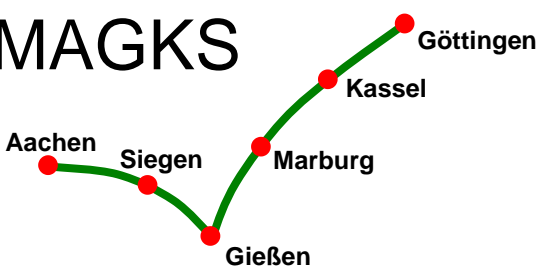


<p>MAGKS</p>  <p>Aachen Siegen Marburg Kassel Göttingen Gießen</p>	<p>Joint Discussion Paper Series in Economics</p> <p>by the Universities of Aachen · Gießen · Göttingen Kassel · Marburg · Siegen</p> <p>ISSN 1867-3678</p>
--	--

No. 10-2009

**Hans-Friedrich Eckey, Reinhold Kosfeld und
Alexander Werner**

**K-Funktionen als Instrument zur Analyse
räumlicher (De-)Konzentrationsprozesse**

This paper can be downloaded from
http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29

Coordination: Bernd Hayo • Philipps-University Marburg
Faculty of Business Administration and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

K-Funktionen als Instrument zur Analyse räumlicher (De-)Konzentrationsprozesse

Hans-Friedrich Eckeyⁱ, Reinhold Kosfeldⁱⁱ, Alexander Wernerⁱⁱⁱ

Zusammenfassung

Zur Messung räumlicher Konzentration ökonomischer Aktivitäten stehen verschiedene Verfahren zur Verfügung. Nachstehend wird mit den distanzbasierten uni- und bivariaten K-Funktionen eine neue Methode vorgestellt, die Vorteile gegenüber den herkömmlichen, auf regionaler Abgrenzung basierenden Verfahren aufweist. Mit Hilfe der univariaten K-Funktionen wird die Clusterung innerhalb der Sektoren im Verarbeitenden Gewerbe in Deutschland untersucht. Mit bivariaten K-Funktionen werden gegenseitige Abhängigkeiten bei der Clusterung dieser Sektoren untersucht.

Gliederung

1. Einführung in die Thematik
2. Traditionelle statistische Verfahren zur Analyse räumlicher (De-)Konzentrationsprozesse
 - 2.1 Betrachtung eines einzelnen Sektors
 - 2.2 Betrachtung der Abhängigkeit zwischen Sektoren
3. Die K-Funktion als alternatives Instrument zur Analyse räumlicher (De-)Konzentrationsprozesse
 - 3.1 Die Abgrenzung von K-Funktionen gegenüber traditionellen Verfahren
 - 3.2 Darstellung der K-Funktion
 - 3.2.1 Univariate K-Funktion: Ballung versus Streuung
 - 3.2.2 Bivariate K-Funktion: Anziehung versus Abstoßung
 - 3.3 Vorteile der K-Funktionen gegenüber traditionellen raumbezogenen Verfahren
4. Darstellung der empirischen Ergebnisse
 - 4.1 Verwendetes Datenmaterial
 - 4.2 Ergebnisse unter Verwendung univariater K-Funktionen
 - 4.3 Ergebnisse unter Verwendung bivariater K-Funktionen
5. Resumée

ⁱ Prof. Dr. Hans-Friedrich Eckey, Universität Kassel, Nora-Platiel-Str. 4, 34109 Kassel, Tel.: +49/561/804/3038, Fax: +49/561/804/3045, E-Mail: eckey@wirtschaft.uni-kassel.de.

ⁱⁱ Prof. Dr. Reinhold Kosfeld, Universität Kassel, Nora-Platiel-Str. 5, 34109 Kassel, Tel.: +49/561/804/3084, Fax: +49/561/804/3045, E-Mail: rkosfeld@wirtschaft.uni-kassel.de.

ⁱⁱⁱ Dipl.-Ök. Alexander Werner, Universität Kassel, Nora-Platiel-Str. 4, 34109 Kassel, Tel.: +49/561/804/3044, Fax: +49/561/804/3045, E-Mail: werner@wirtschaft.uni-kassel.de

1. Einführung in die Thematik

Die Beschäftigung mit der Konzentration ökonomischer Aktivitäten in Raum und Zeit kann auf eine lange Tradition zurückblicken und hat auch heute nichts von ihrer Aktualität und gesellschaftlichen Relevanz eingebüßt. Die entsprechende Literatur baut auf Marshall (1920) auf, der für die Erklärung räumlicher Ballung ökonomischer Aktivitäten bekanntlich zwischen den Vorteilen der industriellen Massenproduktion (betriebsintern) sowie Lokalisations- und Urbanisierungsvorteilen (beide betriebsextern) unterschied. Lokalisationsvorteile beschreiben die externen Ersparnisse, die sich aus der räumlichen Konzentration von mehreren Unternehmen der gleichen Branche ergeben. Die Begründung liegt in der gemeinsamen Nutzung spezifischer regionaler Arbeits-, Beschaffungs- und Informationsmärkte. Der Teil der Ersparnisse, der sich aus der räumlichen Konzentration von nicht branchengleichen Unternehmen ergibt, wird als Urbanitäts- oder Urbanisierungsvorteil bezeichnet. Die Nähe zu Vorlieferanten und Abnehmern, zu Banken, Versicherungen und anderen produktionsorientierten Dienstleistungen sowie zu (Aus- und Weiter-)Bildungs- und Forschungseinrichtungen wirkt standortverbessernd.

Es existiert mittlerweile eine fast unüberschaubare Literatur zu dieser Thematik, die sich in theoretische, empirische und politische Weiterentwicklungen einteilen lässt. Im Rahmen eher theoretisch orientierter Ausführungen kommt den Arbeiten von Romer (1986), Lukas (1988), Porter (2000) und Krugman (1991) ein besonderer Stellenwert zu. Romer betont die nur partielle Ausschließbarkeit jener Wirtschaftssubjekte von neuem Wissen, die keinen Beitrag zu seiner Entstehung geleistet haben („free-rider-position“); der Produktionsfaktor Wissen ist also mit positiven technologischen externen Effekten verbunden, die als Wissensspillovers bezeichnet werden. Lucas argumentiert, dass die Wissensspillovers regionale Bezüge haben, eine räumliche Nähe zum Entwickler neuen Wissens die Partizipation also erleichtert. Dies gilt insbesondere für das „tacit knowledge“, das an Personen gebunden ist und dessen Weitergabe persönlicher Kontakte bedarf. Im Gegensatz dazu trifft das „explicit knowledge“, das kodifizierbar ist und veröffentlicht wird, kaum auf räumliche Barrieren.

Räumliche Konzentration ökonomischer Aktivitäten wird als Cluster bezeichnet; dieser Begriff geht auf Porter zurück, der sich für folgende Definition entschieden hat: „A cluster is a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities“ (Porter 2000, S. 16).

Austauschbeziehungen und Zusammenarbeiten führen zu Verbindungen und Abhängigkeiten von Unternehmen, die räumliche Nähe begünstigen.

Im Modell von Krugman sind die Agglomerationsfaktoren und die aus ihnen resultierende Verteilung ökonomischer Aktivitäten von den Arbeitskosten, den Transportkosten und der Anzahl der Unternehmen abhängig. Eine stärkere Ballung führt zu wachsenden Agglomerationsvorteilen, da Unternehmen nun weitere Vorprodukte und Konsumenten weitere Konsumgüter billiger, nämlich ohne Transportkosten, beziehen können; durch sinkende Preise steigt das Realeinkommen, was zu weiteren Zuzügen führt. Dekonzentrierend wirkt dagegen die naturgegebene gleichmäßige räumliche Verteilung der Landwirtschaft, die bei einer starken räumlichen Konzentration der Industrieproduktion zu steigenden Transportkosten führt. Je nach Konstellation teilt sich die Industrieproduktion gleichmäßig auf beide betrachteten Regionen auf oder konzentriert sich ausschließlich in einer der beiden Regionen; dieses auf neoklassischer Theorie basierende Gleichgewicht kann stabil oder instabil sein und sich bei einer Variation der Modellfaktoren ändern.

Die empirische Clusterforschung beschäftigt sich

- mit der Bestimmung räumlicher Cluster,
- mit den Faktoren, die ihre Entstehung und Entwicklung beeinflussen,
- mit der Frage, ob sich Regionen, die durch Cluster dominiert werden, günstiger entwickeln als Wirtschaftsräume, die durch eine eher zufällige Sektoralstruktur dominiert sind.

Die Antworten auf diese Fragen sind ebenso vielfältig wie heterogen (Menzel 2008)¹. Weitergehender Konsens besteht darüber, dass die Vorteile regionaler Cluster von vielen Faktoren abhängen und nicht für alle Wirtschaftszweige gleich sind. Auch dass Agglomerationseffekte zwar eine Ursache für die Entwicklung eines Clusters darstellen, jedoch nicht seine Gründung entscheidend beeinflussen, weil wachstumsstützende Einflussgrößen wie Netzwerke und spezifische Infrastrukturen bei einer Clusterbildung noch nicht vorhanden sind, sondern sich erst später bilden, ist allgemein akzeptiert. Uneinigkeit herrscht dagegen über die Frage, ob Cluster stabil sind und ob sie das Wachstum von Regionen eher positiv oder eher negativ beeinflussen (Litzenberger 2007)². Ebenso kontrovers fallen die Antworten auf die Frage aus, welche Einflussfaktoren die Bildung und die Entwicklung von Clustern bestimmen, so, ob etwa die Lokalisations- oder die

¹ Ein sehr schöner Überblick über die empirischen Ergebnisse findet sich bei Menzel (2008).

² Litzenberger (2006), insbesondere S. 269 ff.

Urbanisierungsvorteile größeres Gewicht haben. Glaeser u.a. (1992) betonen die große Bedeutung der Lokalisationsvorteile, die auch als MAR-Spillovers bezeichnet werden, während Jacobs (1970, 1986) die Urbanisierungsvorteile betont (Jacobian-Spillovers).

Ähnlich unterschiedlich wie die Ergebnisse der empirischen Analyse fällt die Einschätzung aus, ob die Clusterförderung ein sinnvolles Instrument der regionalen Wirtschaftspolitik darstellt. So wird auf der einen Seite unter Verweis auf entsprechende empirische Ergebnisse eines positiven Zusammenhangs zwischen Clusterbildung und Regionalentwicklung eine Förderstrategie empfohlen, die Clusterbildungen unterstützt (Fritsch, Stephan, Werwatz 2004). Geeignete Instrumente hierfür seien die Errichtung von spezifischen Forschungseinrichtungen und die Förderung des entsprechenden Wissenstransfers in der Region, die Verbesserung der Qualifikation des benötigten Arbeitskräftepotenzials durch einen Ausbau des Bildungsangebots, die Initiierung von Netzwerken zwischen Unternehmen, Bildungs- und Forschungseinrichtungen, das Angebot spezifischer Dienstleistungen und Beratungen für Unternehmen, die dem zu fördernden Cluster angehören, sowie die spezifische Ansiedlung von Unternehmen, die das Cluster sinnvoll ergänzen. Eher neoklassisch ausgerichtete Autoren lehnen eine solche clusterorientierte Regionalpolitik ab und begründen dies vor allem mit zwei Argumenten (Alecke, Untiedt 2007). Zum einen sind die meisten Spillovers pekuniär und nicht technologisch, werden also im Marktprozess berücksichtigt und bedürfen keiner flankierenden staatlichen Maßnahmen. Zum anderen sind Cluster, sofern sie denn existieren, Ergebnisse von Marktprozessen; sie politisch zu antizipieren überfordert die Träger der Regionalpolitik und stellt eine „Anmaßung von Wissen“ dar.

Die folgenden Ausführungen haben ihren Schwerpunkt eindeutig im empirischen Teil der Clusterforschung. Es soll ein Verfahren vorgestellt werden, das nach Auffassung der Autoren den bisher zur Konzentrationsforschung genutzten statistischen Methoden überlegen ist und diese sinnvoll ergänzt. Hierzu werden zunächst die traditionellen Verfahren kurz vorgestellt, um anschließend auf alternative Instrumente einzugehen; in diesem Kontext sind die Vor- und Nachteile der einzelnen Verfahren darzustellen. Anschließend soll die K-Funktion als alternative Methode auf die räumliche Verteilung von Betrieben ausgewählter Sektoren der Industrie in Deutschland angewendet werden.

2. Traditionelle statistische Verfahren zur Analyse räumlicher (De-) Konzentrationsprozesse

2.1 Betrachtung eines einzelnen Sektors

Räumliche Konzentration wird traditionell mit Hilfe des Gini-Koeffizienten gemessen (beispielhaft Krugman 1991, S. 54 ff.)³, der sich aus einer Lorenzkurve ableitet (Eckey, Kosfeld, Türck 2008). Es sei V_i die Ausprägung der interessierenden Variable, etwa die Anzahl der Betriebe eines bestimmten Wirtschaftszweiges, in der Region i ($i = 1, 2, \dots, n$) und B_i die Ausprägung einer sinnvollen Bezugsgröße, so etwa die Fläche oder die Anzahl der Betriebe insgesamt, in der gleichen Region. Wir bilden $\frac{V_i}{B_i}$ für alle i und ordnen sie gemäß

ihrer Ausprägung, beginnend mit dem kleinsten und endend mit dem größten Wert. ($\frac{V_1}{B_1} < \frac{V_2}{B_2} < \dots < \frac{V_n}{B_n}$). Zur Zeichnung der Lorenzkurve werden die relativen Häufigkeiten H_k

der Bezugsgröße B sowie die kumulierten Merkmalsanteile Q_k der Merkmalsvariablen V benötigt; für den Punkt (H_k, Q_k) der Lorenzkurve gilt

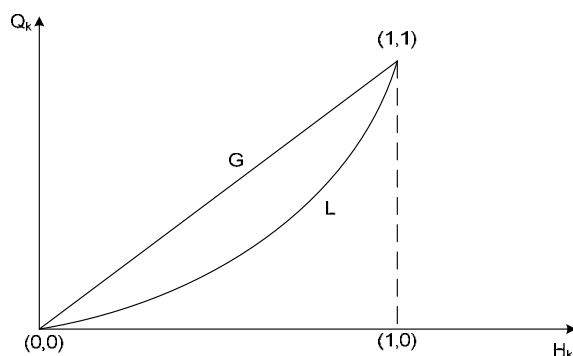
$$2.1-1 \quad H_k = \frac{\sum_{i=1}^k B_i}{\sum_{i=1}^n B_i}$$

$$Q_k = \frac{\sum_{i=1}^k V_i}{\sum_{i=1}^n V_i}$$

Verbinden wir die so gewonnenen Punkte für $k = 1, 2, \dots, n$, so erhalten wir die gesuchte Lorenzkurve L . Sie ist in Abb. 2.1-1 idealtypisch durch einen stetigen Kurvenzug wiedergegeben.

³ Weitere wichtige Konzentrationsmaße finden sich beispielsweise bei Bickenbach und Bode (2008): generalized entropy, Theil index, coefficient of variation, Relative Mean Deviation/Krugman.

Abbildung 2.1-1: Lorenzkurve und Gini-Koeffizient



Gäbe es keinerlei räumliche Korrelation, wäre also die interessierende Variable über alle Regionen gleichverteilt, so würde die Lorenzkurve mit der Gleichverteilungsgeraden G zusammenfallen. Damit bietet es sich an, die Fläche zwischen G und L, bezogen auf die Dreiecksfläche $((0,0), (1,0), (1,1))$, zum Maß der räumlichen Konzentration zu machen; sie heißt Gini-Koeffizient und kann Werte zwischen 0 (Gleichverteilung) und 1 (maximale räumliche Konzentration) annehmen. Mit Hilfe der folgenden Formel kann der Gini-Koeffizient berechnet werden:

$$2.1-2 \quad R = 1 - \sum_{k=1}^n (Q_{k-1} + Q_k) \cdot h_k, \text{ mit } Q_0 = 0 \text{ und } h_k = \frac{B_k}{\sum_{i=1}^n B_i}$$

Eine Weiterentwicklung des Gini-Koeffizienten stellt der Ellison-Glaeser-Index dar (Ellison, Glaeser 1997), der zunehmende Verbreitung findet (Alecke, Untiedt 2006). Mit seiner Hilfe ist es möglich, die räumliche Konzentration um betriebsinterne Agglomerationsvorteile zu bereinigen; Ergebnis ist eine Maßzahl für die räumliche Konzentration, die ausschließlich die betriebsexternen Konzentrationsursachen, also die Lokalisations- und Urbanisierungsvorteile zum Ausdruck bringt. Zur Erläuterung dieses Indexes sind zunächst einige Merkmale zu definieren.

Es ist

$$2.1-3 \quad s_i = \frac{V_i}{\sum_{i=1}^n V_i}$$

Anteil der Region i an der Ausprägung der Variablen insgesamt

$$2.1-4 \quad x_i = \frac{B_i}{\sum_{i=1}^n B_i}$$

Anteil der Region i an der Ausprägung der Bezugsgröße insgesamt

$$2.1-5 \quad \text{Herf} = \sum_{k=1}^m \left(\frac{V_k}{\sum_{k=1}^m V_k} \right)^2$$

k = Betrieb k (k = 1, 2... m mit m = Anzahl der Betriebe)

Herf gibt den Herfindahl-Index wieder. Er liegt im Intervall $1/m < \text{Herf} < 1$; je höher sein Wert, desto stärker die Konzentration der Variablen auf einen oder wenige Betriebe. Herf kann daher als Maß für betriebsinterne Größenvorteile interpretiert werden.

Es gilt nun, die Konzentration um betriebsinterne Größenvorteile zu korrigieren und so eine Kennziffer für betriebsexterne Agglomerationsvorteile zu erhalten. Hierzu ist das Maß der

Gesamtkonzentration $\sum_{i=1}^n (s_i - x_i)^2$ um den Herfindahl-Index zu bereinigen, was zum Index

von Ellison und Glaeser (1997) führt:

$$2.1-6 \quad \gamma = \frac{\sum_{i=1}^n (s_i - x_i)^2 - (1 - \sum_{i=1}^n x_i^2) \cdot \text{Herf}}{(1 - \sum_{i=1}^n x_i^2) \cdot (1 - \text{Herf})}$$

Der Erwartungswert des Ellison-Glaeser-Index γ ist gleich 0. Falls γ größer als 0 ist, liegen betriebsexterne Agglomerationsvorteile vor. Ein Wert von γ kleiner als 0 deutet auf betriebsexterne Agglomerationsnachteile hin. Da die Varianz von γ bestimmbar ist, lässt sich der Ellison-Glaeser-Index (im Gegensatz zum Gini-Koeffizienten) auf Signifikanz testen ($H_0: \gamma = 0$).

2.2 Betrachtung der Abhängigkeit zwischen Sektoren

Gegeben sind zwei Variable V_1 und V_2 , die in einer gegenseitigen räumlichen Abhängigkeit traditionell mit Hilfe des Korrelationskoeffizienten r von Bravais-Pearson berechnet werden.

$$2.2-1 \quad r_{1,2} = \frac{\sum_{i=1}^n (V_{1i} - \bar{V}_1) \cdot (V_{2i} - \bar{V}_2)}{\sqrt{\sum_{i=1}^n (V_{1i} - \bar{V}_1)^2 \cdot \sum_{i=1}^n (V_{2i} - \bar{V}_2)^2}}$$

Ein $r_{1,2} > 0$ (≈ 0 , < 0) deutet auf Anziehung (Unabhängigkeit, Abstoßung) der Sektoren hin.

Nehmen wir eine Unterscheidung nach der abhängigen Variablen V_1 und der unabhängigen Variablen V_2 vor, ist eine Regressionsanalyse sinnvoll; es gilt

$$2.2-2 \quad V_1 = b_0 + b_1 \cdot V_2 + \sum_{\ell=2}^p b_{\ell} \cdot X_{\ell} + u$$

mit b_{ℓ} ($\ell = 0, 1 \dots p$) = Regressionskoeffizienten

X_{ℓ} ($\ell = 2, 3 \dots p$) = übrige auf V_2 Einfluss nehmende Variable

u = Residuum

Geht man plausiblerweise davon aus, dass die Ausprägung von V_{1i} nicht nur von V_{2i} , sondern auch von der Ausprägung von V_2 in den Nachbarregionen abhängt, so ist der Schätzansatz in 2.2-2 um das Konzept der räumlichen Autokorrelation zu erweitern (Anselin 1988). Für diesen Ansatz ist die Gewichtungsmatrix W^* maßgeblich, die im einfachsten Fall auf einer binären 0/1-Matrix basiert.

$$2.2-3 \quad W_{ij}^* = \begin{cases} 1, & \text{falls die Regionen } i \text{ und } j \text{ eine gemeinsame Grenzen haben und } i \neq j \text{ ist} \\ 0 & \text{sonst} \end{cases}$$

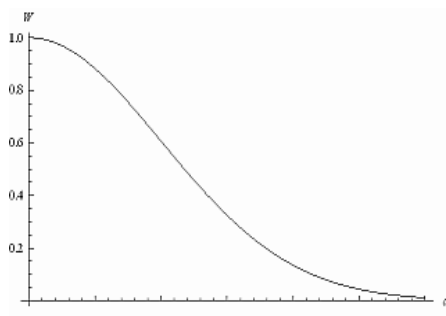
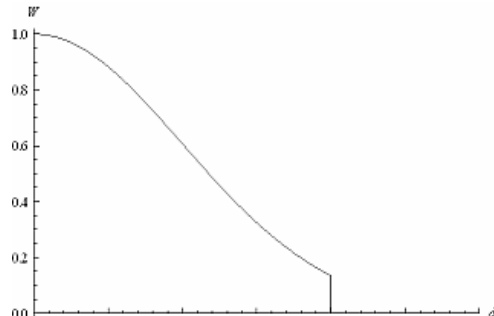
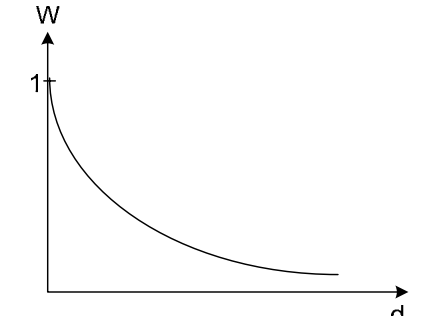
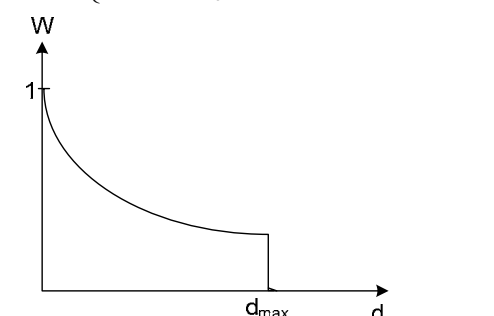
Diese Gewichtungsmatrix W^* wird in der Regel so normiert, dass die Summe der Zeilenelemente eins ergibt; sie hieße nun W . Dadurch wird 2.2-2 zu

$$2.2-4 \quad V_1 = b_0 + b_1 \cdot V_2 + b_2 \cdot W \cdot V_2 + \sum_{\ell=3}^p b_{\ell} \cdot X_{\ell} + u$$

b_2 zeigt an, ob von der Ausprägung von V_2 in den Nachbarregionen ein Einfluss auf V_1 in der betrachteten Region ausgeht.

Bauen wir die Gewichtungsmatrix W auf einer Distanzmatrix D auf, so lässt ihre Verwendung die Schätzung des Einflusses der Entfernung in regionalökonomischen Modellen zu. Bei der Bestimmung der Gewichte d_{ij} wird dabei unterschiedlich vorgegangen, wobei zwischen Decay- bzw. Bi-Square-Funktionen auf der einen Seite und e- bzw. gravitationstheoretischem Ansatz auf der anderen Seite unterschieden werden kann.

Abbildung 2.2-1: Unterschiedliche Ansätze zur Ermittlung von Gewichten in regionalökonomischen Modellen

	Decay-Funktion	Bi-Square-Funktion
e-Funktion	$W_{ij} = e^{-\frac{1}{2} \left(\frac{d_{ij}}{c} \right)^2}$ 	$W_{ij} = \begin{cases} e^{-\frac{1}{2} \left(\frac{d_{ij}}{c} \right)^2} & \text{für } d_{ij} < d_{max} \\ 0 & \text{für } d_{ij} > d_{max} \end{cases}$ 
Gravitations- theoretischer Ansatz	$W_{ij} = \frac{1}{d_{ij}^c}$ 	$W_{ij} = \begin{cases} \frac{1}{d_{ij}^c} & \text{für } d_{ij} < d_{max} \\ 0 & \text{für } d_{ij} > d_{max} \end{cases}$ 

Die so durchgeführte Ermittlung des Einflusses der Entfernung auf die Beziehung zwischen V_1 und V_2 lässt die Ermittlung von Halbwerteffekten und Elastizitäten zu. So ist die Halbwertzeit bei der e-Funktion $c \cdot \sqrt{\ln(4)}$, d.h. in dieser räumlichen Entfernung ist der Einfluss von V_2 auf V_1 nur noch halb so groß wie in unmittelbarer räumlicher Nähe. Die Elastizität $E_{w,d}$ ist $-\frac{d^2}{c^2}$, d.h. bei einer 1%igen Erhöhung von d nimmt w um $\frac{d^2}{c^2}$ % ab.

3. Die K-Funktion als alternatives Instrument zur Analyse räumlicher (De-)Konzentrationsprozesse

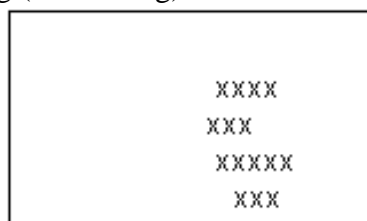
3.1 Die Abgrenzung von K-Funktionen gegenüber traditionellen Verfahren

Die oben geschilderten (und weitere) traditionellen Verfahren basieren auf einer Einteilung eines Gesamttraums, meist einer Volkswirtschaft, in Teilräume, die als Regionen bezeichnet werden. Sie zählen die Merkmalsausprägungen in Regionen und nutzen sie als Dateninputs für die genutzten Algorithmen. Sie können deshalb auch als raumbezogene Verfahren bezeichnet werden.

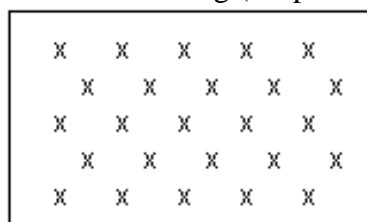
Dagegen ist die K-Funktion punktbezogen. Jede Merkmalsausprägung wird durch ihre Lage im Raum charakterisiert; in der Regel erfolgt diese Positionsbestimmung durch die Verwendung von Längen- und Breitengraden. Stellt man ihre Position im Raum dar, so ergeben sich grundsätzlich drei Möglichkeiten.

Abbildung 3.1-1: Grundsätzliche Möglichkeiten der Positionierung von Beobachtungseinheiten in der Fläche

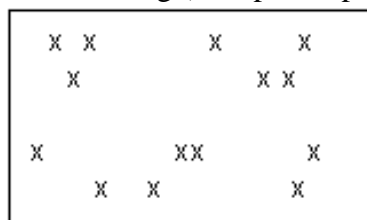
a) räumliche Ballung (Clusterung)



b) gleichmäßige räumliche Verteilung (Dispersion, Regularität)



c) Zufällige räumliche Verteilung (Complete Spatial Randomness (CSR))



In Abbildung 3.1-1 konzentrieren sich die Beobachtungseinheiten in a) auf einen kleinen Teil der Gesamtfläche; ein Beispiel sei die räumliche Verteilung von Industriebetrieben eines Sektors, der durch ausgeprägte Lokalisationsvorteile charakterisiert ist. Dagegen liegen in b)

alle Beobachtungseinheiten gleich weit von den nächsten Nachbarn entfernt; sie meiden räumliche Nähe. Man spricht hierbei von Regularität oder Dispersion. Ein regionalökonomisches Beispiel ist das Zentrale-Orte-Prinzip von Christaller.

Verteilen sich die Merkmalsträger zufällig über die Fläche (3.1-1 c)), wirken also keine Agglomerationsvorteile bzw. -nachteile, oder gleichen diese sich genau aus, so spricht man von einer zufälligen räumlichen Verteilung, im Englischen Complete Spatial Randomness (CSR) genannt. Die Ereignisdichte λ beim Vorliegen eines CSR-Prozesses ist wie folgt definiert: Es sei S der gesamte betrachtete Raum, also etwa eine Volkswirtschaft, $a(S)$ ihre Fläche und n die Anzahl der in ihr beobachteten Merkmalsträger, so ist die durchschnittliche Häufigkeit pro Flächeneinheit

$$3.1-1 \quad \lambda = \frac{n}{a(S)}$$

Im Falle einer zufälligen Verteilung über die Fläche gilt für jede Teilfläche von S, die wir C nennen wollen und die einer Region entsprechen kann, folgender Ausdruck:

$$3.1-2 \quad E(N(C)) = \frac{n}{a(S)} \cdot a(C) = \lambda \cdot a(C)$$

mit

E = Erwartungswert

N = Anzahl der Merkmalsträger in der Teilfläche C.

Die erwartete Anzahl der Merkmalsträger in C entspricht damit genau dem Anteil an n, den C an der Gesamtfläche von S hat.

Mit Hilfe der K-Funktionen ist es u. a. möglich, diese drei Fälle zu unterscheiden und das Ausmaß von Clusterung und Dispersion zu bestimmen (Univariate K-Funktionen). Dies gilt auch für die gleichzeitige Betrachtung von zwei Variablen, die sich zufällig zueinander verhalten können, sich aber auch gegenseitig anziehen (Attraction) oder abstoßen (Repulsion) können (Bivariate K-Funktionen). So mögen zwei Wirtschaftszweige, die gemeinsam Lokalisationsvorteile aufweisen, die räumliche Nähe zueinander suchen, während eine arbeitsintensive und lohnkostensensible Branche die Nähe eines mit hohen Entgelten verbundenen Sektors meidet.

Die K-Funktionen gehen auf Ripley (1977) zurück. Sie finden zunächst vor allem in der Biologie und Geographie Anwendung, so etwa bei der räumlichen Verteilung von Pflanzen.

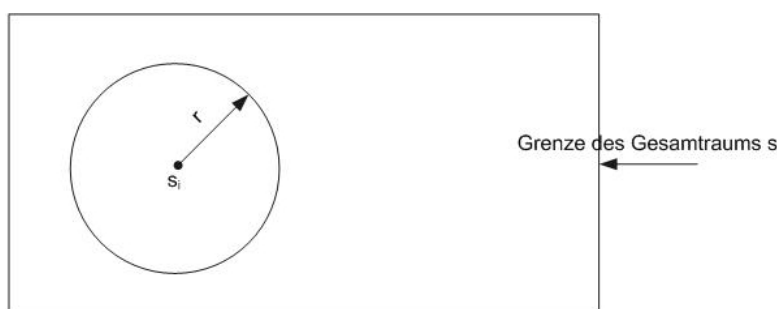
Die geografische Ballung von Industriebetrieben untersuchen Barff (1987) sowie Sweeney und Feser (1998). In die Ökonomie fanden die K-Funktionen zunächst Eingang durch Arbia und Espa (1996). Es folgten wenige weitere Studien, so durch Morcon und Puech (2003), Quah und Simpson (2003) sowie Duranton und Overman (2005), die sich auf die regionale Konzentration von Wirtschaftszweigen sowie die räumliche Verbreitung von Innovationen und damit verbundenen technischen Spillovers beschäftigen. Eine Analyse der Standorte von Patenten in verschiedenen Wirtschaftszweigen Italiens nehmen Arbia, Espa und Quah (2007) vor.

3.2 Darstellung der K-Funktion⁴

3.2.1 Univariate K-Funktion: Ballung versus Streuung

Betrachten wir die geographische Verteilung einer Variablen, so lässt sich ihre räumliche Konzentration mit Hilfe der univariaten K-Funktion darstellen. Die Standorte dieser Variablen lokalisieren sich in den Raumpunkten s_i ($i = 1, 2, \dots, n$), die durch geographische Koordinaten eindeutig bestimmt sind. Zur Ableitung und Definition der univariaten K-Funktion betrachten wir beispielhaft einen dieser Punkte s_i .

Abbildung 3.2.1-1: Beispiel zur Ableitung der univariaten K-Funktion



Wir schlagen um den Raumpunkt s_i einen Kreis mit dem Radius r . Bei zufälliger Verteilung der Standorte im Gesamttraum s erwarten wir $n \cdot \frac{r^2 \cdot \pi}{a(s)} = \lambda \cdot r^2 \cdot \pi$ Raumpunkte. Von dieser erwartenden Anzahl ist die wirkliche Menge der Punkte im Kreis zu unterscheiden. Sie ist

$$3.2.1-1 \quad \sum_{j=1}^n I_r(d_{ij}) \quad \text{für } i \neq j$$

⁴ Eine empfehlenswerte Darstellung der K-Funktion findet sich u.a. bei Bailey/Gatrell (1995), S. 73 ff.

$$\text{für } I_r(d_{ij}) = \begin{cases} 1 & \text{für } d_{ij} \leq r \\ 0 & \text{für } d_{ij} > r \end{cases}$$

Liegt ein anderer Punkt innerhalb des Kreises, ist also die Entfernung zwischen s_i und s_j , die wir mit d_{ij} bezeichnen, kleiner gleich r , so liegt s_j innerhalb des Kreises und erhält eine 1; bei einer Lage außerhalb des Kreises wird eine 0 zugeteilt. Diese Berechnung führen wir für alle Raumpunkte s_i durch und erhalten schließlich die durchschnittliche Anzahl weiterer Raumpunkte innerhalb des Radius r mit Hilfe der Beziehung

$$3.2.1-2 \quad \frac{\sum_{i=1}^n \left(\sum_{j=1}^n I_r(d_{ij}) \text{ für } i \neq j \right)}{n}$$

Damit ist bereits ein Einblick in den Grad der räumlichen Konzentration möglich. Liegen wenige (viele) Punkte in der Nachbarschaft von s_i , so deutet dies auf Dispersion (Clustering) hin. Es gilt also

$$\lambda \cdot r^2 \cdot \pi > 3.2.1-2 \quad \text{Dispersion}$$

$$\lambda \cdot r^2 \cdot \pi \approx 3.2.1-2 \quad \text{zufällige räumliche Verteilung}$$

$$\lambda \cdot r^2 \cdot \pi < 3.2.1-2 \quad \text{Clustering}$$

Diese ersten Beziehungen sind aufgrund von zwei Überlegungen zu modifizieren. Würde man die obigen Formeln unverändert lassen, so wären die Werte zwischen zwei Variablen mit $n_1 \neq n_2$ nicht vergleichbar; ist $n_1 > n_2$ und hat die Variable V_1 bei einem Radius r einen höheren vorläufigen K-Wert 3.2.1-2 als die Variable V_2 , so weiß man nicht, ob dies Ausdruck einer höheren Konzentration ist, oder ob dies einfach an der größeren Anzahl der Standorte von V_1 liegt. Um den zweiten Effekt auszuschalten, erweitern wir 3.2.1-2 um $\frac{1}{\lambda}$ und erhalten eine modifizierte Form der univariaten K-Funktion, die ausschließlich auf räumlicher (De-)Konzentration basiert.

$$3.2.1-3 \quad K(r) = \frac{1}{\lambda \cdot n} \sum_{i=1}^n \left(\sum_{j=1}^n I_r(d_{ij}) \text{ für } i \neq j \right)$$

Damit ist die univariate K-Funktion definiert als die durchschnittliche Anzahl weiterer Punkte innerhalb eines Radius r um einen Standort s_i , dividiert durch die Dichte aller Punkte.

Die zweite Überlegung bezieht sich auf notwendige Randkorrekturen, die wiederum mit Hilfe eines Beispiels erläutert werden sollen.

Abbildung 3.2.1-2 Beispiel für die Notwendigkeit einer Randkorrektur



Nicht die gesamte Fläche des Kreises mit einem Radius von r um s_i liegt innerhalb der Grenze des Gesamttraums. Die Gesamtfläche $F = r^2 \cdot \pi$ lässt sich in die Teilflächen F_1 und F_2 aufteilen; F_1 (F_2) liegt innerhalb (außerhalb) der Grenze des Gesamttraums. Da in F_2 keine Standorte betrachtet werden, würde ohne Randkorrektur die Anzahl der erwarteten im Vergleich zu der Anzahl der realisierten Werte pro Flächeneinheit überschätzt. Die Anzahl der realisierten Werte wird daher in der Regel mit $\frac{F}{F_1}$ multipliziert und damit vergrößert, um dem Randeffect

Rechnung zu tragen; diese Vorgehensweise wird auch als Ripley-Korrektur bezeichnet.

Ist $K(r)$ empirisch geschätzt ($=\hat{K}(r)$), so ergibt ein Vergleich mit der K -Verteilung der zufälligen räumlichen Verteilung K_Z^5 Hinweise auf den Grad von Clusterung bzw. Dispersion bei unterschiedlichen Radien r ; da die empirischen K -Funktionen um die Dichte bereinigt werden, ist der konstante Erwartungswert der zufälligen Verteilung $E(K_Z(r)) = \pi \cdot r^2$.

⁵ Wir unterstellen hier eine homogene Verteilung, also Stationarität des räumlichen Prozesses, wie wir ihn in 3.1-2 definiert haben. Bei einem inhomogenen Prozess ist λ nicht für alle Raumpunkte gleich, sondern divergiert zwischen den Raumpunkten. Unterscheidet sich etwa die Standortqualität zwischen den Raumpunkten aufgrund natürlicher (z.B. Topographie) und geschaffener (z.B. Infrastruktur) Standortfaktoren, so ergeben sich zufällige und berechnete Verteilung aufgrund einer modifizierten K -Funktion. Bisher war sie gemäß 3.2.1-3

$$K(r) = \frac{1}{\lambda \cdot n} \cdot \sum_{i=1}^n \left(\sum_{j=1}^n I_r(d_{ij}) \text{ für } i \neq j \right) = \frac{1}{a(S)} \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^n I_r(d_{ij}) \text{ für } i \neq j \right)}{\lambda^2}$$

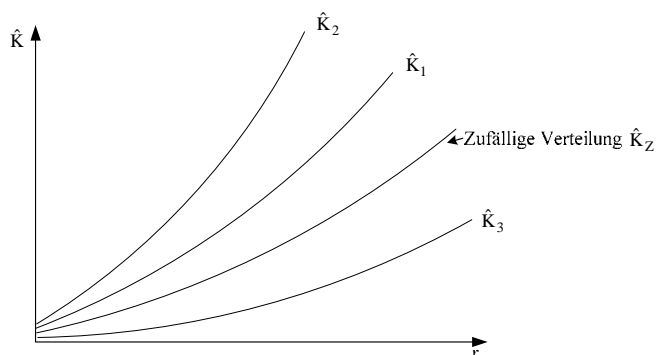
wegen $\lambda = \frac{n}{a(S)}$, wie in 3.1-1 erläutert. Liegt ein inhomogener Prozess vor, so unterscheiden sich die Dichten λ

zwischen den Raumpunkten, so dass nun

$$K(r) = \frac{1}{a(S)} \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^n I_r(d_{ij}) \text{ für } i \neq j \right)}{\lambda(s_i) \cdot \lambda(s_j)}$$

gilt. Da keine räumlich differenzierten Dichten vorliegen, gehen wir von einem stationären räumlichen Prozess aus und schätzen $K(r)$ mit Hilfe von 3.2.1-3.

Abbildung 3.2.1-3: Beispiele für univariate K-Funktionen



\hat{K}_1 liegt oberhalb der zufälligen Verteilung und deutet damit auf eine Clusterung der betrachteten Variablen V_1 bei allen r hin. Die Variable V_2 ist noch stärker räumlich konzentriert, da \hat{K}_2 für alle r oberhalb von \hat{K}_1 liegt. V_3 ist dagegen relativ gleichmäßig über die Fläche verteilt, da \hat{K}_3 für alle r unterhalb der Bezugslinie liegt.

Ob die Unterschiede der geschätzten K-Funktionen zur zufälligen Verteilung statistisch signifikant sind, lässt sich mit Hilfe einer Reihe von Testverfahren überprüfen (Bailey, Gatrell 1995, S. 90). Besonders bekannt ist der Clark-Evans-Test, der auf dem Prinzip des nächsten Nachbarn basiert. Liegen die nächsten Nachbarn näher (weiter) entfernt, als dies bei einer zufälligen Verteilung zu erwarten wäre, deutet dies auf Clusterung (Dispersion) hin. Wir berechnen für alle Raumpunkte s_i die durchschnittliche Entfernung zum nächsten Nachbarn und bezeichnen sie mit \bar{D}_m . Bei zufälliger Verteilung ist der Erwartungswert $E(\bar{D}) = \frac{1}{2\sqrt{\lambda}}$

und die Varianz $\text{VAR}(\bar{D}) = \frac{4 - \pi}{4n\lambda\pi}$. Die standardisierte Variable

$$3.2.1-4 \quad Z_m = \frac{\bar{D}_m - E(\bar{D})}{\sqrt{\text{VAR}(\bar{D})}}$$

konvergiert gegen eine Standardnormalverteilung $N(0, 1)$, die das Ablesen der kritischen Werte erlaubt.

Von besonderem Interesse ist auch die Frage, ob sich räumliche (De-)Konzentrationen von zwei Variablen signifikant voneinander unterscheiden, so etwa \hat{K}_1 und \hat{K}_2 in Abbildung

3.2.1-3. Der hierzu gebräuchliche Test⁶ basiert auf folgender Überlegung: Wir berechnen zunächst die univariaten K-Funktionen für V_1 und V_2 und bilden deren Differenz.

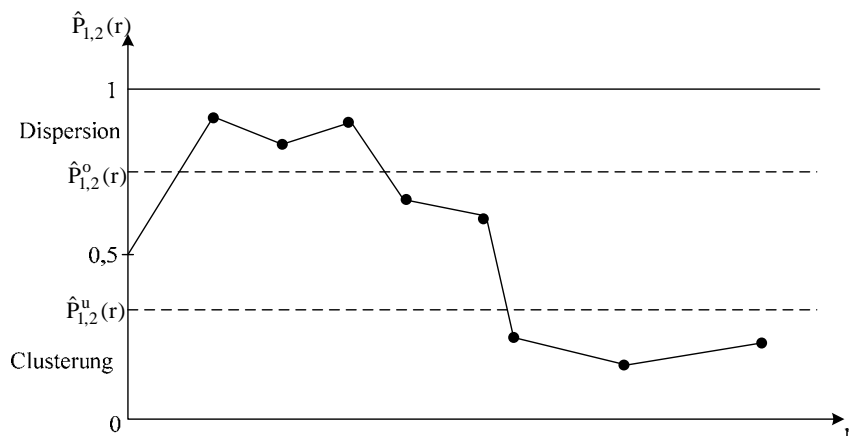
$$3.2.1-5 \quad \Delta^0(r) = \hat{K}_1(r) - \hat{K}_2(r)$$

Anschließend bilden wir aus beiden Standortgruppierungen, bestehend aus n_1 und n_2 Elementen, eine Gruppe, die nun aus $n_1 + n_2$ Elementen besteht. Diese Elemente werden jetzt in eine zufällige Reihenfolge gebracht, wobei anschließend die ersten n_1 Elemente der Gruppe 1 und die folgenden n_2 Elemente der Gruppe 2 zugeordnet werden. Für jedes r wird nun festgelegt, ob die Differenz der wirklichen Verteilung kleiner ist als jene der zufälligen Verteilungen; in diesem Fall wird ℓ um eine Einheit erhöht. Nach insgesamt m Simulationen bilden wir die Relation

$$3.2.1-6 \quad \hat{P}_{1,2}(r) = \frac{\ell(r) + 1}{m + 1}$$

Der Erwartungswert von $P_{1,2}(r)$ bei einem gleichen Grad der Clusterung von V_1 und V_2 ist 0,5; ein Vergleich mit dem Berechneten $\hat{P}_{1,2}(r)$ lässt dann erkennen, ob hinsichtlich der (De-)Konzentration signifikante Unterschiede zwischen V_1 und V_2 bestehen.⁷

Abbildung 3.2.1-4: Beispiel für Dispersions- und Clustereffekte



Wir erhalten drei relevante Intervalle:

⁶ Er umfasst auch den Fall einer inhomogenen Verteilung.

⁷ Für r gegen 0 strebt $\hat{P}_{1,2}(r)$ gegen 0,5, da bei beide Variablen keine Nachbarn innerhalb des Radius r liegen. Die Graphen beginnen daher stets für $r=0$ bei 0,5.

Wir bilden zunächst ein Konfidenzintervall ($\alpha = 95\%$) um $E(P_{1,2}(r)) = 0,5$. Bewegt sich $\hat{P}_{1,2}(r)$ zwischen $\hat{P}_{1,2}^u(r)$ und $\hat{P}_{1,2}^o(r)$, so ergeben sich zwischen V_1 und V_2 keine signifikanten Unterschiede im Hinblick auf den Konzentrationsgrad. Dies ist im obigen Beispiel bei einem mittleren r der Fall.

Liegt dagegen $\hat{P}_{1,2}(r)$ zwischen $\hat{P}_{1,2}^o(r)$ und 1, so ist V_1 gleichmäßiger verteilt als V_2 , hier bei relativ kleinem r . Umgekehrt die Interpretation bei $0 \leq \hat{P}_{1,2}(r) \leq \hat{P}_{1,2}^u(r)$; sie deutet darauf hin, dass V_1 stärker geclustert ist als V_2 , so im obigen Beispiel bei relativ großem r .⁸

3.2.2 Bivariate K-Funktion: Anziehung versus Abstoßung

Bei der bivariaten K-Funktion wird die räumliche Abhängigkeit von zwei Variablen V_1 und V_2 gemessen. In Analogie zur univariaten K-Funktion lässt sich die bivariate K-Funktion als die durchschnittliche Anzahl von Raumpunkten der Variablen 1 innerhalb eines Radius r um einen Standort der Variablen 2, dividiert durch die Dichte der Punkte der Variablen 1, definieren. Hat die Variable 1 n_1 Raumpunkte s_i ($i = 1, 2, \dots, n_1$) und die Variable 2 n_2 Raumpunkte s_j ($j = 1, 2, \dots, n_2$), so gilt für die bivariate K-Funktion $\hat{K}_{12}(r)$ ⁹

$$3.2.2-1 \quad \hat{K}_{12}(r) = \frac{a(S)}{n_1 \cdot n_2} \cdot \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} I_r(d_{ij})$$

$$= \frac{1}{\lambda_2 \cdot n_1} \cdot \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} I_r(d_{ij})$$

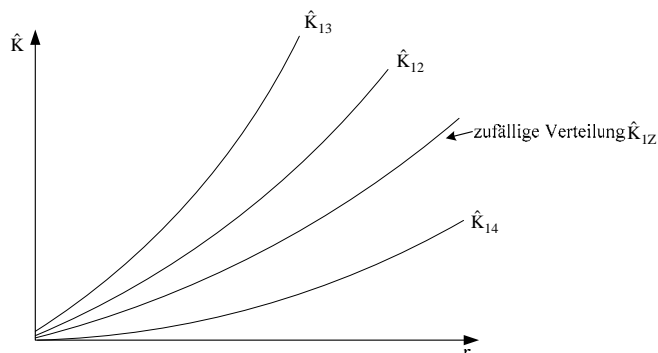
Auch hier ist eine Randkorrektur durchzuführen.

Die Interpretation der Ergebnisse ist ähnlich wie bei der univariaten Darstellung in Abbildung 3.2.1-3.

⁸ Dieses Testverfahren kann in Analogie auch angewendet werden, wenn man die statistische Signifikanz von \hat{K}_1 gegenüber \hat{K}_2 (Test auf homogenes Clustering) bzw. gegenüber einem nicht stationären Prozess (Test auf inhomogenes Clustering) ermitteln will. Im ersteren Fall besteht die Vergleichsvariable aus einem zufällig erzeugten Punktemuster, im zweiten Fall z.B. aus den Punkten einer übergeordneten Variablen, etwa den Gesamtbeschäftigten (im Vergleich zu der Menge der Beschäftigten in einem Sektor).

⁹ Auch hier ist in Analogie zur Fußnote 24 in Kapitel 3.2.1 ein Arbeiten mit einer inhomogenen Funktion möglich.

Abbildung 3.2.2-1: Beispiele für bivariate K-Funktionen



Bezugspunkt ist hier eine bivariate K-Funktion, bei der die zweite Variable zufällig verteilt ist; wir nennen sie \hat{K}_{1Z} . Sie entspricht der zufälligen Verteilung bei der univariaten K-Funktion, da auch hier bei Unabhängigkeit der beiden Variablen ein $K_{ij}(r) = \pi \cdot r^2$ erwartet wird; es gilt $E(K_{1Z}(r)) = E(K_Z) = \pi \cdot r^2$. Da \hat{K}_{12} für alle r oberhalb von \hat{K}_{1Z} liegt, sind die Variablen 1 und 2 positiv miteinander verknüpft; sie ziehen sich an. Dabei ist die Anziehungswirkung zwischen V_1 und V_3 noch größer als die zwischen V_1 und V_2 , da \hat{K}_{13} für alle r oberhalb von \hat{K}_{12} liegt. Dagegen stoßen sich die Variablen 1 und 4 in ihrer räumlichen Verteilung ab, erkenntlich daran, dass \hat{K}_{14} unterhalb von \hat{K}_{1Z} liegt.

Sind nun die Anziehungs- und Abstoßungseffekte statistisch signifikant? Diese Überprüfung kann sich auf drei relevante Fragestellungen beziehen:

- Ziehen sich die zwei Variablen innerhalb ihrer gegebenen geographischen Verteilung gegenseitig an oder stoßen sie sich ab?
- Besteht zwischen den Variablen 1 und 2 ein engerer Zusammenhang als zwischen der Variablen 1 und einer übergeordneten Größe, also etwa zwischen den Betrieben des Sektors 1 und den Betrieben des Sektors 2 im Vergleich zu den Betrieben des Sektors 1 und den Betrieben des Verarbeitenden Gewerbes insgesamt?
- Weisen die Variablen 1 und 2 einen engeren Zusammenhang auf als bei einer rein zufälligen Verteilung?

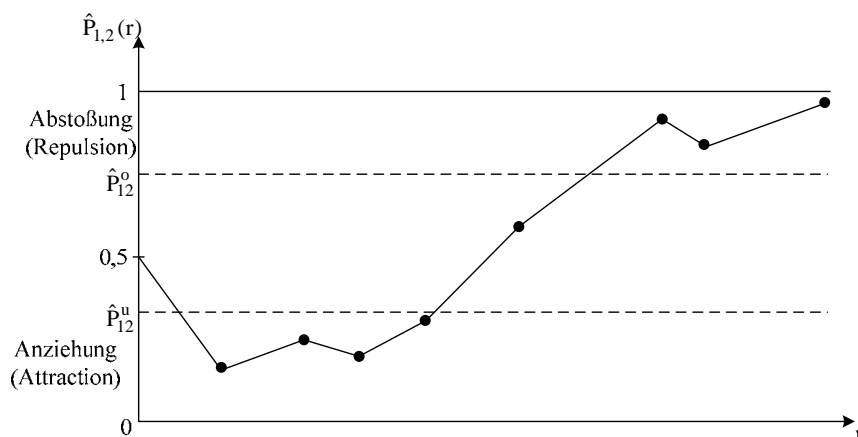
Die Fragestellung bestimmt die Durchführung des Tests. Im Fall a) gehen wir bei der Schätzung \hat{K}_{12}^Z nicht nur von der gegebenen Verteilung der Variablen 1, sondern auch von

der gegebenen Verteilung der Variablen 2 aus. Im Fall b) unterstellen wir für die Variable 2 eine geographische Verteilung, wie sie die übergeordnete Variable aufweist und im Fall c) eine zufällige Verteilung. Die Frage nach Anziehung und Abstoßung beantworten wir nun, indem wir in Analogie zu jenen Ausführungen vorgehen, die zur Abbildung 3.2.1-4 geführt haben. Ist $\hat{K}_{12}^o(r)$ der gemessene Wert der bivariaten K-Funktion beim Radius r , so berechnen wir darüber hinaus m -mal einen bivariaten Wert bei zufälliger Aufteilung der insgesamt $n_1 + n_2$ Merkmalsträger auf zwei Gruppen mit n_1 und n_2 Elementen ($\hat{K}_{12}^Z(r)$). Ist $\hat{K}_{12}^o < \hat{K}_{12}^Z$, so setzen wir $\ell = \ell + 1$ und erhalten schließlich

$$3.2.2-2 \quad \hat{P}_{12}(r) = \frac{\ell(r) + 1}{m + 1}$$

Ist $\hat{P}_{12}(r) > 0,5$ ($< 0,5$), so deutet dies auf Abstoßung (Anziehung) hin.

Abbildung 3.2.2-2: Beispiel für Abstoßungs- und Anziehungseffekte



Hier ziehen sich die Variablen 1 und 2 bei kleinem r an, sind bei mittlerem r weitgehend unabhängig voneinander, um sich bei großem r abzustößeln.

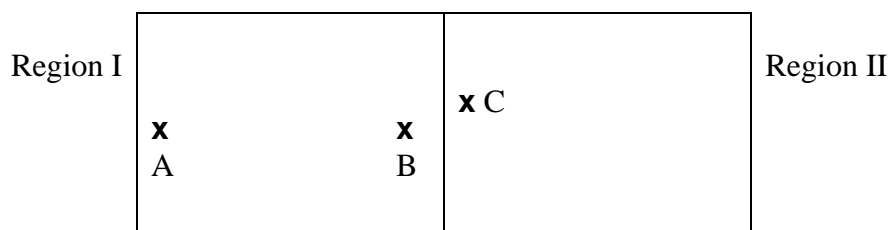
3.3 Vorteile der K-Funktionen gegenüber traditionellen raumbezogenen Verfahren zur Konzentrationsmessung

Die K-Funktionen haben gegenüber den traditionellen raumbezogenen Verfahren zur Konzentrationsmessung eine Reihe von Vorteilen, auf die im Folgenden eingegangen werden soll (Arbia, Espa, Quah 2007, S. 82).¹⁰

Die traditionellen Verfahren verlangen eine Einteilung eines Gesamttraumes in Teilregionen (Regionalisierung); dabei entsteht das Regionalisierungs- oder areal-unit-Problem. Je nach gewählter Unterteilung werden die Ergebnisse der empirischen Analyse unterschiedlich ausfallen. dies gilt sowohl für das Abgrenzungsprinzip (homogen oder funktional, analytisch oder programmatisch) als auch für die gewählte kleinste regionale Ebene (Gemeinden, Stadt- und Landkreise, Regierungsbezirke, Bundesländer) (Eckey 2008, S. 90 ff.). Da bei der Verwendung von K-Funktionen die Untersuchungseinheiten nicht über die Zugehörigkeit zu Wirtschaftsräumen, sondern über ihre geographischen Koordinaten definiert sind, tritt hier das areal-unit-Problem nicht auf.

Die Lage innerhalb eines Wirtschaftsraumes spielt im Rahmen der traditionellen Verfahren keine Rolle; die Zugehörigkeit entscheidet sich über eine ja-nein-Entscheidung. Dies kann zu Verzerrungen führen, wie ein kleines Beispiel zeigen soll.

Abbildung 3.3-1: Regionale Abgrenzung und Entfernung zwischen Raumpunkten



A und B liegen in der gleichen Region I, C aber in der Region II. Die traditionellen Verfahren vernachlässigen die Entfernungen zwischen diesen Raumpunkten und unterstellen für A und B, dass sie räumlich identisch sind, während C sich räumlich von ihnen unterscheidet; dies, obwohl B und C wesentlich enger beieinander liegen als A und B. Die K-Funktionen berücksichtigen dagegen die wirklichen Raumentfernungen d_{AB} und d_{BC} .

Regionale Untersuchungen berücksichtigen in der Regel nicht die stark divergierende Größe der Raumeinheiten, die z.B. auf der Ebene der Stadt- und Landkreise in Deutschland etwa 1:100 beträgt. Oder umgekehrt formuliert: Interessiert man sich z.B. für die räumliche

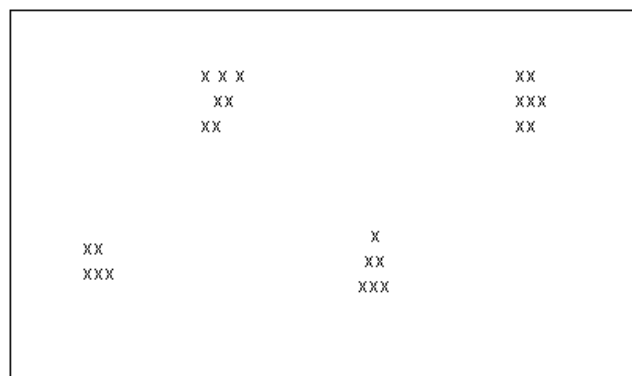
¹⁰ Als Alternative zur Nutzung von K-Funktionen werden auch Erweiterungen der herkömmlichen Indizes vorgeschlagen (Bickenbach, Bode 2008, S. 377).

Verteilung von Betrieben, so würde die Standortwahl einer Arbeitsstätte zugunsten der kleinsten Region das Ergebnis etwa 100-mal so stark beeinflussen wie die Standortwahl zugunsten des größten Kreises. Dagegen werden bei den K-Funktionen alle Standortentscheidungen gleich gewichtet.

Am schwersten wiegt aber das Argument, dass K-Funktionen einen wesentlich größeren Informationsgehalt haben als die traditionellen Verfahren. Während letztere nur eine Kennziffer zum Beleg für (De-)Konzentration ausweisen, stellen die K-Funktionen die Konzentration in Abhängigkeit von der Entfernung dar. Hierdurch entsteht eine Informationsfülle, die traditionelle Verfahren nicht bieten können. Betrachten wir hierzu einige Beispiele:

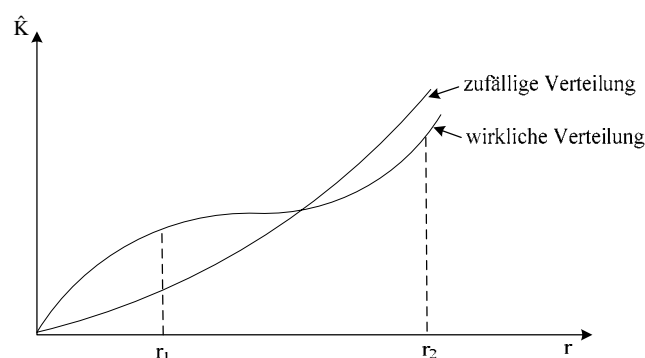
- a) Eine Variable sei kleinräumig stark konzentriert, dagegen großräumig dispers verteilt. Zu denken ist etwa an das Zentrale-Orte-Prinzip von Christaller, nach dem Betriebe des Dienstleistungsgewerbes das Zentrum der Kernstadt bevorzugen, während zwischen den Zentralen Orten eine große und gleichmäßige Entfernung besteht. Die räumliche Verteilung möge folgendes Aussehen haben:

Abbildung 3.3-2: Gleichmäßig verteilte kleinräumige Cluster



Die entsprechende K-Funktion hat dann folgendes Aussehen:

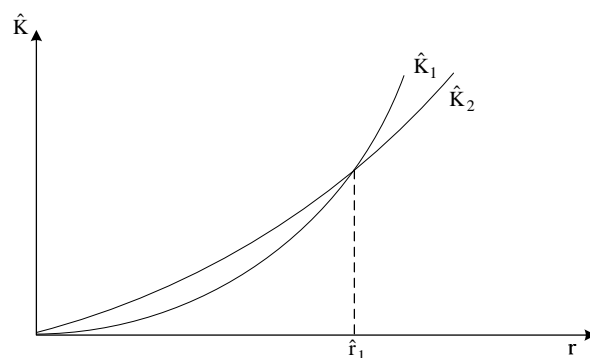
Abbildung 3.3-3: K-Funktionen bei gleichmäßig verteilten kleinräumigen Clustern



Sie weist bei kleiner Entfernung r_1 auf Konzentration, bei einer großen Entfernung r_2 dagegen auf Dekonzentration hin. Eine solche differenzierte Betrachtung ist mit Hilfe traditioneller Verfahren nicht möglich, die in diesem Beispiel eher eine durchschnittliche Clusterung ausweisen würden.

- b) Ein weiteres Beispiel bezieht sich auf die univariate K-Funktion der räumlichen Verteilung von Betrieben zweier Wirtschaftszweige, die wir mit \hat{K}_1 und \hat{K}_2 bezeichnen wollen.

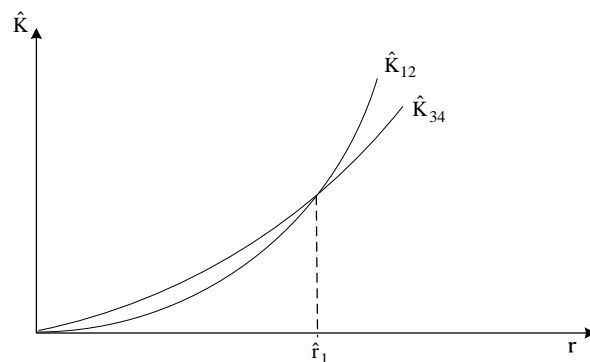
Abbildung 3.3-4: K-Funktionen für zwei Sektoren



Für $r < \hat{r}_1$ weist der Sektor 2 eine höhere Clusterung auf als der Sektor 1; ab $r > \hat{r}_1$ kehrt sich dieses Verhältnis um. Eine solche unterschiedliche Reihung ist bei der Verwendung traditioneller Verfahren nicht zu ermitteln.

- c) Ähnliche Vorteile zeigen sich bei der Nutzung bivariater K-Funktionen. Wir unterstellen beispielhaft folgende Verläufe:

Abbildung 3.3-5: Bivariate K-Funktionen für vier Sektoren



Bis zu $r < \hat{r}_1$ zeigen die Betriebe der Sektoren 3 und 4 einen engeren räumlichen Zusammenhang als die Sektoren 1 und 2; ab $r > \hat{r}_1$ kehrt sich dieser Zusammenhang um. Auch das Aufdecken eines solchen Wechsels überfordert die traditionellen Verfahren.

4. Darstellung der empirischen Ergebnisse

4.1 Verwendetes Datenmaterial

Wir wollen im Folgenden die räumliche Konzentration der Standortwahl von Betrieben in unterschiedlichen Sektoren und deren wechselseitige Abhängigkeit überprüfen. Da eine solche Überprüfung nur für Betriebe Sinn macht, die in ihrer räumlichen Standortwahl frei sind („footless industries“), beziehen wir uns in unserer Analyse auf Betriebe der Industrie und nutzen hierzu die Ergebnisse des Monatsberichts für Betriebe im Verarbeitenden Gewerbe sowie im Bergbau und der Gewinnung von Steinen und Erden, bei dem Daten über sämtliche Betriebe dieses Bereiches mit mindestens 20 tätigen Personen erhoben werden. „Als Betrieb gilt jede örtlich abgegrenzte Produktionseinheit einschließlich der in ihrer unmittelbaren Umgebung liegenden und von ihr abhängigen Einheiten. Art und Anzahl der fachlichen Betriebsteile ergeben sich aufgrund der vom Betrieb ausgeübten Produktions- und sonstigen Wirtschaftstätigkeiten. Maßgebend für die Zuordnung zum Wirtschaftsbereich Verarbeitendes Gewerbe sowie Bergbau und Gewinnung von Steinen und Erden ist die „Klassifikation der Wirtschaftszweige, Ausgabe 2003“ (WZ 2003)“ (Statistisches Bundesamt), die auf der EG-Verordnung zur verbindlichen Einführung der Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft basiert. Diese Statistik ist für die hier verfolgte Fragestellung der räumlichen Konzentration von Betrieben der „footless industries“ am besten geeignet, da sie

- Betriebe ausweist und damit unternehmerische Standortentscheidungen nachvollziehbar macht.
- die hier interessierende Herstellung von Gütern umfassend wiedergibt.
- eine weite sektorale Differenzierung zulässt.
- die Ergebnisse in tiefer räumlicher Gliederung erfasst und publiziert.
- die Bildung von Zeitreihen zulässt und so nicht nur statische, sondern auch dynamische Prozesse ausgewertet werden können.

Allerdings müssen auch Nachteile in Kauf genommen werden. Ein erster Nachteil liegt darin, dass die räumliche Zuordnung nicht aufgrund des Ausweises geographischer Koordinaten erfolgt, wie von den K-Verfahren eigentlich gefordert,¹¹ sondern lediglich eine Zuordnung zu räumlichen Einheiten wie Bundesländern, Kreisen und Gemeinden erfolgt. Hier wird unterstellt, dass sich die Betriebe eines Wirtschaftszweiges in den Stadt- und Landkreisen zufällig verteilen. Dies bedeutet für die ausgewiesenen Ergebnisse, dass sie bei geringen Distanzen von wenigen Kilometern leichte Ungenauigkeiten aufweisen. Ein weiteres Problem besteht darin, dass kleine Betriebe mit wenigen Beschäftigten nicht erfasst sind; da aber vor allem größere Betriebe überregionale Standortentscheidungen treffen und sie mehr als 90 % der Gesamtbeschäftigten umfassen, erscheint dieser Nachteil nicht besonders gravierend. Schließlich ergaben sich in der amtlichen Statistik für die Stadt- und Landkreise Baden-Württembergs Lücken, da für sie für einige Wirtschaftsabteilungen keine Ergebnisse ausgewiesen werden. Diese wenigen Lücken werden durch eigene Schätzungen gefüllt.

4.2 Ergebnisse unter Verwendung univariater K-Verfahren¹²

Wir betrachten zunächst beispielhaft den Wirtschaftszweig „Metallerzeugung und -bearbeitung (DJ 27)“, dessen Betriebe in ihrer räumlichen Verteilung in Deutschland in Abbildung 4.2-1 wiedergegeben sind.

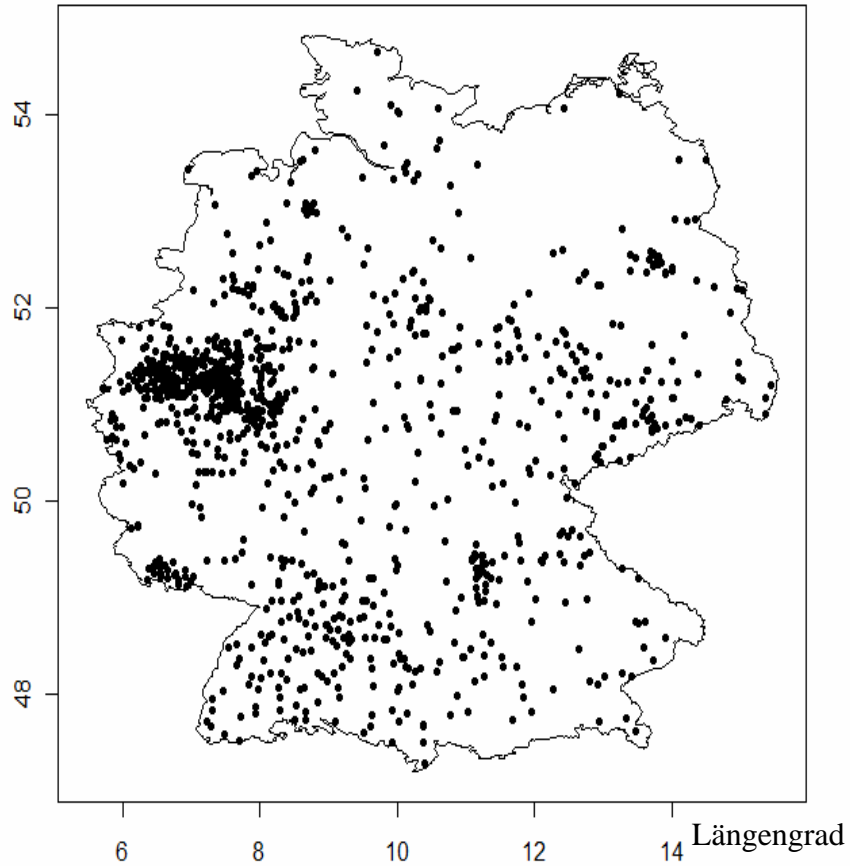
¹¹ Ein solcher Ausweis geographischer Daten ist im Unternehmensregister vorgesehen, so dass sich in Zukunft eine Verwendung der dort erhobenen Daten empfiehlt (vgl. Statistisches Bundesamt, Sturm/Tümmler (2006), insbesondere S. 1030 f.)

¹² Die Berechnungen der folgenden Kapitel wurden mit den „R“-Programmpaketen „Splancs“ und „Spatstat“ sowie eigenen Erweiterungen dieser Software durchgeführt.

Abbildung 4.2-1: Räumliche Verteilung der Betriebe des Sektors „Metallerzeugung und -bearbeitung“ in Deutschland 2006

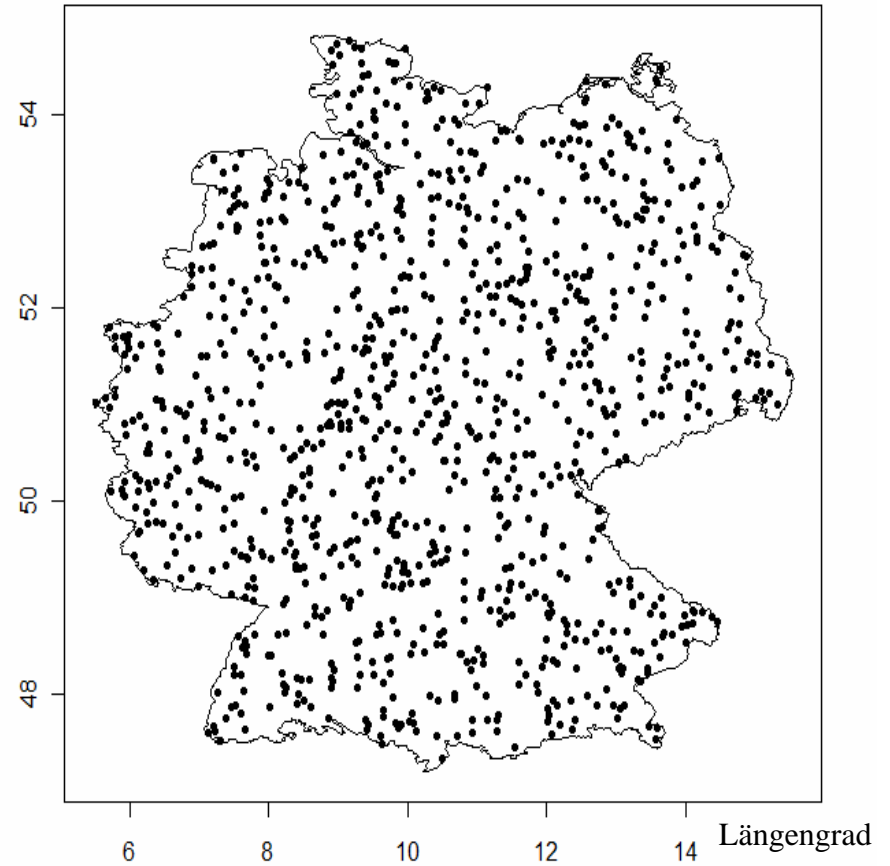
a) Wirkliche Verteilung

Breitengrad



b) Beispiel für eine zufällige Verteilung

Breitengrad



Gegenüber einer rein zufälligen Verteilung in Abb. 4.2-1 b) zeigt sich eine deutliche Clusterung der Betriebe vor allem im weiteren Ruhrgebiet, aber auch im Saarland und den Großräumen Berlin, Chemnitz, Nürnberg und Stuttgart. Die Existenz räumlicher Konzentrationen ist evident und wird durch den Clark-Evans-Test auf räumliche Unabhängigkeit bestätigt. Wir berechnen $\hat{\lambda} = \frac{n}{\text{Area}} = \frac{1068}{50,290^{34}} = 21,237$; pro Flächeneinheit

sind damit im Durchschnitt gut 21 Betriebe ausgewiesen. Damit ist die erwartete Distanz zum nächsten Nachbar $E(\bar{D}) = \frac{1}{2\sqrt{\lambda}} = \frac{1}{2\sqrt{21,237}} = 0,108$ und ihre Varianz

$\text{VAR}(\bar{D}) = \frac{4 - \pi}{4 \cdot 1068 \cdot 21237 \cdot \pi} = 0,00000301$. Da die durchschnittliche Entfernung zum nächs-

ten Nachbarn in der Realität 0,087 beträgt, ist damit der gesuchte Testwert gemäß 3.2.1-3

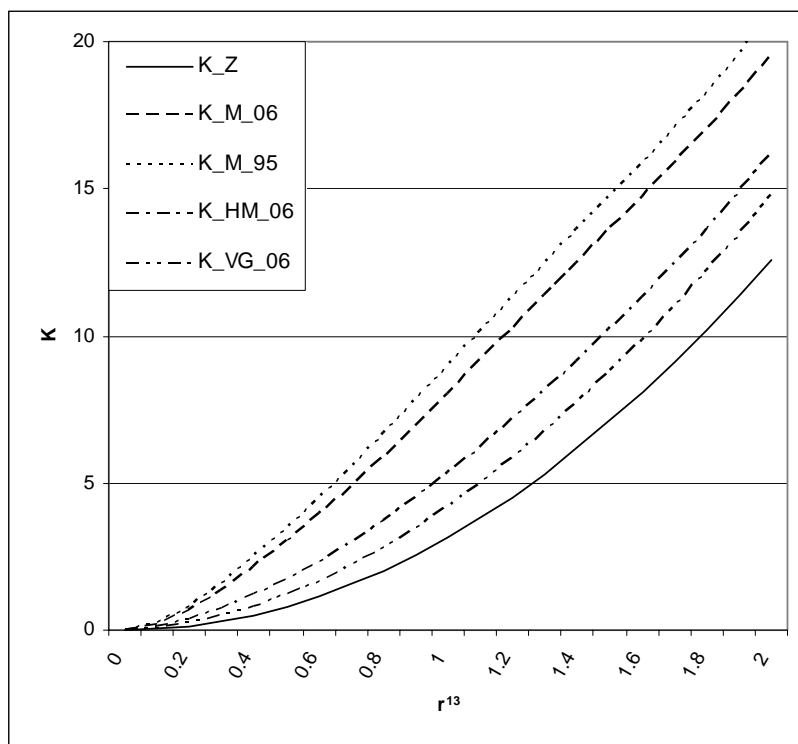
$$4.2-1 \quad Z_m = \frac{0,087 - 0,108}{\sqrt{0,00000301}} = -12,336$$

Das kritische Signifikanzniveau der N(0,1)-Verteilung beträgt (fast) 0 %, so dass mit nahezu 100%-iger Sicherheit davon auszugehen ist, dass sich die Metallherzeugung und -bearbeitung in Deutschland räumlich stark konzentriert.

Dieser Eindruck verfestigt sich bei einem Blick auf Abbildung 4.2-2.

¹³ Aufgrund von Längen- und Breitengraden berechnet und deshalb nicht mit Angaben in qkm zu verwechseln.

Abbildung 4.2-2: Univariate K-Funktionen für ausgewählte Wirtschaftszweige des verarbeitenden Gewerbes in Deutschland 2006



Legende:

- K_Z = zufällige Verteilung der Betriebe
- K_M_06 = Univariate K-Funktion der Betriebe des Sektors „Metallerzeugung und -bearbeitung“ in Deutschland 2006
- K_M_95 = Univariate K-Funktion der Betriebe des Sektors „Metallerzeugung und -bearbeitung“ in Deutschland 1995
- K_HM_06 = Univariate K-Funktion der Betriebe des Sektors „Herstellung von Metall-erzeugnissen“ in Deutschland 2006
- K_VG_06 = Univariate K-Funktion der Betriebe im Verarbeitenden Gewerbe sowie Bergbau und Gewinnung von Steinen und Erden in Deutschland 2006

Sie zeigt, dass

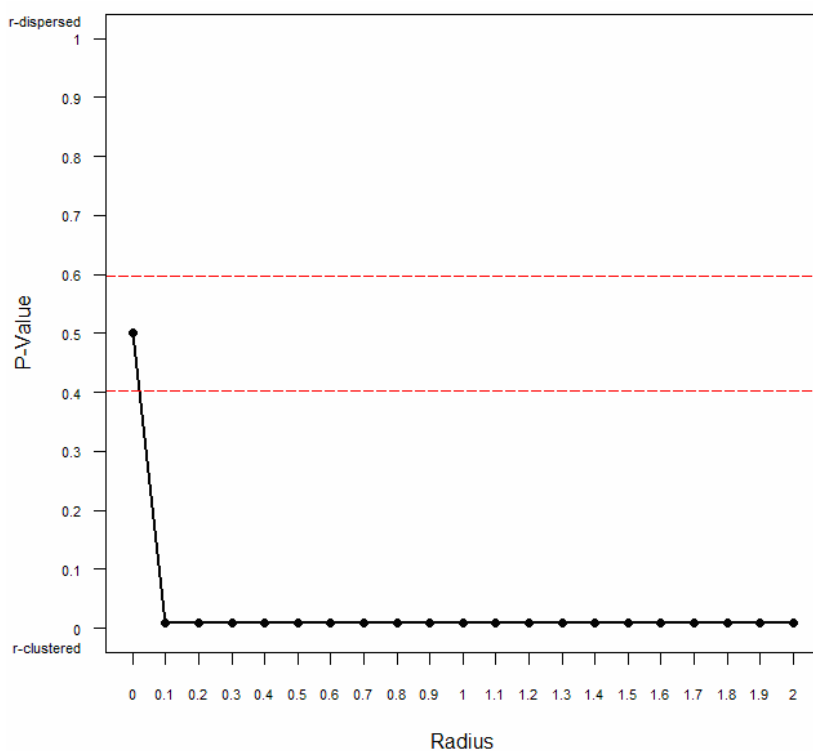
- die Metallerzeugung und -bearbeitung in Deutschland 2006 deutlich räumlich konzentriert ist, da K_{2706} oberhalb der K-Funktion der zufälligen Verteilung Zufall liegt.

¹⁴ Entfernung der Betriebe aufgrund von Längen- und Breitengraden; sie lassen sich in Entfernungen in Kilometern umrechnen.

- diese Clusterung im Zeitablauf rückläufig war, da K2795 oberhalb von K2706 positioniert ist.¹⁵
- die Metallerzeugung und -bearbeitung räumlich stärker konzentriert ist als das Verarbeitende Gewerbe insgesamt, aber auch stärker als der verwandte Wirtschaftszweig „Herstellung von Metallerzeugnissen“.

Wir wenden uns nun der Beantwortung der Frage zu, ob sich die Clusterung des Sektors 27 zum Zeitpunkt 2006 signifikant von jener des gleichen Wirtschaftszweiges im Jahr 1995 und von jener des Verarbeitenden Gewerbes im Jahr 2006 unterscheidet; Antwort auf diese Fragen geben die Abbildungen 4.2-3 und 4.2-4.

Abbildung 4.2-3: Grad der Clusterung der Metallerzeugung und -bearbeitung im Vergleich zum Verarbeitenden Gewerbe insgesamt 2006



¹⁵ Dabei wird davon ausgegangen, dass die leicht geänderte Erhebung zwischen den Jahren 1995 und 2006 durch die Umstellung der WZ 1993 auf die WZ 2003 die empirischen Ergebnisse nur unwesentliche beeinflusst.

Abbildung 4.2-4: Grad der Clustering der Metallerzeugung und -bearbeitung 2006 im Vergleich zum gleichen Sektor 1995

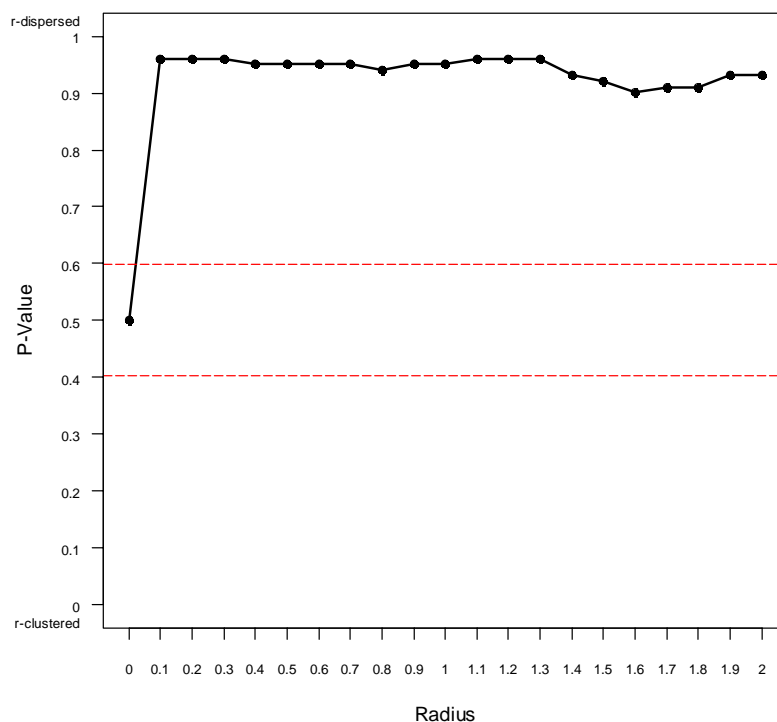


Abbildung 4.2-3 zeigt, dass der Sektor Metallerzeugung und -bearbeitung bei einem kritischen Signifikanzniveau von (fast) 0 % bei allen Entfernungen stärker geclustert ist als das Verarbeitende Gewerbe insgesamt.

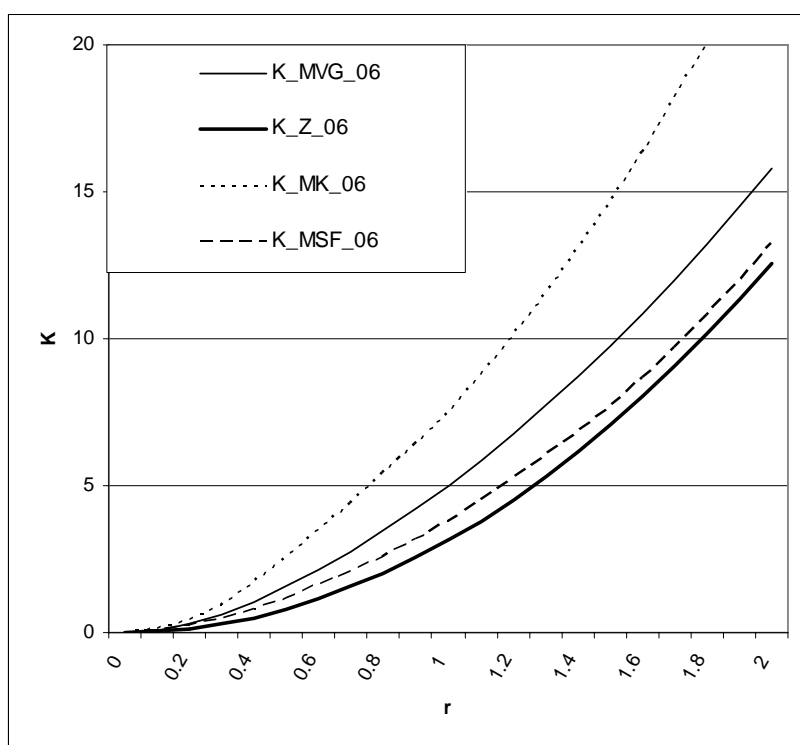
Dagegen ist der Grad der räumlichen Konzentration der Metallerzeugung und -bearbeitung zwischen 1995 und 2006 deutlich zurückgegangen, wie Abbildung 4.2-4 zeigt. Die räumliche Streuung hat bei allen Radien signifikant zugenommen.

4.3 Ergebnisse unter Verwendung bivariater K-Funktionen

Wir berechnen die bivariaten K-Funktionen zwischen den Betrieben des Sektors „Metallerzeugung und -bearbeitung“ auf der einen und den Betrieben der übrigen Wirtschaftszweige des Verarbeitenden Gewerbes auf der anderen Seite. Das entsprechende Ergebnis kann der Tabelle 4.3-1 und der Abbildung 4.3-1 entnommen werden. Der engste Zusammenhang besteht (erwartungsgemäß) mit dem Kohlebergbau; beide Wirtschaftszweige bilden noch immer ein räumliches Cluster, wie es die alten Industriegebiete in den letzten beiden Jahrhunderten geprägt hat. Der Zusammenhang ist deutlich enger als zwischen der Metall-

erzeugung und -bearbeitung und dem Verarbeitenden Gewerbe insgesamt. Dagegen stoßen sich Metallerzeugung und -bearbeitung auf der einen und Rundfunk- und Nachrichtentechnik, Sonstiger Fahrzeugbau und Recycling auf der anderen Seite eher ab, wenn man das Verarbeitende Gewerbe insgesamt zum Vergleichsmaßstab nimmt. Gegenüber einer zufälligen räumlichen Verteilung besteht dagegen eher ein positiver räumlicher Zusammenhang.

Abbildung 4.3-1: Ausgewählte bivariate K-Funktionen für die Betriebe der Metallerzeugung und -bearbeitung



Legende:

- K_MK_06 = Bivariate K-Funktion der Betriebe des Sektors „Metallerzeugung und -bearbeitung“ und der Betriebe des Sektors „Kohlbergbau, Torfgewinnung“
- K_MVG_06 = Bivariate K-Funktion der Betriebe des Sektors „Metallerzeugung und -bearbeitung“ und der Betriebe des Verarbeitenden Gewerbes insgesamt
- K_Z_06 = Bivariate K-Funktion der Betriebe des Sektors „Metallerzeugung und -bearbeitung“ und einer zufälligen Verteilung von Betrieben
- K_MSF_06 = Bivariate K-Funktion der Betriebe des Sektors „Metallerzeugung und -bearbeitung“ und der Betriebe des Sektors „Sonstiger Fahrzeugbau“

Tabelle 4.3-1: Bivariate K-Funktionen für die Betriebe der Metallerzeugung und -bearbeitung

r	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2,0
CA10	0,000	0,118	0,418	0,926	1,687	2,547	3,454	4,388	5,412	6,417	7,521	8,794	10,094	11,509	13,060	14,652	16,362	18,205	20,072	21,916	23,918
CA11	0,000	0,010	0,052	0,118	0,218	0,348	0,522	0,751	1,075	1,481	1,971	2,656	3,662	4,848	6,458	8,416	10,187	11,918	13,854	15,765	17,553
CB14	0,000	0,044	0,169	0,390	0,695	1,095	1,568	2,113	2,735	3,422	4,195	5,035	5,927	6,882	7,884	8,946	10,073	11,244	12,471	13,722	15,034
DA15	0,000	0,061	0,219	0,468	0,797	1,197	1,667	2,211	2,820	3,492	4,228	5,032	5,890	6,796	7,767	8,798	9,901	11,057	12,257	13,532	14,883
DA16	0,000	0,077	0,236	0,399	0,588	0,864	1,162	1,485	1,845	2,281	2,770	3,326	3,938	4,621	5,597	6,864	8,430	9,994	11,710	13,388	15,066
DB17	0,000	0,073	0,289	0,610	1,021	1,515	2,124	2,827	3,633	4,510	5,457	6,455	7,517	8,594	9,694	10,833	11,998	13,148	14,339	15,596	16,904
DB18	0,000	0,060	0,225	0,492	0,852	1,295	1,825	2,429	3,144	3,925	4,788	5,715	6,728	7,789	8,895	10,044	11,210	12,405	13,628	14,929	16,307
DC19	0,000	0,069	0,249	0,534	0,881	1,279	1,753	2,307	3,033	3,808	4,621	5,525	6,518	7,533	8,636	9,805	10,980	12,272	13,700	15,267	17,049
DD20	0,000	0,047	0,183	0,407	0,718	1,103	1,584	2,154	2,803	3,527	4,320	5,179	6,107	7,104	8,143	9,272	10,470	11,713	13,000	14,308	15,669
DE21	0,000	0,081	0,295	0,622	1,045	1,552	2,136	2,786	3,523	4,322	5,208	6,158	7,219	8,316	9,488	10,707	11,946	13,188	14,479	15,831	17,228
DE22	0,000	0,080	0,295	0,607	0,991	1,445	1,967	2,566	3,242	3,981	4,775	5,624	6,523	7,447	8,428	9,481	10,609	11,785	12,999	14,271	15,612
DF23	0,000	0,118	0,368	0,758	1,207	1,733	2,323	2,948	3,616	4,348	5,139	5,960	6,849	7,835	8,850	10,066	11,324	12,577	13,837	15,273	16,718
DG24	0,000	0,082	0,306	0,649	1,087	1,617	2,233	2,919	3,652	4,467	5,336	6,259	7,218	8,220	9,270	10,408	11,615	12,865	14,160	15,502	16,915
DH25	0,000	0,075	0,283	0,600	1,021	1,520	2,087	2,732	3,444	4,213	5,046	5,942	6,891	7,895	8,955	10,078	11,244	12,448	13,695	15,008	16,384
DI26	0,000	0,057	0,213	0,464	0,803	1,223	1,716	2,280	2,910	3,600	4,365	5,193	6,067	6,980	7,946	8,969	10,048	11,167	12,320	13,528	14,796
DJ28	0,000	0,124	0,449	0,918	1,491	2,140	2,846	3,605	4,415	5,270	6,188	7,141	8,119	9,115	10,144	11,216	12,323	13,451	14,628	15,857	17,184
DK29	0,000	0,084	0,314	0,661	1,103	1,631	2,223	2,889	3,613	4,401	5,239	6,135	7,067	8,033	9,055	10,123	11,253	12,429	13,632	14,904	16,255
DL30	0,000	0,074	0,285	0,596	0,961	1,411	1,942	2,524	3,157	3,878	4,671	5,547	6,458	7,419	8,465	9,582	10,766	12,029	13,264	14,564	15,946
DL31	0,000	0,088	0,321	0,679	1,117	1,640	2,230	2,863	3,556	4,307	5,114	5,970	6,868	7,811	8,794	9,834	10,927	12,042	13,197	14,426	15,737
DL32	0,000	0,079	0,270	0,535	0,864	1,237	1,655	2,127	2,655	3,250	3,896	4,604	5,361	6,171	7,034	7,970	8,954	9,991	11,068	12,201	13,424
DL33	0,000	0,081	0,295	0,617	1,018	1,487	2,014	2,588	3,222	3,907	4,636	5,419	6,267	7,146	8,065	9,047	10,082	11,171	12,304	13,510	14,772
DM34	0,000	0,085	0,300	0,621	1,032	1,507	2,045	2,634	3,282	3,990	4,772	5,624	6,507	7,432	8,411	9,433	10,495	11,602	12,759	14,005	15,331
DM35	0,000	0,069	0,246	0,496	0,813	1,188	1,619	2,082	2,584	3,159	3,791	4,485	5,245	6,019	6,824	7,712	8,669	9,679	10,770	11,942	13,211
DN36	0,000	0,062	0,229	0,492	0,840	1,260	1,749	2,313	2,964	3,698	4,509	5,377	6,314	7,310	8,373	9,523	10,757	12,049	13,370	14,727	16,139
DN37	0,000	0,069	0,241	0,505	0,836	1,213	1,645	2,132	2,678	3,301	3,971	4,656	5,399	6,174	6,991	7,869	8,862	9,940	11,001	12,133	13,313
Gesamt	0,000	0,083	0,305	0,638	1,061	1,558	2,124	2,754	3,448	4,200	5,014	5,877	6,785	7,734	8,725	9,776	10,883	12,030	13,222	14,484	15,819
Zufall	0,000	0,032	0,125	0,280	0,501	0,787	1,139	1,557	2,029	2,561	3,155	3,810	4,527	5,315	6,178	7,084	8,063	9,094	10,191	11,346	12,569

Legende:

CA10	Kohlenbergbau, Torfgewinnung
CA11	Gewinnung von Erdöl und Erdgas
CB14	Gewinnung von Steinen und Erden, sonstiger Bergbau
DA15	Ernährungsgewerbe
DA16	Tabakverarbeitung
DB17	Textilgewerbe
DB18	Bekleidungsgewerbe
DC19	Ledergewerbe
DD20	Holzgewerbe (ohne Herstellung von Möbeln)
DE21	Papiergewerbe
DE22	Verlagsgewerbe, Druckgewerbe, Vervielfältigung von bespielten Ton-, Bild- und Datenträgern
DF23	Kokerei, Mineralölverarbeitung, Herstellung und Verarbeitung von Spalt- und Brutstoffen
DG24	Herstellung von chemischen Erzeugnissen
DH25	Herstellung von Gummi- und Kunststoffwaren
DI26	Glasgewerbe, Herstellung von Keramik, Verarbeitung von Steinen und Erden
DJ28	Herstellung von Metallerzeugnissen
DK29	Maschinenbau
DL30	Herstellung von Büromaschinen, Datenverarbeitungsgeräten und -einrichtungen
DL31	Herstellung von Geräten der Elektrizitätserzeugung, -verteilung u. ä.
DL32	Rundfunk- und Nachrichtentechnik
DL33	Medizin-, Mess-, Steuer- und Regelungstechnik, Optik; Herstellung von Uhren
DM34	Herstellung von Kraftwagen und Kraftwagenteilen
DM35	Sonstiger Fahrzeugbau
DN36	Herstellung von Möbeln, Schmuck, Musikinstrumenten, Sportgeräten, Spielwaren und sonstigen Erzeugnissen
DN37	Recycling

Wir testen nun, ob signifikante Anziehungs- und Abstoßungseffekte zwischen ausgewählten Sektoren bestehen und wählen hierzu die Verteilung der Betriebe des Verarbeitenden Gewerbes insgesamt als Bezugspunkt. Dann zeigt Abbildung 4.3-2, dass bei allen Distanzen zwischen den Betrieben der Metallerzeugung und -bearbeitung sowie den Betrieben des Kohlebergbaus ein höchst signifikanter Anziehungseffekt besteht; zwischen dem gleichen Sektor und dem Sonstigen Fahrzeugbau besteht dagegen ein ebenso statistisch gesicherter Abstoßungseffekt (vgl. Abb. 4.3-3).

Abbildung 4.3-2: Anziehungs- und Abstoßungseffekte zwischen den Sektoren Metallerzeugung und -bearbeitung sowie Kohlebergbau

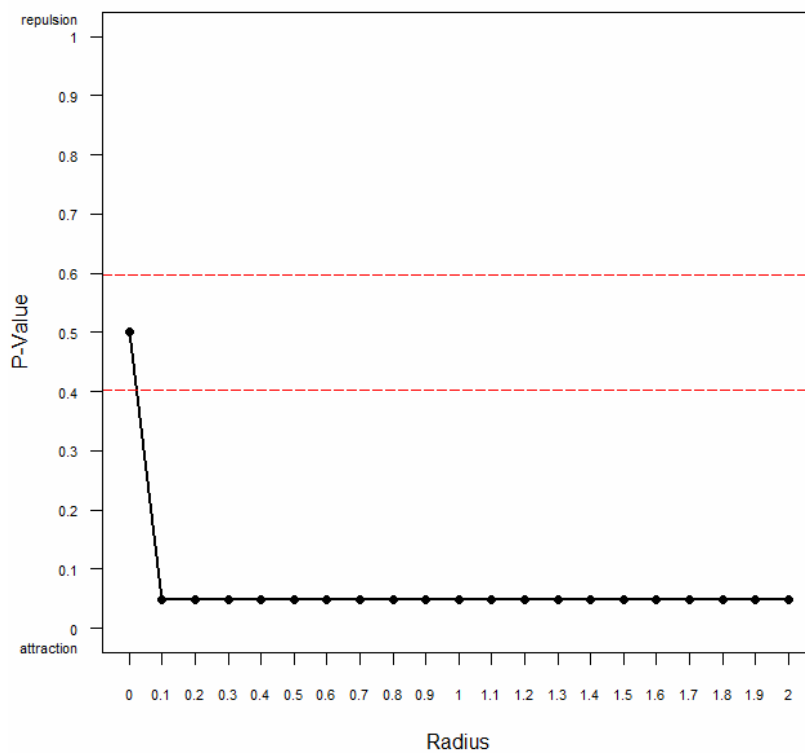
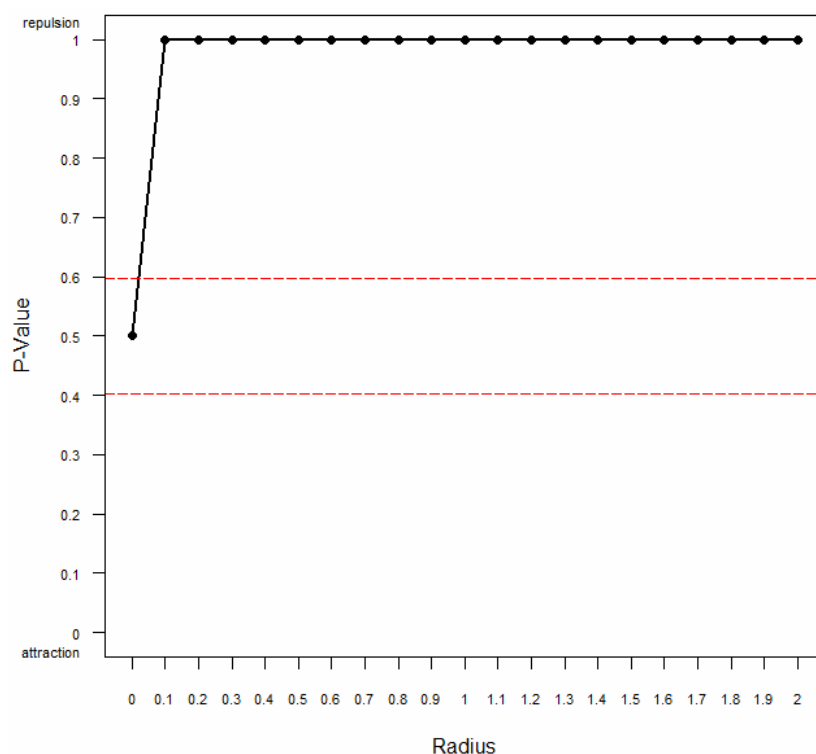


Abbildung 4.3-3: Anziehungs- und Abstoßungseffekte zwischen den Sektoren Metall-
erzeugung und -bearbeitung sowie sonstiger Fahrzeugbau



5. Resümée

Die Messung räumlicher Konzentration ökonomischer Variablen hat im Rahmen der empirischen Regionalforschung eine lange Tradition. Genutzt werden hierzu statistische Methoden, die auf der Einteilung einer Volkswirtschaft in Teilgebiete aufbauen und damit u.a. das areal-unit-Problem aufweisen, d.h. je nach gewählter Regionalisierung können die Ergebnisse höchst unterschiedlich ausfallen. Hier wird mit der uni- und bivariaten K-Funktion ein Verfahren vorgestellt, das dieses höchst relevante Problem umgeht und darüber hinaus weitere Vorteile aufweist. Exemplarisch vorgestellt wird das Verfahren an der räumlichen Verteilung der Betriebe des Sektors „Metallerzeugung und -bearbeitung“. Es führt zu erwarteten und gut interpretierbaren Ergebnissen, die sich bei der univariaten Analyse auf die räumliche Dispersion und Clusterung von Sektoren und bei der bivariaten Analyse auf die räumlichen Anziehungs- und Abstoßungseffekte zwischen ihnen beziehen. So ist die Metallerzeugung und -bearbeitung räumlich relativ stark konzentriert und weist die stärkste Anziehung (erwartungsgemäß) mit dem Kohlebergbau auf.

Literaturverzeichnis

Alecke, B. und Untiedt, G. (2006): Die geographische Konzentration von Industrie- und Dienstleistungen in Deutschland, GEFRA Working Paper 2, Münster.

Alecke, B., Untiedt, G. (2007): Clusterförderung und Wirtschaftspolitik – „Heilsbringer“ oder „Wolf im Schafspelz“, in: List Forum für Wirtschafts- und Finanzpolitik, 33/2, S. 89-105.

Anselin, L. (1988): Spatial Econometrics (Methods and Models), Dordrecht, Boston, London.

Arbia, G. und Espa, G. (1996): Statistica economica territoriale, Cedam, Padua.

Arbia, G.; Espa, G. und Quah, D. (2008): A class of spatial econometric methods in the empirical analysis of clusters of firms in the space, in: Empirical Economics 34/2008, S. 81-103.

Baddeley, A. und Turner, R. (2005): Spatstat: an R package for analyzing spatial point patterns, in: Journal of Statistical Software 12 (6), 1-42, URL: www.jstatsoft.org.

Bailey, T.C. und Gatrell, A.C. (1995): Interactive Spatial Data Analysis, Harlow.

Barff, R. (1987): Industrial Clustering and the Organization of Production: A Point Pattern Analysis of Manufacturing in Cincinnati, in: Annals of the Association of American Geographers 77, S. 89-103.

Bickenbach, F. und Bode, E. (2008): Disproportionality Measures of Concentration, Specialization, and Localization, in: International Regional Science Review, 31, S. 359-388.

Duranton, G. und Overman, H.G. (2005): Testing for localisation using micro-geographic data, in: Review of Economic Studies, 72, S. 1077-1106.

Eckey, H.F. (2008): Regionalökonomie, Wiesbaden.

Eckey, H.F.; Kosfeld, R.; Türck, M. (2008): Deskriptive Statistik, 5. Aufl., Wiesbaden.

Ellison, G. und Glaeser, E.L. (1997): Geographic concentration in U.S. manufacturing industries: A dartboard approach, in: Journal of Political Economy 105, S. 889-917.

Fritsch, M.; Stephan, A.; Werwatz, A. (2004): Regionalisierte Innovationspolitik sinnvoll, in: DIW Wochenbericht 27/04, Berlin.

Glaeser, E.L.; Kallal, H.D.; Scheinkman, J.A.; Shleifer, A. (1992): Growth in Cities, in: Journal of Political Economy 100, S. 1126-1152.

Jacobs, J. (1970): The Economics of Cities, New York/Toronto.

- Jacobs, J. (1986): *Cities and the Wealth of Nations*, Harmondsworth.
- Krugman, P. (1991): *Geography and trade*, Cambridge M.A..
- Litzenberger, T. (2006): *Cluster und die New Economic Geography*, Frankfurt/Berlin/Bern/Brüssel/New York/Oxford/Wien.
- Lucas, R.E. (1988): On the mechanics of economic development, in: *Journal of Monetary Economics* 22, S. 3-42.
- Marcon, E. und Puech, F. (2003): Evaluating the geographic concentration of industries using distance-based methods, in: *Journal of economic geography*, 3, S. 409-428.
- Marshall, A.: *Principles of economics*, London 1920.
- Menzel, M.P. (2008): Zufälle und Agglomerationseffekte bei der Clusterentstehung, in: *Zeitschrift für Wirtschaftsgeographie*, (52/2-3), S. 114-128.
- Porter, M.E. (2000): Location, competition, and economic development: Local clusters in a global economy, in: *Economic Development Quarterly* 14/1, S. 15-34.
- Quah, D. und Simpson, H. (2003): *Spatial cluster empirics*, mimeo.
- Ripley, D.B. (1977): Modelling Spatial Patterns (with discussion), in: *Journal of the Royal Statistical Society* 39, S. 172-212.
- Romer, P. (1986): Increasing returns and long run growth, in: *Journal of Political Economy* 94, S. 1002-1037.
- Rowlingson, B. und Diggle, P. (2008): Adapted, packaged for R by Roger Bivand, pcp functions by Giovanni Petris and goodness of fit by Stephen Eglen. *Splanes: Spatial and Space-Time Point Pattern Analysis*. R package version 2.01-24. <http://www.maths.lancs.ac.uk/~rowlings/Splanes/>.
- Statistisches Bundesamt (2006): *Unternehmensregister-System 95*, Wiesbaden.
- Statistisches Bundesamt:
http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/abisz/Verarbeiten_des_Gewerbe,templateId=renderPrint.psml.
- Sturm, R. und Tümmler, T. (2006): Das statistische Unternehmensregister – Entwicklungsstand und Perspektiven, in: *Wirtschaft und Statistik* 10, S. 1021-1036.
- Sweeney, S. und Feser, E. (1998): Plant Size and Clustering of Manufacturing Activity, in: *Geographical Analysis* 30/1, S. 45-64.