, z. B. für mehrfach klassifizierte Merkmale (Wanderungen nach Regionen und Alter und Beruf). Die Maße werden dann rechnerisch etwas aufwendiger, bleiben aber in ihrer mathematischen Struktur immer noch relativ einfach.

Nachstehend werden zunächst die Ergebnisse der Konzentrationsmessung für BIP-Werte der Kreise Baden-Württembergs erläutert, danach Erweiterungen der Entropieformel eingeführt, soweit sie für die vorliegende Problemstellung notwendig sind.

Die folgenden numerischen Berechnungen wurden auf der IBM 7040 (32 K) – Anlage des Rechenzentrums der Universität Freiburg i. Br. durchgeführt. Das Ausgangsmaterial wurde einer Gemeinschaftsveröffentlichung der Statistischen Landesämter entnommen³. Für die totale Konzentration in

³ Das Bruttoinlandsprodukt der kreisfreien Städte und Land-

Baden-Württemberg (gemessen in nit) ergeben sich gemäß Gleichung (4) folgende Werte:

	1957	1961	1964	1966
Totale Konzentration	0,532	0,454	0,437	0,433
Prozentanteil am Konzentrationsmaximum (= 4,2767)	12,4	10,6	10,2	10,1

Das erste bemerkenswerte Ergebnis ist darin zu sehen, daß der Konzentrationsgrad in allen Jahren ziemlich gering ist. Dies mag zunächst überraschen, schwankt doch der BIP-Anteil der Kreise z. B. 1966 zwischen 14,44‰ (Stuttgart) und 0,29‰ (Münsingen). Man muß jedoch berücksichtigen, daß in das Maß alle 72 Kreise eingehen. Die Prozentsätze zeigen deutlich, wie gering insgesamt die Annäherung an das Maximum der Konzentration in allen vier Jahren ist.

Besonders interessant ist als zweites Ergebnis die zeitliche Entwicklung des Konzentrationskoeffizienten. Man erkennt eine deutliche Tendenz zur Verringerung der Konzentration. Ebenso klar ist aber auch zu beobachten, daß von Querschnitt zu Querschnitt die Abnahmerate kleiner wird.

Der Erhebungsrhythmus der BIP-Werte erlaubt leider nicht eine Beobachtung über die Entwicklung des Konzentrationsmaßes im Verlauf der Konjunktur. Daraus könnte man interessante Aufschlüsse über die Konjunktorempfindlichkeit der Kreise erhalten.

IV. Erste Erweiterung der einfachen Entropieformel

Die erste Verfeinerung der Entropieformel und vertiefte Einsichten über die regionale Konzentration erhält man durch Aufspaltung aller Untersuchungsobjekte in Untergruppen. So lassen sich die Stadt- und Landkreise von Baden-Württemberg zu vier Gruppen, entsprechend den vier Regierungsbezirken, zusammenfassen. Während im vorangegangenen Abschnitt lediglich die totale Konzentration beurteilt werden konnte, ist nun der Frage nachzugehen, inwieweit die gesamte Entropie aus der Ungleichheit innerhalb der Regierungsbezirke einerseits oder zwischen den Regierungsbezirken andererseits resultiert.

Zur Lösung dieser Frage ist davon auszugehen, daß die Gesamtheit der N Objekte in G Gruppen $S_1, \dots, S_G$ aufgespalten wird. N_g ist die Anzahl von Objekten in der Gruppe S_g , so daß $\sum_g N_g = N$. Der Anschaulichkeit halber soll zunächst von der einfachen Entropiegleichung (2) als Maß der Gleichheit ausgegangen werden. Durch einfache Umformung ergibt sich:

$$H(P) = \sum_{g=1}^G \left(\sum_{i \in S_g} p_i \log 1/p_i \right) \quad (6)$$

was lediglich bedeutet, daß die gleichen Berechnungen jetzt gruppenweise durchgeführt wurden. Als neue Größe wird nun $P_g = \sum_{i \in S_g} p_i$ eingeführt, womit der Anteil jeder

Gruppe an der Gesamtheit ausgedrückt wird. Erweitert man jedes p_i in (6) mit P_g , so ergibt sich:

$$H(P) = \sum_{g=1}^G \left(\sum_{i \in S_g} \frac{P_g}{P_g} p_i \log \frac{P_g}{p_i} \frac{1}{P_g} \right) \quad (7)$$

und nach einigen Umformungen:

$$H(P) = \sum_{g=1}^G P_g \left[\sum_{i \in S_g} p_i/P_g \log \left(1/P_g \right) + \sum_{g=1}^G P_g \log (1/P_g) \right] \quad (8)$$

Die totale Entropie ist damit in zwei Summanden aufgeteilt worden. Davon ist der zweite Term besonders einfach zu interpretieren: Er erfaßt die Gleichheit zwischen den Gruppen. Der erste Term beschreibt die Gleichheit innerhalb der Gruppen. Die Besonderheit ist hier darin zu sehen, daß für jedes Objekt der Anteil an der Gesamtheit mit dem Anteil der jeweiligen Gruppe gewogen wird. Für diese Größe (p_i/P_g) wird die Entropie innerhalb der Gruppe berechnet, die Summe einer jeden Gruppe wird noch mit dem Gruppenanteil an der Gesamtheit gewichtet.

Das abgeleitete Maß der Gleichheit besitzt indes noch eine unangenehme Eigenschaft: Die Gleichheit zwischen den Gruppen ist dann am größten, wenn alle Gruppenanteile gleich sind. Im allgemeinen ist jedoch die Anzahl der Objekte in den Gruppen verschieden, so daß aus dieser Eigenschaft von (8) eine Irreführung resultieren kann. Die Anzahl der Objekte ist aus diesem Grunde in das Maß aufzunehmen, was in einfacher Weise durch Transformation in ein Maß der Ungleichheit erreicht wird. Gemäß Ausdruck (4) gilt:

$$\begin{aligned} U(P) &= \log N - H(P) \\ &= \log N - \sum_{g=1}^G P_g \left[\sum_{i \in S_g} (p_i/P_g) \log \left(1/P_g \right) - \sum_{g=1}^G P_g \log (1/P_g) \right], \end{aligned}$$

und nach Umformungen:

$$U(P) = \sum_{g=1}^G P_g \left(\sum_{i \in S_g} \frac{p_i}{P_g} \log \frac{P_i/P_g}{1/N_g} \right) + \sum_{g=1}^G P_g \log \frac{P_g}{N_g/N} \quad (9)$$

Diese Gleichung kann analog zu Ausdruck (8) interpretiert werden, wobei lediglich zu beachten ist, daß es sich jetzt um die Ungleichheit der Verteilung handelt. Zum

kreise 1957 bis 1966. Gemeinschaftsveröffentlichung der Statistischen Landesämter. - Wiesbaden 1968. = Sozialproduktsberechnungen der Länder, H. 3.

zweiten Term in (9) (Ungleichheit zwischen den Gruppen) ist zu bemerken, daß zu seinem Verschwinden (d. h. völlige Gleichheit „zwischen“) die Gleichheit der Gruppenanteile noch nicht genügt. Vielmehr ist das erst dann der Fall, wenn zusätzlich die Anzahl der Objekte in allen Gruppen gleich groß ist. Der erste Term in (9) ist entsprechend als Ungleichheit innerhalb der Gruppen zu interpretieren.

Fassen wir das Ergebnis dieser ersten Verfeinerung der Entropieformel im Hinblick auf das Beispiel der Kreise zusammen. Ausgehend vom Anteil jedes Kreises am gesamten BIP des Landes und einer Gruppeneinteilung der Regionen (z. B. nach Regierungsbezirken), läßt sich exakt feststellen, in welchem Maße die gesamte Ungleichheit der BIP-Verteilung auf die Verhältnisse innerhalb der oder zwischen den Regierungsbezirken zurückgeführt werden kann.

Die Berechnungen ergeben folgende Konzentrationswerte (gemessen in nit):

	1957	1961	1964	1966
Totale Konzentration	0,532	0,454	0,437	0,433
Ungleichheit zwischen den Regierungsbezirken (Prozentualer Anteil am Gesamtwert)	0,039	0,070	0,067	0,066
Ungleichheit innerhalb der Regierungsbezirke	7,3	15,4	15,3	15,1
	0,493	0,384	0,370	0,367

Die Werte der totalen Konzentration wurden schon besprochen und sind hier nur der Übersichtlichkeit halber wiedergegeben. Die Aufspaltung der Gesamtwerte zeigt deutlich, daß die Ungleichheit zwischen den Regierungsbezirken nur in geringem Maße zur gesamten Konzentration beiträgt. Bemerkenswert ist ferner, daß zwar der prozentuale Anteil von 7,3 im Jahr 1957 auf 15,4 im Jahr 1961 gestiegen, seitdem jedoch im wesentlichen konstant geblieben ist. Daraus ist der wichtige Schluß zu ziehen, daß sich die Differenzen im relativen Beitrag der vier Regierungsbezirke zum BIP von Baden-Württemberg seit 1961 nur unerheblich verändert haben. Entsprechend kann man für die gesamte Ungleichheit innerhalb der Regierungsbezirke feststellen, daß eine starke Abnahme zwischen den Jahren 1957 und 1961 stattgefunden hat. Seither ist die Verringerung der Konzentration nur noch geringfügig, insbesondere beim Vergleich der Jahre 1964 und 1966 kann man von einer Konstanz sprechen.

Es würde nun weiter interessieren, wie sich die Ungleichheit innerhalb der einzelnen Regierungsbezirke zeitlich entwickelt hat. Um die Flexibilität des Entropie-Maßes zu demonstrieren, soll dieser Frage bei vermehrter Information im nächsten Abschnitt nachgegangen werden.

V. Erweiterung der Entropieformel bei vermehrter Information

Die Aussagekraft des Maßes kann verfeinert werden, wenn neben den BIP-Werten auch Bevölkerungsangaben

pro Kreis vorhanden sind. Man wird in erster Linie auf die Wirtschaftsbevölkerung zurückgreifen, um eine enge Beziehung zum BIP herzustellen. Nennen wir den Anteil jedes Kreises an der Wirtschaftsbevölkerung b_i , so ergeben sich schon aus der folgenden Definition interessante Aufschlüsse:

$$\frac{p_i}{b_i} = \frac{(BIP)_i / \sum_i (BIP)_i}{B_i / \sum_i B_i} = \frac{(BIP)_i / B_i}{\sum_i (BIP)_i / \sum_i B_i} \tag{10}$$

Aus der letzten Umformung erkennt man deutlich, daß (p_i/b_i) eine Information über das BIP je Kopf der Wirtschaftsbevölkerung des einzelnen Kreises, bezogen auf den Durchschnitt des Landes, liefert.

In der Literatur wird gezeigt, wie man mit der zusätzlichen Information die erweiterte Entropieformel ableiten kann. Hier wird jedoch vereinfacht vorgegangen, und zwar derart, daß in Gleichung (9) statt Angaben über die Zahl der Kreise nun die tatsächlichen Bevölkerungswerte eingesetzt werden. Dann ändert sich (9) folgendermaßen:

$$U(P, B) = \sum_{g=1}^G P_g \left[\sum_{i \in S_g} \frac{P_i}{P_g} \log \left(\frac{P_i/P_g}{b_i/B_g} \right) \right] + \sum_{g=1}^G P_g \log \frac{P_g}{B_g} \tag{11}$$

Der zweite Term in Gleichung (11) gibt wiederum die Ungleichheit zwischen den Gruppen an. Im Gegensatz zu (9) müssen nun sowohl BIP-Anteil als auch Bevölkerungsanteil in allen Gruppen gleich sein, damit dieser Wert Null wird. Die innere Klammer des ersten Terms zeigt die Ungleichheit innerhalb der jeweiligen Gruppe an. Gewogen mit dem spezifischen BIP-Gruppenanteil ergibt deren Summe die Ungleichheit innerhalb der Gruppen insgesamt.

Zur Verdeutlichung der veränderten Fragestellung sei besonders betont, daß es im letzten Abschnitt darum ging, die Konzentration der BIP-Werte, gemessen über die einzelnen Kreise (bzw. Regierungsbezirke), zu untersuchen. In diesem Abschnitt wird aufgrund der vermehrten Information die Analyse verfeinert, und zwar ist nun die Ungleichheit des BIP je Kopf der Wirtschaftsbevölkerung, gemessen über die Kreise, Gegenstand der Untersuchung.

Berechnet man gemäß Gleichung (11) die Konzentration der BIP-Werte per capita, so ergibt sich folgendes:

	1957	1961	1964	1966
Totale Konzentration	0,097	0,025	0,023	0,022
Ungleichheit zwischen den Regierungsbezirken (Prozentualer Anteil)	0,003	0,006	0,005	0,005
Ungleichheit innerhalb der Regierungsbezirke insgesamt	3,1	24,0	21,7	22,6
Ungleichheit innerhalb des Regierungsbezirks	0,094	0,019	0,018	0,017
- Nordwürttemberg (22 Kreise)	0,226	0,026	0,027	0,024
- Nordbaden (13 Kreise)	0,022	0,020	0,016	0,016
- Südbaden (20 Kreise)	0,022	0,012	0,010	0,011
- Südwürttemberg (17 Kreise)	0,004	0,004	0,005	0,005

Zu diesen Ergebnissen ist zunächst zu bemerken, daß die

gesamte Konzentration wesentlich niedriger als im vorangegangenen Abschnitt ist. Dies ist schon durch den Plausibilitätshinweis verständlich, daß bei dem hier verwendeten Maß die hohen BIP-Anteile der großen Kreise auch mit den entsprechenden Bevölkerungsanteilen gewogen werden, wodurch sich die Differenzen erheblich vermindern; z. B. ist 1957 der BIP-Anteil von Stuttgart mit 0,164 etwa 44mal so groß wie derjenige von Backnang (0,0024). Hingegen weist p_i/b_i mit 1,47 nur noch das Sechsfache gegenüber Backnang auf. Die neue Verteilung zeigt also wesentlich geringere Häufungen.

Im zeitlichen Verlauf der Gesamtkonzentration beobachtet man wiederum einen sehr starken Abfall zwischen 1957 und 1961. Danach vermindert sich die Konzentration bis 1966 nur noch geringfügig.

Die Aufspaltung der Gesamtgrößen läßt für die Ungleichheit zwischen den Regierungsbezirken einen umgekehrten Verlauf erkennen. Ihr prozentualer Anteil stieg zwischen 1957 und 1961 ganz erheblich von 3,1 auf 24,0%. Der Prozentsatz schwankt in den weiteren Jahren geringfügig, doch liegt er deutlich höher als bei den Berechnungen unter alleiniger Verwendung der BIP-Werte. Dieses Ergebnis ist interessant, läßt es doch deutlich erkennen, daß die Ungleichheit des BIP je Kopf der Wirtschaftsbevölkerung in den Kreisen erheblich von der Ungleichheit zwischen den Regierungsbezirken beeinflusst wird.

Die gesamte Ungleichheit innerhalb der Regierungsbezirke als zweite Teilgröße bedarf nach den obigen Ausführungen keiner weiteren Erörterung. Den zeitlichen Verlauf kann man aus der obigen Darstellung erkennen.

Interessant ist jedoch die Aufspaltung der Innerhalb-Ungleichheit in die Beiträge der einzelnen Regierungsbezirke. Beginnen wir mit Nordwürttemberg, so ist zwischen 1957 und 1961 eine außerordentlich starke Verminderung der Konzentration von 0,226 auf 0,026 zu beobachten. Der extrem hohe Wert für 1957 ist auf außerordentlich niedrige gewogene BIP-Anteile einiger weniger Kreise zurückzuführen. Berücksichtigt man, daß der Beitrag Nordwürttembergs zur Ungleichheit innerhalb aller Regierungsbezirke $0,226 \cdot 0,3693 = 0,0835$ beträgt, so wird deutlich, daß die hohen Gesamtwerte 1957 weithin durch diesen Regierungsbezirk verursacht werden.

Ein Vergleich zwischen den Regierungsbezirken in bezug auf die Stärke der Konzentration ist nicht ohne weiteres

möglich. Es sei an die obigen Ausführungen erinnert, wo darauf hingewiesen wurde, daß die Entropie auch von der Anzahl der Elemente abhängig ist. Andererseits ist bei der verfeinerten Entropiemessung die Normierung durch die maximale Entropie nicht möglich, weil hier das Maximum gegen den Logarithmus von Unendlich strebt, d. h. das Maximum steigt über alle Grenzen. Man kann jedoch aufgrund anderer Überlegungen, die nicht weiter verfolgt werden sollen, zeigen, daß zumindest ein direkter Vergleich der Größenordnung möglich ist. Unter diesem Aspekt zeigen die Ergebnisse, daß im Regierungsbezirk Nordwürttemberg in allen Untersuchungsjahren die stärkste Konzentration vorliegt. Mit abnehmender Konzentration folgen Nordbaden, Südbaden und Südwürttemberg. Abschließend sei bemerkt, daß in den drei erstgenannten Regierungsbezirken zwischen 1957 und 1961 eine Tendenz zur Verringerung der Konzentration zu beobachten ist. In Südwürttemberg war das nicht der Fall, doch könnte dies angesichts der sehr kleinen Werte auf statistische Ungenauigkeit zurückzuführen sein.

VI. Abschließende Bemerkungen

Die vorstehenden Ausführungen sollten mit den Haupteigenschaften eines Konzentrationsmaßes vertraut machen, das noch verhältnismäßig wenig bekannt ist. Es wurde darauf hingewiesen, daß die Entropie das einzige Konzentrationsmaß ist, wenn Messungen auf einer Nominalskala vorliegen. Ferner zeigten sich im Verlauf der Untersuchung günstige Aggregationseigenschaften des Maßes.

Die numerische Illustration des Konzentrationsmaßes erfolgte am Beispiel der BIP-Konzentration in Baden-Württemberg. Sie wurde bei unterschiedlichem Informationsgrad (mit und ohne Bevölkerungsangaben) durchgeführt. In beiden Fällen zeigte sich eine Abnahme bzw. Stagnation der Konzentration in den Kreisen Baden-Württembergs. Dasselbe Ergebnis war auch bei Aufspaltung in die Regierungsbezirke Baden-Württembergs zu beachten.

Schließlich ist darauf hinzuweisen, daß die Entropiemessung – wie jedes statistische Maß – zunächst lediglich eine Analyse der Gegebenheiten darstellt. Anschlußuntersuchungen, z. B. über die Ursache der dargestellten Entwicklung, wären sicherlich interessant, waren jedoch im Rahmen dieser Arbeit nicht möglich.