Einige grundsätzliche Überlegungen zu zwei Abituraufgaben

P. Laurie Davies, Essen

***Zusammenfassung**: Der Beitrag untersucht aus mathematischer Sicht zwei Aufgaben der zentralen Abiturprüfung 2008 in Nordrhein-Westfalen, und zwar die Aufgaben für den Leistungskurs Mathematik M LK HT 7 und M LK 7. Es geht einerseits um die unzulässige Vermischung von Modellebene und Beobachtungsebene und andererseits um mathematische Präzision. Die Analyse dieser Aufgaben soll dazu anregen, den dabei aufgeworfenen Problemen im Stochastikunterricht mit großer Sorgfalt zu begegnen und Lehrbuchaufgaben kritisch zu hinterfragen.*

1 Leistungskurs M LK HT 7

Ich beschränke mich auf die Teilaufgaben a) und b).

1.1 Worum es geht

Wirft man eine Eineuromünze, beträgt die Wahrscheinlichkeit für einen Zahlwurf $\frac{1}{2}$. Würfelt man mit einem normalen Würfel, beträgt die Wahrscheinlichkeit für eine bestimmte Augenzahl $\frac{1}{6}$. Beim Lotto beträgt die Wahrscheinlichkeit, dass eine bestimmte Zahl gezogen wird, $\frac{1}{49}$. Diese Wahrscheinlichkeiten werden durch Symmetrieüberlegungen ermittelt. In anderen Fällen, wo es keine Symmetrieargumente gibt, z. B. die Erfolgswahrscheinlichkeit bei Freiwürfen im Basketball, greift man auf empirische Ergebnisse zurück und die Wahrscheinlichkeit wird *geschätzt*. Krankenhaus A in Düsseldorf registriert in einem Jahr 514 Geburten, von denen 253 Mädchen sind. Die Mädchenquote beträgt 0,492. Krankenhaus B registriert 358 Geburten mit einer Mädchenquote von 0,45. Krankenhaus C hat eine Mädchenquote von 0,508. Diese drei Zahlen können nicht alle die Wahrscheinlichkeit für ein Mädchen sein, sonst hinge diese Wahrscheinlichkeit vom Krankenhaus ab. Will man ein Mädchen, dann Krankenhaus C, will man einen Jungen, dann Krankenhaus B. In der mathematischen Statistik unterscheidet man zwischen einer wahren Wahrscheinlichkeit und einer geschätzten Wahrscheinlichkeit. Diese begriffliche Unterscheidung ist von grundlegender Bedeutung und unerlässlich. In der Abituraufgabe werden die beiden Begriffe vermischt: Manchmal ist die Quote die wahre Wahrscheinlichkeit, manchmal ist sie ein Schätzwert. Somit ist die Aufgabe schlecht gestellt und die Lösung des Ministeriums falsch.

1.2 Modellierung

Bei der Aufgabe handelt sich um eine Anwendung der Binomialverteilung. Mit der Binomialverteilung modelliert man einen Vorgang, bei dem eine feste Anzahl n von Versuchen durchgeführt wird, wobei es bei jedem Versuch nur zwei Möglichkeiten gibt, die traditioneller Weise „Erfolg" und „Misserfolg" genannt werden. Es wird davon ausgegangen, dass die Erfolgswahrscheinlichkeit p für alle Versuche dieselbe ist, und dass sich die einzelnen Versuche nicht gegenseitig beeinflussen. Die Ergebnisse der einzelnen Versuche werden durch Zufallsgrößen $X_1,..., X_n$ bezeichnet, wobei $X_i = 1$ bzw. $X_i = 0$, falls der i-te Versuch ein Erfolg bzw. ein Misserfolg ist. Bei einer Textaufgabe muss die Modellierung entweder angegeben werden, oder sie muss aus der Beschreibung der Situation eindeutig hervorgehen.

Wenn die Modellierung feststeht, wird die Aufgabe innerhalb des Modells weiter bearbeitet. D. h., es wird angenommen, dass das Modell stimmt. Es ist nicht die Aufgabe des Abiturienten, die Angemessenheit der Modellierung zu hinterfragen. Bei der Binomialverteilung ist die Anzahl n von Versuchen bekannt. Bei der Erfolgswahrscheinlichkeit p gibt es zwei Möglichkeiten: Entweder p ist bekannt, oder p ist nicht bekannt und man will p aus der Versuchsreihe schätzen. In der letzteren Situation muss klar zwischen dem wahren Wert p und einem Schätzwert $\hat{p}_n$ unterschieden werden. Bei der Binomialverteilung ist der Schätzwert immer das arithmetische Mittel $\overline{X}_n = \frac{X_1 + \ldots + X_n}{n}$ der Versuchsreihe. Die klare Unterscheidung zwischen einem wahren Parameterwert und einem Schätzwert hierfür ist fundamental: Die ganze schließende Statistik basiert darauf.

1.3 *p* bekannt

Wenn die Anzahl n von Versuchen und die Erfolgswahrscheinlichkeit p bekannt sind, kann man mithilfe der Binomialverteilung Aussagen über die Anzahl der erfolgreichen Versuche in der Versuchsreihe treffen. Eine Münze wird fünfmal geworfen und wir suchen die Wahrscheinlichkeit dafür, dass wir höchstens drei „Zahl“-Würfe bekommen. Hier ist $n = 5$ und $p = \frac{1}{2}$. Mithilfe der Binomialverteilung ergibt sich eine Wahrscheinlichkeit von 0.8125. Die Wahrscheinlichkeit dafür, dass wir genau drei „Zahl“-Würfe bekommen beträgt 0.3125. Wir haben in diesem Beispiel die Erfolgswahrscheinlichkeit einfach gleich $\frac{1}{2}$ gesetzt. Bei einer normalen Münze scheint dies aus Symmetriegründen plausibel, denn es gibt im Normalfall keinen Grund, die eine oder die andere Seite der Münze vorzuziehen. Ähnliches gilt für einen Würfel oder für die Ziehung der Lottozahlen. Um ganz eindeutig zu sein, kann man von einer „fairen“ oder einer „unverfälschten“ Münze sprechen. Die Wahrscheinlichkeit p ist hier durch Symmetrieüberlegungen bestimmt. In den meisten Anwendungsfällen ist aber eine Symmetrie nicht vorhanden und die Wahrscheinlichkeit p muss anders ermittelt werden. Ein Beispiel ist das Geschlecht eines neugeborenen Kindes. Man kann schreiben „Aus langjähriger Erfahrung ist bekannt, dass die Wahrscheinlichkeit, ein Mädchen zu bekommen, 0,481 beträgt“ oder, noch einfacher „Die Wahrscheinlichkeit, ein Mädchen zu bekommen, beträgt 0,481“, oder „Bei einem gefälschten Würfel beträgt die Wahrscheinlichkeit für die Augenzahl Sechs 0,183“. Es ist nicht Aufgabe des Abiturenten, nachzufragen, wie die Werte 0,481 oder 0,183 ermittelt wurden, er hat sie lediglich bei den weiteren Berechnungen einzusetzen. Auf die Frage, wie solche Werte ermittelt werden, gehen wir später ein, weil dies für die Bearbeitung der Abituraufgabe relevant ist. Auf jeden Fall gibt es bei bekanntem p eine klare Sprachreglung, die einzuhalten ist, und aus der der Wert von p eindeutig hervorgeht.

1.4 *p* unbekannt

Bei Geburten stellt man anhand der Daten ziemlich schnell fest, dass es mehr Jungen als Mädchen gibt. Will man die Wahrscheinlichkeit für eine Mädchengeburt bestimmen, so muss man so viele unverfälschte Daten wie möglich, die über Zeit und Raum homogen sind, bekommen, und daraus die relative Häufigkeit von Mädchengeburten bestimmen. Der Einfachheit halber nehmen wir an, dass wir die Daten für eine Millionen Geburten haben, wovon 481 345 Mädchengeburten waren. Man könnte nun p durch die Zahl 0,481345 *schätzen*, aber man weiß, dass dies nicht der „wahre“ Wert von p ist: Würde man eine faire Münze mit $\frac{1}{2}$ eine Million Mal werfen, dann wäre man sehr überrascht, wenn es genau 500 000 Zahlwürfe gäbe. In der Statistik gibt man deswegen ein sogenanntes Konfidenzintervall von plausiblen Werten für p an. Die Bestimmung eines Konfidenzintervalls in dieser Situation ist Bestandteil des Leistungskurses in der Mathematik. Für die oben angegeben Daten erhalten wir als 95%-Konfidenzintervall für p das Intervall [0.4804, 0.4823]. Wenn man berücksichtigt, dass wir eine Million Datensätze haben, ist es vielleicht überraschend, dass wir sowohl 0.4804 als auch 0.4823 als plausible Werte für p ansehen. Stehen uns nur 1 000 Geburten zur Verfügung, von denen 481 Mädchen sind, so bekommt man als Konfidenzintervall [0.450, 0.512]. Hier ist die Unsicherheit über den wahren p-Wert ziemlich groß.

1.5 Teil a) der Aufgabe

Dieser Aufgabenteil lautet:

Der deutsche Basketball-Profi Dirk Nowitzki spielt in der amerikanischen Profiliga NBA beim Club Dallas Mavericks. In der Saison 2006/2007 erzielte er bei Freiwürfen eine Trefferquote von 90,4 %.

a) Berechnen Sie die Wahrscheinlichkeit, dass er

(*1*) *genau 8 Treffer bei 10 Versuchen erzielt,*

(*2*) *höchstens 8 Treffer bei 10 Versuchen erzielt,*

(*3*) *höchstens viermal nacheinander bei Freiwürfen erfolgreich ist.*

Zuerst ist es aus der Formulierung der Aufgabe klar, dass der Wert 90,4 % aus den Daten der Saison 2006/2007 empirisch festgestellt wurde. Interpretiert man 90,4 % als eine Wahrscheinlichkeit von 0,904, so handelt es sich eindeutig um einen *Schätzwert*, nämlich das arithmetische Mittel der Beobachtungen, und nicht um die wahre unbekannte Wahrscheinlichkeit p. An dieser Stelle wird die Quote 0,904 vom Ministerium als die wahre Wahrscheinlichkeit betrachtet:

„Die Zufallsvariable X für die Anzahl der Treffer bei 10 Versuchen ist $B_{10;0,904}$*-verteilt.“*

Später in Teil b) der Aufgabe wird eine solche Quote als Schätzwert betrachtet. Es liegt also eine Begriffsverwirrung vor, weil nicht konsistent unterschieden wird zwischen der wahren Wahrscheinlichkeit p und einem Schätzwert $\overline{X}_n$ für diese Wahrscheinlichkeit. Wenn der Stichprobenumfang, auf dem der Schätzwert basiert, hinreichend groß ist, ist die Differenz klein, und gerade diese Tatsache rechtfertigt Formulierungen wie „aufgrund langjähriger Erfahrung weiß man, dass die Erfolgswahrscheinlichkeit 0,9 beträgt“. Eine solche Formulierung wurde nicht gewählt, im Gegenteil, der Hinweis auf die Saison 2006/2007 verdeutlicht, dass der Stichprobenumfang eher klein war. In Teil b) der Aufgabe werden genaue Angaben über die Daten gemacht. Danach erzielte Nowitzki 498 Treffer bei 551 Versuchen. Berechnet man hierfür ein 95%-Konfidenzintervall für die wahre Wahrscheinlichkeit p, so bekommt man [0.8792, 0.9284]. Damit ist die einzig plausible Lösung zu (1) die Behauptung, dass die gesuchte Wahrscheinlichkeit zwischen 0.2345 (p = 0.8792) und 0.1273 (p = 0.9284) liegt.

Die Teilaufgabe (3) lässt mehrere Interpretationen zu. Ich analysiere drei. Da wir hierfür einen Wert für die „wahre" Wahrscheinlichkeit brauchen, werde ich trotz der oben angegebenen Kritik den Wert 0,904 einsetzen.

- In der Presse wird behauptet, dass die Anzahl der Versuche fehlt, um die Aufgabe lösen zu können. Da aber die Anzahl 10 in (1) und (2) angegeben wird, liegt es nahe, auch 10 in (3) anzunehmen. Nimmt man an, dass 10 Versuche durchgeführt wurden, ist die Aufgabe nun wohl gestellt, aber die Berechnung der Lösung ist äußerst langwierig: Von den $2^{10} = 1024$ möglichen Versuchsreihen muss man die Anzahl derjenigen bestimmen, in denen die Eins höchstens viermal nacheinander vorkommt. Dies macht man am besten, indem man zuerst die Gesamtanzahl k von Erfolgen festlegt. Bei $k = 10$ und $k = 9$ überlegt man sich schnell, dass es keine Möglichkeit gibt, die Einsen so zu platzieren, dass sie höchstens viermal nacheinander vorkommen. Bei $k = 8$ gibt es 15 Möglichkeiten, die hier aufgelistet sind:

(0111101111)	(1011101111)	(1011110111)
(1101101111)	(1101110111)	(1101111011)
(1110101111)	(1110110111)	(1110111011)
(1110111101)	(1111001111)	(1111010111)
(1111011011)	(1111011101)	(1111011110)

  Nun kann man das Ganze wiederholen mit $k = 7$ usw. Es ist klar, dass der Schwierigkeitsgrad und der dafür notwendige Zeitaufwand einer (Teil-) Abituraufgabe nicht angemessen sind.

- Wir nehmen nun an, dass die Anzahl der Versuche $n = 5$ ist. Das einzige, das für diese Interpretation spricht, ist, dass die Lösung mit der Lösung des Ministeriums übereinstimmt. Zu der Lösung schreibt das Ministerium:

  „Man betrachtet das Gegenereignis, dass er fünf Treffer hintereinander schafft ..."

Das Gegenereignis von „höchstens vier" ist nicht „fünf", sondern „mindestens fünf". Es ist nur dann fünf, wenn er genau fünfmal wirft. Diese Information steht aber nirgends in der Aufgabe und es gibt keinen Grund, dies anzunehmen. Wenn er nur fünfmal wirft, dann ist die Lösung des Ministeriums ($1 - 0{,}904^5 = 1 - 0{,}6037 = 0{,}3963$) sowie die Begründung korrekt. In allen anderen Fällen ist die Begründung falsch.

- In der dritten Interpretation stellen wir uns vor, dass beim Training folgendes Spiel zwischen zwei Spielern gespielt wird: Einer fängt an (entschieden durch einen Münzwurf) und macht Freiwürfe bis zum ersten Fehlwurf. Der zweite Spieler ist nun an der Reihe und macht Freiwürfe bis zum ersten Fehlwurf usw. Nowitzki spielt dieses Spiel gegen einen Teamkameraden und fängt an. Man bestimme die Wahrscheinlichkeit dafür, dass er bei seinem ersten Versuch höchstens viermal erfolgreich war. Die möglichen Versuchsfolgen sind

  0, 10, 110, 1110, 11110

  mit den Wahrscheinlichkeiten

  $0{,}096,\quad 0{,}904 \cdot 0{,}096,\quad 0{,}904^2 \cdot 0{,}096,$

  $0{,}904^3 \cdot 0{,}096\quad 0{,}904^4 \cdot 0{,}096.$

  Man rechnet nun aus, dass die gewünschte Wahrscheinlichkeit $1 - 0{,}904^5 = 0{,}3963$ beträgt. Man stellt fest, dass auch diese Lösung mit der Lösung des Ministeriums übereinstimmt. Die Begründung ist aber eine ganz andere. Die Interpretation ist die einzige, bei der die Anzahl der Würfe nicht von vornherein festgelegt ist. Ein sehr guter Schüler hätte durchaus diese Variante wählen können.

1.6 Teil b) der Aufgabe

Wir zitieren diesen Aufgabenteil:

Bei Heimspielen hatte er eine Freiwurfbilanz von 267 Treffern bei 288 Versuchen, bei Auswärtsspielen lag die Quote bei 231:263. Ein Sportreporter berichtet, dass Dirk Nowitzki auswärts eine deutlich schwächere Freiwurfquote habe.

Untersuchen Sie auf einem Signifikanzniveau von 5 %, ob die Trefferanzahl bei Auswärtsspielen

(*1*) *signifikant unter dem Erwartungwert für Heim- und Auswärtsspiele liegt,*

(*2*) *signifikant unter dem Erwartungswert für Heimspiele liegt.*

Gehen wir Teilaufgabe (1) durch. Hier wird der Begriff „Erwartungswert" erwähnt. Führt man n Versuche durch mit einer konstanten Erfolgswahrscheinlichkeit p, so ist der Erwartungswert der Trefferanzahl $T_n = X_1 + ... + X_n$ genau np: In der Schreibweise der Statistik $E(T_n) = np$. Dabei ist p der wahre unbekannte Wert für die Wahrscheinlichkeit. Aus der Aufgabe geht hervor, dass wir zwischen Heim- und Auswärtsspielen unterscheiden müssen. Wir bezeichnen die Erfolgswahrscheinlichkeit für Heimspiele mit p_H und für Auswärtsspiele mit p_A. Dann ist der Erwartungswert der Trefferanzahl für Heimspiele $n_H p_H$ und für Auswärtsspiele $n_A p_A$. In der Schreibweise der Statistik, nun etwas umständlicher, um Heim- und Auswärtsspiele unterscheiden zu können

$$E\left(T_{n_H}^H\right) = n_H p_H \ , \ E\left(T_{n_A}^A\right) = n_A p_A.$$

Nun, die wahren Werte p_H und p_A kennen wir nicht; Das Ministerium ersetzt sie einfach durch die Schätzwerte $\overline{X}_{n_H}^H$ und $\overline{X}_{n_A}^A$. Es schreibt dazu

„Es sei X $\quad B_{263;0,904}$-verteilt;“ ...,

woraus die Verwechslung eines Schätzwertes mit dem wahren Parameterwert sichtbar ist. Damit ist der „Erwartungswert“ für Heim- und Auswärtsspiele nun nichts anderes als $T_{n_H}^H + T_{n_A}^A$. Die Trefferanzahl bei Auswärtsspielen ist $T_{n_A}^A$ und wir müssen testen, ob $T_{n_A}^A$ signifikant kleiner ist als $T_{n_H}^H + T_{n_A}^A$. Dies ist nun ja immer der Fall, insofern man wenigstens einen Treffer in einem Heimspiel erzielt hat. Setzt man die Daten ein, müssen wir testen, ob 231 < 267 + 231, was ja stimmt. Hier liegt vermutlich ein Tippfehler vor: Man meint nicht die „Trefferanzahl“, sondern die Trefferquote. Wiederholen wir nun das Ganze mit Trefferquote.

Die Trefferquote für Heimspiele bzw. Auswärtsspiele ist $\overline{X}_{n_H}^H$ bzw. $\overline{X}_{n_A}^A$ mit den Erwartungswerten

$$E\left(\overline{X}_{n_H}^H\right) = p_H, \quad E\left(\overline{X}_{n_A}^A\right) = p_A$$

Der Erwartungswert der Trefferquote für Heim- und Auswärtsspiele beträgt

$$\frac{n_H p_H + n_A p_A}{n_H + n_A}$$

und wir sollen nun anhand der Daten testen, ob

$$p_A < \frac{n_H p_H + n_A p_A}{n_H + n_A}.$$

Da aber die wahren p_H und p_A nicht bekannt sind, hat das Ministerium sie in dem Ausdruck $\frac{n_H p_H + n_A p_A}{n_H + n_A}$ einfach durch $\overline{X}_{n_H}^H$ und $\overline{X}_{n_A}^A$ ersetzt und tat so, als ob sie die wahren Werte sind (s. o.). Gleichzeitig wird aber in der Lösung des Ministeriums $\overline{X}_{n_A}^A$ als Schätzwert für p_A behandelt. Mit anderen Worten, $\overline{X}_{n_A}^A$ wird manchmal als der wahre Wert betrachtet und manchmal als ein Schätzer hierfür und noch dazu im selben Satz. Die Begriffsverwirrung ist komplett.

Die vom Ministerium angebotene Lösung zeugt von einer völligen Begriffsverwirrung. Das Ministerium unterscheidet nicht zwischen dem wahren Wert p und dem Schätzwert $\overline{X}_n$ für p. Die Verwirrung geht soweit, dass in ein und derselben Aufgabe $\overline{X}_n$ sowohl als der wahre Wert als auch als der Schätzwert behandelt wird. Die Lösung des Ministeriums ist unsinnig und falsch. Die Unterscheidung von Schätzwerten und wahren Werten ist von grundlegender Bedeutung in der Statistik: Die ganze Schätz- und Testtheorie, d. h. die ganze schließende Statistik, basiert darauf. Ein Schüler, der richtig gelernt hat, diese Begriffe zu trennen, kann diese Aufgabe nicht lösen. Die Versuche, die Wahrscheinlichkeitstheorie und die Statistik als Teilgebiete der Mathematik an der Schule zu etablieren, etwa dadurch, dass die Stochastik inzwischen Pflichtbereich für Lehramtskandidaten des Faches Mathematik ist, werden durch solche fehlerhafte Aufgabenstellungen konterkariert.

2 Leistungskurs M LK 7

2.1 Worum es geht

Es geht um den Umkehrschluss. Die Mathematik zeichnet sich durch die Strenge ihrer Argumentation aus.

Ein Ingenieur, ein Physiker und ein Mathematiker fahren mit dem Zug durch England. Durch das Fenster sehen sie ein einzelnes Schaf. Sagt der Ingenieur: „Alle Schafe in England sind schwarz.“ Sagt der Physiker: „Es gibt schwarze Schafe in England.“ Sagt der Mathematiker: „Es gibt mindestens ein Schaf in England, von dem mindestens eine Seite schwarz ist.“

Im Leben ist eine solche Strenge nicht immer ratsam, weil sie zusätzliche Information, die für die Behauptung relevant ist, ausschließt. Jeder, der in England war, weiß, dass es viele Schafe gibt, dass die Mehrheit weiß ist, und dass ein Schaf, ob weiß oder schwarz, dieselbe Farbe auf beiden Seiten hat. So verhält es sich auch mit dem Umkehrschluss: Im Leben ist der Umkehrschluss oft berechtigt, obwohl er weder logisch noch empirisch zwingend ist. Wenn es regnet, wird die Straße nass. Im Umkehrschluss, ist die Straße nass, dann hat es geregnet. Im Allgemeinen stimmt dieser Umkehrschluss, und wenn er nicht stimmt, sucht man eine Erklärung, z. B. eine geplatzte Wasserleitung. Ein Umkehrschluss kann aber nicht nur logisch, sondern auch empirisch daneben liegen. Alle Menschen sind Lebewesen. Der Umkehrschluss: Alle Lebewesen sind Menschen.

In der Mathematik gibt es aber kein Pardon. Entweder stimmt der Umkehrschluss oder er stimmt nicht, und für den letzteren Fall reicht eine einzige Ausnahme oder ein Gegenbeispiel aus. Im Mathematikunterricht unterscheidet man zwischen notwendig (das brauchen wir) und hinreichend (das genügt). Um ein Mensch zu sein, ist es notwendig, ein Lebewesen zu sein. Um ein Lebewesen zu sein, ist es hinreichend, ein Mensch zu sein. Der Umkehrschluss ist nur dann gerechtfertigt, wenn hinreichend und notwendig zusammenfallen. Manchmal wissen die Mathematiker nicht, ob der Umkehrschluss richtig ist. Hier ist ein berühmtes Beispiel. Die Summe zweier Primzahlen $p_1 + p_2$ mit $p_1, p_2 \geq 3$ ergibt eine gerade Zahl $n \geq 6$. Dies stimmt und ist leicht zu zeigen. Der Umkehrschluss lautet: Jede gerade Zahl ist die Summe zweier Primzahlen $p_1 + p_2$ mit $p_1, p_2 \geq 3$. Ob dies stimmt, ist nicht bekannt: Es gibt weder einen Beweis noch ein Gegenbeispiel.

Die Unterscheidung zwischen notwendig und hinreichend ist für die Mathematik essenziell und wird nicht nur im Leistungskurs, sondern auch im Grundkurs Mathematik gelehrt, z. B. im Zusammenhang mit der Kurvendiskussion, wo der Unterschied sehr stark betont wird. Ist eine Funktion zweimal differenzierbar, dann ist eine erste Ableitung von null eine notwendige Bedingung für ein lokales Maximum. Einfache Beispiele zeigen, dass die Bedingung nicht hinreichend ist. Eine erste Ableitung von null und die zweite Ableitung kleiner als null ist eine hinreichende Bedingung für ein lokales Maximum. Wieder zeigen einfache Beispiele, dass diese Bedingung nicht notwendig ist. Ein Schüler, der behauptet, es liege ein lokales Maximum vor, weil die erste Ableitung null ist, würde hierfür keine Punkte bekommen. Dies gilt nicht nur für die Kurvendiskussion und die Analysis. Es gilt für die Mathematik im Ganzen, also auch für die Wahrscheinlichkeitstheorie: Entweder wird der Umkehrschluss bewiesen oder er wird durch ein Gegenbeispiel widerlegt oder man schüttelt den Kopf und gibt zu, es nicht zu wissen.

2.2 Teil a) der Aufgabe

Die Aufgabe lautet:

Begründen Sie, dass nur eine der beiden Zufallsvariablen binomialverteilt ist.

Hier ist ein Satz aus der Wahrscheinlichkeitstheorie.

Satz 1: Es seien $X_1,..., X_n$ unabhängig verteilte Zufallsvariable mit

$P(X_i = 0) = 1 - p, P(X_i = 1) = p, i = 1,..., n.$

Dann genügt die Summe $S_n = X_1 + ... + X_n$ der Binomialverteilung $B_{n;p}$.

Mithilfe dieses Satzes folgt, dass die Zufallsvariable X_2 der Aufgabe der $B_{10;1/6}$-Verteilung genügt. Die Argumentation des Ministeriums ist in diesem Falle korrekt. Um die Aufgabe vollständig zu lösen, muss noch gezeigt werden, dass die Zufallsvariable X_1 der Binomialverteilung *nicht* genügt. Das Ministerium argumentiert mit dem Umkehrschluss:

Die Zufallsvariable X_1 ist nicht binomialverteilt, da die Wahrscheinlichkeit, eine Kugel abzulegen, vom Zustand der Ablegeplätze abhängig und somit bei den einzelnen Würfen unterschiedlich ist.

Dieses Argument ist nur richtig, wenn der Umkehrschluss zu Satz 1 stimmt. Er lautet:

„Satz 2“: Es seien $X_1, ..., X_n$ Zufallsvariable. Genügt die Summe $S_n = X_1 + ... + X_n$ der Binomialverteilung $B_{n;p}$, dann sind $X_1, ..., X_n$ unabhängig verteilt und es gilt

$P(X_i = 0) = 1 - p, P(X_i = 1) = p, i = 1, ..., n.$

Diesen Satz finden Sie in keinem Buch und aus gutem Grund, er stimmt nämlich nicht. Hier ist ein Gegenbeispiel mit $n = 2$:

$$P(X_1 = 0) = \frac{1}{3}, \quad P(X_1 = 1) = \frac{2}{3},$$

$$P(X_2 = 0 \mid X_1 = 0) = \frac{3}{4}, \quad P(X_2 = 1 \mid X_1 = 0) = \frac{1}{4},$$

$$P(X_2 = 0 \mid X_1 = 1) = \frac{5}{8}, \quad P(X_2 = 1 \mid X_1 = 1) = \frac{3}{8}.$$

Die Zufallsvariablen X_1 and X_2 sind nicht unabhängig verteilt und es gilt

$$P(X_1 = 0) = \frac{1}{3}, \quad P(X_1 = 1) = \frac{2}{3},$$

$$P(X_2 = 0) = \frac{2}{3}, \quad P(X_2 = 1) = \frac{1}{3},$$

d. h., die Erfolgswahrscheinlichkeit ist nicht konstant. Trotzdem genügt $X_1 + X_2$ der Binomialverteilung $B_{2;0,5}$.

Obwohl das Argument des Ministeriums falsch ist, ist die Behauptung selbst richtig. Nach jedem Wurf hat man eine Kugel mehr oder eine Kugel weniger. Nach zwei Würfen hat man entweder 2 Kugeln mehr, dieselbe Anzahl von Kugeln oder 2 Kugeln weniger,

d. h., der Unterschied ist entweder 0 oder eine gerade Zahl. Nach fünfmal zwei Würfen ist der Unterschied somit entweder null oder eine gerade Zahl. Mit anderen Worten, die Wahrscheinlichkeit, dass X_1 eine ungerade Zahl ist, ist null und somit ist X_1 nicht binomialverteilt. Falle ich mit dieser Lösung durch? Die Frage ist nicht rhetorisch. Ein Schüler schreibt:

„X_1 ist nicht binomialverteilt, weil X_1 immer eine gerade Zahl ist."

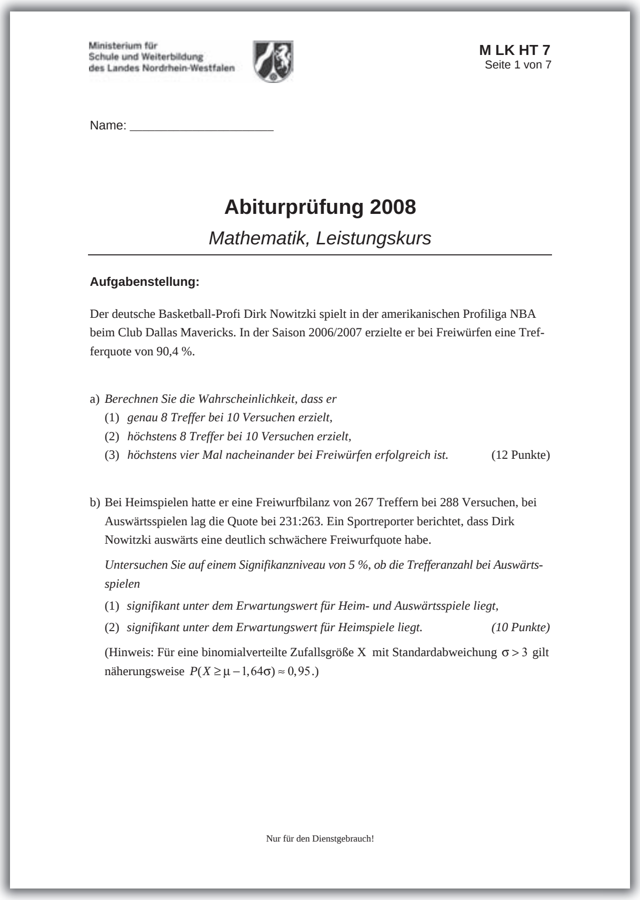
Die Antwort besticht durch Einfachheit, Klarheit und, obendrein, Korrektheit. Wie aber reagiert ein Lehrer bei der Korrektur? Die Begründung hat nichts, aber gar nichts mit der „Lösung" des Ministeriums zu tun. Es ist gut möglich, dass der Lehrer nicht weiter darüber nachdenkt und die Lösung als falsch bewertet. Immerhin hat die „Lösung" des Ministeriums alle Kontrollinstanzen durchlaufen, ohne dass es jemandem aufgefallen ist, dass X_1 immer eine gerade Zahl ist.

3 Fazit

Im Abitur 2008 waren beide Aufgaben im Bereich der Stochastik fehlerhaft.

Anschrift des Verfassers

P. Laurie Davies
Universität Duisburg-Essen
Fachbereich Mathematik
45117 Essen
laurie.davies@uni-due.de



Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen

M LK HT 7
Seite 1 von 7

Name: ____________________

Abiturprüfung 2008

Mathematik, Leistungskurs

Aufgabenstellung:

Der deutsche Basketball-Profi Dirk Nowitzki spielt in der amerikanischen Profiliga NBA beim Club Dallas Mavericks. In der Saison 2006/2007 erzielte er bei Freiwürfen eine Trefferquote von 90,4 %.

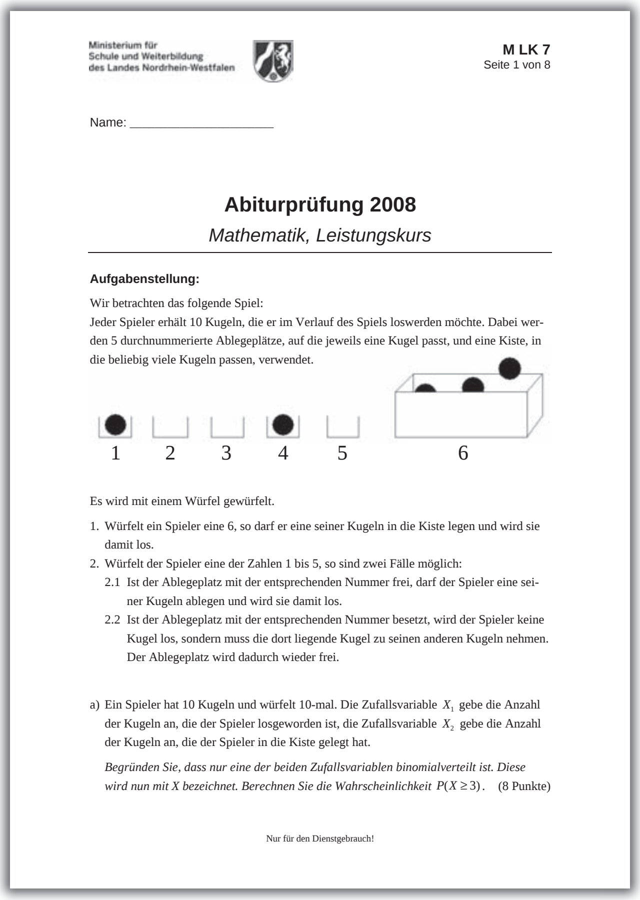
a) *Berechnen Sie die Wahrscheinlichkeit, dass er*
(1) *genau 8 Treffer bei 10 Versuchen erzielt,*
(2) *höchstens 8 Treffer bei 10 Versuchen erzielt,*
(3) *höchstens vier Mal nacheinander bei Freiwürfen erfolgreich ist.* (12 Punkte)

b) Bei Heimspielen hatte er eine Freiwurfbilanz von 267 Treffern bei 288 Versuchen, bei Auswärtsspielen lag die Quote bei 231:263. Ein Sportreporter berichtet, dass Dirk Nowitzki auswärts eine deutlich schwächere Freiwurfquote habe.

Untersuchen Sie auf einem Signifikanzniveau von 5 %, ob die Trefferanzahl bei Auswärtsspielen
(1) *signifikant unter dem Erwartungswert für Heim- und Auswärtsspiele liegt,*
(2) *signifikant unter dem Erwartungswert für Heimspiele liegt.* *(10 Punkte)*

(Hinweis: Für eine binomialverteilte Zufallsgröße X mit Standardabweichung $\sigma > 3$ gilt näherungsweise $P(X \geq \mu - 1{,}64\sigma) \approx 0{,}95$.)

Nur für den Dienstgebrauch!



Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen

M LK 7
Seite 1 von 8

Name: ____________________

Abiturprüfung 2008

Mathematik, Leistungskurs

Aufgabenstellung:

Wir betrachten das folgende Spiel:
Jeder Spieler erhält 10 Kugeln, die er im Verlauf des Spiels loswerden möchte. Dabei werden 5 durchnummerierte Ablegeplätze, auf die jeweils eine Kugel passt, und eine Kiste, in die beliebig viele Kugeln passen, verwendet.

1 2 3 4 5 6

Es wird mit einem Würfel gewürfelt.

1. Würfelt ein Spieler eine 6, so darf er eine seiner Kugeln in die Kiste legen und wird sie damit los.
2. Würfelt der Spieler eine der Zahlen 1 bis 5, so sind zwei Fälle möglich:
   2.1 Ist der Ablegeplatz mit der entsprechenden Nummer frei, darf der Spieler eine seiner Kugeln ablegen und wird sie damit los.
   2.2 Ist der Ablegeplatz mit der entsprechenden Nummer besetzt, wird der Spieler keine Kugel los, sondern muss die dort liegende Kugel zu seinen anderen Kugeln nehmen. Der Ablegeplatz wird dadurch wieder frei.

a) Ein Spieler hat 10 Kugeln und würfelt 10-mal. Die Zufallsvariable X_1 gebe die Anzahl der Kugeln an, die der Spieler losgeworden ist, die Zufallsvariable X_2 gebe die Anzahl der Kugeln an, die der Spieler in die Kiste gelegt hat.

Begründen Sie, dass nur eine der beiden Zufallsvariablen binomialverteilt ist. Diese wird nun mit X bezeichnet. Berechnen Sie die Wahrscheinlichkeit $P(X \geq 3)$. (8 Punkte)

Nur für den Dienstgebrauch!