

Schätz- und Prüfstatistik - Aufgabe 4

Bei den ausländischen Touristen Heidelbergs bilden Asien und Nordamerika die Hauptquellgebiete. Im Rahmen von standardisierten Interviews wurde u.a. gefragt, wie sich diese beiden wichtigsten Touristengruppen vorab über die Stadt Heidelberg informiert haben. Bei der Frage, ob man sich durch Reiseliteratur informiert habe, ergab sich folgendes Ergebnis:

	Keine Information durch Reiseliteratur	Information durch Reiseliteratur
Herkunft: Asien	144	83
Herkunft: Nordamerika	87	19

- a) Sind die Unterschiede bezüglich der „Information durch Reiseliteratur“ zwischen den nordamerikanischen und asiatischen Touristen statistisch hoch signifikant? Nennen Sie die entsprechenden Hypothesen, führen Sie die notwendigen Berechnungen durch, treffen Sie die richtige Entscheidung.
- b) Wie stark ist der gemessene Zusammenhang? Berechnen Sie den passenden Koeffizienten.

Lösung:

a)

Schritt 1: Formulierung der Hypothesen:

Hypothese H0: Die Unterschiede bezüglich der „Information durch Reiseliteratur“ zwischen den nordamerikanischen und asiatischen Touristen sind so gering, dass sie nicht als hoch signifikant angesehen werden können.

Hypothese H1: Es gibt hoch signifikante Unterschiede bezüglich der Information durch Reiseliteratur zwischen Touristen aus Asien und Nordamerika.

Schritt 2: Bestimmung des Signifikanzniveaus:

Die Unterschiede müssen laut Aufgabenstellung **hoch signifikant** sein
→ Irrtumswahrscheinlichkeit = 0,1%

Schritt 3: Bestimmung des kritischen Chi²:

Freiheitsgrad = (Anzahl der Spalten - 1) x (Anzahl der Zeilen - 1)

Hier: FG = (2-1) x (2-1) = 1 x 1 = 1

FG = 1, Irrtumswahrscheinlichkeit = 0,1% → kritisches Chi² = 10,83
(ablesbar in der Chi²-Tabelle in der Formelsammlung)

Schritt 4: Bestimmung des empirischen Chi²:

Die Erwartungswerte f_{ei} müssen berechnet und in die Indifferenztabelle eingetragen werden.

$$f_{bi} = \text{beobachteter Wert} \quad f_{ei} = \text{erwarteter Wert} = \frac{\text{Zeilensumme} \times \text{Spaltensumme}}{\text{Gesamtsumme}}$$

Beispielrechnung für den erwarteten Wert in der linken oberen Zelle $f_{ei} = \frac{231 \times 227}{333} = 157,5$

	Keine Information durch Reiseliteratur	Information durch Reiseliteratur	Σ
Herkunft: Asien	f_{bi} 144 / f_{ei} 157,5	f_{bi} 83 / f_{ei} 69,5	227
Herkunft: Nordamerika	f_{bi} 87 / f_{ei} 73,5	f_{bi} 19 / f_{ei} 32,5	106
Σ	231	102	333

$$\chi^2_{emp} = \sum_{i=1}^k \frac{(f_{bi} - f_{ei})^2}{f_{ei}}$$

$$\Rightarrow \chi^2_{emp} = \frac{(144 - 157,5)^2}{157,5} + \frac{(83 - 69,5)^2}{69,5} + \frac{(87 - 73,5)^2}{73,5} + \frac{(19 - 32,5)^2}{32,5}$$

$$\Leftrightarrow \chi^2_{emp} = \frac{182,25}{157,5} + \frac{182,25}{69,5} + \frac{182,25}{73,5} + \frac{182,25}{32,5}$$

$$\Leftrightarrow \chi^2_{emp} = 1,15 + 2,62 + 2,48 + 5,61$$

$$\Rightarrow \chi^2_{emp} = 11,81$$

5. Schritt: Vergleich des empirischen und des kritischen Chi²:

$$11,81 > 10,83 \quad \chi^2_{emp} > \chi^2_{krit.}$$

→ Da das empirische Chi² größer als das kritische Chi² ist, wird H₀ abgelehnt und H₁ angenommen:

Damit gibt es statistisch hoch signifikante Unterschiede bezüglich der Information durch Reiseliteratur zwischen Touristen aus Asien und Nordamerika

b)

Um die Stärke des Zusammenhangs zu messen muss der Kontingenzkoeffizient C berechnet werden.

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

$$C = \sqrt{\frac{11,81}{11,81 + 333}} = 0,19$$