

Fragen und Antworten zu Kapitel 21

(1) Zur Auswertung welcher Fragestellungen braucht man log-lineare Modelle?

Log-lineare Modelle dienen zur Analyse multidimensionaler Kontingenztafeln. Mit log-linearen Modellen können Scheinkorrelationen und maskierte Zusammenhänge bei kategorialen Variablen aufgedeckt werden und bedingte Zusammenhänge zwischen zwei Variablen, gegeben die Ausprägungen anderer Variablen, geschätzt werden.

(2) Was versteht man unter dem Simpson-Paradox?

Das Simpson-Paradox bezeichnet den Sachverhalt, dass sich der Zusammenhang zweier Variablen verändern kann, wenn er als bedingter Zusammenhang, gegeben die Ausprägungen einer dritten Variablen, bestimmt wird. Es beschreibt das Phänomen, dass der Zusammenhang in der Gesamtpopulation selbst dann nicht dem mittleren Zusammenhang in zwei Teilpopulationen entsprechen muss, wenn der Zusammenhang in beiden Gruppen gleich ist und die Gesamtpopulation sich vollständig in die beiden Teilpopulationen zerlegen lässt.

(3) Wie ist das multiplikative log-lineare Modell für eine 2×2-Tabelle definiert?

Das multiplikative Modell ist auf Stichprobenebene wie folgt definiert:

$$n_{ij} = \hat{\gamma} \cdot \hat{\gamma}_i^A \cdot \hat{\gamma}_j^B \cdot \hat{\gamma}_{ij}^{AB}.$$

Auf Populationsebene lautet es:

$$\pi_{ij} = \gamma^* \cdot \gamma_i^A \cdot \gamma_j^B \cdot \gamma_{ij}^{AB}.$$

(4) Wie ist das additive log-lineare Modell für eine 2×2-Tabelle definiert?

Auf Stichprobenebene ist das multiplikative Modell wie folgt definiert:

$$m_{ij} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_{ij}^{AB}.$$

Das Populationsmodell lautet:

$$\ln(\epsilon_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

(5) Wie erhält man das additive log-lineare Modell für eine 2 × 2-Tabelle aus dem multiplikativen Modell?

Man erhält die Koeffizienten des additiven Modells, indem man die Koeffizienten des multiplikativen Modells logarithmiert.

(6) Wie hängt der Koeffizient $\hat{\gamma}_{ij}^{AB}$ im log-linearen Modell für eine 2 × 2-Tabelle mit dem Kreuzproduktverhältnis zusammen?

Der Koeffizient $\hat{\gamma}_{ij}^{AB}$ entspricht der vierten Wurzel des Odds Ratio:

$$\hat{\gamma}_{11}^{AB} = \sqrt[4]{\frac{n_{11} \cdot n_{22}}{n_{12} \cdot n_{21}}} = \sqrt[4]{OR}$$

(7) Was versteht man unter einem

- (a) multinomialen Erhebungsschema,**
- (b) produkt-multinomialen Erhebungsschema,**
- (c) Poisson-Erhebungsschema?**

- (a) Beim multinomialen Erhebungsschema zieht man eine Stichprobe vorgegebener Größe und bestimmt die Häufigkeiten in den einzelnen Zellen.
- (b) Das produkt-multinomial Erhebungsschema unterscheidet sich vom multinomialen Erhebungsschema darin, dass die Randverteilung einer der beiden Variablen nicht frei variieren kann, sondern vorher festgelegt wird.
- (c) Beim Poisson-Erhebungsschema liegt die Größe der Stichprobe vorher nicht fest, sondern ist das Ergebnis einer Untersuchung, bei der man in einem Zeitabschnitt Daten erhebt.

(8) Wie lautet das Gleichverteilungsmodell für eine 2 × 2-Tabelle?

Das Gleichverteilungsmodell lautet: $\ln(\varepsilon_{ij}) = \lambda$.

(9) Wie lautet das Unabhängigkeitsmodell für eine 2 × 2-Tabelle?

Das Unabhängigkeitsmodell lautet: $\varepsilon_{ij} = \gamma \cdot \gamma_i^A \cdot \gamma_j^B$.

(10) Wie kann man anhand des Likelihood-Ratio-Differenzen-Tests Modellvergleiche durchführen?

Anhand des Likelihood-Ratio-Tests können Modelle gegeneinander getestet werden, die ineinander geschachtelt sind. Ein (restriktiveres) Modell ist in ein anderes (allgemeineres) Modell geschachtelt, wenn es durch spezifische Restriktionen aus dem allgemeineren Modell hervorgeht. Die Nullhypothese geht von der Gleichheit beider Modelle aus. Verwirft man die Nullhypothese, geht man davon aus, dass das weniger restriktive Modell die Daten besser beschreibt als das restriktivere Modell.

(11) Was versteht man unter hierarchischen log-linearen Modellen?

Hierarchische log-lineare Modelle zeichnen sich durch folgende Eigenschaft aus: Befindet sich ein Parameter im Modell, der eine spezifische Variable enthält, so müssen alle Parameter niedrigerer Ordnung ebenfalls im Modell enthalten sein, die diese Variable enthalten.

(12) Was versteht man unter Ockhams Rasiermesser (Occam's Razor)?

Unter Ockhams Rasiermesser versteht man das Sparsamkeitsprinzip in den Wissenschaften. Diesem Prinzip zufolge ist diejenige von mehreren Erklärungen (Theorien) vorzuziehen, die ein Phänomen am einfachsten erklärt.

(13) Wie ist das multiplikative log-lineare Modell für eine 2 × 2 × 2-Tabelle definiert?

Auf Stichprobenebene lautet das Modell:

$$n_{ijk} = \hat{\gamma} \cdot \hat{\gamma}_i^A \cdot \hat{\gamma}_j^B \cdot \hat{\gamma}_k^C \cdot \hat{\gamma}_{ij}^{AB} \cdot \hat{\gamma}_{jk}^{BC} \cdot \hat{\gamma}_{ik}^{AC} \cdot \hat{\gamma}_{ijk}^{ABC}$$

(14) Wie ist das additive log-lineare Modell für eine 2 × 2 × 2-Tabelle definiert?

Auf Stichprobenebene lautet das Modell:

$$m_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ij}^{AB} + \hat{\lambda}_{jk}^{BC} + \hat{\lambda}_{ik}^{AC} + \hat{\lambda}_{ijk}^{ABC}$$

(15) Wie kann man mit einem log-linearen Modell für eine $2 \times 2 \times 2$ -Tabelle Scheinzusammenhänge aufdecken?

Im log-linearen Modell für eine $2 \times 2 \times 2$ -Tabelle repräsentiert der Parameter höchster Ordnung Unterschiede im Zusammenhang zweier Variablen in Abhängigkeit von der Ausprägung einer dritten Variablen. Die Parameter, die den Zusammenhang zweier Variablen widerspiegeln, hängen vom gemittelten Zusammenhang zweier Variablen ab, wobei die (geometrische) Mittelung über die Ausprägungen der dritten Variablen erfolgt. Ist der Parameter höchster Ordnung und der Parameter zweithöchster Ordnung für ein Variablenpaar in der Population im multiplikativen Modell gleich 1 bzw. im additiven Modell gleich 0, obwohl ohne Berücksichtigung der Drittvariablen zwischen ihnen ein Zusammenhang besteht, dann handelt es sich bei diesem Zusammenhang um einen Scheinzusammenhang.

(16) Was versteht man unter strukturellen Nullzellen, was unter Stichproben-Nullzellen?

Strukturelle Nullzellen entstehen dadurch, dass bestimmte Kombinationen von Kategorien nicht möglich sind. Stichproben-Nullzellen entstehen dadurch, dass eine bestimmte Merkmalskombination in der Stichprobe nicht vorgekommen ist, obwohl diese Merkmalskombination möglich ist.