

# Statistik Teil 5: Kontingenzkoeffizienten

M. Kohl, F. Münch

Die KARDIOTECHNIK stellt in der Rubrik Tutorials relevante Methoden für wissenschaftliche Arbeiten zur klinischen Perfusion vor.

## EINFÜHRUNG

Neben quantitativen und ordinalen Merkmalen spielen auch nominale Merkmale und deren Zusammenhänge eine wichtige Rolle in der Medizin. Beispiele hierfür sind etwa das Geschlecht, Grunderkrankungen, verabreichte Medikamente, durchgeführte Operationen oder auch das Outcome der Patient:innen. Die Stärke des Zusammenhangs zweier nominaler Merkmale lässt sich jedoch nicht mit Hilfe von Korrelationen bestimmen, da man für die Definition eines Korrelationskoeffizienten zumindest eine Anordnung, also zwei ordinale Merkmale, benötigt. Stattdessen werden hierfür sogenannte Kontingenzkoeffizienten herangezogen, mit denen wir uns im Folgenden genauer beschäftigen werden. Wie im Fall der Korrelationen darf Kontingenz nicht mit einem Kausalzusammenhang gleichgesetzt werden [1].

## KONTINGENZTAFEL

Der erste Schritt in der Untersuchung des Zusammenhangs zweier nominaler Merkmale besteht in der Aufstellung der sogenannten Kontingenztafel oder Kreuztabelle (vgl. Abschnitt 7.7.11 in [2]). Hierbei handelt es sich um eine Tabelle, welche die (absoluten) Häufigkeiten für die Merkmalskombinationen der beiden Merkmale beinhaltet (vgl. Abb. 1).

Gibt es keinen Zusammenhang zwischen X und Y bzw. mathematisch genauer ausgedrückt, sind X und Y stochastisch unabhängig, so gilt im Fall der  $r \times s$ -Kontingenztafel aus Abb. 1:

$$P(X = x_i \text{ und } Y = y_j) = P(X = x_i) \cdot P(Y = y_j) = \frac{n_{i\bullet}}{n} \cdot \frac{n_{\bullet j}}{n}$$

Prof. Dr. Matthias Kohl  
Department of Medical and Life Sciences  
Institute of Precision Medicine  
Hochschule Furtwangen  
Jakob-Kienzle-Str. 17,  
78054 Villingen-Schwenningen (Germany)  
E-Mail: kohl@hs-furtwangen.de  
www.hs-furtwangen.de  
www.life-data-science.org

## Fazitbox

### PRO UND CONTRA KONTINGENZKOEFFIZIENTEN:

#### Pro

- Mit Kontingenzkoeffizienten lassen sich auf einfache Weise Zusammenhänge zwischen zwei nominalen Variablen untersuchen.
- Der am besten standardisierte Kontingenzkoeffizient ist Cramér's V, welcher daher für praktische Anwendungen zum Einsatz kommen sollte.

#### Contra

- Kontingenz darf nicht mit einem Kausalzusammenhang gleichgesetzt werden.
- Für Cramér's V hängt die Bewertung der Effektstärke von der Größe der Kontingenztafel ab.
- Das Vorliegen von Kontingenz, also einer stochastischen Abhängigkeit, ist deutlich schwieriger zu veranschaulichen als das Vorliegen einer Korrelation. Es eignen sich hierfür zum Beispiel Mosaikdiagramme.

für  $i = 1, 2, \dots, r$  und  $j = 1, 2, \dots, s$ . Hieraus folgt für die erwartete Anzahl von Beobachtungen in Zelle  $ij$

$$e_{ij} = n \cdot \frac{n_{i\bullet}}{n} \cdot \frac{n_{\bullet j}}{n} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}$$

Diese erwartete Anzahl, die man unter der Annahme erhält, dass X und Y stochastisch unabhängig sind, kann nun mit der tatsächlich beobachteten Anzahl  $n_{ij}$  verglichen werden. Hierfür wird meist die folgende  $\chi^2$ -Statistik herangezogen:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

Sind X und Y stochastisch unabhängig, so ist dies äquivalent zu  $n_{ij} = e_{ij}$  für alle  $i = 1, 2, \dots, r$  und  $j = 1, 2, \dots, s$  und folglich auch äquivalent zu  $\chi^2 = 0$ . Ergibt sich demnach für zwei Merkmale X und Y, dass  $\chi^2 > 0$  ist, so müssen X und Y stochastisch abhängig sein. Hierbei ist zu beachten, dass bei realen Datensätzen gewisse zufällige Abweichungen von den erwarteten Werten auftreten, weshalb wir nahezu immer  $\chi^2 > 0$  bekommen. Wollen wir uns sicher sein, dass tatsächlich eine stochastische Abhän-

gigkeit vorliegt, sollten wir zusätzlich ein Konfidenzintervall für die  $\chi^2$ -Statistik oder den zugehörigen  $\chi^2$ -Test berechnen. Ist der linke Rand des 95 %-Konfidenzintervalls größer als 0, so können wir von einem signifikanten Zusammenhang zwischen X und Y sprechen. Der Wert der  $\chi^2$ -Statistik sagt außerdem nichts über die Stärke des Zusammenhangs aus, da der Wert von der vorliegenden Stichprobengröße n abhängt. Ein Ausweg besteht darin, die  $\chi^2$ -Statistik

Gegeben sei eine Stichprobe der Größe n, in der die beiden nominalen Merkmale X und Y die Ausprägungen  $x_1, x_2, \dots, x_r$  ( $r \geq 2$ ) und  $y_1, y_2, \dots, y_s$  ( $s \geq 2$ ) annehmen können. Weiter sei  $n_{ij}$  die Anzahl an Beobachtungen (Patienten), welche die Merkmalskombination  $x_i$  ( $i = 1, 2, \dots, r$ ) und  $y_j$  ( $j = 1, 2, \dots, s$ ) aufweisen und es seien

$$n_{\bullet j} = n_{1j} + n_{2j} + \dots + n_{rj} \quad (\text{Spaltensumme für } j = 1, 2, \dots, s)$$

und

$$n_{i\bullet} = n_{i1} + n_{i2} + \dots + n_{is} \quad (\text{Zeilensumme für } i = 1, 2, \dots, r)$$

Entsprechend folgt:

$$n_{\bullet 1} + n_{\bullet 2} + \dots + n_{\bullet s} = n = n_{1\bullet} + n_{2\bullet} + \dots + n_{r\bullet}$$

Mit Hilfe dieser Bezeichnungen ergibt sich die folgende allgemeine Form einer  $r \times s$ -Kontingenztafel ( $r \geq 2, s \geq 2$ )

		Y				
		$y_1$	$y_2$	$\dots$	$y_s$	Summe
X	$x_1$	$n_{11}$	$n_{12}$	$\dots$	$n_{1s}$	$n_{1\bullet}$
	$x_2$	$n_{21}$	$n_{22}$	$\dots$	$n_{2s}$	$n_{2\bullet}$
	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$	$\vdots$
	$x_r$	$n_{r1}$	$n_{r2}$	$\dots$	$n_{rs}$	$n_{r\bullet}$
Summe	$n_{\bullet 1}$	$n_{\bullet 2}$	$\dots$	$n_{\bullet s}$	$n$	

Abb. 1: Definition der allgemeinen  $r \times s$ -Kontingenztafel

Effektstärke	$\phi$ -Koeffizient	
stark	$\phi \geq 0,5$	
moderat	$0,3 \leq \phi < 0,5$	
klein	$0,1 \leq \phi < 0,3$	
vernachlässigbar	$\phi < 0,1$	

Tab. 1: Effektstärken für den  $\phi$ -Koeffizienten

zu standardisieren. Dies führt uns auf die sogenannten Kontingenzkoeffizienten.

### KONTINGENZKOEFFIZIENT VON PEARSON

Bereits im Jahr 1904 führte Pearson einen ersten Kontingenzkoeffizienten CC ein [3]. Er definierte seine Koeffizienten wie folgt:

$$CC = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

Leider ist auch dieser Koeffizient etwas schwierig zu interpretieren, da das Maximum gleich  $\sqrt{(M-1)/M}$  mit  $M = \min\{r, s\}$  ist. Das Maximum hängt somit von der minimalen Anzahl von Merkmalsausprägungen von X und Y ab. Alternativ sollte daher besser der folgende normierte Pearson-Koeffizient verwendet werden (vgl. Abschnitt 7.7.11.1 in [2]):

$$CC_{\text{norm}} = \sqrt{\frac{M}{M-1}} \cdot CC$$

Damit ergibt sich ein Koeffizient, welcher die Werte von 0 bis 1 annehmen kann, wobei X und Y genau dann stochastisch unabhängig sind, wenn  $CC_{\text{norm}} = 0$  und genau dann vollständig stochastisch abhängig sind, wenn  $CC_{\text{norm}} = 1$ .

### $\phi$ -KOEFFIZIENT

Pearson [3] beschreibt außerdem den folgenden  $\phi$ -Koeffizienten, welchen man als mittlere quadratische Kontingenz bezeichnen kann (vgl. Abschnitt 7.7.11.1 in [2]):

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

Der  $\phi$ -Koeffizient liegt immer im Intervall  $[0, 1]$ , wobei X und Y genau dann stochastisch unabhängig sind, wenn  $\phi = 0$ . Der Wert 1 wird aber nur angenommen, falls  $r = 2$  und/oder  $s = 2$ . Betrachtet man den Fall  $r = s = 2$  und kodiert die nominalen Variablen mit 0 und 1, so ist der  $\phi$ -Koeffizient gerade identisch zur Pearson-Korrelation (vgl. [1]) und zum Korrelationskoeffizienten von Matthews (MCC) [4], welcher im Kontext binärer Klassifikationen verwendet wird. Der

$\phi$ -Koeffizient ist außerdem identisch zu Cohen's  $w$  [5]. Die Effektstärken, die Cohen für sein  $w$  und damit für den  $\phi$ -Koeffizienten vorschlägt, finden sich in Tab. 1.

### KONTINGENZKOEFFIZIENT VON CRAMÉR

Cramér [6] führt den folgenden Kontingenzkoeffizienten ein, der auch als Cramér's V bezeichnet wird (vgl. Abschnitt 7.7.11.1 in [2]):

$$V = \sqrt{\frac{\chi^2}{n \cdot (M-1)}} = \frac{\phi}{\sqrt{M-1}}$$

wobei  $M = \min\{r, s\}$ . Der Nenner  $n \cdot (M-1)$  entspricht gerade dem Maximum der  $\chi^2$ -Statistik, weshalb Cramér's V alle Werte zwischen 0 und 1 annehmen kann. X und Y sind genau dann stochastisch unabhängig, wenn  $V = 0$  und genau dann vollständig stochastisch abhängig, wenn  $V = 1$ . Obwohl dies auch für den normierten Pearson-Koeffizienten  $CC_{\text{norm}}$  gilt, sind V und  $CC_{\text{norm}}$  trotzdem nicht identisch und V sollte  $CC_{\text{norm}}$  vorgezogen werden. Für die Bewertung der Stärke des Zusammenhangs von X und Y auf der Basis von V gibt es verschiedenste Vorschläge, die wie im Fall der Korrelation auch vom entsprechenden Kontext abhängen. Überträgt man den Vorschlag von Cohen [5] für den  $\phi$ -Koeffizient auf Cramér's V, so ergeben sich aufgrund des oben dargestellten Zusammenhangs zwischen V und  $\phi$  die Effektstärken in Tab. 2. Die Bewertung der Effektstärke hängt folglich von der Größe der Kontingenztafel ab. Für größere Kontingenztafeln gelten demnach kleinere Grenzen.

Effektstärke	Cramér's V
stark	$0,5 / \sqrt{M-1} \leq V$
moderat	$0,3 / \sqrt{M-1} \leq V < 0,5 / \sqrt{M-1}$
klein	$0,1 / \sqrt{M-1} \leq V < 0,3 / \sqrt{M-1}$
vernachlässigbar	$V < 0,1 / \sqrt{M-1}$

Tab. 2: Effektstärken für Cramér's V mit  $M = \min\{r, s\}$

### BEISPIEL

Wir untersuchen den Zusammenhang zwischen dem Operateur ( $r = 5$ ) und der gewählten Myokardprotektion ( $s = 4$ ) bei kardioplegischem Herzstillstand auf der Basis von 627 Operationen (eigene Daten). Wir

erhalten die in Tab. 3 dargestellte Kontingenztafel.

Wir veranschaulichen den Zusammenhang mit einem sogenannten Mosaikdiagramm (vgl. Abb. 2).

Auf der linken Seite von Abb. 2 sehen wir ein Mosaikdiagramm der beobachteten Häufigkeiten, auf der rechten Seite ein Mosaikdiagramm der erwarteten Häufigkeiten. Hierbei sind die rechteckigen Felder proportional zu den Häufigkeiten. Den auffälligsten Unterschied sehen wir bei Operateur 3. Insgesamt sehen wir aber nur relativ geringe Verschiebungen zwischen dem linken und dem rechten Diagramm. Wir gehen daher von einem eher schwachen Zusammenhang zwischen dem Operateur und dem Myokardprotektionsverfahren aus. Dies wollen wir nun mit Hilfe der oben vorgestellten Kontingenzkoeffizienten ausdrücken und verwenden hierfür die Statistiksoftware R [7]. Zur Berechnung der 95 %-Konfidenzintervalle (CI95) verwenden wir die Bootstrap-Bc<sub>a</sub>-Methode mit 9999 Wiederholungen (vgl. doi Supplement <https://doi.org/10.47624/kt.031.ZZUR4312> von [10]). Wir erhalten eine  $\chi^2$ -Statistik von  $\chi^2 = 33,31$  (CI95: 11,96 – 45,61). Da der linke Rand des CI95 größer 0 ist, können wir schließen, dass ein signifikanter Zusammenhang zwischen dem Operateur und dem Verfahren zur Myokardprotektion besteht. Wir wissen jedoch noch nicht, wie stark dieser Zusammenhang tatsächlich ist. Der Kontingenzkoeffizient von Pearson liegt bei  $CC = 0,225$  (CI95: 0,133 – 0,261) und der normierte Pearson-Koeffizient bei  $CC_{\text{norm}} = 0,259$  (CI95: 0,163 – 0,299). Als  $\phi$ -Koeffizient ergibt sich  $\phi = 0,230$  (CI95: 0,141 – 0,269), was einer kleinen Effektstärke und demnach einem schwachen Zusammenhang entspricht (vgl. Tab. 1). Für Cramér's V schließlich erhalten wir  $V = 0,133$  (CI95:

0,079 – 0,156). Auch dies entspricht einem schwachen Zusammenhang, da es zwischen  $0,1/\sqrt{3} = 0,057$  und  $0,3/\sqrt{3} = 0,173$  liegt (vgl. Tab.2). Da die linken Ränder der CI95 aller Kontingenzkoeffizienten größer 0 sind, können wir ana-

log zur  $\chi^2$ -Statistik schließen, dass ein signifikanter Zusammenhang zwischen dem Operateur und dem Myokardprotektionsverfahren vorliegt. Aufgrund der Werte des  $\phi$ -Koeffizienten und von Cramér's V ist dieser Zusammenhang, wie bereits anhand

Operateur	Myokardprotektionsverfahren				Summe
	Beating Heart	Custodiol	Mikroplegie	Custodiol + Mikroplegie	
1	4	15	45	6	70
2	19	58	175	25	277
3	0	3	22	0	25
4	24	49	86	13	172
5	16	22	42	3	83
Summe	63	147	370	47	627

Tab. 3: Kontingenztafel für die Untersuchung des Zusammenhangs zwischen Operateur und dem gewählten Myokardprotektionsverfahren (Beating Heart, Custodiol, Mikroplegie und Kombiverfahren aus Custodiol + Mikroplegie)

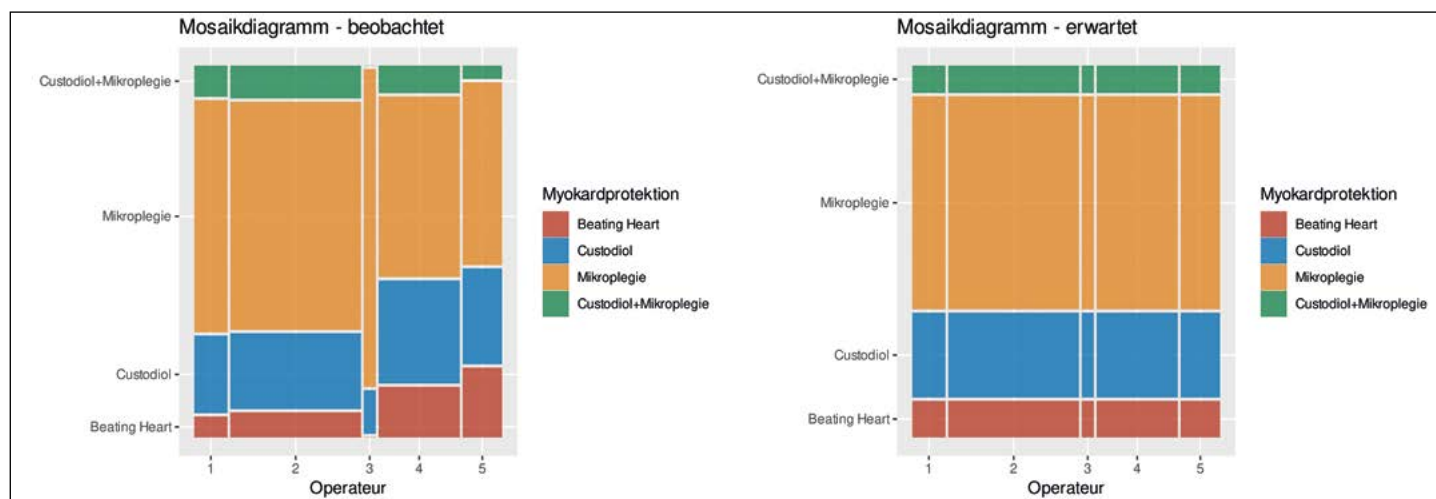


Abb. 2: Mosaikdiagramme der beobachteten (links) und erwarteten (rechts) Häufigkeiten für die nominalen Merkmale Operateur und Myokardprotektionsverfahren auf Basis von 627 Operationen (erstellt mit der Statistiksoftware R [7] und den R Paketen ggplot2 [8] und ggmosaic [9])

von Abb. 2 vermutet, aber nur schwach und daher vermutlich ohne klinische Relevanz.

## ZUSAMMENFASSUNG

Mit Kontingenzkoeffizienten lassen sich auf einfache Weise Zusammenhänge zwischen zwei nominalen Variablen untersuchen. Von den verschiedenen vorgestellten Kontingenzkoeffizienten eignet sich Cramér's V für praktische Anwendungen am besten, da dieser die geeignetste Standardisierung besitzt. Jedoch muss im Fall von Cramér's V für die Bewertung der Effektstärke auch die Größe der Kontingenztafel berücksichtigt werden. Im Unterschied zur Korrelation ist das Vorliegen einer Kontingenz (stochastischen Abhängigkeit) schwierig graphisch zu veranschaulichen. Man kann hierfür zum Beispiel sogenannte Mosaikdiagramme verwenden.

## LITERATUR

1. Kohl M, Münch F. Statistik Teil 4: Korrelationen. *Kardiotechnik* 2022; 31(4):146-149.
2. Hedderich J, Sachs L (2020). *Angewandte Statistik. Methodensammlung mit R*. 17. Auflage, Springer-Verlag.
3. Pearson K. *On the Theory of Contingency and its Relation to Association and Normal Correlation*. London 1904: Drapers' Co. *Memoirs (Biometric Series No. 1; pp 1-35)*.
4. Matthews BW. *Comparison of the predicted and observed secondary structure of T4 phage lysozyme*. *Biochim Biophys Acta* 1975 Oct 20; 405(2):442-51.
5. Cohen, J. *Statistical Power Analysis for the Social Sciences (2nd. Edition)*. Hillsdale 1988, New Jersey, Lawrence Erlbaum Associates.
6. Cramér H. *Mathematical Methods of Statistics*. Princeton 1946, NJ: Princeton Univ. Press; pp. 592.
7. R Core Team R: *A language and environment for statistical computing*. R Foundation for Statistical Computing 2022, Vienna, Austria. URL <https://www.R-project.org/>.
8. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag 2016, New York.
9. Jeppson H, Hofmann H, Cook D. *ggmosaic: Mosaic Plots in the 'ggplot2' Framework*. 2021; R package version 0.3.3.
10. Kohl M, Münch F. Statistik Teil 3: Konfidenzintervalle. *Kardiotechnik* 2022; 31(3):95-98. doi: 10.47624/kt.031.QQOV9624