Anmerkungen zum Telekolleg Stochastik von BR αlpha - eine Diskussionsgrundlage für den Stochastikunterricht

Frank Marohn, Würzburg

Zusammenfassung: *Der vorliegende Artikel beinhaltet kritische Anmerkungen zum Telekolleg Stochastik von BR αlpha, dem Fernseh-Bildungskanal des Bayerischen Rundfunks. Konkret geht es dabei um die Berechnung der Wahrscheinlichkeit für sechs Richtige im Zahlenlotto „6 aus 49", um das Testen von Hypothesen im Binomialmodell und um die Begriffe „Unwahrscheinlichkeit" und „unwahrscheinlich". Die angesprochenen Themen können im Schulunterricht behandelt und diskutiert werden.*

1 Einleitung

Mit der vorliegenden Arbeit soll die vom Autor gegenüber dem Bayerischen Rundfunk (BR) in einem Brief vorgebrachte Kritik einem größeren Publikum bekannt gemacht werden. Die kritischen Anmerkungen beziehen sich dabei auf einige Ausführungen im Telekolleg (TK) Stochastik. Die Kritikpunkte beinhalten fachliche wie didaktische Aspekte. Die Notwendigkeit einer Veröffentlichung ist der Tatsache geschuldet, dass – so das Fazit im Antwortschreiben des BR – „kein Anlass gesehen wird, die kritisierte Sendung zu bearbeiten" (auf die Argumente, die zu diesem Fazit führen, wird im Einzelnen eingegangen). Sicherlich: Eine Sendung kann nicht von heute auf morgen geändert werden. Aber eine Überarbeitung in näherer Zukunft wäre aus Sicht des Autors mehr als wünschenswert gewesen. Möge der an Stochastik interessierte Leser sich selbst ein Urteil darüber bilden.

Die Kritik bezieht sich auf die folgenden drei Themen:

- Wahrscheinlichkeit für 6 Richtige in Zahlenlotto „6 aus 49" (Abschnitt 2.1)
- Entscheidungsvorschrift beim Binomialtest (Abschnitt 2.2)
- Die Begriffe „Unwahrscheinlichkeit" und „unwahrscheinlich" (Abschnitt 2.3).

Die beiden erstgenannten Punkte betreffen das Telekolleg Stochastik mit dem Thema *Hypothesentest*; der dritte Punkt bezieht sich auf das Thema *Zufallsexperimente*. Die Sendungen des Telekollegs sind komplett online als Video verfügbar.

2 Die Kritikpunkte

Die folgenden Anmerkungen nehmen Bezug auf die entsprechenden Internetseiten des Telekollegs, im Folgenden zitiert mit TK Stochastik (2013),

2.1 Wahrscheinlichkeit für 6 Richtige im Zahlenlotto „6 aus 49"

Ausgangspunkt ist die Frage nach der Wahrscheinlichkeit für 6 Richtige im Zahlenlotto „6 aus 49". Nach der Feststellung, dass es $\binom{49}{6}$ mögliche Gewinnreihen gibt, heißt es im TK Stochastik (2013):

„Diese Trefferwahrscheinlichkeit beim Zahlenlotto lässt sich über die hypergeometrische Verteilung ermitteln. Sie führt zu der Formel für die Wahrscheinlichkeit $\binom{6}{r}/\binom{49}{6}$, wobei klein r für die Anzahl der Richtigen steht. Suchen wir nach sechs Richtigen, ist für r sechs einzusetzen, also $p = \binom{6}{6}/\binom{49}{6} = 1/13983816$."

Und hier setzt die Kritik an. Es wird zwar richtigerweise die hypergeometrische Verteilung genannt, die Wahrscheinlichkeit $p = \binom{6}{r}/\binom{49}{6}$ für r richtige Gewinnzahlen im Zahlenlotto „6 aus 49" ist aber falsch. Dass dies nicht die richtige Wahrscheinlichkeit sein kann, ist leicht einzusehen (natürlich nicht, wenn man nur $r = 6$ einsetzt). Für $r = 0$ und $r = 6$ impliziert $\binom{6}{0} = \binom{6}{6}$ die Aussage (mit der üblichen Konvention $0! = 1$), dass die Wahrscheinlichkeit, keine richtige Gewinnzahl zu haben, gleich der Wahrscheinlichkeit für sechs Richtige ist. Oder für $r = 1$ und $r = 5$ ergibt sich wegen $\binom{6}{1} = \binom{6}{5}$ die Folgerung, dass die Wahrscheinlichkeit für eine richtige Gewinnzahl gleich der Wahrscheinlichkeit ist, fünf richtige Gewinnzahlen zu ziehen. Die richtige Lösung unter Verwendung der hypergeometrischen Verteilung ist

$$p_r = \frac{\binom{6}{r}\cdot\binom{43}{6-r}}{\binom{49}{6}}, \quad r = 0,\ldots,6 \qquad (1)$$

Die Antwort des BR auf diese Kritik war, dass in einer Fernsehsendung nur eine begrenzte Sendezeit zur Verfügung steht, die nicht überzogen werden kann. Daher wurde in der Sendung klar darauf hingewiesen, dass die hypergeometrische Verteilung zu der vereinfachten angewendeten Formel führt, die zur Bestimmung von sechs Richtigen verwendet werden kann. Ferner war die Aufgabenstellung nicht für andere Trefferwahrscheinlichkeiten gedacht.

Dazu ist Folgendes anzumerken:

Ein Hinweis kostet höchstens so viele Sekunden Sendezeit wie die Ergänzung um den Binomialkoeffizienten „43 über 6 minus r", zumal weitere Erklärungen zu diesem Binomialkoeffizienten nicht notwendig gewesen wären (die Aussage $\binom{43}{0} = 1$ ist dem Lernenden bereits bekannt). Das obige Zitat aus dem TK überträgt sich wörtlich auf die Sendung. Mit keinem Wort findet sich der Hinweis, von dem im Antwortschreiben des BR die Rede ist.

Die Formulierungen im TK suggerieren eindeutig, dass hier r als „Platzhalter" dient. Der Lernende wird hier völlig zu recht für r auch fünf oder eins einsetzen wollen und kommt dann zu falschen Wahrscheinlichkeiten.

Um die Wahrscheinlichkeit p_6 zu berechnen, wird die hypergeometrische Verteilung nicht benötigt. Bei einem Laplace-Experiment, und dies wird unterstellt, sind alle möglichen Gewinnreihen gleichwahrscheinlich. Also $p_6 = 1/\binom{49}{6}$. Fertig! Die hypergeometrische Verteilung setzt im übrigen in der stochastischen Modellierung mittels einer Zählvariablen eine Laplace-Verteilung voraus (für Details sei auf Henze (2012), Kapitel 13, verwiesen).

Es stellt sich die Frage, warum bei der Berechnung von p_6 die hypergeometrische Verteilung überhaupt verwendet wird, kommt sie doch erst dann ins Spiel, wenn es um die Berechnung von p_r für $r = 1, \ldots, 5$ geht. Wenn die hypergeometrische Verteilung als bekannt voraus gesetzt wird, besteht keine Notwendigkeit, einen Binomialkoeffizienten (aus Zeitmangel) wegzulassen. Ist diese Verteilung nicht bekannt, so kann und sollte auf den Umweg über die hypergeometrische Verteilung verzichtet werden, zumal es ja (angeblich) nur um sechs Richtige geht. Wie auch immer: Der Lernende, der sich die hypergeometrische Verteilung anschaut, wird bei Anwendung der „vereinfachten Formel" ins Straucheln geraten. Darüber hinaus wird er sicherlich von selbst zu der Feststellung gelangen, dass die gesuchte Wahrscheinlichkeit gerade 1 geteilt durch die Anzahl der möglichen Gewinnreihen ist. Und er wird sich zu Recht fragen, warum man hier die hypergeometrische Verteilung braucht. Die gesuchte Wahrscheinlichkeit ergibt sich doch unmittelbar aus der Laplace-Annahme.

Folgende Lösungen hätten sich für das Telekolleg angeboten:

(a) Berechnung über die hypergeometrische Verteilung mit Platzhalter r unter Verwendung von (1)

(b) Berechnung über die hypergeometrische Verteilung ohne Platzhalter

$$p = \frac{\binom{6}{6} \cdot \binom{43}{0}}{\binom{49}{6}} = \frac{1 \cdot 1}{\binom{49}{6}}$$

(c) Berechnung direkt über die Laplace-Annahme.

Lösung (c) wird vom Autor präferiert. Sie ist unkompliziert, dem Lernenden verständlich und berücksichtigt im besonderen Maße den Aspekt einer begrenzten Sendezeit.

2.2 Entscheidungsvorschrift beim Binomialtest

In dem vorliegenden Beispiel des TK geht es um eine herkömmliche und eine neue Lehrmethode für Skianfänger. Von der alten Methode sei bekannt, dass die Anzahl der Stürze bei 15% liegt. Nach der neuen Methode stürzten nach einem Kurs 4 von 60 Teilnehmern. Lässt sich aus diesem Ergebnis schließen, dass die neue Lehrmethode besser ist als die alte oder handelt es sich um einen Zufall? Die Alternative im TK Stochastik (2013) lautet: „Die neue Methode ist besser als die alte Methode." Modelliert man wie im TK die vorliegende Situation durch Binomialverteilungen $B_{n;p}$ mit $n = 60$ und „Trefferwahrscheinlichkeiten" $p \in (0,1)$, so lässt sich das Testproblem mit Nullhypothese H_0 und Alternativhypothese H_1 formal schreiben als

$$H_0 : p \geq 0,15 \quad vs. \quad H_1 : p < 0,15$$

Zur Testentscheidung im TK Stochastik (2013) heißt es dazu, dass „...die Anzahl der gestürzten Skifahrer mit über 90-prozentiger Wahrscheinlichkeit zwischen 5 und 13 liegt, wenn die Nullhypothese als Basis dient. In der Studie sind aber nur vier gestürzt, somit ist die Hypothese H_0 zu verwerfen. Die Alternativhypothese H_1 hat Gültigkeit".

Hier besteht der folgende Einwand gegen diese Entscheidungsvorschrift: Zunächst ist festzustellen, dass es sich hier um ein *einseitiges* Testproblem handelt (zweiseitig bedeutet $H_1 : p \neq 0,15$). Wenn also die Anzahl der Stürze nur klein genug ist, so spricht dies für die Alternative und gegen die Nullhypothese. Die im TK vorgestellte Testentscheidung lautet nun: Keine Ablehnung von H_0 zum Signifikanzniveau von 0,1, falls die Anzahl der Stürze zwischen 5 und 13 liegt, sonst Ablehnung. Damit lautet der Annahmebereich $\{5,\ldots,13\}$ bzw. der Ablehnungsbereich (auch kritischer Bereich genannt) $\{0,\ldots,4\} \cup \{14,\ldots,60\}$. Da die beobachtete Anzahl der Stürze von 4 im Ablehnungsbereich liegt, wird H_0 zum Niveau $\alpha = 0.1$ verworfen. Aber, und dies gilt es zu beachten: Das Testproblem ist einseitig. Folglich muss der Ablehnungsbereich ebenfalls einseitig, und nicht zweiseitig sein! Auch ein beobachteter Wert von 14 (oder größer) würde nach dieser Entscheidungsregel zu einer Ablehnung von H_0 führen! Der Ablehnungsbereich muss folglich vom Typ $\{0,\ldots,c_{\text{krit}}\}$ sein, wobei der kritische Wert $c_{\text{krit}} \in \{0,\ldots,n\}$ von dem vorgegebenen Signifikanzniveau $0,1$ abhängt. Genauer ist c_{krit} so zu wählen, dass die Bedingungen

$$B_{60;0,15}(\{0,\ldots,c_{\text{krit}}\}) = \sum_{k=0}^{c_{\text{krit}}} \binom{60}{k} \cdot 0,15^k \cdot 0,85^{60-k} \le 0,1$$

und

$$B_{60;0,15}(\{0,\ldots,c_{\text{krit}}+1\}) > 0,1$$

erfüllt sind. Es stellt sich heraus, dass $c_{\text{krit}} = 5$ ist, der Ablehnungsbereich zum Signifikanzniveau $\alpha = 0,1$ also $\{0,\ldots,5\}$ ist. Demnach würde man sich auch bei einer beobachteten Anzahl von 5 Stürzen für die Alternative entscheiden.

Auch hier konnte die vorgebrachte Kritik nicht überzeugen. Die Vorgehensweise, so im Antwortschreiben des BR, sei geeignet und die vorgebrachte Kritik an dem einführenden Beispiel „Skikurs” dürfte eher eine Geschmacksfrage sein.

Auf die berechtigte Frage, warum man die neue Lehrmethode auch bei 14 beobachteten Stürzen als besser bewertet, bleibt der BR die Antwort schuldig. Die Wahrscheinlichkeit, 14 Stürze zu beobachten, ist im übrigen positiv:

$$B_{60;0,15}(\{14\}) = \binom{60}{14} \cdot 0,15^{14} \cdot 0,85^{46} = 0,028...$$

Plausibel und intuitiv einsichtig ist nur, sich für die (einseitige) Alternative zu entscheiden, wenn die Anzahl der Stürze nur „klein genug” ist. Eine andere Entscheidungsvorschrift macht keinen Sinn und ist nicht vermittelbar. Der am Binomialtest interessierte Leser sei auf Henze (2012) und Kütting und Sauer (2011) verwiesen.

2.3 Die Begriffe „Unwahrscheinlichkeit” und „unwahrscheinlich”

Im TK wird der Begriff „Unwahrscheinlichkeit” motiviert durch die Fragestellung: „Wir haben die Wahrscheinlichkeit berechnet - können wir auch die Unwahrscheinlichkeit berechnen?” Dann kommt die Definition der Unwahrscheinlichkeit:

$$P(\bar{E}) = 1 - P(E)$$

(Anmerkung: $\bar{E}$ bezeichnet das Gegenereignis.) Danach werden die Begriffe *wahrscheinlich* - *unwahrscheinlich* erläutert, dies am Beispiel des einfachen Würfelwurfs. Im TK Stochastik (2013) heißt es dazu:

„Auf das Beispiel mit dem Würfelwurf angewendet ist also die Wahrscheinlichkeit, eine bestimmte Zahl zu würfeln, $P(E) = 1/6$, ein Wert, der kleiner ist als 0,5, und somit ist das Eintreffen des Ereignisses *unwahrscheinlich*.” In der betreffenden Sendung heißt es wortwörtlich: „Ist $P(E)$ größer als $1/2$, so ist das Eintreffen des Ereignisses wahrscheinlich. Ist $P(E)$ kleiner als $1/2$, so ist das Eintreffen des Ereignisses unwahrscheinlich.”

Zunächst fällt auf, dass zwischen dem Substantiv *Unwahrscheinlichkeit* und dem Adjektiv *unwahrscheinlich* kein Zusammenhang besteht (den der Lernende aber zu Recht vermutet). Zu welchen Irritationen diese Begriffe führen, zeigt sich am Beispiel des einfachen Würfelwurfes. Die Wahrscheinlichkeit, eine bestimmte Augenzahl zu würfeln, ist $P(E) = 1/6$. Jetzt berechnet man die Unwahrscheinlichkeit: $P(\bar{E}) = 5/6$. Das Ereignis $\bar{E}$ ist aber wahrscheinlich. Das ist Hokuspokus!

In der Stochastik hat sich der Begriff „Unwahrscheinlichkeit” aus gutem Grund nicht etabliert. Man spricht sinnvollerweise von einer Gegenwahrscheinlichkeit oder Komplementärwahrscheinlichkeit. Diese Begriffsbildung ist auch ganz natürlich. Das Gegenereignis $\bar{E}$, man sagt auch Komplementärereignis, wird mengentheoretisch beschrieben durch die Komplementärmenge $\bar{E} = \Omega \setminus E := \{\omega \in \Omega : \omega \notin E\}$. Dabei bezeichnet Ω die Ergebnismenge des zugrundeliegenden Zufallsexperimentes. Der Begriff „Unwahrscheinlichkeit” ist in der Stochastik deplatziert.

Auch der im TK verwendete Begriff *unwahrscheinlich* ist mehr als problematisch, assoziiert man diesen Begriff üblicherweise mit „kaum möglich” oder „kleine Wahrscheinlichkeit”. Ein Ereignis E als unwahrscheinlich zu bezeichnen, falls $P(E) < 0,5$ widerspricht dieser Intuition.

Völlig unproblemtisch sind die Sprechweisen: Im Fall $P(E) < 1/2$ ist das Eintreten von E weniger wahrscheinlich als das Eintreten von $\bar{E}$ oder das Eintreten von $\bar{E}$ ist wahrscheinlicher als das Eintreten von E. Im üblichen Sprachgebrauch wird aber das Eintreten von E als (nahezu) unwahrscheinlich angesehen, falls $P(E)$ „hinreichend klein" ist. Auf eine nähere Quantifizierung durch die Angabe einer „benchmark" wird aus gutem Grunde verzichtet. Konsens besteht sicherlich darüber, dass es unwahrscheinlich ist, mit nur einem Tippfeld sechs Richtige zu bekommen, beträgt die Wahrscheinlichkeit nur $0,0000000715....$ Dagegen ein Ereignis mit einer Eintrittswahrscheinlichkeit von $1/6 \approx 16\%$ als unwahrscheinlich zu bezeichnen, ist nicht nachvollziehbar. Aus frequentistischer Sicht würde bei 100 Würfen eines echten Würfels „im Mittel" 16, 17 „Sechsen" beobachtet werden. Und dies soll unwahrscheinlich sein? Auch das Ereignis „Augensumme beim zweifachen Würfelwurf ist höchstens 5" wäre demnach ein unwahrscheinliches Ereignis, denn dieses besitzt eine Wahrscheinlichkeit von $10/36 = 0.2777... < 1/2$. Und diese Wahrscheinlichkeit ist kaum größer als $0,2775...$, jener Wahrscheinlichkeit für das Auftreten zweier identischer Gewinnreihen bei den Ziehungen der Lottozahlen am 20.12.1986 und 21.06.1995 (siehe Henze (2012), Kapitel 10). Alles andere als eine Lottosensation! Ein weiteres Beispiel aus dem Zahlenlotto „6 aus 49": Bei nahezu jeder zweiten Ziehung beobachtet man mindestens einen Zwilling, d.h. ein Paar $(i, i+1)$ von zwei aufeinander folgenden Gewinnzahlen. Alles andere als unwahrscheinlich, liegt die Wahrscheinlichkeit doch bei $1 - \binom{44}{6}/\binom{49}{6} = 0,495...$ (siehe z. B. Daume und Schmitz (2013)

Die Antwort des BR auf diese Kritik war, dass man in der Stochastik über angemessene Beispiele für Wahrscheinlichkeiten streiten kann und das Beispiele zumindest anschaulich und für den Lernenden interessant sein sollen.

Das Beispiel für 6 Richtige im Zahlenlotto „6 aus 49" wäre, ohne Angabe einer benchmark, sicherlich ein geeignetes Beispiel gewesen. Der einfache Würfelwurf im Zusammenhang mit „unwahrscheinlich" ist weder interessant noch anschaulich.

In der *subjektivistischen* Sichtweise, Wahrscheinlichkeit als (subjektiven) Gewissheitsgrad (Vertrauensgrad) zu intepretieren, mag ein Zahlenwert unter $0,5$ einen Gewissheitsgrad von „unwahrscheinlich" signalisieren, vgl. Kütting und Sauer (2011), Seite 119. Aber in diesem Sinne ist der im TK verwendete Wahrscheinlichkeitsbegriff nicht zu verstehen. Dagegen spricht der Würfelwurf, ist doch dieses Zufallsexperiment schlechthin der Prototyp eines beliebig oft wiederholbaren Zufallsexperimentes. Bei einer *frequentistischen* Auffassung des Wahrscheinlichkeitsbegriffes sollte auf eine Quantifizierung des Begriffes „unwahrscheinlich" verzichtet werden.

Einen ersten Überblick über die verschiedenen Bedeutungen des Begriffes Wahrscheinlichkeit gibt Stahel (2008), Abschnitt 4.10. Der Leser, der mehr über die Interpretationen des Wahrscheinlichkeitsbegriffes wissen möchte, sei auf den Artikel von Hájek (2012) und die dort zitierte Literatur verwiesen.

3 Schlussbemerkungen

Stochastik gilt gemeinhin als schwierig. Es sind die spezifischen Begriffe und Denkweisen der Stochastik, die dem Lernenden (anfangs) große Schwierigkeiten bereiten. Die Einführung einer neuen Begrifflichkeit wie „Unwahrscheinlichkeit" kann diese Schwierigkeiten nur vergrößern. [Ausgewiesene Didaktiker der Mathematik äußerten sich zum Begriff Unwahrscheinlichkeit gegenüber dem Autor von „noch nie gehört" bis „pervers".] Es ist kaum zu erwarten, dass der Lernende ein Verständnis für die Philosophie des statistischen Testens entwickelt, wenn inadäquate Testentscheidungen propagiert werden. Formeln nicht richtig wieder zu geben, aus welchen Gründen auch immer, ist inakzeptabel. Für Lehrende sollte es ein selbstverständliches Anliegen sein, korrektes Wissen zu vermitteln. Andernfalls tut Aufklärung Not.

Die angesprochenen Themen können sicherlich im Schulunterricht behandelt und diskutiert werden:

- Herleitung der Wahrscheinlichkeit für r richtige Lottozahlen, $0 \leq r \leq 6$
- (Logischer) Zusammenhang zwischen Alternativhypothese und Ablehnungsbereich
- Der Begriff „unwahrscheinlich".

Dabei sollten die entsprechenden Sendungen des Telekollegs mit einbezogen werden. Es wäre interessant zu erfahren, wie Schüler die Dinge sehen und einordnen. Studierende des Lehramts, die vom Autor auf diese Punkte aufmerksam gemacht wurden, konnten nur mit dem Kopf schütteln.

Literatur

Bayerischer Rundfunk (2013): Telekolleg Stochastik: http://www.br.de/telekolleg/faecher/mathematik/ (Abgerufen am 15.1.2014)

Daume, P. und Schmitz, M. (2013): Zwillinge und andere Mehrlinge beim Lotto: Stochastik in der Schule, Jahrgang 33, Heft 1

Hájek, A. (2012): Interpretations of Probability: The Stanford Encyclopedia of Philosophy (Winter Edition 2012) Ed.: E. N. Zalta http://plato.stanford.edu/archives/win2012/entries/probability-interpret/ (Abgerufen am 15.1.2014)

Henze, N. (2012): Stochastik für Einsteiger, 9. Auflage: Vieweg + Teubner, Stuttgart.

Kütting, H. und Sauer, M. J. (2011): Elementare Stochastik, 3. Auflage: Spektrum Akademischer Verlag, Heidelberg.

Stahel, W.A. (2008): Statistische Datenanalyse, 5. Auflage: Vieweg Verlag, Wiesbaden.

Anschrift des Verfassers
Frank Marohn
Institut für Mathematik
Universität Würzburg
Emil-Fischer-Str. 30
97074 Würzburg
marohn@mathematik.uni-wuerzburg.de