

Lösung zu Kapitel 4: Beispiel 4

Autoklasse und Unfallhäufigkeit

Zulassungsdaten aus einem Land zeigen, dass 15% der Autos Kleinwagen, 25% Kompaktmodelle, 40% Mittelklassemodelle und der Rest größere oder Sondermodelle sind. Eine Zufallsstichprobe von Verkehrsunfällen mit Autos wird gezogen. Kann man schließen, dass bestimmte Größenklassen von Autos häufiger in Verkehrsunfälle verwickelt sind, als es die Zulassungszahlen vermuten lassen?

Daten: **unfaelle.sav**

Variable: **auto**

Die Nullhypothese lautet somit: Es gibt keine Unterschiede in der Unfallhäufigkeit zwischen den verschiedenen Autoklassen.

Ein erster Blick auf die Daten wird durch ein Balkendiagramm ermöglicht.

PASW

Diagramme

Diagrammerstellung

in Galerie auswählen von Balken

Doppelklick auf Galeriediagramm

Variable **auto** in das Feld X-Achse? ziehen

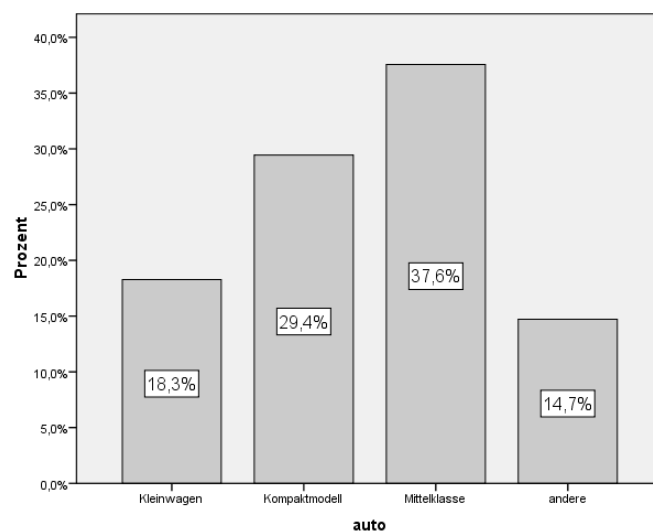


Abbildung 1: einfaches Balkendiagramm mit der Variable **auto**

Das Balkendiagramm lässt vermuten, dass die tatsächlichen Werte nicht augenscheinlich stark von den erwarteten Werten abweichen. Im Anschluss werden die erhobenen Daten den erwarteten Werten gegenüber gestellt. Dies kann gleich im Zuge des Chi-Quadrat-Tests erfolgen:

Analysieren

Nichtparametrische Tests

Chi-Quadrat ...

markieren der Variable (auto) ☒ Testvariable(n):

Optionen ...

Statistiken: Deskriptive Statistik

Weiter

Erwartete Wert Werte:

15 Hinzufügen

25 Hinzufügen

40 Hinzufügen

20 Hinzufügen

OK

auto			
	Beobachtetes N	Erwartete Anzahl	Residuum
Kleinwagen	36	29,6	6,5
Kompaktmodell	58	49,3	8,8
Mittelklasse	74	78,8	-4,8
andere	29	39,4	-10,4
Gesamt	197		

Abbildung 2: Deskriptive Statistik für die Variable `auto`

Die Häufigkeitstabelle zeigt, dass Kleinwagen und Kompaktmodelle in der Stichprobe häufiger in Verkehrsunfälle verwickelt sind als dies die Zulassungsdaten vermuten lassen. Man kann nun mit Hilfe des Chi-Quadrat-Tests auf eine vorgegebene Verteilung feststellen, ob es sich dabei um einen Zufall oder um systematische Abweichungen handelt.

Statistik für Test	
	auto
Chi-Quadrat	6,000 ^a
df	3
Asymptotische Signifikanz	,112

a. Bei 0 Zellen (0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 29,6.

Abbildung 3: Chi-Quadrat-Test auf eine erwartete Verteilung: (15%, 25%, 40%, 20%) mit der Variable `auto`

Der durchgeführte Test zeigt, dass die Unterschiede in den einzelnen Gruppen im Vergleich mit den Erwartungswerten durchaus auf Zufall basieren können. Die Wahrscheinlichkeit, die Nullhypothese zu verwerfen, obwohl sie richtig ist, ist zu groß ($p\text{-Wert} = 0,112$), sodass die Nullhypothese beizubehalten ist. Die Beteiligung an Verkehrsunfällen ist daher als unabhängig von der Größenklasse des Autos zu betrachten.