Die Wahrscheinlichkeit liegt auf der Hand

Hypothesentests beim Boule-Spiel

Lösungsvorschlag

Hinweis: Diese Lösung basiert auf bespielhaften Werte und stellt lediglich einen Lösungsvorschlag dar. Je nach Spielverlauf weichen die Ergebnisse ab.

A2 In diesem Beispiel wurde 32 Mal geworfen. Von allen Würfen waren 18 Treffer und 14 Kugeln landeten im Aus.

A3 Die Binomialverteilung beschreibt die Anzahl der Erfolge in einer Serie von gleichartigen und unabhängigen Versuchen, die jeweils genau zwei mögliche Ergebnisse haben. Die Schülerinnen und Schüler können darüber diskutieren, ob die Würfe tatsächlich als gleichartig und unabhängig voneinander aufgefasst werden dürfen. Da das Experiment als Gruppenexperiment durchgeführt wird, kann es sein, dass nicht alle Schülerinnen und Schüler gleich gut werfen, weshalb die Wahrscheinlichkeit, einen Treffer zu erzielen, nicht bei jedem Wurf gleich groß wäre. Hinzu kommt, dass sich die einzelnen Schülerinnen und Schüler von Wurf zu Wurf verbessern können oder auch verschlechtern, wenn sie zum Beispiel nach einigen Würfen erschöpft sind.

Unter der Annahme, dass die einzelnen Würfe der Kugeln unabhängig voneinander sind, kann man jeden einzelnen Kugelwurf als Bernoulli-Experiment betrachten, bei dem ein Abstand von höchstens siebzig Zentimetern zur Zielkugel als Erfolg und ein Abstand von mehr als siebzig Zentimetern als Misserfolg gewertet wird. Dann ist die Testgröße X, welche die Anzahl der Erfolge nach n Würfen angibt, binomialverteilt. Dabei ist der Stichprobenumfang $n = 32$, da in Teilaufgabe **A2** insgesamt 32-mal geworfen wurde. Außerdem leiten wir aus Teilaufgabe **A2** die Trefferwahrscheinlichkeit $p = \frac{18}{32} \approx 0,56$ ab, indem wir die Anzahl der Treffer (18) durch die Anzahl der Gesamtwürfe (32) teilen.

B1 In Teilaufgabe **A3** haben wir schon die Hypothese aufgestellt, dass $p = 0,56$ ist. Somit gilt:

Nullhypothese $H_0: \; p = 0,56$ Alternative $H_1: p \neq 0,56$

Den Stichprobenumfang n können wir für unseren Signifikanztest wie in Teilaufgabe **A2** wählen: $n = 32$. Es ist natürlich auch ein größerer Stichprobenumfang möglich, da dieser die Genauigkeit des Ergebnisses erhöht. Der Stichprobenumfang sollte jedoch nicht weniger als 32 Würfe umfassen, um ein aussagekräftiges Ergebnis zu erhalten.

B2 Da hier das Signifikanzniveau $\alpha = 5\%$ beträgt, bestimmen wir nun das kleinste a, sodass $P(X \leq a) > \frac{5\%}{2} = 2,5\% = 0,025$ und das kleinste b, sodass $P(X \leq b) > 1 - \frac{5\%}{2} = 0,975$.

Wir suchen also das kleinste a, bei dem in der Tabelle der kumulierten Wahrscheinlichkeiten von X der Wert von 0,025 überschritten wird und das kleinste b, bei dem der Wert von 0,975 überschritten wird.
Mit dem Taschenrechner erstellen wir eine Tabelle der kumulierten Wahrscheinlichkeiten von X. Einen Auszug davon (mit gerundeten Zahlen) findet man in der Tabelle. Aus dieser Tabelle kann man ablesen, dass $a = 12$ und $b = 23$ ist. Somit ist der Annahmebereich $[12; 23]$.

r	$P(X \leq r)$
11	0,011
12	0,027
⋮	⋮
22	0,951
23	0,979

Die Irrtumswahrscheinlichkeit, also die Wahrscheinlichkeit, die Hypothese zu verwerfen, obwohl sie richtig ist, entspricht der Wahrscheinlichkeit des Ablehnungsbereiches. Der Ablehnungsbereich ist hier $[0;11] \cup [24;32]$. Da $P(X \leq 11) \approx 0,011$ und $P(X \geq 24) = 1 - P(X \leq 23) \approx 1 - 0,979 = 0,021$, beträgt die Irrtumswahrscheinlichkeit $0,011 + 0,021 = 0,032 = 3,2\%$. Das Signifikanzniveau gibt an, wie groß die Wahrscheinlichkeit des Ablehnungsbereiches höchstens sein kann. Da das Signifikanzniveau in diesem Fall 5% beträgt, kann die Irrtumswahrscheinlichkeit höchstens 5% betragen.

B3 Die Kugeln werden erneut geworfen und die Abstände zur Zielkugel werden überprüft. Das Ergebnis ist:

Würfe insgesamt: 32 davon: Würfe mit Treffer: 17 Würfe ins Aus: 15

Da 17 im Annahmebereich der Nullhypothese liegt ($17 \in [12;23]$), untermauert dieser Signifikanztest die aufgestellte Nullhypothese $H_0 : p = 0,56$ und wir können sie annehmen. (Bei anderen Daten oder einer anderen Hypothese kann es natürlich auch vorkommen, dass die Hypothese nach dem Signifikanztest verworfen werden muss.)
Wäre die Anzahl der Treffer im Ablehnungsbereich gewesen (also kleiner als 12 oder größer als 23), dann hätten wir die Nullhypothese verwerfen müssen. Ein möglicher Grund hierfür wäre, dass sich die Bedingungen in dieser Teilaufgabe gegenüber den vorherigen Bedingungen in Teilaufgabe **A2** verändert haben. Beispielsweise wurde die Zielkugel weiter oder weniger weit weggeworfen oder ein anderer Standort gewählt, was die Trefferwahrscheinlichkeit beeinflusst. Außerdem ist es auch möglich, dass sich die Gruppenmitglieder in Teilaufgabe **A2** noch eingespielt haben und jetzt durch die Übung eine höhere Trefferwahrscheinlichkeit haben als zuvor.

C1 Die Nullhypothese H_0 und die Trefferwahrscheinlichkeit $p = 0,56$ wurden für das Werfen mit der ersten Hand aufgestellt. Da man in der Regel zuerst mit seiner stärkeren Hand wirft, ist zu erwarten, dass die Würfe im zweiten Versuch ungenauer sind und damit die Trefferwahrscheinlichkeit p' geringer ist als p. Aufgrund dieser Überlegung stellt die Hypothese

$$H_1' : p' < p$$

eine sinnvolle Alternative zur Nullhypothese $H_0 : p = 0,56$ dar.

C2 Als Stichprobenumfang wählen wir 40 ($n' = 40$) und als Signifikanzniveau 6% ($\alpha' = 6\%$). (Natürlich können n' und α' auch beliebig anders gewählt werden.) Um den Annahmebereich $[a'; b']$ zu bestimmen, wird wie in Teilaufgabe **B2** mit dem Taschenrechner eine Tabelle der kumulierten Wahrscheinlichkeiten von X erstellt.

Da wir dieses Mal keinen zweiseitigen, sondern einen einseitigen Signifikanztest machen wollen, und zwar einen linksseitigen, suchen wir ein a' mit $P(X \leq a') > 0,06 = \alpha'$ und verwenden $b' = n'$. Aus der Tabelle der kumulierten Wahrscheinlichkeiten von X können wir ablesen, dass 18 die kleinste Trefferanzahl a' ist, die eine kumulierte Wahrscheinlichkeit von mindestens 6% hat (siehe Tabelle).

r	$P(X \leq r)$
17	0,0598
18	0,1074

Somit ist $[18; 40]$ unser Annahmebereich. Dementsprechend ist $[0; 17]$ der Ablehnungsbereich für die Hypothese H_0 in diesem linksseitigen Signifikanztest.

C3 Die Kugeln werden mit der jeweils anderen Hand geworfen und die Abstände zur Zielkugel werden überprüft. Das Ergebnis ist:

Würfe insgesamt: 40 davon: Würfe mit Treffer: 18 Würfe ins Aus 22

Da 18 noch im Annahmebereich liegt ($18 \in [18; 40]$), muss die Hypothese nicht verworfen werden. (Bei anderen Daten oder einer anderen Hypothese kann es natürlich auch vorkommen, dass die Hypothese nach dem Signifikanztest abgelehnt werden muss.)

D1 Der Fehler erster Art tritt auf, wenn der Signifikanztest das Ergebnis hervorbringt, dass die Nullhypothese verworfen werden muss, obwohl sie eigentlich richtig war. Die Wahrscheinlichkeit für den Fehler erster Art entspricht der Irrtumswahrscheinlichkeit. Diese wurde für den ersten Signifikanztest bereits in Teilaufgabe **B2** berechnet und beträgt $3,2$%.
Im zweiten Signifikanztest war der Ablehnungsbereich $[0; 17]$. Die Wahrscheinlichkeit für den Fehler erster Art im zweiten Signifikanztest ist daher $P(X \leq 17) \approx 0,0598 = 5,98$%.
Bei keinem der beiden Signifikanztests kann der Fehler erster Art aufgetreten sein, da die Nullhypothese beide Male nicht verworfen wurde. (Bei anderen Daten kann die Nullhypothese natürlich zuvor auch verworfen worden sein und daher auch ein Fehler erster Art vorkommen.)

D2 Der Fehler zweiter Art kommt vor, wenn die Nullhypothese nach dem Signifikanztest angenommen wird, obwohl sie eigentlich falsch ist. Wenn die Trefferwahrscheinlichkeit p' für die schwächere Hand 20% geringer ist als für die stärkere Hand, dann beträgt sie

$$p' = (1 - 0,2) \cdot p = 0,8 \cdot 0,56 = 0,448$$

und damit $44,8$ Prozent. Die Wahrscheinlichkeit für den Fehler zweiter Art, wenn $p' = 0,448$, entspricht der Wahrscheinlichkeit des Annahmebereiches $[18; 40]$ für $p' = 0,448$ statt für $p = 0,56$, also:

$$P(X_{p'} \geq 18) = 1 - P(X_{0,448} \leq 17) \approx 1 - 0,449 = 0,551 = 55,1\%,$$

wobei die kumulierte Wahrscheinlichkeit $P(X_{0,448} \leq 17)$ mit dem Taschenrechner berechnet werden kann.
Wie man sieht, ist die Wahrscheinlichkeit für den Fehler zweiter Art unter der Annahme, dass die Trefferwahrscheinlichkeit für die zuletzt benutzte Hand 20% geringer ist als für die zuerst genutzte Hand, mit $55,1$% sehr hoch. Daher ist es gut möglich, dass es in diesem Signifikanztest zu einem Fehler zweiter Art gekommen ist. Wie man in Teilaufgabe **C3** sieht, lag die Trefferanzahl mit 18 auch nur am äußersten Rand des Annahmebereiches ($[18; 40]$), was ein Indiz dafür ist, dass es zu einem Fehler zweiter Art gekommen sein kann.

D3 Die Wahrscheinlichkeit für den Fehler 1. Art können wir verringern, indem wir den Annahmebereich vergrößern. Dies vergrößert allerdings die Wahrscheinlichkeit für den Fehler 2. Art und ist daher, vor allem weil diese sowieso schon sehr hoch ist, nicht sinnvoll. Analog kann man die Wahrscheinlichkeit für den Fehler 2. Art verringern, indem man den Annahmebereich verkleinert. Dies vergrößert allerdings die Wahrscheinlichkeit für den Fehler 1. Art und ist daher ebenfalls nicht sinnvoll.
Eine sinnvolle Möglichkeit, die Wahrscheinlichkeit für beide Fehler gering zu halten, besteht darin, den Stichprobenumfang zu erhöhen. Beispielsweise könnte man statt 32 Würfe beziehungsweise 40 Würfe auch 80 Würfe machen. Um die Wahrscheinlichkeit für den Fehler zweiter Art sehr niedrig zu bekommen, müsste man den Stichprobenumfang jedoch sehr stark erhöhen. (Selbst bei 200 Würfen liegt die Wahrscheinlichkeit für den Fehler 2. Art noch bei gut 6 Prozent.)

Da es sich hier aber nur um ein Freizeitspiel ohne finanziellen Einsatz handelt, ist die genaue Trefferwahrscheinlichkeit nicht von hoher Bedeutung. Daher lohnt es sich nicht, den Stichprobenumfang auf über 200 Würfe zu erhöhen, da dies zu zeitaufwändig wäre und keinen hohen Nutzen bringen würde.

Didaktischer Kommentar

Dieser mathematische Spaziergang kann mit Schülerinnen und Schülern der Sekundarstufe II im Rahmen der Unterrichtseinheit *Testen von Hypothesen* bearbeitet werden. Das Lernziel der Aufgabe ist es, dass die Schülerinnen und Schüler passend zu ihren erhobenen Daten eine Hypothese aufstellen und diese mithilfe eines ein- und zweiseitigen Signifikanztests überprüfen. Sie bestimmen außerdem die Fehler erster und zweiter Art und erkennen sinnvolle Möglichkeiten, diese zu minimieren.
Die einzelnen Schritte eines Signifikanztests werden in den Teilaufgaben dieses Spazierganges beschrieben. Grundsätzlich wird aber vorausgesetzt, dass die Lernenden bereits in der Lage sind, sowohl einen zwei- als auch einen einseitigen Signifikanztest durchzuführen. Des Weiteren sollen für diesen Spaziergang das Konzept und die Berechnung der Fehler erster und zweiter Art bekannt sein. Falls diese im Unterricht noch nicht behandelt wurden, besteht auch die Möglichkeit Aufgabenteil **D** wegzulassen.
Da weite Teile der Unterrichtseinheit *Testen von Hypothesen* für diesen Spaziergang bereits behandelt sein müssen, eignet sich der Spaziergang vor allem zur Anwendung und Verfestigung des bereits Erlernten und kann zum Abschluss der Unterrichtseinheit sowie zur Klausurvorbereitung genutzt werden.

Als Lernort dient bei diesem mathematischen Spaziergang ein Boule-Platz oder alternativ eine öffentliche Wiese, beispielsweise in einem Park. Dort spielen die Schülerinnen und Schüler in Kleingruppen Boule, wofür jede Gruppe ein Boule-Spiel benötigt. Falls nicht genug Boule-Spiele zur Verfügung stehen, können stattdessen Sandsäckchen geworfen werden. Bei dieser Alternative sollte jedoch bei der Datenerhebung in Teilaufgabe **A2** erst ein Abstand von etwa 30 Zentimetern statt 70 Zentimetern als Treffer gezählt werden, da die Sandsäckchen genauer geworfen werden können als die Boule-Kugeln. In beiden Varianten können die Gruppen eigenständig die Schwierigkeit ihres Boule-Spiels variieren, indem sie die Entfernung der Werfenden zur Zielkugel entsprechend ihres Könnens festlegen oder einen größeren oder kleineren Abstand als Treffer werten.
In Aufgabenteil **A** erheben die Schülerinnen und Schüler die ersten Daten an ihrem Boulespiel und stellen damit eine Hypothese für die Trefferwahrscheinlichkeit auf, die als Grundlage für die Signifikanztests der folgenden Aufgabenteile dient. Dabei fassen sie ihre erhobenen Daten als binomialverteilt auf. In Aufgabenteil **B** testen sie ihre Hypothese mittels eines zweiseitigen Signifikanztests. Dabei bedenken die Lernenden auch mögliche Ursachen für das Verwerfen ihrer Hypothese. Für Aufgabenteil **C** wird die Situation ein wenig verändert, sodass die Schülerinnen und Schüler daraufhin einen einseitigen Signifikanztest durchführen. Zum Abschluss werden in Aufgabenteil **D** die Fehler erster und zweiter Art von den Lernenden berechnet und deren Relevanz in der Sachsituation sowie Möglichkeiten zur Verringerung der beiden Fehlerarten diskutiert.