Lösungen der Übungsaufgaben aus Kapitel 20

(1) Zeigen Sie, dass im log-linearen Modell basierend auf einer Referenzkategorie $\hat{\delta}_{11}^{AB} = \ln\left(\frac{n_{11} \cdot n_{22}}{n_{12} \cdot n_{21}}\right)$ gelten muss.

Für die Zellkombination (1,1) gilt nach Gleichung F 20.19:

$$\ln(n_{11}) = \hat{\delta} + \hat{\delta}_1^A + \hat{\delta}_1^B + \hat{\delta}_{11}^{AB}$$

Hieraus erhält man durch Umformen

$$\hat{\delta}_{11}^{AB} = \ln(n_{11}) - \hat{\delta} - \hat{\delta}_1^A - \hat{\delta}_1^B$$

Ersetzt man die geschätzten log-linearen Parameter durch ihre Schätzungen, die auf den beobachteten Häufigkeiten n_{ij} basieren, erhält man:

$$\begin{aligned}
\hat{\delta}_{11}^{AB} &= \ln(n_{11}) - \ln(n_{22}) - \ln(n_{12}) + \ln(n_{22}) - \ln(n_{21}) + \ln(n_{22}) \\
&= \ln(n_{11}) - \ln(n_{12}) - \ln(n_{21}) + \ln(n_{22}) \\
&= \ln(n_{11}) + \ln(n_{22}) - \ln(n_{12}) - \ln(n_{21}) \\
&= \ln\left(\frac{n_{11} \cdot n_{22}}{n_{12} \cdot n_{21}}\right)
\end{aligned}$$

(2) Zeigen Sie, dass im Unabhängigkeitsmodell für eine 2×2-Tabelle alle Koeffizienten $\hat{\gamma}_{ij}^{AB}$ den Wert 1 annehmen müssen.

Der Koeffizient $\hat{\gamma}_{ij}^{AB}$ ist im Unabhängigkeitsmodell wie folgt definiert:

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

Dies folgt aus Gleichung F 20.8 und dem Sachverhalt, dass die log-linearen Parameter in einem restringierten Modell auf Grundlage der geschätzten erwarteten Häufigkeiten bzw. Wahrscheinlichkeiten bestimmt werden. Im Unabhängigkeitsmodell gilt: $e_{ij} = n_{i\bullet} \cdot n_{\bullet j}$. Hieraus ergibt sich:

$$\hat{\gamma}_{11}^{AB} = \sqrt[4]{\frac{e_{11} \cdot e_{22}}{e_{12} \cdot e_{21}}} = \sqrt[4]{\frac{n_{1\bullet} \cdot n_{\bullet 1} \cdot n_{2\bullet} \cdot n_{\bullet 2}}{n_{1\bullet} \cdot n_{\bullet 2} \cdot n_{2\bullet} \cdot n_{\bullet 1}}} = 1$$

Aufgrund von Gleichung F 20.9 müssen auch alle anderen Koeffizienten $\hat{\gamma}_{ij}^{AB}$ gleich 1 sein.

(3) Berechnen Sie die Werte der Pearson-Teststatistik und des Likelihood-Ratio-Tests für das Modell $\ln(\varepsilon_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} + \lambda_{jk}^{AC}$ und die Daten in Tabelle 20.3.

Die erwarteten Häufigkeiten unter dem Modell $\ln(\varepsilon_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} + \lambda_{jk}^{AC}$ sind in Tabelle 20.10 zusammengestellt. Setzt man diese erwarteten Häufigkeiten und die beobachteten Häufigkeiten aus Tabelle 20.3 in die Formeln F 20.30 und F 20.31, die man für den dreidimensionalen Fall erweitern muss, ein, erhält man:

$$PE=\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\left(n_{ijk}-e_{ijk}\right)^2}{e_{ijk}}=\frac{(14-18{,}188)^2}{18{,}188}+\frac{(25-20{,}812)^2}{20{,}812}+\frac{(30-25{,}812)^2}{25{,}812}+\frac{(45-49{,}188)^2}{49{,}188}$$
$$+\frac{(122-117{,}812)^2}{117{,}812}+\frac{(81-85{,}188)^2}{85{,}188}+\frac{(87-91{,}188)^2}{91{,}188}+\frac{(114-109{,}812)^2}{109{,}812}$$
$$=3{,}550$$

und

$$LR=2\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}n_{ij}\cdot\ln\frac{n_{ij}}{e_{ij}}=\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}2\cdot n_{ij}\cdot\ln\frac{n_{ij}}{e_{ij}}$$
$$=2\cdot 14\cdot\ln\frac{14}{18{,}188}+2\cdot 25\cdot\ln\frac{25}{20{,}812}+2\cdot 30\cdot\ln\frac{30}{25{,}812}+2\cdot 45\cdot\ln\frac{45}{49{,}188}$$
$$+2\cdot 122\cdot\ln\frac{122}{117{,}815}+2\cdot 81\cdot\ln\frac{81}{85{,}188}+2\cdot 87\cdot\ln\frac{87}{91{,}188}+2\cdot 114\cdot\ln\frac{114}{109{,}812}$$
$$=3{,}562$$

(4) Schätzen Sie die Parameter λ_i^A, λ_j^B und λ_k^C für das Modell $\ln\left(\varepsilon_{ijk}\right)=\lambda+\lambda_i^A+\lambda_j^B+\lambda_k^C+\lambda_{ij}^{AB}+\lambda_{jk}^{BC}+\lambda_{jk}^{AC}$ anhand der Daten in Tabelle 20.10.

Übertragt man Gleichung F 20.38 auf den Fall erwarteter Häufigkeiten, auf die man in restringierten Modellen zurückgreift, gilt:

$$\hat{\gamma}_1^A=\sqrt[8]{\prod_{j=1}^{2}\prod_{k=1}^{2}\frac{e_{1jk}}{e_{2jk}}}=\sqrt[8]{\frac{e_{111}}{e_{211}}\cdot\frac{e_{112}}{e_{212}}\cdot\frac{e_{121}}{e_{221}}\cdot\frac{e_{122}}{e_{222}}}$$
$$=\sqrt[8]{\frac{18{,}188}{117{,}812}\cdot\frac{20{,}812}{85{,}188}\cdot\frac{25{,}812}{91{,}188}\cdot\frac{49{,}188}{109{,}812}}=0{,}513$$

Für $\hat{\lambda}_1^A$ ergibt sich $\hat{\lambda}_1^A=\ln\left(\hat{\gamma}_1^A\right)=\ln(0{,}513)=-0{,}667$ und hieraus $\hat{\lambda}_2^A=-\hat{\lambda}_1^A=0{,}667$. Für $\hat{\gamma}_1^B$ erhält man in analoger Weise:

$$\hat{\gamma}_1^B=\sqrt[8]{\prod_{i=1}^{2}\prod_{k=1}^{2}\frac{e_{i1k}}{e_{i2k}}}=\sqrt[8]{\frac{e_{111}}{e_{121}}\cdot\frac{e_{112}}{e_{122}}\cdot\frac{e_{211}}{e_{221}}\cdot\frac{e_{212}}{e_{222}}}$$
$$=\sqrt[8]{\frac{18{,}188}{25{,}812}\cdot\frac{20{,}812}{49{,}188}\cdot\frac{117{,}812}{91{,}188}\cdot\frac{85{,}188}{109{,}812}}=0{,}860$$

Somit ist $\hat{\lambda}_1^B=\ln(0{,}860)=-0{,}151$ $\hat{\lambda}_2^B=0{,}151$. Für $\hat{\gamma}_1^C$ gilt:

$$\hat{\gamma}_1^C=\sqrt[8]{\prod_{i=1}^{2}\prod_{j=1}^{2}\frac{e_{ij1}}{e_{ij2}}}=\sqrt[8]{\frac{e_{111}}{e_{112}}\cdot\frac{e_{121}}{e_{122}}\cdot\frac{e_{211}}{e_{212}}\cdot\frac{e_{221}}{e_{222}}}$$
$$=\sqrt[8]{\frac{18{,}188}{20{,}812}\cdot\frac{25{,}812}{49{,}188}\cdot\frac{117{,}812}{85{,}188}\cdot\frac{91{,}188}{109{,}812}}=0{,}923$$

Daher ist $\hat{\lambda}_1^C=\ln(0{,}860)=-0{,}080$ und $\hat{\lambda}_2^C=0{,}080$.

(5) Bestimmen sie alle hierarchischen additiven log-linearen Modelle für die 2×2×2-Tabelle. Stellen Sie die Modelle als Populationsmodelle dar.

(a) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} + \lambda_{ik}^{AC} + \lambda_{ijk}^{ABC}$

(b) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} + \lambda_{ik}^{AC}$

(c) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC}$

(d) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$

(e) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC} + \lambda_{ik}^{AC}$

(f) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$

(g) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC}$

(h) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC}$

(i) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$

(j) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_k^C + \lambda_{ik}^{AC}$

(k) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC}$

(l) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C$

(m) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_j^B$

(n) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A + \lambda_k^C$

(o) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_j^B + \lambda_k^C$

(p) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_i^A$

(q) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_j^B$

(r) $\ln\left(\varepsilon_{ijk}\right) = \lambda + \lambda_k^C$

(s) $\ln\left(\varepsilon_{ijk}\right) = \lambda$