Mit deM AUTo unterwegs

Signifikanztests im Straßenverkehr

Lösungsvorschlag

Hinweis: Diese Lösung wurde an einem Montagnachmittag in der Meckenheimer Allee in Bonn erstellt und stellt lediglich einen Lösungsvorschlag dar. Je nach Lernort weichen die Ergebnisse ab.

A1 In einem Zeitraum von zehn Minuten wurden folgende Daten erhoben:

Straßenseite 1 (Richtung Innenstadt):	Autos mit anderen Kennzeichen:	29
	Autos mit Bonner Kennzeichen:	26
Straßenseite 2 (von der Innenstadt kommend):	Autos mit anderen Kennzeichen:	10
	Autos mit Bonner Kennzeichen:	26

A2 Gesamtzahl der vorbeigefahrenen Autos: 91, davon:

- Autos mit anderem Kennzeichen: 39
- Autos mit Bonner Kennzeichen: 52

Anteilig hatten $\frac{39}{91} \approx 43\%$ der beobachteten Autos ein fremdes Kennzeichen, während $\frac{52}{91} \approx 57\%$ ein Bonner Kennzeichen hatten.

A3 In Teilaufgabe **A2** wurde berechnet, dass der Anteil der Autos mit einem Kennzeichen einer anderen Stadt/eines anderen Kreises annähernd 43% beträgt. Dies nehmen wir als Annäherung für die Wahrscheinlichkeit p und stellen daher die Hypothese $H_0: \; p = 43\% = 0,43$ auf.
Unsere Beobachtungen aus Teilaufgabe **A1** lassen sich als Binomialverteilung auffassen, wenn man jedes vorbeifahrende Auto als Bernoulli-Experiment betrachtet, bei dem ein Erfolg vorliegt, wenn das Auto ein Kennzeichen einer anderen Stadt/eines anderen Kreises hat und ein Misserfolg vorliegt, wenn das Auto ein Bonner Kennzeichen hat. Dann ist $p = 0,43$ unsere Trefferwahrscheinlichkeit und die Anzahl aller beobachteten Autos ist unser Stichprobenumfang. Die Testgröße X gibt an, wie viele der vorbeigefahrenen Autos ein Kennzeichen einer anderen Stadt/eines anderen Kreises hatten.

A4 Nullhypothese H_0: $p = 0,43$ (vgl. Teilaufgabe **A3**)
Alternative H_1: $p \neq 0,43$
Wir entscheiden uns für $n = 90$ als Stichprobenumfang und $\alpha = 5\%$ als Signifikanzniveau.

Wir bestimmen den Annahmebereich unserer Hypothese, indem wir (mit dem Taschenrechner) eine Tabelle der kumulierten Wahrscheinlichkeiten von X zu unserem n und p erstellen und daraus die kleinsten Zahlen a und b ablesen, sodass $P(X \leq a) > \frac{\alpha}{2} = 0,025$ und $P(X \leq b) > 1 - \frac{\alpha}{2} = 0,975$.
Wir suchen also das kleinste a, für das in der Tabelle der kumulierten Wahrscheinlichkeiten der Wert von $0,025$ überschritten wird und das kleinste b, für das der Wert von $0,975$ überschritten wird. Einen Auszug der Tabelle findet man in der nebenstehenden Tabelle. Aus dieser Tabelle kann man $a = 30$ und $b = 48$ ablesen. Somit ist der Annahmebereich $[30; 48]$.

r	$P(X \leq r)$
29	$0,024$
30	$0,039$
$\vdots$	$\vdots$
47	$0,969$
48	$0,981$

Die neue Datenerhebung an der Straße ergibt:

Gesamtzahl der vorbeigefahrenen Autos:		90
davon:	Autos mit anderen Kennzeichen:	31
	Autos mit Bonner Kennzeichen:	59

Da hierbei 31 Autos kein Bonner Kennzeichen hatten und 31 im Annahmebereich liegt ($31 \in [30, 48]$), wird unsere Nullhypothese durch den Test bestätigt und wir können sie annehmen.

B1 Es werden nun die beiden Signifikanzniveaus $\alpha_1 = \frac{\alpha}{2} = 2,5\%$ und $\alpha_2 = 2\alpha = 10\%$ betrachtet.

Analog zu Teilaufgabe **A4** bestimmen wir den Annahmebereich, indem wir (mit dem Taschenrechner) eine Tabelle der kumulierten Wahrscheinlichkeiten von X für unser $p = 0,43$ und den Stichprobenumfang $n = 90$, den wir in Teilaufgabe **A4** gewählt haben, erstellen. Daraus lesen wir (wie in Teilaufgabe **A4**) die Grenzen unseres Annahmebereiches ab. Dabei suchen wir diesmal für $\alpha_1 = 2,5\%$ das kleinste a_1, für das in der Tabelle der kumulierten Wahrscheinlichkeiten der Wert von $\frac{\alpha_1}{2} = 0,0125$ überschritten wird und das kleinste b_1, für das der Wert von $1 - \frac{\alpha_1}{2} = 0,9875$ überschritten wird.

Analog suchen wir für $\alpha_2 = 10\%$ das kleinste a_2, für das der Wert von $\frac{\alpha_2}{2} = 0,05$ überschritten wird und das kleinste b_2, für das der Wert von $1 - \frac{\alpha_2}{2} = 0,95$ überschritten wird.
Wie der Tabelle zu entnehmen ist, ergibt dies den Annahmebereich $[28; 49]$ für das Signifikanzniveau $\alpha_1 = 2,5\%$ und den Annahmebereich $[31; 46]$ für das Signifikanzniveau $\alpha_2 = 10\%$.

r	$P(X \leq r)$
27	$0,00776$
28	$0,01394$
$\vdots$	$\vdots$
30	$0,03921$
31	$0,06155$
$\vdots$	$\vdots$
45	$0,92562$
46	$0,95102$
$\vdots$	$\vdots$
48	$0,98108$
49	$0,98892$

B2 Die drei unterschiedlichen Annahmebereiche sind:

- für $\alpha_1 = 2,5\%$: $[28; 49]$,
- für $\alpha = 5\%$: $[30; 48]$,
- für $\alpha_2 = 10\%$: $[31; 46]$.

Wie man erkennt, wird der Annahmebereich für ein steigendes Signifikanzniveau immer kleiner, er umfasst also immer weniger Werte und ist damit immer schmaler um den Erwartungswert verteilt.
Die Nullhypothese $H_0 : \ p = 0,43$ aus Teilaufgabe **A3** kann für alle drei Signifikanzniveaus angenommen werden.

B3 Allgemeine Regel: Je größer das Signifikanzniveau ist, desto kleiner ist der Annahmebereich. Somit darf die Datenerhebung mit steigendem Signifikanzniveau nur noch weniger vom Erwartungswert abweichen, damit die Hypothese noch angenommen wird.

B4 Ein niedriges Signifikanzniveau führt, wie wir gesehen haben, zu einem größeren Annahmebereich. Wenn die Hypothese richtig gewählt wurde, ist dies vorteilhaft, da auch kleinere „Ausreißer“ in der Datenerhebung toleriert werden und nicht zum (fälschlichen) Verwerfen der

Hypothese führen. Ist die Hypothese jedoch falsch, so steigt mit einem niedrigen Signifikanzniveau die Wahrscheinlichkeit, dass die Hypothese durch den Signifikanztest bestätigt wird, obwohl sie eigentlich falsch ist, sie also fälschlicherweise akzeptiert wird.
Bei einem hohen Signifikanzniveau hingegen ist der Annahmebereich kleiner. Daher kann es schneller passieren, dass eine richtige Hypothese wegen kleinerer „Ausreißer" in der Datenerhebung fälschlicherweise verworfen wird. Andersherum ist dies bei einer falschen Hypothese vorteilhaft. Denn in diesem Fall fällt der erhobene Wert häufiger in den recht großen Ablehnungsbereich, wodurch falsche Hypothesen zuverlässiger erkannt und verworfen werden können.

C1 Es werden nun die beiden Stichprobenumfänge $n_1 = \frac{n}{2} = 45$ und $n_2 = 2n = 180$ betrachtet.

Analog zu Teilaufgabe **A4** bestimmen wir den Annahmebereich, indem wir (mit dem Taschenrechner) eine Tabelle der kumulierten Wahrscheinlichkeiten von X für die beiden verschiedenen Stichprobenumfänge n_1 und n_2 und unsere Trefferwahrscheinlichkeit $p = 0,43$ erstellen und daraus die Grenzen unseres Annahmebereiches ablesen.
Wie in der linken Tabelle zu erkennen ist, ergibt dies den Annahmebereich $[13; 26]$ für den Stichprobenumfang $n_1 = 45$. Der rechten Tabelle kann man entnehmen, dass der Annahmebereich für den Stichprobenumfang $n_2 = 100$ $[64; 90]$ ist.

r	$P(X \leq r)$
12	$0,01784$
13	$0,03717$
$\vdots$	$\vdots$
25	$0,96735$
26	$0,98390$

r	$P(X \leq r)$
63	$0,01750$
64	$0,02531$
$\vdots$	$\vdots$
89	$0,96528$
90	$0,97528$

C2 Bei unserer Nullhypothese aus Teilaufgabe **A3**: $H_0 : p = 0,43$ wäre der Erwartungswert der Testgröße X:

- $p \cdot n_1 = 0,43 \cdot 45 \approx 19$ für den Stichprobenumfang $n_1 = 45$,
- $p \cdot n = 0,43 \cdot 90 \approx 39$ für $n = 90$,
- $p \cdot n_2 = 0,43 \cdot 180 = 77$ für $n_2 = 180$.

Wenn das Testergebnis an der Straße ergäbe, dass der Erwartungswert um 10 Autos übertroffen wird, dann wäre dies:

- $19 + 10 = 29$ Autos aus anderen Städten/Kreisen für den Stichprobenumfang $n_1 = 45$,
- $39 + 10 = 49$ für $n = 90$,
- $77 + 10 = 87$ für $n_2 = 180$.

Damit würde die Nullhypothese:

- für (den Stichprobenumfang) $n_1 = 45$ verworfen werden, denn $29 \notin [13; 26]$,
- für (den Stichprobenumfang) $n = 90$ verworfen werden, denn $49 \notin [30; 48]$,
- für (den Stichprobenumfang) $n_2 = 180$ angenommen werden, denn $87 \in [64; 90]$.

Wie man sieht, müsste die Nullhypothese bei einer Abweichung um 10 vom Erwartungswert bei einem geringeren Stichprobenumfang (wie $n_1 = 45$ und $n = 90$) verworfen werden, während sie bei einem größeren Stichprobenumfang (wie $n_2 = 180$) trotz derselben absoluten Abweichung beibehalten werden könnte.

C3 Allgemeine Regel: Bei einem kleinen Stichprobenumfang genügt schon eine kleinere absolute Abweichung des Testergebnisses vom Erwartungswert, damit die Hypothese verworfen wird. Bei einem großen Stichprobenumfang hingegen kann die Hypothese bei der gleichen absoluten Abweichung vom Erwartungswert noch akzeptiert werden. Hier ist eine viel größere Abweichung vom Erwartungswert nötig, damit die Hypothese verworfen wird.

C4 Wenn das Testergebnis um 20% nach oben vom Erwartungswert abweichen würde, dann entspräche dies:

- $19 + (0,2 \cdot 19) \approx 23$ Autos aus anderen Städten/Kreisen für den Stichprobenumfang $n_1 = 45$,
- $39 + (0,2 \cdot 39) \approx 47$ für $n = 90$,
- $77 + (0,2 \cdot 77) \approx 92$ für $n_2 = 180$.

Für das Ergebnis des Signifikanztests würde dies bedeuten, dass die Hypothese:

- für den Stichprobenumfang $n_1 = 45$ angenommen werden kann, denn $23 \in [13; 26]$,
- für den Stichprobenumfang $n = 90$ angenommen werden kann, denn $47 \in [30; 48]$,
- für den Stichprobenumfang $n_2 = 180$ verworfen werden muss, denn $92 \notin [64; 90]$.

C5 Allgemeine Regel: Die gleiche prozentuale Abweichung des Testergebnisses vom Erwartungswert wirkt sich auf einen Signifikanztest mit großem Stichprobenumfang stärker aus als auf einen Signifikanztest mit kleinerem Stichprobenumfang. So genügt bei einem großen Stichprobenumfang schon eine kleinere prozentuale Abweichung, um die Hypothese zu verwerfen, während dafür bei einem kleineren Stichprobenumfang eine größere prozentuale Abweichung nötig ist.

D1 Wie in Teilaufgabe **B4** zu sehen war, sind sowohl ein sehr hohes als auch ein sehr niedriges Signifikanzniveau nachteilig. Daher ist etwa $\alpha = 5\%$ ein sinnvolles Mittelmaß und wird deshalb auch häufig als Signifikanzniveau verwendet.
Der Stichprobenumfang sollte recht hoch sein, damit kleinere zufällige absolute Abweichungen vom Erwartungswert den Test nicht fälschlicherweise beeinflussen und dazu führen, dass die Hypothese verworfen wird, obwohl sie eigentlich richtig gewählt war. Ein hoher Stichprobenumfang ist darüber hinaus hilfreich, wenn eine falsche Hypothese vorliegt. Denn dann kommt es aufgrund der fehlerhaften Hypothese zu prozentualen Abweichungen vom Erwartungswert, die bei einem hohen Stichprobenumfang eher das Verwerfen der Hypothese bewirken als bei einem niedrigen Stichprobenumfang. Zu hohe Stichprobenumfänge sind jedoch schwierig umzusetzen, da dafür viel Zeit und Arbeit aufgewendet werden muss. Daher genügt für unsere Aufgabe ein Kompromiss von etwa $n = 100$ als Stichprobenumfang. Würde man die Hypothese jedoch wirklich genau testen wollen, so müsste man einen deutlich größeren Stichprobenumfang wählen.

D2 Bisher haben wir nur eine Datenerhebung und damit auch nur einen Signifikanztest durchgeführt. Dabei haben wir lediglich einen Stichprobenumfang von 90 Autos betrachtet. Um sicherzugehen, dass die Hypothese wirklich stimmt und sie nicht nur aufgrund des niedrigen Stichprobenumfangs angenommen wurde, sollten wir eine weitere Datenerhebung mit einem sehr viel größeren Stichprobenumfang durchführen und daran unsere Hypothese erneut testen. Erst dann können wir uns sicher sein, dass sie stimmt.

D3 Entsprechend unserer Erkenntnisse aus den vorherigen Teilaufgaben können wir sagen:
Wenn wir in unserem Signifikanztest möglichst das Ergebnis erhalten wollen, dass wir die Hypothese annehmen können, dann sollten wir ein kleines Signifikanzniveau (z.B. $\alpha = 2,5\%$) verwenden.
Würden wir hingegen wollen, dass die Hypothese möglichst verworfen wird, dann sollten wir ein großes Signifikanzniveau (z.B. $\alpha = 10\%$) verwenden.
Je größer der Stichprobenumfang, desto sicherer arbeitet der Test und desto eher erkennt der Test, ob die Nullhypothese stimmt. Ein nicht zu großer Stichprobenumfang ist also sinnvoll, wenn wir unser Wunschergebnis erhalten wollen. (Wenn wir beispielsweise die Nullhypothese annehmen wollen, obwohl sie eigentlich falsch ist, dann würde es zu prozentualen Abweichungen vom Erwartungswert kommen, die bei kleinen Stichprobenumfängen nicht so ins Gewicht fallen.)

D4 „Lügen“ ist vielleicht das falsche Wort, aber nur weil ein Signifikanztest das Ergebnis bringt, dass eine Hypothese angenommen bzw. verworfen werden kann, bedeutet dies nicht mit Sicherheit, dass die Hypothese wirklich richtig bzw. falsch ist. Wie wir oben gesehen haben, kann ein Signifikanztest bei den gleichen zugrunde liegenden Daten für unterschiedliche Signifikanzniveaus und Stichprobenumfänge verschiedene Ergebnisse bringen. Dementsprechend kann man das Ergebnis des Signifikanztests durch die Wahl des Signifikanzniveaus und Stichprobenumfangs durchaus beeinflussen. Deshalb sollte man dem Ergebnis eines Signifikanztests nicht blind vertrauen, sondern erstmal hinterfragen, welche Parameter im Signifikanztest gewählt wurden und wie der Test durchgeführt wurde.

Didaktischer Kommentar

Dieser Mathematische Spaziergang richtet sich an Schülerinnen und Schüler der Sekundarstufe 2 und lässt sich in den Themenkomplex Stochastik einordnen. Idealerweise wird der Spaziergang im Rahmen der Unterrichtsreihe *Testen von Hypothesen* durchgeführt.

Für die Bearbeitung der Aufgaben wird vorausgesetzt, dass die Schülerinnen und Schüler Binomialverteilungen kennen und dazu in der Lage sind, einen zweiseitigen Signifikanztest durchzuführen. In diesem Spaziergang wird bewusst nicht auf einseitige Signifikanztests und Fehler erster und zweiter Art eingegangen, damit er bereits vor deren Behandlung im Unterricht zur Vertiefung der zweiseitigen Signifikanztests durchgeführt werden kann. Vor allem in Teilaufgabe **B4** aber auch in **C3** und **C5** machen sich die Schülerinnen und Schüler allerdings schon Gedanken, die in Richtung der Fehlerarten und deren Minimierung gehen. Deshalb können diese Teilaufgaben auch noch einmal aufgegriffen werden, wenn die Fehler erster und zweiter Art später im Unterricht behandelt werden.

Am Ende des Spaziergangs sollen die Schülerinnen in der Lage sein, zweiseitige Signifikanztests durchzuführen, um Hypothesen zu testen. Weiter sollen sie die Bedeutung des Signifikanzniveaus und des Stichprobenumfangs verstehen und die Auswirkungen einer Veränderung dieser beiden Parameter kennen.

Als Lernort eignet sich eine Straße, welche von genügend Autos aus dem Ort und aus anderen Städten/Kreisen befahren wird und gleichzeitig von außen gut und sicher beobachtet werden kann (beispielsweise von einem breiten Bürgersteig oder einer angrenzenden Wiese aus). Eine Ortsdurchgangsstraße eignet sich somit besser als beispielsweise eine Straße in einem Wohngebiet. In der Nachbesprechung der Aufgabe kann thematisiert werden, dass die Wahrscheinlichkeit p, dass ein auf der Straße vorbeifahrendes Auto ein Kennzeichen einer anderen Stadt/eines anderen Kreises hat, abhängig von der Wahl des Standpunktes ist. In allen Aufgabenteilen, in denen die Straße nicht benötigt wird, dient ein Park, eine Wiese oder eine Parkbank in der Nähe als Lernort.