Friedrich-Alexander-Universität Erlangen-Nürnberg
Wirtschafts-und Sozialwissenschaftliche Fakultät

Diskussionspapier
37 / 2000

Bivariate Abhängigkeitsmessung für ordinalskalierte Merkmale im Falle fixierter Randverteilungen und zahlreicher Bindungen

Ingo Klein

Lehrstuhl für Statistik und Ökonometrie
Lehrstuhl für Statistik und empirische Wirtschaftsforschung
Lange Gasse 20 · D-90403 Nürnberg

Bivariate Abhängigkeitsmessung für ordinalskalierte Merkmale im Falle fixierter Randverteilungen und zahlreicher Bindungen

Ingo Klein
Lehrstuhl für Statistik und Ökonometrie
Universität Erlangen–Nürnberg
Lange Gasse 20
D–90403 Nürnberg
Germany
E-mail: ingo.klein@wiso.uni-erlangen.de

1 Einleitung

Es gibt in der statistischen Literatur eine Vielzahl von Vorschlägen für bivariate Abhängigkeitsmaße. Einige von diesen fußen auf Rängen oder Funktionen von Rängen und sollen für Paare von ordinalskalierten Merkmalen geeignet sein. Ob sie geeeignet sind, ist allerdings nur auf dem Hintergrund eines plausiblen und allgemein akzeptierten Forderungskataloges, also einer Axiomatik, zu beurteilen.
Seit den Arbeiten von Rényi (1959) und Lehmann (1966) wird versucht, Maßzahlen zur Messung der bivariaten Abhängigkeit durch eine Axiomatisierung der Konzepte der positiven und allgemeiner der monotonen Abhängigkeit zu fundieren. Konkret wird zum einen eine einstellige Relation auf der Menge der bivariaten Verteilungen eingeführt, die festlegt, wann zwei Merkmale oder Zufallsvariablen positiv abhängig sind (siehe z.B. Lehmann (1966)). Kimeldorf & Sampson (1989) sprechen in diesem Zusammenhang von der "Positive Dependence Property" (PDP). Zum anderen wird eine zweistellige Relation auf dem Raum der bivariaten Verteilungen betrachtet, die angibt, wann zwei Merkmale oder Zufallsvariablen positiver abhängig sind als zwei andere (siehe z.B. Rényi (1959)). Kimeldorf & Sampson (1987, 1989) benutzen dafür die Terminologie "Positive Dependence Ordering" (PDO). Dabei läßt sich untersuchen, wie durch eine PDP eine PDO festgelegt werden kann und umgekehrt.
Diese axiomatische Vorgehensweise hat den Vorteil, sich zunächst grundsätzlich und maßzahlunabhängig über die PDP bzw. die PDO Klarheit verschaffen zu müssen. In der meßtheoretischen Literatur wird davon gesprochen, daß zunächst die qualitativen Anforderungen (die sog. qualitative relationale Struktur) festgelegt werden, bevor eine Maßzahl versucht, diese durch eine strukturerhaltende (homomorphe) Abbildung in die reellen Zahlen (in ein sog. quantitatives relationales System) zu übertragen (siehe z.B. Pfanzagl (1968), Krantz et al. (1971), Roberts (1979), Klein (1984), Dabrowska (1985)).

Die in der Literatur diskutierten axiomatischen Ansätze beziehen sich vor allem auf den Fall qualitativ-ungeordneter und quantitativer Merkmale (siehe Rényi (1959), Dabrowska (1985)). Wird der Fall qualitativ-geordneter Merkmale diskutiert, so wird im allgemeinen von stetigen Randverteilungsfunktionen ausgegangen, so daß keine Bindungen auftreten können (siehe z.B. Schweizer & Wolff (1981), Capéraà & Genest (1990), Averous & Dortet-Bernadet (2000)). Der Fall qualitativ-geordneter (ordinalskalierter) Merkmale, für die Bindungen auftreten können, wird zumeist vernachlässigt. Dieser Fall ist allerdings für die praktische Anwendungen in der empirischen Sozialforschung und in der Markt- und Meinungsforschung von großer Bedeutung. So werden Einstellungen zu Politikern, Produkten oder Medien häufig mittels Rating-Skalen gemessen, die nur eine kleine Anzahl (3, 5 oder 7) von beantwortbaren Kategorien ermöglichen. Wenn nun beispielsweise 2000 Personen befragt werden und nur diese wenigen Kategorien ankreuzen können, ist klar, daß Bindungen in großer Zahl auftreten müssen. Somit ist die Behandlung von Paaren qualitativ-geordneter Merkmale mit großen Bindungszahlen für die statistische Praxis äußerst wichtig.
Es soll deshalb im folgende aufbauend auf den Vorschlägen der Literatur ein axiomatischer Zugang zur bivariaten Abhängigkeitsmessung ordinalskalierter Merkmale mit Bindungen vorgestellt werden. Die Vorgehensweise ist ähnlich wie bei Kimeldorf & Sampson (1989). Es wird zunächst die PDP und dann die PDO diskutiert, um schließlich für gängige Abhängigkeitsmaße (z.B. Spearmanscher Rangkorrelationskoeffizient ρ, Kendalls τ, γ-Maß von Goodman & Kruskal, Maß von Voegel & Wiede und ein kovarianzbasiertes Maß) zu überprüfen, inwieweit sie eine strukturerhaltende Abbildung der PDP bzw. PDO ermöglichen.
Zuvor wird allerdings ausgehend von meßtheoretischen Invarianzüberlegungen für sog. partiell-unabhängige Ordinalskalen gezeigt, daß bivariate Abhängigkeitsmaße für ordinalskalierte Merkmale funktional nur von kumulierten Häufigkeiten und nicht von den einzelnen Merkmalswerten abhängen können. Die Argumentation folgt im wesentlichen Klein (1994), wo auch die formalen Beweise geführt werden.

2 Notation

Die im folgenden eingeführte Notation ist eine Verallgemeinerung der in Klein (1999a, 1999b) für die Messung von Streuung und Schiefe ordinalskalierter Merkmale eingeführten Notation.
Es sollen zwei qualitativ-geordnete Merkmale U bzw. V betrachtet werden, die nur die möglichen Ausprägungen $u_1 < u_2 < \ldots u_k$ bzw. $v_1 < v_2 < \ldots < v_l$. besitzen können. Die Festlegung dieser Merkmalsausprägungen geschieht vor jeder statistischen Erhebung, so daß nicht sichergestellt ist, daß in einer solchen Erhebung, alle Merkmalsausprägungen auch vorkommen müssen. O.b.d.A legen wir fest, daß reelle Zahlen als Merkmalsausprägungen verwendet werden.
Es sollen nun in einer Erhebung n paarweise Beobachtungen (x_p, y_p), $p = 1, 2, \ldots, n$ vorliegen, wobei es für alle $p = 1, 2, \ldots, n$ genau ein $i \in \{1, 2, \ldots, k\}$ und ein $j \in \{1, 2, \ldots, l\}$ gibt, so daß $x_p = u_i$ und $y_p = v_j$ gilt. (x_p, y_p), $p = 1, 2, \ldots, n$ heißen Merkmalswerte.
Für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$ bezeichnen wir mit $n_{U,V}(i, j)$ die Anzahl der Beob-

achtungspaare mit den Merkmalsausprägungen (u_i, v_j). Die relative Häufigkeit, mit der (u_i, v_j) auftritt wird mit $f_{U,V}(i,j) = n_{U,V}(i,j)/n$ bezeichnet. Die absoluten bzw. relativen Randhäufigkeiten sind

$$\begin{aligned} n_U(i) &= \sum_{j=1}^{l} n_{U,V}(i,j) = n_{i.}, \quad n_V(j) = \sum_{i=1}^{k} n_{U,V}(i,j) = n_{.j}, \\ f_U(i) &= \sum_{j=1}^{l} f_{ij} = n_U(i)/n = f_{i.}, \quad f_V(j) = \sum_{i=1}^{k} f_{ij} = n_V(j)/n = f_{.j} \end{aligned}$$

für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, k$. Summiert man die absoluten bzw. relativen Häufigkeiten entsprechend der Ordnung von U und V auf, so ergeben sich die kumulierten absoluten oder relativen Häufigkeiten

$$\begin{aligned} N_{U,V}(i,j) &= \sum_{p \le i, q \le j} n_{UV}(p,q) = N_{ij}, \\ F_{U,V}(i,j) &= \sum_{p \le i, q \le j} f_{UV}(p,q) = N_{UV}(i,j)/n = F_{ij} \end{aligned}$$

mit den kumulierten absoluten und relativen Randhäufigkeiten

$$\begin{aligned} N_U(i) &= \sum_{p=1}^{i} \sum_{j=1}^{l} n_{U,V}(p,j) = N_{il}, \quad N_V(j) = \sum_{i=1}^{k} \sum_{q=1}^{j} n_{U,V}(i,q) = N_{kj}, \\ F_U(i) &= \sum_{p=1}^{i} \sum_{j=1}^{l} f_{U,V}(p,j) = F_{il}, \quad F_V(j) = \sum_{i=1}^{k} \sum_{q=1}^{j} f_{U,V}(i,q) = F_{kj} \end{aligned}$$

für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$. Es wird im folgenden sowohl die längere Notation mit den Merkmalen U und V als Index und (i,j) als Argument als auch die Kurznotation mit ij als Index verwendet.
Sämtliche $n_{U,V}(i,j)$ bzw. $f_{U,V}(i,j)$ lassen sich in einer $(k \times l)$-Kontingenztabelle (oder Kontingenzmatrix) der absoluten Häufigkeiten $\mathbf{n}_{U,V}$ (oder kurz $\mathbf{n}$) bzw. der relativen Häufigkeiten $\mathbf{f}_{U,V}$ (oder kurz $\mathbf{f}$) anordnen. $\mathbf{N}_{U,V}$ (oder kurz $\mathbf{N}$) bzw. $\mathbf{F}_{U,V}$ (oder kurz $\mathbf{F}$) bezeichnen die entsprechenden $(k \times l)$-Tabellen (oder Matrizen) der kumulierten absoluten bzw. relativen Häufigkeiten. Damit es sich um "echte" $(k \times l)$-Tabellen handelt, soll im folgenden angenommen werden, daß $f_U(k) \neq 0$ und $f_V(l) \neq 0$ sind. Mit dieser Einschränkung bezeichnet $\mathcal{F}_{k \times l}$ die Menge aller $(k \times l)$-Matrizen kumulierter relativer Häufigkeiten.

3 Funktionale Abhängigkeit von Abhängigkeitsmaßen

Im Falle der bivariaten Abhängigkeitsmessung werden zwei Merkmale betrachtet, die zwar auf identischem Skalenniveau, aber mittels unterschiedlicher Skalen gemessen werden. Möchte man den Zusammenhang zwischen der Außentemperatur und den Ausgaben für z.B. Eiscreme quantifizieren, so spielt es keine Rolle, welche konkrete Intervallskala für die Temperatur (z.B. Celsius oder Fahrenheit-Skala) und welche Verhältnisskala für die Ausgaben (z.B. US-Dollar oder Euro) verwendet werden. Die Temperaturskalen können beliebig mit den Ausgabenskalen kombiniert werden, ohne daß dies die Aufgabenstellung der Abhängigkeitsmessung stört. In diesem Sinne sind die beiden Skalen unabhängig. Hat man sich aber für zwei konkrete Skalen entschieden (z.B. Celsius und Euro), dann sind diese Skalen für sämtliche Beobachtungspaare zu verwenden. In Klein (1994) wird für diesen Fall der Begriff der partiell unabhängigen Skalen eingeführt. D.h. die Skalen für beide Merkmale können unabhängig voneinander variiert werden, sind aber für alle Paare von Messungen identisch.

Für zwei ordinalskalierte Merkmale U und V betrachten wir analog den Fall, daß jeweils für jedes Merkmal n Werte vorliegen, die derselben Ordinalskala entstammen, U und V aber aus unterschiedlichen Ordinalskalen kommen können. Erfaßt ein Merkmal die Einstellungen zu einem Produkt und das andere Merkmal die Einstellung zu einer TV-Sendung, so können (müssen aber nicht) für beide 5-Punkt Rating-Skalen vorgelegt werden. Die numerischen Werte auf diesen beiden Rating-Skalen können aber völlig unterschiedlich sein und können isoliert mittels beliebiger sog. zulässiger Skalentransformationen streng monoton transformiert werden.

Formalisiert bedeutet dies: Liegen die Beobachtungspaare $(x_1, y_1), \ldots, (x_n, y_n)$ vor, dann führen zulässige Skalentransformationen g und h zu den Paaren $(g(x_1), h(y_1), \ldots, g(x_n), h(y_n)$, wobei es sich um unterschiedliche Funktionen g und h handelt. Im Falle zweier Ordinalskalen sind g und h streng monoton zunehmende Funktionen.

In der Tradition von Pfanzagl (1968) läßt sich als eine Minimalforderung für sog. sinnvolle statistische Maßzahlen die sog. Vergleichsinvarianz fordern, d.h. für eine Maßzahl t muß

$$\begin{aligned} & t((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)) = t((x'_1, y'_1), (x'_2, y'_2), \ldots, (x'_n, y'_n)) \\ \Leftrightarrow \quad & t((g(x_1), h(y_1)), (g(x_2), h(y_2)), \ldots, (g(x_n), h(y_n))) \\ & = t((g(x'_1), h(y'_1)), (g(x'_2), h(y'_2)), \ldots, (g(x'_n), h(y'_n))) \end{aligned}$$

für alle $x_i, x'_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$ und $y_j, y'_j \in \mathbb{R}$, $j = 1, 2, \ldots, n$ und streng monoton zunehmenden Funktionen g und h sein. Zulässige Skalentransformationen verbinden "informations"-äquivalente Skalen und sinnvolle Maßzahlen sollen die Gleichheit zweier Maßzahlwerte unabhängig von der willkürlichen Festlegung einer Skala ermöglichen.

Im Fall streng monoton zunehmender Transformationen g und h läßt sich zeigen (Klein, (1994), daß die Maßzahl sogar absolutinvariant sein muß, d.h. es gilt

$$\begin{aligned} & t((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)) \\ & = t((g(x_1), h(y_1)), (g(x_2), h(y_2)), \ldots, (g(x_n), h(y_n))) \end{aligned}$$

für alle $x_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$ und $y_j \in \mathbb{R}$, $j = 1, 2, \ldots, n$ und streng monoton zunehmenden Funktionen g und h, wenn es mindestens ein Paar streng monotoner Transformationen g^* und h^* gibt mit

$$\begin{aligned} & t((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)) \\ & = t((g^*(x_1), h^*(y_1)), (g^*(x_2), h^*(y_2)), \ldots, (g^*(x_n), h^*(y_n))) \end{aligned}$$

für alle $x_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$ und $y_j \in \mathbb{R}$, $j = 1, 2, \ldots, n$. Da Abhängigkeitsmaße von Translationen nicht berührt werden sollen, existiert solch ein Paar streng monoton zunehmender Transformationen g^* und h^*. Somit ist die Suche auf absolutinvariante Abhängigkeitsmaße zu beschränken.
Bei dieser Suche spielen bekanntlich die sog. Maximalinvarianten eine große Rolle. Bezeichne G die Menge der auf $\mathbb{R}$ streng monoton zunehmenden Transformationen. Dann enthält $G \times G$ die zulässigen Skalentransformationen für zwei partiell unabhängige Ordinalskalen. Eine Maßzahl u heißt maximalinvariant bezüglich $G \times G$, wenn aus

$$u((x_1, y_1), \ldots (x_n, y_n)) = u((x'_1, y'_1), \ldots, (x'_n, y'_n))$$

folgt, daß es $(g, h) \in G \times G$ gibt mit $(x'_i, y'_i) = (g(x_i), h(y_i))$ für $i = 1, 2, \ldots, n$. Jede $(G \times G)$-absolutinvariante Maßzahl ist Funktion dieser Maximalinvarianten (siehe z.B. Lehmann (1959), S. 216, Theorem 1).
Klein (1994), S. 200ff. hat u.a. die Maximalinvarianten bezüglich der Transformationsgruppe $G \times G$ untersucht und folgendes festgestellt:

1. Wenn Ties vorkommen können, d.h. es kann $x_i = x_j$ oder $y_i = y_j$ für $i, j \in \{1, 2, \ldots, n\}$ sein, so ist eine $G \times G$-Maximalinvariante $(I(\mathbf{x}), I(\mathbf{y}))$, wobei

   $$I(\mathbf{x}) = (\text{sign}(x_i - x_j))_{i,j=1,2,\ldots,n} \quad \text{bzw.} \quad I(\mathbf{y}) = (\text{sign}(y_i - y_j))_{i,j=1,2,\ldots,n}$$

   sind für $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ und $\mathbf{y} = (y1, \ldots, y_n) \in \mathbb{R}^n$. Die Matrix $I(\mathbf{x})$ enthält alle Angaben darüber, ob beim paarweisen Vergleich der Elemente des Vektors $\mathbf{x}$ die Elemente gleich sind, oder ein Element größer bzw. kleiner als ein anderes ist. So gibt z.B. $\text{sign}(x_i - x_j)\text{sign}(y_i - y_j) > 0$ an, daß (x_i, y_i) und (x_j, y_j) ein konkordantes Paar darstellen. Ist z.B. $\text{sign}(x_i - x_j)\text{sign}(y_i - y_j)) < 0$, dann handelt es sich bei (x_i, y_i) und (x_j, y_j) um ein diskordantes Paar.

   Raveh (1986) gibt eine Klasse von Maßzahlen auf der Basis konkordanter und diskordanter Paare an, deren Zähler dem Bildungsgesetz

   $$\sum_{i>j} \omega_{ij} \text{sign}(x_i - x_j)\text{sign}(y_i - y_j)$$

   gehorcht, wobei unterschiedliche Gewichte ω_{ij} auch unterschiedliche Maßzahlen festlegen.

2. Wird zusätzlich noch die Permutationsinvarianz der Maßzahl verlangt, d.h. ist diese Maßzahl invariant bezüglich einer Umordnung der Argumente $(x_1, y_1), \ldots (x_n, y_n)$, so ist die $G \times G$-Maximalinvariante durch $\mathbf{F}$, d.h. durch die $(k \times l)$-Matrix kumulierten relativen Häufigkeiten gegeben. Die Permutationsinvarianz ist für die Klasse der Maßzahlen der monotonen Assoziation von Raveh z.B. erfüllt, wenn $\omega_{ij} = 1$ für $i, j = 1, 2, \ldots, n$ ist. Dann steht im Zähler die Differenz aus der Anzahl der konkordanten und diskordanten Paare. Maßzahlen mit diesem Zähler sollen im folgenden Korkordanzmaße genannt werden. Es wird späterhin gezeigt, daß diese Konkordanzmaße nur Funktionen von $\mathbf{F}$ sind.

3. Wenn Ties ausgeschlossen werden und Permutationsinvarianz gefordert wird, kann o.B.d.A. $x_1 < x_2 < \ldots < x_n$ betrachtet werden. Sei R_i der Rang von y_i für $i = 1, 2, \ldots, n$, dann ist der Vektor der Ränge $(R_1, \ldots, R_n)$ von $(y_1, \ldots, y_n)$ maximalinvariant. Da keine Ties vorliegen, muß $k = l = n$ sein, und es ist $R_j = N_{kj} = nF_{kj}$ für $j = 1, 2, \ldots, n$. D.h. die kumulierten absoluten Randhäufigkeiten des Merkmals U bestimmen die Ränge. Beispiele für Abhängigkeitsmaße die nur von dem Vektor der Ränge abhängen sind der Spearmansche Rangkorrelationskoeffizient

$$\frac{12}{n(n^2-1)} \sum_{i=1}^{n} (i - (n+1)/2)(R_i - (n+1)/2).$$

oder allgemein lineare Rangordnungsstatistken (siehe z.B. Hájek & Šidák (1967)).

Im Rahmen der deskriptiven Statistik können Ties nicht ausgeschlossen werden, wohingegen die Auszeichnung einer bestimmten Reihenfolge keinen Sinn macht, so daß Permutationsinvarianz gefordert werden kann. Wir werden uns deshalb im folgenden auf permutationsinvariante bivariate Abhängigkeitsmaße beschränken, die Funktionen kumulierter relativer Häufigkeiten sind.

4 Positive Abhängigkeitseigenschaft (PDP)

Es besteht zunächst die Aufgabe anzugeben, wann zwei Merkmale U und V die Eigenschaft der positiven monotonen Abhängigkeit besitzen, d.h. positiv abhängig sind. Es geht nicht darum, die Stärke der positiven Abhängigkeit zu quantifizieren. Kimeldorf & Sampson (1984) sprechen von der "Positive dependence property" (PDP), also der Eigenschaft der positiven Abhängigkeit. PDP-Konzepte sind die positive Regressionsabhängigkeit von Tukey (1958), die positive Quadrantenabhängigkeit von Lehmann (1966), die LTD-Eigenschaft von Lehmann (1966), die RTI-Eigenschaft von Esary & Proschan (1972) und die positive Assoziationseigenschaft von Esary, Proschan & Walkup (1967). Esary & Proschan (1972) haben die Beziehungen zwischen diesen Konzepten der positiven Abhängigkeit untersucht, und festgestellt, daß die positive Quadrantenabhängigkeit aus allen anderen positiven Abhängigkeitskonzepten folgt und damit die schwächste Form der positiven Abhängigkeit formuliert. Hinreichend für alle anderen Abhängigkeitskonzepte ist die positive Regressionsabhängigkeit. Zusätzlich wurde die TP_2-Eigenschaft eingeführt (siehe z.B. Block et al. (1982)), wobei TP_2 für "total positive of order 2" steht (siehe auch Karlin (1968)). Barlow & Proschan (1981) haben nachgewiesen, daß diese Eigenschaft ebenfalls hinreichend für die positive Quadranteneigenschaft ist.
Kimeldorf & Sampson (1989) konstruieren einen axiomatischen Überbau für die genannten Konzepte der positiven Abhängigkeit. Insbesondere fordern sie, daß jede PDP die positive Quadrantenordnung zu implizieren habe (Axiom (C1). Im folgenden soll deshalb mit der positiven Quadranteneigenschaft die schwächste PDP-Form explizit eigeführt und in dem dann folgenden Abschnitt darauf aufbauend die positive Quadrantenabhängigkeit als Abhängigkeitsordnung in den Mittelpunkt gerückt werden.

Wir beschränken uns auf die bivariate Abhängigkeitsmessung und damit auf die PDP für zwei Merkmale. Verallgemeinerungen auf mehr als zwei Merkmale sind möglich und werden z.B. von Shaked (1982) diskutiert.
Zwei Merkmale U und V heißen positiv quadrantenabhängig, wenn

$$F_{UV}(i,j) \geq F_U(i)F_V(j) \text{ für } i = 1,2,\ldots,k \text{ und } j = 1,2,\ldots,l$$

gilt. Dividiert man diese Beziehung durch $F_V(j)$, so ergibt

$$F_{UV}(i,j)/F_V(j) \geq F_U(i)$$

für $i = 1,2,\ldots,k$ und $j = 1,2,\ldots,l$ eine äquivalente Bedingung für die positive Quadrantenabhängigkeit von U und V. Faßt man U und V als Zufallsvariablen auf, dann lautet diese Bedingung in der Sprache der bedingten Wahrscheinlichkeiten

$$P(U \leq i | V \leq j) \geq P(U \leq i)$$

für $i = 1,2,\ldots,k$ und $j = 1,2,\ldots,l$. Die Kenntnis, daß die Werte von V "klein" sind, steigert mithin bei positiver Quadrantenabhängigkeit die Wahrscheinlichkeit, daß auch U "kleine" Werte annimmt.

5 Positive Abhängigkeitsordnung (PDO)

Es geht nun darum, zwei Paare von Merkmalen (U,V) und (U',V') bezüglich der Eigenschaft der positiven Abhängigkeit zu ordnen. Generell gibt es zwei unabhängige Zugänge zu einem solchen Ordnungskonzept. Der erste geht von fixierten Randverteilungen aus, d.h. die Verteilungen von U und U' bzw. V und V' sind identisch. Der zweite Ansatz versucht, auf diese einschränkende Bedingung zu verzichten.
Tchen (1980) geht von fixierten Randverteilungen aus und führt auf der Grundlage der positiven Quadranteneigenschaft die positive Quadrantenordnung ein. Zusätzlich diskutiert er eine Konkordanzordnung, von der sich nachweisen läßt, daß sie zu der positiven Quadrantenordnung äquivalent ist. Yanagimoto & Okamato (1969) schlagen auf der Basis von Tukeys positiver Regressionsabhängigkeit die sog. SI-Ordnung vor. Sie gehen explizit von identischen und stetigen Randverteilungen aus. Kimeldorf & Sampson (1987) verallgemeinern die TP_2-Eigenschaft zu einer TP_2-Ordnung. Gleichzeitig geben sie ein Axiomensystem an, das jede Ordnung der positiven Abhängigkeit (PDO) erfüllen sollte, wenn identische Randverteilungen betrachtet werden. Wiederum stellt ein Axiom sicher, daß die positive Quadrantenordnung aus jeder PDO gefolgert werden kann, d.h. die schwächste Abhängigkeitsordnung darstellt. Einbettbar in die Axiomatik von Kimeldorf & Sampson (1989) sind u.a. die positive Quadrantenordnung, die SI-Ordnung , die TP_2-Ordnung und eine Ordnung aufn Basis der positiven Assoziationseigenschaft von Esary et al. (1967).
Mosler & Scarsini (1993) stellen die bis dahin verfügbare Literatur auch zur Axiomatik der Abhängigkeitsmessung zusammen. Nur wenige Ansätze betreffen den Fall, daß auf die Fixierung identischer Randverteilungen verzichtet wird. Scarsini (1984) diskutiert die positive Quadrantenordnung im Fall nicht-identischer Randverteilungen. Er betrachtet insbesondere auch diskrete Verteilungen. Capéraà & Genest (1990) verallgemeinern die von Yanagimoto & Okamoto (1969) betrachtete SI-Ordnung auf den Fall nicht-identischer

Randverteilungen. Sie verwenden dabei die inverse Verteilungsfunktion. Diese wird schließlich ebenfalls von Averous & Bernardot (2000) benutzt, um auf der Basis der LTD- und RTI-Eigenschaft Ordnungsrelationen der positiven Abhängigkeit ohne Fixierung der Randverteilungen vorzuschlagen.
Generell ist aber anzumerken, daß alle Ansätze, die auf die Fixierung von Randverteilungen verzichten, von Konzepten ausgehen, die für ordinalskalierte Merkmale nicht zur Verfügung stehen. So werden entweder inverse Verteilungsfunktionen betrachtet oder sogar explizit die Stetigkeit der Randverteilungsfunktionen unterstellt, so daß keine Ties zugelassen werden. Wir wollen deshalb im folgenden nur die positive Quadrantenordnung und die Konkordanzordnung betrachten, die auch die Basis für die Axiomatisierung bivariater Abhängigkeitsmaße für Paare ordinalskalierter Merkmale darstellen werden.
Statt Paare von Merkmalen (U, V) und (U', V') bezüglich ihrer positiven Abhängigkeitseigenschaft zu vergleichen, können alternativ die zugehörigen Matrizen kumulierter relativer Häufigkeiten $\mathbf{F}$ und $\mathbf{F}'$ für die Ordnungsrelationen betrachtet werden.
Auf der Basis der positiven Quadranteneigenschaft von Lehmann (1966) wird definiert:

Definition 5.1 *Seien $\mathbf{F}, \mathbf{F}' \in \mathcal{F}_{k\times l}$ mit $F_{il} = F'_{il}$ für $i = 1, 2, \ldots, k$ und $F_{kj} = F'_{kj}$ für $j = 1, 2, \ldots, l$. $\mathbf{F}'$ ist positiver quadrantenabhängig als $\mathbf{F}$ (kurz: $\mathbf{F} \preceq \mathbf{F}'$), wenn $F_{ij} \leq F'_{ij}$ für $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, l$ gilt.*

D.h. $\mathbf{F}$ wird stochastisch von $\mathbf{F}'$ dominiert.
Tchen (1981) betrachtet die sog. Konkordanzordnung, die sich dadurch ergibt, daß man, ohne die Randverteilungen zu verletzen, Häufigkeitsumschichtugen derart vornimmt, daß sich die Anzahl konkordanter Paare erhöht.

Definition 5.2 *Betrachte die Matrizen*

$$\mathbf{N} = \mathbf{N}^1, \mathbf{N}^2, \ldots, \mathbf{N}^m = \mathbf{N}'$$

der absoluten kumulierten Häufigkeiten mit denselben kumulierten Randhäufigkeiten N_U und N_V. Bezeichnen

$$\mathbf{n} = \mathbf{n}^1, \mathbf{n}^2, \ldots, \mathbf{n}^m = \mathbf{n}'$$

die zugehörigen Matrizen der absoluten Häufigkeiten. Für $p = 1, 2, \ldots, m-1$ ergibt sich $\mathbf{n}^{p+1}$ aus $\mathbf{n}^p$, in dem an Stellen (i, j) und (i', j') ganzzalige Häufigkeitsmasse ϵ addiert und an den Stellen (i, j') und (i', j) ganzzahlige Häufigkeitsmasse ϵ substrahiert wird mit $i < i'$ und $j < j'$. Dann heißt $\mathbf{N}'$ konkordanter als $\mathbf{N}$ (bzw. $\mathbf{F} = \mathbf{N}/n \preceq_c \mathbf{F}' = \mathbf{N}'/n$).

Tchen (1981) zeigt sogar, daß

$$\mathbf{F} \preceq_c \mathbf{F}' \iff F_{ij} \leq F'_{ij} \text{ für } i = 1, 2, \ldots, k,\ j = 1, 2, \ldots, l$$

gilt, d.h die Konkordanzordnung und die positive Quadrantenordnung äquivalent sind.

Beispiel 5.1 *Betrachte*

$$\mathbf{n} = \begin{pmatrix} 2 & 4 & 0 \\ 0 & 0 & 4 \\ 4 & 2 & 2 \end{pmatrix} \quad \textit{und } \mathbf{n}' = \begin{pmatrix} 2 & 4 & 0 \\ 6 & 0 & 0 \\ 0 & 2 & 6 \end{pmatrix}.$$

Seien $i = 2,\ j = 1,\ i' = 3,\ j' = 3$. Dann sind

$$\begin{aligned} n'_{ij} &= n_{ij} + 4,\ \ n'_{i',j'} = n_{i',j'} + 4 \\ n'_{i',j} &= n_{i',j} - 4,\ \ n'_{i,j'} = n_{i,j'} - 4, \end{aligned}$$

so daß $\mathbf{N}/n \preceq_c \mathbf{N}'/n$ ist.

6 Extreme Abhängigkeit

Hoeffding (1948) und Fréchet (1951) haben untersucht, durch welche bivariaten Verteilungen mit identischen Randverteilungen sich bivariate Verteilungen nach unten und oben abschätzen lassen. Das Ergebnis nennt das folgende Lemma:

Lemma 6.1 *Sei $\mathbf{F} \in \mathcal{F}_{k\times l}$ mit den mit den Randhäufigkeiten F_{il}, $i = 1, 2, \ldots, k$ und F_{kj}, $j = 1, 2, \ldots, l$. Dann gelten*

$$\max(0, F_{il} + F_{kj} - 1) \leq F_{ij} \leq \min(F_{il}, F_{kj})$$

für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$.

Setzt man $F^u_{ij} = \min(F_{il}, F_{kj}$ und $F^l_{ij} = \max(0, F_{il} + F_{kj} - 1)$ für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$, dann sind $\mathbf{F}^l, \mathbf{F}^u \in \mathcal{F}_{k\times l}$ mit denselben relativen Randhäufigkeiten wie $\mathbf{F}$. Somit gilt im Sinne der positiven Quadrantenordnung $\mathbf{F}^l \preceq \mathbf{F} \preceq \mathbf{F}^u$. D.h. $\mathbf{F}^l$ und $\mathbf{F}^u$ sind für fixierte Randverteilungen im Sinne der positiven Quadrantenordnung diejenigen bivariater mit kleinst- und mit größtmöglicher positiver Abhängigkeit.

Beispiel 6.1 *Seien*

$$\mathbf{f} = 1/19 \begin{pmatrix} 4 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 6 \end{pmatrix} \quad \textit{und damit} \ \mathbf{F} = 1/19 \begin{pmatrix} 4 & 4 & 4 \\ 5 & 10 & 10 \\ 5 & 12 & 13 \\ 5 & 12 & 19 \end{pmatrix}$$

vorgegeben. Die Verteilungen mit maximaler negativer Abhängigkeit sind

$$\mathbf{F}^l = 1/19 \begin{pmatrix} 0 & 0 & 4 \\ 0 & 3 & 10 \\ 0 & 6 & 13 \\ 5 & 12 & 19 \end{pmatrix} \quad \textit{und} \ \mathbf{f}^l = 1/19 \begin{pmatrix} 0 & 0 & 4 \\ 0 & 3 & 3 \\ 0 & 3 & 0 \\ 5 & 1 & 0 \end{pmatrix}.$$

Bei maximaler positiver Abhängigkeit gilt:

$$\mathbf{F}^u = 1/19 \begin{pmatrix} 4 & 4 & 4 \\ 5 & 10 & 10 \\ 5 & 12 & 13 \\ 5 & 12 & 19 \end{pmatrix} \quad \textit{und} \ \mathbf{f}^u = 1/19 \begin{pmatrix} 4 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 6 \end{pmatrix}.$$

7 Spiegelungen bivariater Verteilungen

Es soll von Spiegelung bivariater Verteilungen gesprochen werden, wenn die Reihenfolge der Ausprägungen mindestens eines der betrachteten Merkmale umgedreht wird.
Zweiseitige Spiegelung bedeutet, daß die Reihenfolge der Ausprägungen beider Merkmale U und V umgedreht wird. Dadurch ergibt sich die relative Häufigkeitsverteilung $\mathbf{f}_{-U,-V}$ mit

$$\mathbf{f}_{-U,-V} = \begin{pmatrix} f_{kl} & f_{k,l-1} & \dots & f_{k1} \\ f_{k-1,l} & f_{k-1,l-1} & \dots & f_{k-1,1} \\ \dots & \dots & \dots & \dots \\ f_{1l} & f_{1,l-1} & \dots & f_{11} \end{pmatrix}.$$

Für die kumulierten relativen Häufigkeiten gilt dann

$$\begin{aligned} F_{-U,-V}(i,j) &= \sum_{p=1}^{i} \sum_{q=1}^{j} f_{U,V}(k+1-p, l+1-q) \\ &= \sum_{p=1}^{i} \sum_{q=1}^{j} [F_{U,V}(k+1-p, l+1-q) - F_{U,V}(k-p, l+1-q) \\ &\quad - F_{U,V}(k-p+1, l-q) + F_{U,V}(k-p, l-q)] \end{aligned}$$

für $i = 1, 2, \dots, k$ und $j = 1, 2, \dots, l$ mit $F_{U,V}(p,q) = 0$ für $p \leq 0$ oder $q \leq 0$.
Wird nur die Reihenfolge der Ausprägungen von Merkmal U umgedreht, ergibt sich

$$\mathbf{f}_{-U,V} = \begin{pmatrix} f_{k1} & f_{k2} & \dots & f_{kl} \\ f_{k-1,1} & f_{k-1,2} & \dots & f_{k-1,l} \\ \dots & \dots & \dots & \dots \\ f_{11} & f_{12} & \dots & f_{1l} \end{pmatrix}$$

bzw.

$$\begin{aligned} F_{-U,V}(i,j) &= \sum_{p=1}^{i} \sum_{q=1}^{j} f_{U,V}(k+1-p, q) \\ &= \sum_{p=1}^{i} \sum_{q=1}^{j} [F_{U,V}(k+1-p, q) - F_{U,V}(k-p, q) \\ &\quad - F_{U,V}(k-p+1, q-1) + F_{U,V}(k-p, q-1)] \end{aligned}$$

für $i = 1, 2, \dots, k$ und $j = 1, 2, \dots, l$, wobei wiederum $F_{U,V}(p,q) = 0$ für $p \leq 0$ oder $q \leq 0$ zu setzen ist.
Analog erhält man im Fall, daß die Reihenfolge der Ausprägungen von V umgedreht wird

$$\mathbf{f}_{U,-V} = \begin{pmatrix} f_{1l} & f_{1,l-1} & \dots & f_{11} \\ f_{2l} & f_{2,l-1} & \dots & f_{21} \\ \dots & \dots & \dots & \dots \\ f_{kl} & f_{k,l-1} & \dots & f_{k1} \end{pmatrix}$$

bzw.

$$
\begin{aligned}
F_{U,-V}(i,j) &= \sum_{p=1}^{i}\sum_{q=1}^{j} f_{U,V}(p,l-q+1) \\
&= \sum_{p=1}^{i}\sum_{q=1}^{j} [F_{U,V}(p,l-q+1) - F_{U,V}(p,l-q) \\
&\quad - F_{U,V}(p-1,l-q+1) + F_{U,V}(p-1,l-q)]
\end{aligned}
$$

für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$ mit $F_{U,V}(p,q) = 0$ für $p \leq 0$ oder $q \leq 0$.

8 Bivariate Abhängigkeitsmaße

8.1 Überblick

Neben der Axiomatisierungen der Eigenschaft der positiven Abhängigkeit und der Ordnung der positiven Abhägigkeit werden in der Literatur einige wenige Versuche der Axiomatisierung von bivariaten Abhängigkeitsmaßen vorgenommen. Die Arbeit von Rényi (1959) kann als ein Ausgangspunkt gelten. Rényis Axiome werden von Schweizer & Wolff (1981) erweitert und für einzelne gängige Maßzahlen geprüft, ob sie die Axiome erfüllen. Weitere axiomatische Anforderungen stammen von Kowalczyk & Pleszczyńska (1977) und Dabrowska (1985), die regressionsanalytische asymmetrische Meßkonzepte in den Mittelpunkt rückt. Scarsini (1984) gibt eine Axiomatisierung auf der Basis der positiven Quadrantenordnung an.
Kimeldorf & Sampson (1989) kritisieren alle diese Ansätze, da die zugrundeliegende Abhängigkeitsstruktur nicht explizit gemacht wird, d.h. es fehlt eine Axiomatisierung des Ordnungskonzeptes der positiven Abhängigkeit. Im Sinne der Meßtheorie fehlt die Spezifikation des qualitativen relationalen Systems.
Auffällig ist, daß seit der Arbeit von Rényi (1959) von einem Abhängigkeitsmaß gefordert wird, daß es für eine bivariate Normalverteilung in den Bravais-Pearson-Korrelationskoeffizienten übergeht. Diese Forderung ist für ordinalskalierte Merkmale nicht sinnvoll. Weiterhin wird gefordert, daß Unabhängigkeit hinreichend und notwendig dafür sein soll, daß eine Maßzahl den Wert 0 annimmt. Kimeldorf & Sampson (1989) haben nachgewiesen, daß diese Forderung im Widerspruch steht zu den Forderungen, daß die Maßzahl möglichst nur Werte aus dem Intervall $[-1, 1]$ annehmen und ein gewisses Stetigkeitsverhalten besitzen soll.
Da wir uns auf das Konzept der positiven Quadrantenabhängigkeit beschränkt haben, kommen wir für ordinalskalierte Merkmale mit einem recht einfachen Forderungskatalog an ein bivariates Abhängigkeitsmaß für ordinalskalierte Merkmale aus.

8.2 Anforderungen an ein bivariates Abhängigkeitsmaß

Sei $\mathbf{F}^0 = (F(il) \cdot F(kj))_{i=1,2,\ldots,k;j=1,2,\ldots,l}$ die Verteilung bei Unabhängigkeit und gegebenen Randverteilungen. Die folgende Definition nennt minimale Anforderungen, die ein bivariates Abhängigkeitsmaß für ordinalskalierte Merkmale erfüllen sollte:

Definition 8.1 *Ein bivariates Abhängigkeitsmaß für ordinalskalierte Merkmale ist ein Abbildung S sämtlicher $(k \times l)$-Matrizen kumulierter relativer Häufigkeiten in die Menge der reellen Zahlen, d.h. $S : \mathcal{F}_{k \times l} \to \mathbb{R}$, so daß folgende Bedingungen gelten:*

1. *S nimmt den Wert 0 an, wenn U, V unabhängig sind, d.h. $S(\mathbf{F}^0) = 0$.*

2. *Besitzen $\mathbf{F}$ und $\mathbf{F}'$ dieselbe Randverteilung, dann ist S ordnungserhaltend in bezug auf die positive Quadrantenordnung, d.h. $\mathbf{F} \preceq \mathbf{F}' \Longrightarrow S(\mathbf{F}) \leq S(\mathbf{F}')$.*

3. *Einfache Spiegelung verändert lediglich das Vorzeichen und nicht den Absolutbetrag von S, d.h. $S(\mathbf{F}_{-U,V}) = S(\mathbf{F}_{U,-V}) = -S(\mathbf{F}_{U,V})$.*

Durch zweifaches Anwenden der einfachen Spiegelung folgt dann, daß sich bei doppelter Spiegelung der Wert der Maßzahl nicht ändert. Die Unabhängigkeit der Axiome ist zum Teil bereits von Rényi (1959) und später von Kimeldorf & Sampson (1989) untersucht worden. Die Anforderungen sind leicht zu interpretieren. Die implizite Forderung, nur Funktionen kumulierter relativer Häufigkeiten zu betrachten, basiert auf der meßtheoretischen Invarianzeigenschaft und der Permutationsinvarianz. Der "Nullpunkt" der Abhängigkeitsmessung ist durch den Fall der Unabhängigkeit geeicht. Einfache Spiegelungen tangieren das Vorzeichen eines Abhängigkeitsmaßes und doppelte Spiegelungen berühren ein Abhängigkeitsmaß nicht. Entscheidend ist aber die Eigenschaft ordnungserhaltend zu sein bez. der positiven Quadrantenordnung.
Für fixierte Randverteilungen ist der Wertebereich eines Abhängigkeitsmaßes leicht angebbar: Wenn $\mathbf{F}$ eine bivariate kumulierte relative Häufigkeitsmatrix und S ein bivariates Abhängigkeitsmaß sind, dann gilt $S(\mathbf{F}^l) \leq S(\mathbf{F}) \leq S(\mathbf{F}^u)$, wobei die Randverteilungen von $\mathbf{F}$, $\mathbf{F}^u$ und $\mathbf{F}^l$ identisch sind. Dieses Ergebnis folgt wegen $\mathbf{F}^l \preceq \mathbf{F} \preceq \mathbf{F}^u$.
Gelegentlich werden an bivariate Abhängigkeitsmaße Normierungsanforderungen gestellt, die die Extremwerte für maximale Abhängigkeiten ohne und mit fixierten Randverteilungen betreffen.

Definition 8.2 *Sei S ein bivariates Abhängigkeitsmaß, dann heißt S*

1. *randverteilungsnormiert, wenn $S(\mathbf{F}^l) = -1$ und $S(\mathbf{F}^u) = 1$,*

2. *normiert, wenn $S(\mathbf{F}^n) = -1$ und $S(\mathbf{F}^p) = 1$ für alle $\mathbf{F}^n \in \mathcal{F}^n$ und $\mathbf{F}^p \in \mathcal{F}^p$*

gelten.

Es ist nun leicht zu zeigen, daß jedes randverteilungsnormiertes Abhängigkeitsmaß auch normiert ist.

Lemma 8.1 *Sei S ein randverteilungsnormiertes bivariates Abhängigkeitsmaß, dann ist S auch normiert.*

Beweis: Da in $\mathbf{f}^n$ bzw. in $\mathbf{f}^p$ in jeder Spalte $(k \leq l)$ oder jeder Zeile $k \geq l$ genau ein Element von Null verschieden ist, sind $(\mathbf{F}^n)^l = \mathbf{F}^n$ und $(\mathbf{F}^p)^u = \mathbf{F}^p$ und damit $S(\mathbf{F}^n) = S((\mathbf{F}^n)^l) = -1$ bzw. $S(\mathbf{F}^p) = S((\mathbf{F}^p)^u) = 1$. $\square$

9 Untersuchung gängiger Assoziationsmaße

9.1 Überblick

Bereits Lehmann (1966) hat nachgewiesen, daß Maßzahlen wie Kendalls τ (siehe Kendall (1938), Kendall (1970)), Spearmans ρ (siehe Spearman (1904)) und das Quadrantenmaß von Blomqvist (siehe Blomqvist (1950)) die positive Quadranteneigenschaft einhalten, d.h. für eine bivariate Verteilung, dien positiv quadrantenabhängig ist, einen positiven Wert liefern. Tchen (1981) hat dieses Ergebnis auf die positive Quadrantenordnung erweitert, d.h. bivariate Verteilungen, die positiver quadrantenabhängig sind, liefern größere Werte der angesprochenen Maßzahlen, sofern Verteilungen mit identischen Randverteilungen verglichen werden. Der Nachweis wird aber indirekt geführt. Yanagimoto & Okamoto (1969) untersuchen generelle Rangordnungsstatistiken. Allerdings betrachten sie nur den Fall, daß keine Bindungen vorliegen.
Ansonsten diskutierte Maßzahlen betreffen vorzugsweise quantitative Merkmale wie z.B. die Maximalkorrelation (siehe Rényi (1959)).
Wir wollen nun mit elementaren Mitteln direkt nachweisen, daß einige wichtige Maßzahlen, zu denen auch der Spearmans ρ und Kendalls τ gehören, die geforderten Eigenschaften eines bivariaten Abhängigkeitsmaßes für ordinalskalierte Merkmale erfüllen. Es werden lediglich einfache algebraischen Umformungen und partielle Ableitungen bemüht. Die Abhängigkeitsmaße lassen sich durchgängig als Funktion der kumulierten relativen Häufigkeiten F_{ij} darstellen, sind also auf einer Teilmenge der Menge der rationalen Zahlen definiert. Diese Funktion kann aber problemlos auf eine auf der Menge der reellen Zahlen differenzierbare Funktion erweitert werden, so daß die Monotonieeigenschaften mittels partieller Ableitungen überprüft werden können.

9.2 Spearmanscher Rangkorrelationskoeffizient

Der Spearmansche Rangkorrelationskoeffizient wird zumeist nur betrachtet, wenn keine Bindungen vorliegen. Dann ist er wie eingangs erwähnt durch

$$\frac{12}{n(n^2-1)}\sum_{i=1}^{n}(i-(n+1)/2)(R_i-(n+1)/2).$$

gegeben und damit nur Funktion der maximalinvarianten Ränge $R_1, \ldots, R_n$.
Liegen Bindungen vor, so wird empfohlen Mid-Ranks zu bilden. Wir wollen den Spearmanschen Rangkorrelationskoeffizienten als Bravais-Pearson-Korrelationskoeffizient der kumulierten Randhäufigkeiten einführen, d.h.

$$\rho(\mathbf{F}_{U,V}) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{l}(F_U(i)-\overline{F}_U)(F_V(j)-\overline{F}_V)f_{U,V}(i,j)}{\sqrt{\sum_{i=1}^{k}(F_U(i)-\overline{F}_U)^2 f_U(i)\sum_{j=1}^{l}(F_V(j)-\overline{F}_V)^2 f_V(j)}}$$

mit $\overline{F}_U = \sum_{i=1}^{k} F_U(i)f_U(i)$ und $\overline{F}_V = \sum_{j=1}^{l} F_V(j)f_V(j)$. $\rho(\mathbf{F}_{U,V})$ ist eine Funktion der bei Vorliegen von Bindungen und unter der Annahme der Permutationsinvarianz maximalinvarianten Matrix der kumulierten relativen Häufigkeiten $\mathbf{F}_{U,V}$. Diese Formel geht im Falle, daß keine Ties vorliegen, in die eingangs erwähnte Form des Spearmanschen Rangkorrelationskoeffizenten über.

Unabhängigkeit: Da bei Unabhängigkeit $f_{U,V}(i,j) = f_U(i) f_V)j)$ für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$ ist und $\sum_{i=1}^{k} (F_U(i) - \overline{F}_U) f_U(i) = 0$ ist, gilt $\rho(\mathbf{F}^0) = 0$.

Funktion von Maximalinvarianten: $\rho(\mathbf{F}_{U,V})$ läßt sich deutlicher als Funktion von $\mathbf{F}_{U,V}$ darstellen, wenn man die folgende Umformung vornimmt:

$$\begin{aligned}
\rho(\mathbf{F}_{U,V}) &= \sum_{i=1}^{k} \sum_{j=1}^{l} \alpha_i \beta_j f_{U,V}(i,j) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{l} \alpha_i \beta_j \left(F_{U,V}(i,j) - F_{U,V}(i,j-1) \right. \\
&\quad \left. - F_{U,V}(i-1,j) + F_{U,V}(i-1,j-1) \right)
\end{aligned}$$

mit

$$\alpha_i = \frac{F_U(i) - \overline{F}_U}{\sqrt{\sum_{i=1}^{k} (F_U(i) - \overline{F}_U)^2 f_U(i)}} \quad \text{und} \quad \beta_j = \frac{F_V(j) - \overline{F}_V}{\sqrt{\sum_{j=1}^{l} (F_V(j) - \overline{F}_V)^2 f_V(j)}}$$

Offensichtlich sind $\alpha_1 < \alpha_2 < \ldots \alpha_k$ und $\beta_1 < \beta_2 < \ldots < \beta_l$.
Eine nochmalige Umformung stellt $\rho(\mathbf{F}_{U,V})$ nur noch als Funktion von $F_{U,V}(i,j)$ dar:

$$\rho(\mathbf{F}_{U,V}) = \sum_{i=1}^{k} \sum_{j=1}^{l} (\alpha_i - \alpha_{i+1})(\beta_j - \beta_{j+1}) F_{U,V}(i,j)$$

mit $\alpha_i = 0$ für $i > k$ und $b_j = 0$ für $j > l$.

Monotonie: Betrachtet man nun die Funktion

$$\sum_{i=1}^{k} \sum_{j=1}^{l} (\alpha_i - \alpha_{i+1})(\beta_j - \beta_{j+1}) x_{ij}$$

für $x_{ij} \in \mathbb{R}$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, l$, dann ist diese Funktion offensichtlich differenzierbar nach x_{ij} mit der partiellen Ableitung $(\alpha_i - \alpha_{i+1})(\beta_j - \beta_{j+1}) > 0$. D.h. ρ ist monoton in $F_{U,V}(i,j)$ und erhält somit die positive Quadrantenordnung $\preceq$ (siehe auch Tchen (1981)).

Spiegelung und Normiertheit: An einem Gegenbeispiel läßt sich nun leicht zeigen, daß der Spearmansche Rangkorrelationskoeffizient nicht normiert ist.

Beispiel 9.1 *Mögliche zugehörige* (3×3)*-Kontingenztabellen mit extremer Abhängigkeit, wenn keine Randverteilungen fixiert sind, lauten z.B.*

$$\mathbf{f}^n = \begin{pmatrix} 0 & 0 & 7/24 \\ 0 & 10/24 & 0 \\ 7/24 & 0 & 0 \end{pmatrix} \quad \text{und } \mathbf{f}^p = \begin{pmatrix} 7/24 & 0 & 0 \\ 0 & 10/24 & 0 \\ 0 & 0 & 7/24 \end{pmatrix}.$$

Bezeichnen $\mathbf{F}$, $\mathbf{F}^n$ *und* $\mathbf{F}^p$ *die zugehörigen Matrizen kumulierter relativer Häufigkeiten, dann sind* $\rho(\mathbf{F}^n) = -0.9744$, $\rho(\mathbf{F}) = -0.0183$ *und* $\rho(\mathbf{F}^p) = 1.0$.

Da sich im vorstehenden Beispiel $\mathbf{F}^n$ durch einfache Spiegelung aus $\mathbf{F}^p$ ergibt, verletzt der Spearmansche Rangkorrelationskoeffizient auch die Forderung, daß eine einfache Spiegelung nur das Vorzeichen eines Abhängigkeitsmaßes tangieren darf. Da der Spearmansche Rangkorrelationskoeffizient ein Bravais-Pearson-Korrelationskoeffizient für die kumulierten relativen Randhäufigkeiten darstellt und diese für $\mathbf{F}^p$ auch auf der Winkelhalbierenden und damit auf einer Geraden liegen, dies aber für $\mathbf{F}^n$ nicht der Fall ist, verwundern die schlechten Eigenschaften von ρ nicht. Midranks können diese Eigenschaft mildern, aber nicht beseitigen.
Daß der Spearmansche Rangkorrelationskoeffizient nicht randverteilungsnormiert ist, zeigt das folgende Gegenbeispiel:

Beispiel 9.2 *Seien*

$$\mathbf{f} = 1/10 \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 3 \\ 3 & 0 & 0 & 0 \end{pmatrix} \quad \textit{und} \quad \mathbf{F} = 1/10 \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 7 \\ 4 & 5 & 6 & 10 \end{pmatrix},$$

dann sind $\rho(\mathbf{F}^l) = -0.6415$*,* $\rho(\mathbf{F}) = -0.6415$ *und* $\rho(\mathbf{F}^u) = 0.6677$*.*

9.3 Allgemein: Rangkorrelationskoeffizienten

Neben dem Spearmanschen Rangkorrelationskoeffizient werden in der nichtparametrischen Statistik als Testgrößen auf monotone bivariate Abhängigkeit lineare Rangordnungsstatistiken der Form

$$S_{a,b} = \sum a(i/(n+1))b(R_i/(n+1))$$

benutzt, falls keine Ties vorliegen. a und b sind geeignet zu wählende Scores. Diese Scores lassen sich auch über scoreserzeugende Funktionen $a : (0,1) \to \mathbb{R}$ bzw. $b : (0,1) \to \mathbb{R}$ einführen. Gebräuchliche scoreserzeugende Funktionen sind u.a.

1. $a(u) = b(u) = u$ für $u \in [0,1]$ ergibt den Spearmanscher Rangkorrelationskoeffizenten,

2. $a(u) = b(u) = 1/2 + \text{sign}(u - 1/2)$ für $u \in [0,1]$ liefert Blomqvists Quadrantenkoeffzient (siehe Blomquist (1950) und Lehmann (1966)) und

3. $a(u) = b(u) = \Phi^{-1}(u)$ führt zu dem Van der Waerdenschen Rangkorrelationskoeffizienten (siehe Hájek & Šidák (1967), S. 113).

Prinzipiell sind diese scoreserzeugenden Funktionen auch im Falle von Bindungen einsetzbar. In Analogie zur Schreibweise des Spearmanschen Rangkorrelationskoeffizienten erhält man allgemeine Rangkorrelationskoeffizienten mit Scoresfunktionen a und b als

$$\rho_{a,b}(\mathbf{F}_{U,V}) = A/B$$

mit

$$
\begin{aligned}
A &= \sum_{i=1}^{k}\sum_{j=1}^{l} a(nF_U(i)/(n+1))b(nF_V(j)/(n+1))f_{U,V}(i,j) \\
&\quad -\sum_{i=1}^{k} a(nF_U(i)/(n+1))f_U(i)\sum_{j=1}^{l} b(nF_V(j)/(n+1))f_V(j)
\end{aligned}
$$

und

$$
\begin{aligned}
B &= \sqrt{\sum_{i=1}^{k} a(nF_U(i)/(n+1))^2 f_U(i) - \left(\sum_{i=1}^{k} a(nF_U(i)/(n+1))f_U(i)\right)^2} \\
&\quad \cdot\sqrt{\sum_{j=1}^{l} b(nF_V(j)/(n+1))^2 f_V(j) - \left(\sum_{j=1}^{l} b(nF_V(j)/(n+1))f_V(j)\right)^2}\,.
\end{aligned}
$$

Ganz analog zur Diskussion des Spearmanschen Rangkorrelationskoeffizienten läßt sich nun leicht zeigen, daß auch allgemeine Rangkorrelationskoeffizienten monoton zunehmend in F_{ij} sind, wenn die Scoresfunktionen a und b monoton zunehmend sind.

9.4 Konkordanzmaße

Konkordante und diskordante Paare wurden bereits bei der Konkordanzordnung betrachtet. Für die absolute Häufigkeit konkordanter Paare gilt (siehe z.B. Schneider, Kornrumpf & Mohr (1995), S. 66)

$$
N_c = n^2\sum_{i=1}^{k-1}\sum_{j=1}^{l-1} f_{ij}^c f_{ij} \quad \text{mit} \quad f_{ij}^c = \sum_{p=i+1}^{k}\sum_{q=j+1}^{l} f_{pq}
$$

für $i = 1, 2, \ldots, k-1$ und $j = 1, 2, \ldots, l-1$. Analog erhält man die absolute Häufigkeit diskordanter Paare als

$$
N_d = n^2\sum_{i=1}^{k-1}\sum_{j=2}^{l} f_{ij}^d f_{ij} \quad \text{mit} \quad f_{ij}^d = \sum_{p=i+1}^{k}\sum_{q=1}^{j-1} f_{pq}
$$

für $i = 1, \ldots, k-1$ und $j = 2, \ldots, l$.

Funktion von Maximalinvarianten: Die Anzahl der konkordanten und diskordanten Paare lassen sich direkt als Funktion der kumulierten relativen Häufigkeiten F_{ij} darstellen. Zunächst sind:

$$
\begin{aligned}
h_{ij} &= n(F_{ij} - F_{i,j-1} - F_{i-1,j} + F_{i-1,j-1}) \\
h_{ij}^d &= n(F_{k,j-1} - F_{i,j-1}), \quad h_{ij}^c = n(1 - F_{kj} - F_{il} + F_{ij})
\end{aligned}
$$

für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$, wobei $F_{0,0} = F_{i,0} = F_{0,j} = 0$ für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$ gesetzt werden.

Setzt man diese Beziehungen in die Formel für die Anzahl der konkordanten Paare ein, so lassen sich diese als Funktion von F_{ij} darstellen:

$$
\begin{aligned}
N_c &= n^2 \sum_{i=1}^{k-1} \sum_{j=1}^{l-1} f_{ij}^c f_{ij} \\
&= n^2 \sum_{i=1}^{k-1} \sum_{j=1}^{l-1} (F_{ij} - F_{i,j-1} - F_{i-1,j} + F_{i-1,j-1}) \\
&\quad -n^2 \sum_{j=1}^{l-1} F_{kj} \sum_{i=1}^{k-1} (F_{ij} - F_{i,j-1} - F_{i-1,j} + F_{i-1,j-1}) \\
&\quad -n^2 \sum_{i=1}^{k-1} F_{il} \sum_{j=1}^{l-1} (F_{ij} - F_{i,j-1} - F_{i-1,j} + F_{i-1,j-1}) \\
&\quad +n^2 \sum_{i=1}^{k-1} \sum_{j=1}^{l-1} (F_{ij}^2 - F_{ij}F_{i,j-1} - F_{ij}F_{i-1,j} + F_{ij}F_{i-1,j-1}).
\end{aligned}
$$

Die entsprechende Darstellung der Anzahl der diskordanten Paare liefert:

$$
\begin{aligned}
N_d &= n^2 \sum_{i=1}^{k-1} \sum_{j=2}^{l} f_{ij}^d f_{ij} \\
&= n^2 \sum_{j=2}^{l} F_{kj} \sum_{i=1} k-1 (F_{ij} - F_{i,j-1} - F_{i-1,j} + F_{i-1,j-1} \\
&\quad -n^2 \sum_{i=1}^{k-1} \sum_{j=2}^{l} F_{ij}F_{i,j-1} - F_{i,j-1}^2 - F_{i-1,j}F_{i,j-1} + F_{i-1,j-1}F_{i,j-1}).
\end{aligned}
$$

Monotonie: Faßt man in einem Zwischenstritt N_c als differenzierbare Funktion von F_{ij} auf, so ist lediglich die partielle Ableitung des letzten Summanden von Null verschieden, so daß

$$
\begin{aligned}
\frac{\partial N_c}{\partial F_{ij}} &= n^2(2F_{ij} - F_{i,j-1} - F_{i,j+1} - F_{i-1,j} - F_{i+1,j} + F_{i-1,j-1} + F_{i+1,j+1}) \\
&= n^2(f_{i+1,j+1} + f_{ij}) \geq 0
\end{aligned}
$$

und N_c eine monoton zunehmende Funktion in F_{ij} ist. Die partielle Ableitung von N_d nach F_{ij} ist wegen

$$
\frac{\partial N_d}{\partial F_{ij}} = -n^2(f_{i,j+1} + f_{i+1,j}) \leq 0
$$

nicht-positiv und damit N_d monoton abnehmend in F_{ij}.
Die meisten permutationsinvarianten Konkordanzmaße sind durch Quotienten definiert, wobei der Zähler durch die Differenz der Anzahl der konkordanten und der Anzahl der

diskordaten Paare $N_c - N_d$ gegeben ist. Offenbar ist diese Differenz mit den vorstehenden Überlegungen monoton zunehmend in F_{ij}.
Daraus folgt, daß Kendalls τ

$$\tau(F) = \frac{N_c - N_d}{n(n-1)/2}$$

monoton zunehmend in F_{ij} ist und damit die positive Quadrantenordnung einhält.
Etwas komplizierter verhält es sich mit Goodman & Kruskals γ

$$\gamma(F) = \frac{N_c - N_d}{N_c + N_d},$$

da im Nenner die Summe aus konkordanten und diskordanten Paaren auftaucht. Mit

$$\frac{\partial \gamma(F)}{\partial F_{ij}} = \frac{2}{(N_c + N_d)^2} \left(\frac{\partial N_c}{\partial F_{ij}} N_d - \frac{\partial N_d}{\partial F_{ij}} N_c \right) \geq 0$$

ist auch $\gamma(F)$ monoton zunehmend in F_{ij} und hält die positive Quadrantenordnung ein.

Unabhängigkeit: Bekanntlich (siehe z.B. Schneider, Kornrumpf & Mohr (1995), S.66) ist die Unabhängigkeit von U und V nur im Falle $k = l = 2$ (Vierfeldertafel) hinreichend und notwendig dafür, daß das Maß von Goodman & Kruskal bzw. das Maß von Kendall den Wert 0 annehmen. In allen anderen Fällen ist die Unabhängigkeit nur hinreichend, was aber auch nur gefordert wird.

Spiegelung:

Lemma 9.1 *Seien $N_c(U,V))$, $N_c(-U,V)$, $N_c(-U,-V)$, $N_c(U,-V)$ bzw. $N_d(U,V))$, $N_d(-U,V)$, $N_d(-U,-V)$, $N_d(U,-V)$ die zu $\mathbf{F}_{U,V}$, $\mathbf{F}_{-U,V}$, $\mathbf{F}_{-U,-V}$, $\mathbf{F}_{U,-V}$ gehörenden Anzahlen von konkordanten bzw. diskordanten Paaren, dann gelten:*

1. $N_c(U,-V) = N_d(U,V)$ *2.* $N_d(U,-V) = N_c(U,V)$ *3.* $N_c(-U,-V) = N_c(U,V)$
4. $N_d(-U,-V) = N_d(U,V)$ *5.* $N_c(-U,V) = N_d(U,V)$ *6.* $N_d(-U,V) = N_c(U,V)$

Beweis: Da die Beweise alle nach dem gleichen Schema ablaufen, greifen wir den dritten Fall exemplarisch heraus: Es ist $f_{-U,-V}(i,j) = f_{U,V}(k+i-1, l+1-j)$ für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$. Dann gilt:

$$\begin{aligned} N_c(-U,-V) &= n^2 \sum_{i=1}^{k-1} \sum_{j=1}^{l-1} f_{-U,-V}(i,j) \sum_{p=i+1}^{k} \sum_{q=j+1}^{l} f_{-U,-V}(p,q) \\ &= n^2 \sum_{i'=2}^{k} \sum_{j'=2}^{l} f_{U,V}(i',j') \sum_{p'=1}^{i'-1} \sum_{q'=1}^{(j'-1} f_{U,V}(p',q') \end{aligned}$$

mit $i' = k+1-i$, $j' = l+1-j$, $p' = k+1-p$ und $q' = l+1-q$. Wegen

$$p' \leq i'-1 \iff i' \geq p'+1 \;\text{ und }\; q' \leq j'-1 \iff j' \geq q'+1$$

folgt

$$N_c(-U,-V) = n^2 \sum_{p'=1}^{k-1} \sum_{q'=1}^{l-1} f_{U,V}(p',q') \sum_{i'=p'+1}^{k} \sum_{j'=q'+1}^{l} f_{U,V}(i',j') = N_c(U,V). \ \Box$$

Als Konsequenz dieses Lemmas ergibt sich sofort, daß die Summe der Anzahlen konkordanter und diskordanter Paare durch einseitge oder zweiseitige Spiegelung nicht tangiert wird. Da für die Differenz der Anzahlen konkordanter und diskordanter Paare folgt

1. $N_c(U,-V) - N_d(U,-V) = N_d(U,V) - N_c(U,V) = -(N_c(U,V) - N_d(U,V))$,
2. $N_c(-U,-V) - N_d(-U,-V) = N_c(U,V) - N_d(U,V)$ und
3. $(N_c(-U,V) - N_d(-U,V)) = N_d(U,V) - N_c(U,V) = -(N_c(U,V) - N_d(U,V)$,

erfüllen das Maß von Goodman & Kruskal und das Maß von Kendall die axiomatischen Anforderungen bezüglich der einseitigen und zweiseitigen Spiegelung.

Normiertheit: Bekanntlich nehmen das Maß von Goodman & Kruskal und das Maß von Kendall ihr Minimum an, wenn es keine konkordanten Paare und ihr Maximum, wenn es keine diskordanten Paare gibt. Das Maß von Goodman & Kruskal ist zusätzlich normiert, was bei Vorliegen von Bindungen nicht für das Maß von Kendall gilt.
Das folgende Lemma zeigt nun, daß die Anzahl der konkordanten Paare für $\mathbf{F}^l$ und die Anzahl der diskordanten Paare für $\mathbf{F}^u$ Null sind. D.h. das Maß von Goodman & Kruskal ist nicht nur normiert, sondern auch randverteilungsnormiert.

Lemma 9.2 *1. Sei $F_{ij} = \max(0, F_{il} + F_{kj} - 1)$, dann ist $f_{ij}^c f_{ij} = 0$.*

2. Sei $F_{ij} = \min(F_{kj}, F_{il})$, dann ist $f_{ij}^d f_{ij} = 0$.

Beweis: Es sind

$$f_{ij}^c = 1 - F_{kj} - F_{il} + F_{ij} \text{ und } f_{ij}^d = F_{k,j-1} - F_{i,j-1}.$$

1. Sei $F_{ij} = \max(0, F_{il} + F_{kj} - 1)$.
   1. Fall: Wenn $F_{ij} = 0$ ist, dann ist auch $f_{ij} = 0$ und damit $f_{ij}^c f_{ij} = 0$.
   2. Fall: Wenn $F_{ij} = F_{il} + F_{kj} - 1$ ist, dann ist

$$f_{ij}^c = 1 - F_{kj} - F_{il} + F_{ij} = 1 - F_{kj} - F_{il} + (F_{il} + F_{kj} - 1) = 0.$$

2. Sei $F_{ij} = \min(F_{kj}, F_{il}$. Es ist $f_{ij}^d => 0$. Dann ist $F_{k,j-1} \geq F_{i,j-1} = \min(F_{k,j-1}, F_{il}$.
   1. Fall: Wenn $F_{i,j-1} = F_{k,j-1}$ ist, folgt unmittelbar $f_{ij}^d = 0$.
   2. Fall: Wenn $F_{i,j-1} = F_{il}$ ist, ist auch $F_{ij} = F_{il}$, da $F_{ij} \leq F_{il}$ sein muß. Somit ist

$$\begin{aligned} f_{ij} &= F_{ij} - F_{i,j-1} - F_{i-1,j} + F_{i-1,j-1} \\ &= F_{il} - F_{il} - F_{i-1,j} + F_{i-1,j-1} = 0 \end{aligned}$$

wegen $F_{i-1,j} \geq F_{i-1,j-1}$. $\Box$

9.5 Ein kovarianzbasiertes Maß

Das Maß

$$\nu(\mathbf{F}) = \sum_{i=1}^{k}\sum_{j=1}^{l}(F_{ij} - F_{kj}F_{il})$$

ist offensichtlich für gegebene $F_{kj},\ j = 1, 2, \ldots, l$ und $F_{il},\ i = 1, 2, \ldots, k$ eine monoton zunehmende Funktion der Maximalinvarianten $\mathbf{F}$ und mißt den "Abstand" zwischen $\mathbf{F}$ und der Matrix der kumulierten relativen Häufigkeit bei Unabhängigkeit $\mathbf{F}^0$. Da negative Summanden für negative Abhängigkeit und positive Summanden für positive Abhängigkeit sprechen, werden die "Abstände" nicht absolutbetragsmäßig sondern mit ihrem Vorzeichen erfaßt.
Mittels einiger elementarer Umformungen kann man $\nu(\mathbf{F})$ auch als Kovarianz der Indizes der Merkmalsausprägungen $u_1, \ldots, u_k$ und $v_1, \ldots, v_l$ darstellen (siehe auch Lehmann (1966)). Wegen

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{j=1}^{l} F_{ij} &= klf_{11} + (k-1)lf_{12} + k(l-1)f_{21} + \ldots + f_{kl} \\
&= \sum_{i=1}^{k}\sum_{j=l}(k-i+1)(l-j+1)f_{ij} \\
&= kl - k\sum_{j=1}^{l} jf_{.j} - l\sum_{i=1}^{k} if_{i.} + \sum_{i=1}^{k}\sum_{j=1}^{l} ijf_{ij},
\end{aligned}$$

$$\begin{aligned}
\sum_{j=1}^{l} F_{kj} &= \sum_{j=1}^{l}\sum_{q=1}^{j} f_{.q} = lf_{.1} + (l-1)f_{.2} + \ldots + f_{.l} \\
&= \sum_{j=1}^{l}(l-j+1)f_{.j} = l + 1 - \sum_{j=1}^{l} jf_{.j}
\end{aligned}$$

und

$$\sum_{i=1}^{k} F_{il} = k + 1 - \sum_{i=1}^{k} if_{i.}$$

folgt insgesamt:

$$\nu(\mathbf{F}) = \sum_{i=1}^{k}\sum_{j=1}^{l} ijf_{ij} - \sum_{i=1}^{k} if_{i.}\sum_{j=1}^{l} jf_{.j}.$$

Unabhängigkeit und Spiegelung: Aus den bekannten Eigenschaften der Kovarianz ist sofort ersichtlich, daß im Falle der Unabhängigkeit $\nu(\mathbf{F}^0) = 0$ gilt, umgekehrt aber aus $\nu(\mathbf{F}) = 0$ nicht folgt, daß $\mathbf{F} = \mathbf{F}^0$ sein muß. Für die Kovarianz gilt weiterhin

$$\sum_{i=1}^{k}\sum_{j=1}^{l} aibjf_{ij} - \sum_{i=1}^{k} aif_{i.}\sum_{j=1}^{l} bjf_{.j} = ab\left(\sum_{i=1}^{k}\sum_{j=1}^{l} ijf_{ij} - \sum_{i=1}^{k} if_{i.}\sum_{j=1}^{l} jf_{.j}\right),$$

für $a, b \in \mathbb{R}$. Wählt man $a \in \{-1, 1\}$ und $b \in \{-1, 1\}$, so erhält man sofort die in der Axiomatik geforderte Invarianz bezüglich Spiegelungen.

Normiertheit: Das Maß ν ist offensichtlich weder normiert noch randverteilungsnormiert. Wegen $\max(0, F_{il} + F_{kj} - 1) \leq F_{ij} \leq \min(F_{il}, F_{kj})$ ist

$$\max(-F_{il}F_{kj}, -(1 - F_{il})(1 - F_{kj}) \leq F_{ij} - F_{il}F_{kj} \leq \min(F_{il}(1 - F_{kj}), F_{kj}(1 - F_{il})$$

für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$. Summiert man diese Ungleichung über $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, k$, so erhält man nur von den kumulierten Randhäufigkeiten abhängende Grenzen für die Maßzahl $\nu(\mathbf{F})$.

9.6 Normierte Kovarianz

Bezeichnen U^* und V^* quantitative Merkmale mit den Ausprägungen $u_i = i$, $i = 1, 2, \ldots, k$ und $v_j = j$, $j = 1, 2, \ldots, l$, so ist der Bravais-Pearson-Korrelationskoeffizient für (U^*, V^*)

$$\rho_{BP}(U^*, V^*) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{l} ijf_{ij} - \sum_{i=1}^{k} fif_{i.}\sum_{j=1}^{l} jf_{.j}}{\sqrt{(\sum_{i=1}^{k} i^2 f_{i.} - (\sum_{i=1}^{k} if_{i.})^2)(\sum_{j=1}^{l} j^2 f_{.j} - (\sum_{j=1}^{l} jf_{.j})^2)}}.$$

Der Zähler entspricht exakt $\nu(\mathbf{F})$ und dieser Bravais-Pearson-Korrelationskoeffizient angewendet auf die Indizes der Ausprägungen von U und V kann nur Werte zwischen -1 und $+1$ annehmen kann. In der praktischen Anwendung werden häufig als Merkmalsausprägungen für ordinalskalierte Merkmale die ersten k bzw. l natürlichen Zahlen verwendet und auf die Zahlen die bekannten Analyseverfahren für quantitative Merkmale angewendet. Die vorstehende Überlegung scheint dieser Vorgehensweise recht zu geben. Bedauerlicherweise läßt sich aber die im Nenner verwendete Varianz der Indizes nicht als alleinige Funktion von F_{ij} darstellen, sondern ist auch explizit eine Funktion der Indizes, so daß dieser Bravais-Pearson-Koeffizient nicht die erforderliche Vergleichsinvarianz für ordinalskalierte Merkmale erfüllt. Die Anwendung auf die Indizes ist damit zwar für die Kovarianz, aber nicht für den Korrelationskoeffizienten zulässig. Für geeignete Streuungsmaße für ordinalskalierte Merkmale siehe z.B. Klein (1999a).
Über die im letzten Abschnitt hergeleiteten Grenzen des kovarianzbasierten Maßes läßt sich eine randverteilungsnormierte Maßzahl herleiten, die den Nachteil des Bravais-Pearson-Koeffizienten, nicht für ordinalskalierte Merkmale zulässig zu sein, nicht besitzt.

Setze

$$
\begin{aligned}
U(\mathbf{F}) &= \sum_{i=1}^{k}\sum_{j=1}^{l} \max\left(-F_{il}F_{kj}, -(1-F_{il})(1-F_{kj})\right), \\
O(\mathbf{F}) &= \sum_{i=1}^{k}\sum_{j=1}^{l} \min\left(F_{il}(1-F_{kj}), F_{kj}(1-F_{il})\right),
\end{aligned}
$$

dann ist

$$
\nu^*(\mathbf{F}) = \begin{cases} \nu(\mathbf{F})/(-U(\mathbf{F})) & \text{für } \nu(\mathbf{F}) < 0 \\ \nu(\mathbf{F})/O(\mathbf{F}) & \text{für } \nu(\mathbf{F}) > 0 \end{cases}
$$

ein randverteilungsnormiertes Maß. Es hält die positive Quadrantenordnung ein, da $U(\mathbf{F})$ und $O(\mathbf{F})$ nur durch die kumulierten Randhäufigkeiten bestimmt sind und sich damit die Ordnungseigenschaften von der Kovarianz ν auf die normierte Kovarianz ν^* übertragen. Da sich die Grenzen bei Spiegelung aber verändern, muß separat überprüft werden, wie sich ν^* bei doppelter und einfacher Spiegelung verhält. Es läßt sich nun mit wenigen Umformungen zeigen, daß

1. $U(\mathbf{F}_{U,-V}) = U(\mathbf{F}_{-U,V}) = -O(\mathbf{F}_{U,V})$,
2. $O(\mathbf{F}_{U,-V}) = O(\mathbf{F}_{-U,V}) = -U(\mathbf{F}_{U,V})$ und
3. $U(\mathbf{F}_{-U,-V}) = U(\mathbf{F}_{U,V})$ bzw. $O(\mathbf{F}_{-U,-V}) = O(\mathbf{F}_{U,V})$

sind. Damit ist offensichtlich $\nu^*(\mathbf{F}_{-U,-V}) = \nu^*(\mathbf{F}_{U,V})$. Sei $\nu(\mathbf{F}_{U,-V}) < 0$. Dann ist

$$
\nu^*(\mathbf{F}_{U,-V}) = \nu(\mathbf{F}_{U,-V})/(-U(\mathbf{F}_{U,-V})) = -\nu(\mathbf{F}_{U,V})/O(\mathbf{F}_{U,V}) = -\nu^*(\mathbf{F}_{U,V}),
$$

da $\nu^*(\mathbf{F}_{U,V}) > 0$ ist. Analog können die Fälle $\nu(\mathbf{F}_{U,-V}) > 0$ bzw. $\nu(\mathbf{F}_{-U,V}) > 0$ und $\nu(\mathbf{F}_{-U,V}) > 0$ diskutiert werden. Somit erfüllt ν^* die bezüglich Spiegelungen geforderten Eigenschaften und damit alle Axiome und die Randverteilungsnormiertheit.

9.7 Verallgemeinerung der Kovarianz

Wiederum lassen sich monoton zunehmende Scorefunktionen $a(.)$ und $b(.)$ einführen und die Kovarianz

$$
\sum_{i=1}^{k}\sum_{j=1}^{l} a(i)b(j)f_{ij} - \sum_{i=1}^{k} a(i)f_{i.} \sum_{j=1}^{l} b(j)f_{.j}
$$

der gewichteten Indizes der Merkmalsausprägungen von U und V betrachten. Bereits diese Kovarianz kann die Vergleichsinvarianz für ordinalskalierte Merkmale nicht erfüllen. Dies gilt erst recht für den zugehörigen Bravais-Pearson-Korrelationskoeffizient angewendet auf $a(1), \ldots, a(k)$ und $b(1), \ldots, b(l)$. Für diesen schreiben wir kurz $\rho_{BP}(a(U^*), b(V^*))$. Er ist monoton zunehmend in F_{ij} wegen $(a_{i+1} - a_i)(b_{j+1} - b_j) \geq 0$ für $i = 1, 2, \ldots, k-1$ und $j = 1, 2, \ldots, l-1$ und hält somit die positive Quadrantenordnung ein.

Eng verbunden sind diese auf Scorefunktionen basierenden Maßzahlen mit der Eigenschaft der positiven Assoziation von Esary, Proschan & Walkup (1967). Diese definieren

$$(U,V) \text{ positiv assoziiert } :\Longleftrightarrow \quad \mathrm{Cov}(g(U,V),h(U,V)) \geq 0$$

für jede in beiden Komponenten monoton zunehmende Funktion g und h und beweisen, daß aus dieser Ordnung die positive Quadranteneigenschaft folgt (siehe Esary & Proschan (1972), S. 653).
Aus der Assoziationseigenschaft haben Kimeldorf & Sampson (1984) eine positive Assoziationsordnung abgeleitet:

$$\begin{aligned}(U,V) \text{ positiver assoziiert als } (U',V') :\Longleftrightarrow \quad & \mathrm{Cov}(g(U,V),h(U,V)) \\ \geq \quad & \mathrm{Cov}(g(U',V'),h(U',V'))\end{aligned}$$

für jede in beiden Komponenten monoton zunehmende Funktion g und h, von der sich allerdings zeigen läßt, daß sie die von Fréchet und Hoeffding hergeleiteten Grenzen nicht einhalten.
Setzt man $g(U^*,V^*) = a(U^*)$ und $h(U^*,V^*) = b(V^*)$ mit a und b monoton zunehmende Scorefunktionen, dann folgt aus der Assoziationsordnung unmittelbar, daß

$$\begin{aligned}(U^*,V^*) \text{ positiver assoziiert als } ((U^*)',(V^*)') \Longrightarrow \quad & \rho_{BP}(a(U^*),b(V^*)) \\ \geq \quad & \rho_{BP}(a((U^*)'),b((V^*)'))\end{aligned}$$

gilt.

9.8 Randverteilungsnormiertes Maß von Vogel und Wiede

Von Vogel & Wiede (1994) wurde ein Maß vorgeschlagen, das explizit die Bedingung der Randverteilungsnormiertheit erfüllen muß.
Als Ausgangspunkt für die Herleitung der Maßzahl dient ein Abstandsmaß d zwischen bivariaten kumulierten Häufigkeitsverteilungen. Sie schlagen als Abstandsmaß den L_1-Abstand

$$d(\mathbf{F},\mathbf{G}) = 1/(kl-1)\sum_{i=1}^{k}\sum_{j=1}^{l}|F_{ij} - G_{ij}|$$

vor. Dann nimmt

$$\eta(\mathbf{F}) = \begin{cases} 1 - d(\mathbf{F},\mathbf{F}^u)/d(\mathbf{F}^0,\mathbf{F}^u) & \text{für } d(\mathbf{F},\mathbf{F}^u) \leq d(\mathbf{F}^0,\mathbf{F}^u) \\ d(\mathbf{F},\mathbf{F}^l)/d(\mathbf{F}^0,\mathbf{F}^l) - 1 & \text{für } d(\mathbf{F},\mathbf{F}^u) > d(\mathbf{F}^0,\mathbf{F}^u) \end{cases}$$

per Konstruktion für die Fréchet-Hoeffding-Grenzen die extremen Werte -1 und $+1$ an, d.h. η ist randverteilungsnormiert.

Unabhängigkeit: $\eta(\mathbf{F}) = 0$ ist lediglich eine notwendige, aber keine hinreichende Bedingung für Unabhängigkeit, da aus

$$\sum_{i=1}^{k}\sum_{j=1}^{l}|F_{ij} - \max(0, F_{il} + F_{kj} - 1)| = \sum_{i=1}^{k}\sum_{j=1}^{l}|F_{il}F_{kj} - \max(0, F_{il} + F_{kj} - 1)|$$

nicht geschlossen werden kann, daß $F_{ij} = F_{il}F_{kj}$ für $i = 1,2,\ldots,k$ und $j = 1,2,\ldots,l$ gilt.

Monotonie: Daß dieses Maß auch monoton zunehmend in F_{ij} bei gegebenen F_{il} und F_{kj} ist, zeigt die folgende Überlegung:
Sei $F_{ij} \leq F'_{ij}$ mit $F_{kj} = F'_{kj}$ und $F_{il} = F'_{il}$ für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$. Dann sind $F_{ij} - F^u_{ij} \leq F'_{ij} - F^u_{ij} \leq 0$ und $0 \leq F_{ij} - F^l_{ij} \leq F'_{ij} - F^l_{ij}$ für $i = 1, 2, \ldots, k$ und $j = 1, 2, \ldots, l$, so daß

$$\sum_{i=1}^{k} \sum_{j=1}^{l} |F_{ij} - F^u_{ij}| \geq \sum_{i=1}^{k} \sum_{j=1}^{l} |F'_{ij} - F^u_{ij}|$$

bzw.

$$\sum_{i=1}^{k} \sum_{j=1}^{l} |F_{ij} - F^l_{ij}| \leq \sum_{i=1}^{k} \sum_{j=1}^{l} |F'_{ij} - F^l_{ij}|,$$

so daß

$$\eta(\mathbf{F}) \leq \eta(\mathbf{F}')$$

folgt.

Spiegelung: Wir betrachten stellvertretend einen Fall, da sich alle anderen analog nachweisen lassen.
Sei $d(\mathbf{F}_{U,-V}, \mathbf{F}^u{}_{U,-V}) \leq d(\mathbf{F}^0{}_{U,-V}, \mathbf{F}^u{}_{U,-V})$. Dann ist

$$\begin{aligned}
\eta(\mathbf{F}_{U,-V}) &= 1 - \frac{\sum_i \sum_j |F_U(i) - F_{U,V}(i, l-j) - \min(F_U(i), 1 - F_V(l-j))|}{\sum_i \sum_j |F_U(i)(1 - F_V(l-j) - \min(F_U(i), 1 - F_V(l-j))|} \\
&= 1 - \frac{\sum_i \sum_j |-F_{U,V}(i, l-j) + \max(0, F_U(i) + F_V(l-j) - 1)|}{\sum_i \sum_j |-F_U(i)F_V(l-j) + \max(0, F_U(i) + F_V(l-j) - 1)|} \\
&= 1 - \frac{d(\mathbf{F}_{U,V}, \mathbf{F}^l{}_{U,V})}{d(\mathbf{F}^0{}_{U,V}, \mathbf{F}^l{}_{U,V})} = -\left(\frac{d(\mathbf{F}_{U,V}, \mathbf{F}^l{}_{U,V})}{d(\mathbf{F}^0{}_{U,V}, \mathbf{F}^l{}_{U,V})} - 1\right) = -\eta(\mathbf{F}_{U,V}).
\end{aligned}$$

Mithin ist auch das Maß von Vogel & Wiede ein bivariates Abhängigkeitsmaß für ordinalskalierte Merkmale.

10 Zusammenfassung

Es wurden elementare axiomatische Anforderungen für die bivariate Abhängigkeitsmessung ordinalskalierter Merkmale zusammengestellt und für gängige Maßzahlen überprüft, ob diese den Anforderungen genügen. Dabei wurden bewußt Ratingskalen mit wenigen Ausprägungen in den Mittelpunkt gerückt, da für diese große Mengen von Bindungen in größeren Datensätzen unvermeidlich sind. Das Fehlen von Bindungen ist nur dann erklärlich, wenn Ränge aus den Merkmalswerten stetiger quantitativer Merkmale gebildet werden. Dies ist bei Rating-Skalen explizit nicht der Fall.
Es werden lediglich Maßzahlen diskutiert, die Funktionen der kumulierten relativen Häufigkeiten sind. Diese Maßzahlen dürfen nicht explizit von den Merkmalsausprägungen,

sondern nur von ihren Häufigkeiten abhängen. Diese grundlegende Einschränkung folgt aus der meßtheoretischen Forderung, daß mit "sinnvollen" Maßzahlen unabhängig von willkürlich ausgewählten Skalen, auf denen die Merkmalsausprägungen gemessen werden, Vergleiche vorgenommen werden können. Dann nämlich kann diese Maßzahl nur von sog. Maximalinvarianten abhängen, die im Falle zweier unabhängiger Ordinalskalen für zwei Merkmale, deren Abhängigkeit beurteilt werden soll, die erwähnten kumulierten relativen Häufigkeiten sind.

Im Rahmen der Axiomatik stand die positive Quadrantenordnung im Mittelpunkt, die allerdings nur den Vergleich von bivariaten Häufigkeitsverteilungen mit identischen Randverteilungen ermöglicht. Im Falle nicht-identischer Randverteilungen gibt es zur Zeit noch kein überzeugendes Ordnungskonzept, so daß diese zugegebenermaßen eingeschränkt verwendbare Ordnung benutzt wurde.

Es hat sich gezeigt, daß insbesondere die Konkordanzmaße (z.B. das Maß von Goodman & Kruskal und das Maß von Kendall) alle Anforderungen erfüllen. Der Spearmansche Rangkorrelationskoeffizient und generell die aus der nichtparametrischen Statistik als Testgrößen bekannten linearen Rangordnungsstatistiken erfüllen alle Anforderungen, wenn keine Bindungen vorliegen. Sie erhalten zwar auch im Falle des Vorliegens von Bindungen die positive Quadrantenordnung, scheitern aber, wenn es um die Spiegelung von Verteilungen geht. Midranks mildern dieses Problem, können es aber nicht gänzlich beseitigen. Das in der Praxis gerne vorgenommene Verfahren, Methoden für quantitative Merkmale (wie den Bravais-Pearson-Korrelationskoeffizienten) auf die Indizes der ordinalskalierten Merkmale anzuwenden, führt zwar für die Kovarianz zum Erfolg. Der Bravais-Pearson-Korrelationskoeffizient der Indizes erfüllt aber nicht die entscheidende meßtheoretische Anforderung, daß mit statistischen Maßzahlen "sinnvolle' Vergleiche möglich sein müssen. Ein in jüngster Zeit von Vogel und Wiede vorgeschlagenes Maß zielt zwar auf die Randverteilungsnormiertheit ab, d.h. nimmt die extremen Werte -1 und $+1$ für die Kontingenztabellen an, für die bei gegebenen Randverteilungen im Sinne der positiven Quadrantenordnung minimale bzw. maximale positive Abhängigkeit besteht (sog. Fréchet-Hoeffding-Grenzen). Diese Eigenschaften besitzen aber auch die erwähnten und hinreichend in die statistische Praxis eingeführten Konkordanzmaße. Dasselbe gilt auch für ein hier vorgeschlagenes, normiertes kovarianzbasiertes Maß.

Literatur

1. Averous, J. & Dortet-Bernadet, J.-L. (2000). LTD and RTI dependence orderings. *Canadian Journal of Statistics* **28**, 151-157.

2. Barlow, R.E. & Proschan, F. (1975). *Statistical theory of reliability and life testing.* Holt, Rhinehart & Winston, New York.

3. Block, H.W., Savitis, T. & Shaked, M. (1982). Some conceps of negative dependence. *Annals of Probability* **10**, 765-772.

4. Blomqvist, N. (1950). On a measure of dependence between two random variables. *Annals of Mathematical Statistics.* **21**, 593-600.

5. Capéràa, P. & Genest, C. (1990). Concepts de dependence et ordres stochastiques pour des lois bidimensionelles. *Canadian Journal of Statistics* **18**, 315-326.

6. Dabrowska, D. (1985). Descriptive parameters of location, dispersion and stochastic dependence. *Statistics* **16**, 63-88.

7. Esary, J.D. & Proschan, F. (1972). Relationships between some concepts of bivariate dependence. *Annals of Mathematical Statistics* **43**, 651-655.

8. Esary, J.D., Proschan, F. & Walkup, D.W. (1967). Association of random variables. *Annals of Mathematical Statistics* **38**, 1466-1474.

9. Fréchet, M. (1951). Sur les tableaux de correlation dont les marges sont données. *Ann. Univ.: Lyon Sect. A* **14**, 53-77.

10. Hájek, J. & Šidák, Z. (1969). *Theory of rank tests.* Academic Press, New York.

11. Hoeffding, W. (1940). Maßstabsinvariante Korrelationstheorie. *Schriften des Mathematischen Instituts für Angewandte Mathematik der Universität Berlin* **5**, 179-233.

12. Karlin, S. (1968). *Total positivity, Vol. I.* Stanford University Press, Stanford.

13. Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika* **30**, 81-93.

14. Kendall, M.G. (1970). *Rank correlation methods.* Griffin, London.

15. Kimeldorf, G. & Sampson, A.R. (1978). Monotone dependence. *Annals of Statistics* **6**, 895-903.

16. Kimeldorf, G. & Sampson, A.R. (1984). Concordant and discordant monotone correlations and their evaluation by nonlinear optimization. *Studies in the Management Sciences (19): Optimization in Statistics,* (eds. Zanakis and Rustagi), 117-130, North-Holland, Amsterdam.

17. Kimeldorf, G. & Sampson, A.R. (1987). Positive dependence orderings. *Annals of the Institute of Statistical Mathematics* **39**, 113-128.

18. Kimeldorf, G. & Sampson, A.R. (1989). A framework for positive dependence. *Annals of the Institute of Statistical Mathematics* **41**, 31-45.

19. Klein, I. (1984). *Das Problem der Auswahl geeigneter Maßzahlen in der deskriptiven Statistik.* Physica-Verlag, Würzburg.

20. Klein, I. (1994). *Mögliche Skalentypen, invariante Relationen und wissenschaftliche Gesetze.* Verlag Vandenhoeck & Ruprecht, Göttingen.

21. Klein, I. (1999a). Rangordnungsstatistiken als Verteilungsmaßzahlen für ordinalskalierte Merkmale. I. Streuungsmessung. *Diskussionspapier der Lehrstühle für Statistik der Universität Erlangen-Nürnberg* **27/1999**.

22. Klein, I. (1999b). Rangordnungsstatistiken als Verteilungsmaßzahlen für ordinalskalierte Merkmale. II. Schiefemessung. *Diskussionspapier der Lehrstühle für Statistik der Universität Erlangen-Nürnberg* **28/1999**.

23. Kowalczyk, T. & Pleszczyńska, E. (1977). Monotonic dependence functions of bivariate distributions. *Annals of Statistics* **5**, 1221-1227.

24. Krantz, D.H., Luce, R.D., Suppes, P. & Tversky, A. (1971). *Foundations of measurement. Volume I: Additive and polynomial representations.* Academic Press, San Diego.

25. Lehmann, E.L. (1959). *Testing statistical hypothesis.* Wiley, New York.

26. Lehmann, E.L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics* **37**, 1137-1153.

27. Mosler, K.C. & Scarsini, M. (1993). *Stochastic orders and application: a classified bibliography.* Springer, Berlin.

28. Nelsen, R.B. (1997). Dependence and order in families of archimedian copulas. *Journal of Multivariate Analysis* **60**, 111-127.

29. Pfanzagl, J. (1968). *Theory of measurement.* Physica-Verlag, Würzburg.

30. Raveh, A. (1986). On measures of association. *American Statistician* **40**, 117-123.

31. Rényi, A. (1959). On measures of dependence. *Acta Math. Acad. Sci. Hungar.* **10**, 441-451.

32. Roberts, F.S. (1979). *Measurement theory.* Addison-Wesley, Reading.

33. Schneider, W., Kornrumpf, J. & Mohr, W. (1995). *Statistische Methodenlehre.* Oldenbourg-Verlag, München.

34. Schweizer, B. & Wolff, E.F. (1981). On nonparametric measures of dependence for random variables. *Annals of Statistics* **9**, 879-885.

35. Spearman, C. (1904). The proof and measurement of association between things. *American Journal of Psychology* **15**, 72-101.

36. Tchen, A.H. (1980). Inequalities for distributions with given marginals. *Annals of Probability* **8**, 814-827.

37. Tukey, J.W. (1958). A problem of Berkson, and minimum variance oderly estimators. *Annals of Mathematical Statistics* **29**, 588-592.

38. Vogel, F. & Wiede, T. (1994). Ein neues Zusammenhangsmaß für ordinalskalierte Merkmale. *Jahrbücher für Nationalökonomie und Statistik* **213**, 1-30.

39. Yanagimoto, T. & Okamato, M. (1969). Partial orderings of permutations and monotonicity of a rank correlation statistic. *Annals of the Institute of Statistical Mathematics* **21**, 451-506.