

Kein Witz!? Zur Nowitzki-Aufgabe im NRW Zentralabitur 2008

RAPHAEL DIEPGEN, BOCHUM

Zusammenfassung: Politisch skandalisierende Kritik an einer Stochastikaufgabe im Zentralabitur Nordrhein-Westfalen 2008 zwang die Ministerialbürokratie zur Wiederholung der entsprechenden schriftlichen Prüfungen. Der Beitrag zeigt aus stochastikdidaktischer Perspektive, inwiefern die publikumswirksame Kritik in den Medien viel zu kurz gegriffen hat.

1 Die Nowitzki-Aufgabe und ihr politisches Schicksal

Obwohl Hauptautor eines seit Jahrzehnten weit verbreiteten Stochastikschulbuchs (Diepgen, Kuypers, Rüdiger, 1989) erfahre ich nur ganz selten etwas über die Wirklichkeit des nicht zuletzt wohl auch auf dieses Buch gestützten Stochastikunterrichts, denn ich verlor schon vor vielen Jahren meine Stelle im NRW-Schuldienst. Im Frühjahr 2008 drang dann aber doch immer wieder etwas über diese Wirklichkeit zu mir durch, und zwar just bei der morgendlichen Lektüre der Tagespresse: Eine Stochastikaufgabe im ohnehin umstrittenen NRW-Zentralabitur war dort über Wochen Gegenstand öffentlicher Kritik – so lange und so heftig, bis NRW-Schulministerin Sommer (CDU) nach wochenlanger Leugnung der Probleme gezwungen war oder wurde, den betroffenen Abiturienten das Neuschreiben anzubieten. Ein Sommerloch ganz neuer Art! Wenn das vornehmste Lehrziel des Stochastikunterrichtes sein sollte, zwischen Wissen und Vermuten, zwischen Sicherheit und Wahrscheinlichkeit zu unterscheiden, so schien jedenfalls Ministerin Sommer dieses Lernziel nicht erreicht zu haben.

Die Aufgabe

Und so lautete die inkriminierte Aufgabe für den Leistungskurs – ähnlich etwas verkürzt auch für den Grundkurs:

Der deutsche Basketballprofi 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 $P(X \geq \mu - 1,64\sigma) \approx 0,95$.)

- c) In der Vorbereitung zur nachfolgenden Saison vermutet der Trainer, dass die Quote seines Schützlings gesunken ist. Bevor er mit dieser Vermutung an die Öffentlichkeit geht, möchte er aber anhand der ersten 50 Freiwürfe in der neuen Saison überprüfen, ob diese Aussage auf einem Signifikanzniveau von 10 % gesichert ist.

Bestimmen Sie eine Entscheidungsregel. Gehen Sie dabei von der gerundeten (Vorjahres-)Trefferquote von 90 % aus. (10 Punkte)

- d) Dirk Nowitzki bestritt 41 Heim- und 37 Auswärtsspiele mit den bei b) gegebenen Trefferbilanzen. Es sei bekannt, dass er in einem Spiel der Saison 10 von 10 Freiwürfen getroffen hat.

Ermitteln Sie die Wahrscheinlichkeit dafür, dass es sich bei diesem Spiel um ein Heimspiel handelte. (8 Punkte)

- e) Es wird nun allgemein ein Spieler betrachtet, der einen Freiwurf mit der Wahrscheinlichkeit p trifft. Ganz am Ende eines Spieles kann es zu der Situation kommen, dass der Spieler noch 2 Freiwürfe erhält und das Spiel unmittelbar danach beendet ist. Für einen verwandelten Freiwurf erhält die Mannschaft einen Punkt.

(1) Beurteilen Sie in Abhängigkeit des vorherigen Spielstandes die Siegchancen für die Mannschaft des Spielers, der die letzten 2 Freiwürfe erhält. Gehen Sie dabei davon aus, dass im Fall eines Unentschiedens in der folgenden Verlängerung die Siegchance für beide Mannschaften 50 % beträgt. Geben Sie für die einzelnen Fälle die Wahrscheinlichkeiten in Abhängigkeit von p an.

(2) Vor den beiden letzten Freiwürfen sei Gleichstand. Bestimmen Sie p so, dass die Gewinnwahrscheinlichkeit für die Mannschaft des werfenden Spielers mindestens 99,5 % beträgt. (10 Punkte)

Als Hilfsmittel zugelassen waren ein wissenschaftlicher Taschenrechner, eine mathematische Formelsammlung und ein Wörterbuch zur deutschen Rechtschreibung. Als Anlagen waren der Aufgabe beigelegt Tabellen zur Binomialverteilung zu ausgewählten Parametern und eine Tabelle zur Standardnormalverteilung.

Das politische Ende der Aufgabe

Erzwungen wurde der Rückzug der Ministerialbürokratie durch die öffentliche Kritik an der offensichtlichen Unklarheit der Aufgabenstellung in a) (3): Ohne Angabe der Anzahl n der Freiwürfe, innerhalb derer dann immer höchstens vier Erfolge nacheinander vorkommen dürfen, lässt sich die gestellte Frage offensichtlich nicht beantworten. Intuitiv ist klar: Für n gegen unendlich geht die gesuchte Wahrscheinlichkeit sicherlich gegen Null – auch wenn dies formal durch einen Abiturienten nicht ganz einfach zu beweisen sein dürfte. Ist vielleicht $n = 10$ gemeint – wie in den vorangehenden Aufgabenteilen? Dann wäre die Lösung für einen Abiturienten zwar wohl machbar, allerdings ziemlich komplex – und mit vermutlich nur vier Punkten auch ziemlich unterbewertet. Dass hier von den Aufgabenautoren schlicht die Wahrscheinlichkeit $1 - 0,904^5$ gemeint war, dass Dirk Nowitzki bei den jeweils nächsten Freiwürfen spätestens beim fünften Versuch Misserfolg (also nicht fünfmal hintereinander Erfolg) hat, war der missglückten Aufgabenformulierung nicht klar zu entnehmen – allenfalls aufgrund unterstellter Schlichtheit der Aufgabenautoren zu vermuten. Bei einem Zentralabitur aber hat der Kandidat zu derartigen Unterstellungen keinen vernünftigen Grund, anders als beim traditionellen Abitur mit Aufgabenstellung durch den ihm bekannten Lehrer. Unterstellt man, dass die prognostische Validität von Prüfungen auch daraus resultiert, dass die Fähigkeit des Prüflings mitgeprüft wird, sich der subjektiven Sprach- und Denkwelt seines Prüfers anzupassen, zeigt sich hier exemplarisch ein Problem für das Zentralabitur, dessen Verfechter ja grundsätzlich etwas naiv dazu neigen, den vermeintlichen Objektivitätsgewinn der zentralistischen Einheitlichkeit mit einem Validitätsgewinn gleichzusetzen.

2 Die eigentlichen Probleme der Nowitzki-Aufgabe

Shit happens. Unklar gestellte oder nicht zu Ende gedachte Aufgaben kommen vor, gerade in der Stochastik, auch bei den vermeintlichen Expertenteams für das Zentralabitur – und bemerkenswert wäre hier allenfalls die selbst für Politiker ungewöhnliche Hartnäckigkeit, mit der der offensichtliche Lapsus in a) (3) über viele Wochen gelegnet wurde.

Die eigentlichen Probleme dieser Aufgabe liegen woanders – und diese eröffnen einen eher ernüchternden Blick auf den inzwischen augenscheinlich etablierten Stochastikunterricht, der sich an solchen Abituraufgaben sowohl offenbart als auch notgedrungen orientieren muss – Anlass genug für einen ehemaligen Schulbuchautor, sich die Hände in Unschuld zu waschen.

Relevanz?

Dass es sich bei dieser Aufgabe um eine (schlecht) „eingekleidete“ Textaufgabe alter Tradition und nicht etwa um eine sinnvolle ergebnisoffene Problemstellung handelt – geschenkt! Das Zentralabitur wird hier absehbar alle neuen Ansätze ihrer kaum prüfungstauglichen Offenheit wegen zerstören. Dass hier Abiturienten gezwungen werden, sich ausgerechnet mit den „Problemen“ eines ultra-reichen Sportstars zu beschäftigen – auch dies sei geschenkt. Jeder Realitätsbezug von Aufgaben zur Stochastik (entstanden schließlich aus der Analyse moralisch fragwürdiger Glücksspiele reicher Müßiggänger) trifft notwendig auf eine allemal heikle und ganz und gar nicht wertfreie Welt. Aber vielleicht darf ich mir angesichts der auch politischen Konnotationen der Nowitzki-Aufgabe – Stichworte: Starkult, Werbefigur, Extremreichtum, gesellschaftlicher Nutzen des Durchwerfens hoch gehängter Körbe mit roten Bällen – eine Erinnerung erlauben: Seinerzeit führte eine Aufgabe in meinem Schulbuch zu einer empörten kleinen Anfrage im NRW-Landtag aus den Reihen der CDU, weil man die Moral gefährdet sah: In der Aufgabe ging es lediglich um die Wahrscheinlichkeit für ein Mädchen, nach einem positiven Schwangerschaftstest tatsächlich schwanger zu sein!

Grundsätzlicher und fachimmanenter sind die folgenden Probleme:

Binomialverteilung?

Augenscheinlich wird eine Lösung auf Basis der Binomialverteilung erwartet. Es wird also vom

Schüler erwartet, dass er selbstverständlich die Freiwürfe von Dirk Nowitzki als strukturgleich zu den Würfeln einer Münze interpretiert. Insbesondere wird damit unterstellt, dass es in gleicher Weise für die Nowitzkifrei- wie für Münzwürfe als konstante Trefferwahrscheinlichkeiten interpretierbare Parameter p dieser Binomialverteilung gibt, und zwar insbesondere folgende vier:

Trefferwahrscheinlichkeit bei Heimspielen in der Saison 2006/2007: p_1

Trefferwahrscheinlichkeit bei Auswärtsspielen in der Saison 2006/2007: p_2

Trefferwahrscheinlichkeit (bei Heim- und Auswärtsspielen) in der Saison 2006/2007: p_3

Trefferwahrscheinlichkeit (bei Heim- und Auswärtsspielen) in der Saison 2007/2008: p_4

Wie wären diese sich jeweils über eine ganze Saison nicht ändernden Trefferwahrscheinlichkeiten zu verstehen? Als objektive, jeweils für genau eine Saison stabile Eigenschaft von Dirk Nowitzki – so wie die Trefferwahrscheinlichkeit einer Münze als deren objektive Eigenschaft gilt? Wohl kaum. In diesem Sinne wäre die über Heim- wie Auswärtsspiele entsprechend den jeweiligen (zufälligen?) Freiwurfhäufigkeiten gewichtet „summierende“ Wahrscheinlichkeiten p_3 kaum noch als Trefferwahrscheinlichkeit eines Freiwurfs interpretierbar. Geht es doch wohl eher – ganz anders als beim Münzwurf – um subjektive Wahrscheinlichkeiten? Wie groß erscheint die betreffende Wahrscheinlichkeit dem Subjekt, das lediglich weiß oder eben noch nicht einmal das weiß, um welche Saison und ggf. welche Art von Spiel es sich handelt, und welches lediglich die in der Aufgabe präsentierten Häufigkeiten kennt? Wer mehr weiß, etwa den aktuellen Trainingszustand, den Tabellenstand, den Spielstand, die Tages- und Jahreszeit, die Stimmung, den Gesundheitszustand, den Kontostand, ..., der würde hier zu anderen Wahrscheinlichkeiten gelangen: Würde man Dirk Nowitzki selbst vor jedem Freiwurf nach seiner subjektiv eingeschätzten Trefferwahrscheinlichkeit für diesen Freiwurf fragen, so dürfte man ganz andere, insbesondere stark variierende Werte erhalten. Und trotzdem soll der Schüler selbstverständlich die Strukturgleichheit zum Münzwurf mit seiner konstanten objektiven Trefferwahrscheinlichkeit unterstellen?!

Oder bezieht sich die Aufgabe implizit auf ein Metaexperiment, etwa folgender Art? Die Ergebnisse aller (also nicht etwa nur einiger zufällig gezogener – dies wird unten noch wichtig!) Nowitzkifreiwürfe einer Saison werden auf Zettelchen

notiert (für jeden Freiwurf ein Zettelchen) und diese dann – ggf. noch getrennt nach Heim- und Auswärtsspielen – in eine Urne gelegt; aus dieser Urne wird dann „blind“ (mit Zurücklegen) jeweils ein Zettelchen gezogen, und die obigen Trefferwahrscheinlichkeiten sind dann die Wahrscheinlichkeiten, dass auf dem gezogenen Zettelchen „Treffer“ steht. Anders als durch dieses oder ein strukturgleiches fiktives Metaexperiment scheint die von den Aufgabenautoren nur implizit gemeinte stochastische Situation kaum rekonstruierbar. Solange die Trefferwahrscheinlichkeit (nur) als eine Disposition von Dirk Nowitzki gilt, wie es der Aufgabentext nahe legt, ist die mit der Binomialverteilung unterstellte stochastische Unabhängigkeit der Würfe gänzlich unplausibel: Die Trefferwahrscheinlichkeit auch eines coolen Profis wie Dirk Nowitzki dürfte im 101. Freiwurf nicht dieselbe sein, wenn er die hundert Freiwürfe vorher alle verwandelt, wie wenn er sie alle nicht verwandelt hat. Ganz grundsätzlich dürfte im Bereich menschlichen, also jeweils auf vorangehende Ereignisse hochkomplex reagierenden Handelns, Verhaltens oder Tuns – auch Sport gehört dazu – stochastische Unabhängigkeit die seltene Ausnahme, jedenfalls nicht die Regel sein.

Kurzum: Die vom Schüler als ganz selbstverständlich erwartete, aber explizit überhaupt nicht legitimierte Anwendung der Binomialverteilung ist alles andere als trivial, sondern ergibt sich nur bei ganz bestimmten Rekonstruktionen der Aufgabenstellung. Dass Stochastik immer nur in mehr oder minder vereinfachenden Modellen arbeiten kann, bedeutet für den Stochastikunterricht aber gerade, die Voraussetzungen und Restriktionen solcher Modelle explizit – und damit kritisierbar und revidierbar – zu machen: Solche Voraussetzungen implizit zu lassen, verbietet sich daher insbesondere für eine Stochastikaufgabe im abschließenden Abitur. Konkret: Der eher problembewusste Schüler, der bei der Nowitzki-Aufgabe unter Hinweis auf die augenscheinliche stochastische Abhängigkeit der Freiwürfe die gestellten Fragen für unbeantwortbar erklärt, hat möglicherweise mehr verstanden als der Schüler, der die Fragen routinemäßig beantwortet und sich der dabei implizit gemachten Annahmen gar nicht bewusst wird.

Die Nichtexplizierung impliziter, ggf. unplausibler Modellannahmen hat nicht nur ideologischen Charakter, insofern sie die Möglichkeiten der Mathematisierung von Realität größer erscheinen lässt, als sie wohl tatsächlich ist. Sie führt bisweilen auch zu folgenreichen inhaltlichen Fehleinschätzungen. Beispiel: Luftfahrtingenieure gingen lange Zeit von

einem faktisch zu vernachlässigenden Restrisiko aus, dass bei einem zweistrahligen Verkehrsflugzeug gleich beide Triebwerke ausfallen, bis sie ein tragischer Unfall eines Besseren belehrte, nämlich dass der Ausfall eines Triebwerks die Ausfallwahrscheinlichkeit des anderen Triebwerks erhöht – vermittelt über die durch den ersten Triebwerksausfall erhöhte psychische Belastung und Fehlerträchtigkeit der Piloten –, die bequemer Weise unterstellte stochastische Unabhängigkeit in der beruhigenden Risikokalkulation also eine irrije Unterstellung war.

Zweites Beispiel: Vor einigen Jahren ging es in der Politik um die Frage, ob und welche Psychotherapien von den gesetzlichen Krankenkassen zu finanzieren seien. Es war klar, dass hier nicht alle, sondern nur die „nachweislich“ besten Psychotherapien eine Chance hatten. Die vorliegenden Vergleichsstudien hatten aber viel zu geringe Stichprobengrößen, um bei den realistischerweise zu erwartenden mäßigen Unterschieden zwischen den verschiedenen Psychotherapien Signifikanz erwarten zu können. In dieser Situation publizierte der „Papst“ der deutschen Psychotherapieforschung, zugleich Mitglied der entsprechenden Enquetekommission des Deutschen Bundestages, einen programmatischen Artikel mit folgender statistischer Argumentation bezüglich einer vorher publizierten Vergleichsstudie: In dieser Studie wurden – relativ wenig – Patienten per Zufall den Psychotherapieverfahren A bzw. B zugeordnet. Im Verlauf der Therapie und auch noch danach wurden an den Patienten zu mehreren Messzeitpunkten in mehreren – mehr oder minder hoch korrelierten – Variablen (wie Depressivität, Selbstvertrauen, ...) Messungen vorgenommen und die Mittelwerte dieser Messungen jeweils für beide Gruppen verglichen. Dabei zeigten sich kaum signifikante Unterschiede. Bei 90 von den insgesamt 113 Vergleichen lag aber der Mittelwert der Therapiegruppe A – ggf. nur einen Deut – über dem der Gruppe B. Dies – so nun die frappierende Argumentation des renommierten Psychotherapieforschers – müsse man aber als einen „hochsignifikanten“ Beleg für die Überlegenheit der Psychotherapie A interpretieren: Wären nämlich beide Therapien gleich gut, so müsste im Zweigruppenvergleich jeweils mit einer Wahrscheinlichkeit von 50 % der Mittelwert bei Therapie A höher als der bei Therapie B sein. Diese Nullhypothese müsse aber im Binomialtest zurückgewiesen werden, denn dass eine $B(113, 0,5)$ -verteilte Prüfgröße mindestens den Wert 90 erreiche, habe eine Wahrscheinlichkeit von weniger als 0,001. Abgesehen davon, dass diese „metaanalytische“ Konstruktion eines „Signifikanztestes“ erkenntlich völlig unab-

hängig von der Größe der untersuchten Stichprobe war (und auch noch für $n = 2$ gelten würde), wurde von diesem Experten (und seinen Gutachtern) völlig die hochkomplexe stochastische Abhängigkeitsstruktur zwischen den vielen Zweigruppenvergleichen übersehen, welche hier die Anwendbarkeit der Binomialverteilung ausschloss: Hat Therapiegruppe A zu einem Messzeitpunkt eine geringere Depressivität als Gruppe B, dann steigt dadurch natürlich die Wahrscheinlichkeit, dass dies auch noch für den nächsten Messzeitpunkt gilt – und dass Gruppe A zu beiden Messzeitpunkten auch höheres Selbstvertrauen hat. Auch dieses Beispiel aus der Forschung zeigt, wie gefährlich es wird – bis hin zu politischen Fehlentscheidungen –, wenn man sich nicht mehr bewusst bleibt, an entscheidender Stelle stochastische Unabhängigkeit vorausgesetzt zu haben.

Während die bisherigen Einwände schon gegen Teil a) der Nowitzki-Aufgabe aus eher didaktisch-pädagogischer Perspektive die mangelnde Explikation (und Fraglichkeit) von impliziten Voraussetzungen monierten – also gleichsam mangelnde intellektuelle Reflektiertheit und Redlichkeit –, geht es in den nächsten beiden Punkten um krasse innermathematische Fehler und Widersprüche, und zwar in Teil b).

Schätzer und/oder Parameter?

Zur Bestimmung bzw. empirischen Schätzung der oben genannten Parameter werden in Aufgabenteil b) folgende relativen Häufigkeiten genannt:

267 Treffer bei 288 Versuchen bei Heimspielen in der Saison 2006/2007 (gleich bzw. Schätzer für p_1)

231 Treffer bei 263 Versuchen bei Auswärtsspielen in der Saison 2006/2007 (gleich bzw. Schätzer für p_2)

$267 + 231 = 498$ Treffer von $288 + 263 = 551$ Versuchen (bei Heim- und Auswärtsspielen) in der Saison 2006/2007 (gleich bzw. Schätzer – etwa 90,4 % – für p_3)

Es bleibt etwas unklar, ob die genannten Parameter p_1 , p_2 , und p_3 , also Wahrscheinlichkeiten, gleich den genannten relativen Häufigkeiten und damit grundsätzlich bekannt sind (Interpretation 1) oder durch diese nur mehr oder minder zufallsfehlerbehaftet geschätzt werden, letztlich also grundsätzlich unbekannt sind (Interpretation 2).

Für die Interpretation 1 (Trefferwahrscheinlichkeit gleich beobachteter relativer Trefferhäufigkeit, also auch Erwartungswert für die Trefferhäufigkeit gleich beobachteter Trefferhäufigkeit) spricht eigentlich der gesamte Aufgabenkontext, unbeschadet der oben

skizzierten grundsätzlichen Schwierigkeiten mit dem Wahrscheinlichkeitskonzept. Schon Aufgabenteil a) lässt sich nur bei dieser Interpretation lösen, denn die dort genannte und als Wahrscheinlichkeit zu interpretierende „Trefferquote“ von 90,4 % entspricht ja gerade der berichteten relativen Häufigkeit $(267 + 231) : (288 + 263)$. Es ist auch nirgends die Rede davon, die 288 bzw. 263 Freiwürfe seien nicht alle der Saison 2006/2007, sondern nur eine (zufällige?) Stichprobe daraus. Aber in dieser Interpretation machte die inferenzstatistische Fragestellung unter b) („signifikant“) keinen Sinn: Da sich dort die Worte „Trefferanzahl“ und „Erwartungswert“ äquivalent durch die Worte „relative Trefferhäufigkeit“ und „Trefferwahrscheinlichkeit“ ersetzen lassen, wäre in b) angesichts der unterstellten Gleichheit jeweils von relativer Trefferhäufigkeit und zugehöriger Trefferwahrscheinlichkeit danach gefragt, ob $p_2 = 231 : 263 = 0,878$ „signifikant“ (auf dem 5 %-Niveau) unter $p_1 = 267 : 288 = 0,927$ bzw. $p_3 = (231 + 267) : (263 + 288) = 0,904$ liege. Eine sinnlose Frage! Trivialerweise unterschieden sich dann diese Parameter – und die Frage nach der „Signifikanz“ dieser Unterschiede machte überhaupt keinen Sinn. Anders formuliert: Da in dieser Interpretation die berichteten relativen Häufigkeiten jeweils ihrem Erwartungswert entsprächen, wäre in b) danach gefragt, ob der Erwartungswert bei Auswärtsspielen „signifikant“ unter dem Erwartungswert für Heim- und Auswärtsspiele bzw. für Heimspiele liege. Auch dies eine zutiefst sinnlose Frage! Oder noch einmal in der klassischen Begrifflichkeit der Inferenzstatistik formuliert: Interpretation 1 bedeutete, dass sich in den relativen Häufigkeiten jeweils die gesamte Population repräsentierte, dass sie jeweils aus einer Vollerhebung stammten. Die Frage aber, ob sich die aus einer Vollerhebung resultierenden Kennwerte zweier Populationen „signifikant“ unterschieden, wäre völlig sinnlos: Sie unterschieden sich nicht oder eben doch – aber niemals „signifikant“. Ein Signifikanztest macht überhaupt nur Sinn, wenn es Stichproben aus Populationen gibt, wenn es also des zufälligen „Stichprobenfehlers“ wegen überhaupt irgendeine Unsicherheit gibt. Genau diese gäbe es aber in der Interpretation 1 nicht: Ja, es gäbe noch nicht einmal irgendeine Zufallsvariable, auf die sich die als Signifikanzniveau benannte Wahrscheinlichkeit von 5 % überhaupt beziehen könnte. Die Aufgabenautoren hätten – sich – über diese Unsinnigkeit mit ihrer suggestiven Aufgabenformulierung hinweggetäuscht, die den Eindruck erweckt, als sei die Trefferanzahl 231 (von 263) bei Auswärtsspielen eine zufällige Größe und kein – nicht zufälliger – Erwartungswert, die Trefferanzahlen 498 (von 551) bzw. 267 (von 288)

bei Heim- und Auswärtsspielen bzw. Heimspielen im Unterschied dazu aber durchaus nicht zufällige Erwartungswerte.

Also zur Interpretation 2, die eigentlich dem übrigen Aufgabenkontext wider-, aber der üblichen inferenzstatistischen Perspektive entspricht: Die Parameter selbst sind und bleiben hier grundsätzlich unbekannt, sie werden durch die genannten relativen Häufigkeiten nur geschätzt, und zwar – insbesondere die ersten beiden – in etwa gleich gut. In der Aufgabenstellung von b) wird dann aber plötzlich implizit mit der Rede von den entsprechenden Erwartungswerten unterstellt, dass die beiden Parameter p_1 und p_3 doch bekannt, nämlich durch die oben genannten relativen Häufigkeiten fehlerlos geschätzt seien, einzig der – tatsächlich genau so gut geschätzte – Parameter p_2 weiterhin unbekannt, seine Schätzung also nicht perfekt, sondern mit einem zufälligen Stichprobenfehler behaftet sei. Dies ist nun wirklich Irrwitz und entbehrt jeder tieferen Logik! Wenn man aufgrund der genannten relativen Häufigkeiten unterstellt, man kenne die Parameter p_1 und p_3 und könne daher entsprechende Erwartungswerte berechnen, dann muss man auch konsequenter Weise unterstellen, man kenne p_2 . Dann wären allerdings die Fragestellungen von b) (1) und b) (2) wieder sinnlos, denn bei Kenntnis aller Parameter kann man zwar formal, nicht aber im wohlverstandenen Sinne eines Signifikanztestes danach fragen, ob sich der nun willkürlich als Zufallsgröße interpretierte Schätzer für den einen (bekannten!) Parameter „signifikant“ von dem anderen (bekannten!) Parameter unterscheidet. Weiß man schon, dass zwei interessierende Parameter verschieden sind, dann ist die Frage, ob der Schätzer für den einen Parameter „signifikant“ von dem anderen Parameter abweicht, keine sinnvolle Frage für einen Signifikanztest mehr, sondern allenfalls Anlass für formale Spielerei: Denn ein Signifikanztest prüft ungewisse Hypothesen, nie aber Hypothesen, derer man sich schon gewiss ist. Ein reflektierter Schüler hätte beide Fragen b) (1) und b) (2) als zutiefst „sinnlos“, nämlich mit dem Sinn eines Signifikanztests unvereinbar, zurückweisen müssen.

Anders formuliert: In der Interpretation 2 wird von dem Schüler eine geradezu groteske Inkonsequenz verlangt. Er muss die Zufallsabhängigkeit der Schätzung von p_1 und p_3 ignorieren, sonst kommt er nicht weiter, die gänzlich gleichartige Zufallsabhängigkeit des Schätzers für p_2 darf er aber gleichzeitig nicht ignorieren, sonst kommt er auch nicht weiter.

Das Ganze mündet in eine Karikatur eines Signifikanztestes: Ganz gleich, ob man nun die Konzep-

tion von Fisher oder aber die Konzeption von Neyman und Pearson zugrunde legt, in jedem Fall muss sich die Nullhypothese eines Signifikanz- oder Hypothesentests unabhängig von den für den Test gewonnenen bzw. zu gewinnenden empirischen Daten formulieren lassen, darf also eine dort erhobene zufällige relative Häufigkeit nie und nimmer nullhypothetischer Parameterwert sein. Ganz krass wird dies bei dem in b) (1) verlangten „Signifikanztest“: Hier nämlich bestimmt die zufällige „Prüfgröße“ dieses „Tests“ etwa zur Hälfte den – eigentlich immer nichtzufälligen (!) – nullhypothetischen Parameterwert mit. Dass in einer Abituraufgabe tatsächlich ein Signifikanztest verlangt wird, bei dem die Nullhypothese von der zu ihrer Prüfung erhobenen Zufallsgröße (mit)bestimmt wird – dies wäre ein netter Witz, wären sich die Autoren dieser Absurdität bewusst gewesen; tatsächlich wollten sie aber – als mutmaßliche „No-Witz-kis“ – bestimmt keinen Witz machen, sondern haben das ernst gemeint. (Hätten sie ihre Fragestellung b) (1) durchdacht für das Szenario, dass Dirk Nowitzki seine Freiwürfe fast nur bei Auswärtsspielen geworfen hätte, dass also die relative Trefferhäufigkeit bei Auswärtsspielen als Schätzer für p_2 fast vollständig auch den nullhypothetischen Parameterwert p_3 bestimmt hätte, wäre ihnen vielleicht die Absurdität ihrer Fragestellung aufgegangen.)

Fazit aus der Diskussion beider Interpretationen: Die Aufgabenautoren scheinen im Kern nicht verstanden zu haben, worum es beim statistischen Testen eigentlich geht: nicht um eine formalistische Spielerei oder Rechnerei, sondern um ein für bestimmte Situationen sinnvolles Verfahren zur Beurteilung von oder Entscheidung über statistische, also ungewisse Hypothesen. Mit wohlverstandener Stochastik oder Inferenzstatistik hat die Aufgabenstellung b) – gleich in welcher Interpretation – nichts mehr zu tun.

Vermutlich hat diese Absurdität einen Grund auch darin, dass sich die interessierenden Nullhypothesen $p_2 = p_3$ und $p_2 = p_1$ angesichts einer Zufallsfehlerbehaftetheit der Schätzer aller drei Parameter schlicht nicht auf Basis der Binomialverteilung testen ließen, die halt nur die Zufälligkeit jeweils eines dieser Schätzer zu modellieren gestattet. Hier hätte es statt des einzig auf der Schule verfügbaren Binomialtestes eines anderen Testes – nämlich für eine Vier-Felder-Tafel, nicht nur eine Zwei-Felder-Tafel – bedurft, etwa des exakten Fisher-Yates-Tests auf Basis der hypergeometrischen Verteilung oder des asymptotischen Chi-Quadrat-Tests auf Basis der Normalverteilung.

Verwerfen und zugleich beibehalten?

Geradezu verrückt wird die Aufgabenstellung b), wenn man schließlich bedenkt, dass die in (1) zu prüfende Nullhypothese $p_2 = p_3$ und die in (2) zu prüfende Nullhypothese $p_2 = p_1$ äquivalent sind, also denselben hypothetischen Zustand der Welt (und darin Dirk Nowitzkis) beschreiben: Da die über alle Spiele gemittelte Trefferquote p_3 definitionsgemäß ein gewichteter Durchschnitt der Heimspieltrefferquote p_1 und der Auswärtstrefferquote p_2 ist, gilt $p_2 = p_3$ dann und nur dann, wenn $p_2 = p_1$ gilt. Damit wird aber in (1) und (2) ein und dieselbe Nullhypothese geprüft, und zwar zu ein und demselben Signifikanzniveau von 5 % mit ein und demselben Binomialtest auf der Basis ein und desselben Datums, nämlich dass die Trefferquote bei Auswärtsspielen 231 : 263 war. Es ist nun für jeden, der auch nur ansatzweise die Logik des Testens verstanden hat, von vornherein schier undenkbar, dass man hier die einheitliche Nullhypothese in (2) verwirft, in (1) aber nicht – wie es sich aber als „Lösung“ ergibt, wenn man ebenso brav wie sinnentleert jeweils das von den Aufgabenautoren vorgezeichnete μ und σ berechnet und in die vorgegebene Formel einsetzt. Spätestens hier hätte sich den Aufgabenautoren doch beim Durchrechnen ihrer Aufgabe die Frage aufdrängen müssen: Hoppla, wie kann es denn sein, dass wir auf der Basis eines identischen empirischen Datums in einem identischen Signifikanztest auf dem identischen Signifikanzniveau von 5 % eine identische Nullhypothese in Teil (1) verwerfen und zugleich in Teil (2) nicht verwerfen? Dass die Aufgabenautoren aber selbst an dieser grotesk absurden Stelle nicht innegehalten und ihre Wahnsinnskonstruktion abgebrochen haben, zeigt ganz krass, wie sehr sie in ihrer bloß formalen Spielerei das Verständnis für die zu modellierende stochastische Situation gänzlich verloren haben. Hätten sie, und darum geht's nun mal beim statistischen Test einer Hypothese, von der zu testenden Hypothese aus gedacht – und nicht vom Datum oder vom Testalgorithmus her –, dann hätten sie wohl gemerkt, dass es sich in (1) und (2) um dieselbe Hypothese handelt, die kaum in (1) und (2) unterschiedlich bewertet werden kann. Es scheint mir eine typische Deformation schulischen Stochastikunterrichtes zu sein, dass zumeist nicht stochastische Situationen mit interessierenden Hypothesen den Unterricht strukturieren, sondern – weitgehend losgelöst oder entfremdet davon – die innermathematische Logik von einfachen Stichprobenverteilungen und ihrer algorithmischen Beherrschung.

Nach dem mathematisch zwingenden Aufweis der – nun tatsächlich wohl skandalösen – logischen

Widersprüche in Teil b) der Nowitzki-Aufgabe seien nun gegen die folgenden Aufgabenteile einige weitere Einwände vorgebracht, die zwar mathematisch nicht zwingend sein mögen, mir aber dennoch stochastikdidaktisch dringend geboten scheinen.

Fehler zweiter Art?

Während sich Aufgabenteil b) – wenn auch grotesk verunglückt – an der Fisher'schen Idee des Signifikanztestes orientiert („Halte eine Nullhypothese für unplausibel, wenn die nullhypothetische Überschreitungswahrscheinlichkeit beobachteter Daten hinreichend klein ist“), orientiert sich nun Aufgabenteil c) – zumindest auf den ersten Blick – an der „moderneren“ Idee des Hypothesentestens als eines vor der Datenerhebung im Hinblick auf seine Fehlerwahrscheinlichkeiten beider Art fixierten regelhaften Entscheidungsverfahrens im Sinne von Neyman und Pearson. Aber statt nun die hier testkonstitutive Fehlerwahrscheinlichkeit zweiter Art ins Spiel zu bringen, wird der Test durch eine völlig unmotivierte Stichprobengröße fixiert (warum um Gottes willen soll der Trainer genau 50 Versuche abwarten?) – und so landet man wieder trotz entscheidungstheoretischer Rhetorik beim traditionellen Fisher'schen Signifikanztest. Dass den Aufgabenautoren der elementare Unterschied beider inferenzstatistischer Konzepte gar nicht bewusst zu sein scheint, sie vielmehr – wie vermutlich immer noch der Großteil der Mathematiklehrer – ganz und gar im Fisher'schen Konzept verhaftet scheinen, darauf deutet auch die für ein Entscheidungsverfahren wenig passende Rede von einem „Signifikanzniveau von 10 %“ statt von einer „Fehlerwahrscheinlichkeit erster Art von 10 %“ hin. So ist auch schließlich kaum erkennbar, inwiefern Aufgabenteil c) irgendetwas nennenswert anderes als Aufgabenteil b) überprüfte.

Wenn denn historisch der wesentliche Rationalitätsgewinn der Testtheorie von Neyman und Pearson gegenüber der älteren Signifikanzkonzeption von Fisher gerade darin bestand, dass nun die Wahl der Stichprobengröße begründet werden konnte – nämlich im Hinblick auf die erwünschten Fehlerwahrscheinlichkeiten beider Art –, dann wirkt es befremdlich, wenn dieser grundlegende Ansatz in c) durch die völlig unmotivierte Vorgabe einer Stichprobengröße konterkariert wird.

Eigentlich „schreit“ übrigens die in c) entworfene Situation nach einem sequentiellen Test, denn es soll ja offensichtlich möglichst schnell entschieden werden. Augenscheinlich aber haben meine damaligen Versuche, das sequentielle Testen auch seiner didaktisch-aufklärerischen Vorzüge wegen im Sto-

chastikcurriculum zu verankern (vgl. Diepgen, 1987), leider überhaupt nichts gefruchtet.

Bayes?

Auch Aufgabenteil d) bleibt in einer rein formalistischen Anwendung – nun der „Formel“ von Bayes – stecken; es geht nicht etwa ernsthaft um eine bayessche Argumentation. So bleibt die von den Aufgabenautoren augenscheinlich als selbstverständlich suggerierte A-priori-Verteilung $41/78$ zu $37/78$ unbegründet: Warum sollte der Schüler die für diese A-priori-Verteilung notwendige Zusatzannahme machen, das Spiel mit den 10 Treffern bei 10 Versuchen sei rein zufällig („gleichwahrscheinlich“) aus den 78 Spielen gezogen? Auch die von den Aufgabenautoren augenscheinlich für die Lösung unterstellten Likelihoods $(267/288)^{10}$ und $(231/263)^{10}$ sind nicht zwingend, denn sie repräsentieren lediglich die Wahrscheinlichkeit des Datums: „Dirk Nowitzki erzielt bei dem Spiel bei 10 Freiwürfen 10 Treffer.“ Ignoriert wird hier die weitere Information, dass Dirk Nowitzki in dem Spiel überhaupt 10 Freiwürfe bekommt: Denn die Wahrscheinlichkeit auch dafür dürfte zwischen Heim- und Auswärtsspielen – zumindest ein bisschen – differieren, schließlich bekam Dirk Nowitzki bei einem Heimspiel im Schnitt $288/41 = 7,02$, bei einem Auswärtsspiel aber durchschnittlich $263/37 = 7,11$ Freiwürfe. Das in der bayesschen Argumentation zu berücksichtigende Datum müsste also eigentlich lauten: „Dirk Nowitzki bekommt in dem Spiel 10 Freiwürfe und verwandelt sie dann alle.“ Die von den Aufgabenautoren vermutlich unterstellten Likelihoods $(267/288)^{10}$ und $(231/263)^{10}$ müssten also jeweils noch mit der entsprechenden Wahrscheinlichkeit multipliziert werden, dass Dirk Nowitzki in dem Spiel 10 Freiwürfe bekommt. Die Kalkulation dieser Wahrscheinlichkeit – aufgrund der genannten mittleren Freiwurfanzahlen im Prinzip und unter nicht unbedingt plausiblen Zusatzannahmen etwa zur stochastischen Unabhängigkeit möglich – überforderte aber vermutlich das für das Abitur verlangte stochastische Wissen der Schüler (Stichwort: Poisson-Verteilung).

Natürlich könnte man hier einwenden, eine bayessche Argumentation erfordere es nicht unbedingt, alle in einem Datum enthaltene Information zu berücksichtigen. Es dürfte aber allemal heikel sein, im Abitur von dem Schüler zu erwarten, im Datum offensichtlich enthaltene relevante Information zu ignorieren, nur weil sein mathematisches Wissen zu deren Modellierung nicht ausreicht. In jedem Falle dürfte der Aufgabenteil d) für die hochkompetenten

Schüler, die diese Problematik wenigstens bemerken und sich dann vergeblich – oder vielleicht bei Kenntnis der Poisson-Verteilung auch erfolgreich – an ihr abarbeiten, allemal viel schwieriger sein als für die weniger reflektierten Schüler. Der Hinweis übrigens auf die letztlich sehr geringe und daher nur wenig informationshaltige Differenz der Wahrscheinlichkeiten für 10 Freiwürfe zwischen Heim- und Auswärtsspielen entlastet nicht: Schließlich dürfen die Schüler auch in Aufgabenteil a) – anders als in c) – trotz minimaler Effekte nicht mit dem tabellierten Parameter 0,90 statt mit dem exakten Parameter 0,904 arbeiten.

Fazit

Die „stochastische“ Nowitzki-Aufgabe im NRW-Zentralabitur 2008 hat kaum etwas mit den spezifischen Lehrzielen eines wohlverstandenen Stochastikunterrichtes zu tun: Es geht nicht um Annahmen und Voraussetzungen explizierende Argumentation, nicht um transparente und in ihren Grenzen reflektierte Modellbildung, nicht um vernünftiges Razonieren in einer glaubwürdigen Situation der Ungewissheit. Vielmehr geht es um mehr oder minder unreflektiertes Abspulen von Standardroutinen auf Stichwort hin, um algebraische und algorithmische Operationen, um technische Aspekte, die in der statistischen Praxis ohnehin dem Computer überlassen bleiben. Sollten derartige Abituraufgaben „Schule“ machen – und warum sollten sie dies nicht –, dann dürfte vom schulischen Stochastikunterricht nur noch wenig zu erwarten sein, schon gar nicht ein tragfähiger Widerstand gegen eine degenerierte Praxis hoch entfremdeter und unverstandener Statistikrituale in universitärer Forschung und Lehre.

Die Qualifikationen, auf die ein wohlverstandener Stochastikunterricht eigentlich zielen sollte, dürften kaum einfach zu messen sein, schon gar nicht durch eine hochstandardisierte landeseinheitliche Klausur. Es steht zu befürchten, dass die geradezu naiv test- und messgläubige Evaluationsfixiertheit heutiger Schulpolitik unter der Hand anspruchsvolle Lehrziele trivialisieren wird: Die Nowitzki-Aufgabe dürfte hier ein typisches Beispiel sein.

Sie dürfte auch ein Beispiel sein für die sinnentstellend simplifizierende, mechanisierende und schematisierende prüfungstaugliche „Verschulung“ komplexer Themen, ihre nur oberflächliche Aneignung durch dilettierende Lehrer: Ein professioneller Statistiker dürfte in der Nowitzki-Aufgabe kaum mehr sein professionelles Gegenstandsverständnis repräsentiert sehen, von den geistigen Vätern der Inferenzstatistik – etwa Fisher, Neyman und Pear-

son – ganz zu schweigen. Aus dem Neyman-Motto „life is complicated but not uninteresting“ macht hier die Schule ein „statistics is simple but uninteresting“. Fachliche Expertise professioneller Statistiker scheint bei der Konstruktion der Nowitzki-Aufgabe übrigens gänzlich gefehlt zu haben. Zu teuer? Unerwünscht? Oder überflüssig nach dem Bürokratenmotto „Wem Gott gibt ein Amt, dem gibt er auch Verstand“?

Die missglückte Nowitzki-Aufgabe dürfte schließlich auch Ergebnis des geradezu vermessenen Versuchs sein, gemäß bürokratischer Rahmenvorgaben in nur einer einzigen (!) Aufgabe unterschiedlichste Themen (Bernoulli-Experimente, Binomialverteilung, Pfadregeln, Signifikanztest im Sinne von Fisher, Hypothesentesten im Sinne von Neyman und Pearson, Hypothesenbewertung nach Bayes, ...) unterzubringen: Dieser Versuch – so scheint mir – musste nachgerade in eine hoch artifiziell durchkonstruierte „eingekleidete“ Textaufgabe mit weitgehend sinnentleerten Fragen zu Pseudoproblemen bemerkenswerter Irrelevanz münden, sich trotz oberflächlicher Einheit in inneren Widersprüchen verheddern, in denen dann derselbe Wert willkürlich mal fester Parameter, mal Realisation einer Zufallsvariable sein soll, ein empirischer Befund zugleich „signifikant“ und „nicht signifikant“ sein soll. Schade!

Post scriptum

Ein Gutachter dieses Artikels regte an, man könne auch darauf eingehen, „warum die Aufgabe von vielen Schülern so ‚richtig‘ gelöst wird, wie die Modelllösungen des Ministeriums ‚richtig‘ sind.“ Nun, dazu wenigstens eine skeptische Spekulation: Viele Schüler erleben den Mathematikunterricht vermutlich als Einübung in mehr oder minder sinnfreie oder unverstandene Sprachspiele oder Regelsysteme, und dies reproduzieren sie dann halt auch im Abitur. Dass vermutlich Tausende von Schülern zum Abschluss eines 9-jährigen gymnasialen Mathematikunterrichtes völlig bedenkenlos eine Nullhypothese in Aufgabe b) (1) nicht verworfen und zugleich in Aufgabe b) (2) doch verworfen haben – und für diesen (von ihnen wohl gar nicht bemerkten) Widerspruch dann auch noch mit voller Punktzahl belohnt wurden –, spricht jedenfalls nicht dafür, dass es im schulischen Mathematikunterricht um die zentrale Kategorie von Mathematik und Logik geht, nämlich Widerspruchsfreiheit. Nur Witze leben vom Spiel mit Widersprüchen.

Und was ist mit den vermutlich wenigen Schülern, die den Widerspruch (und die anderen Probleme) in der Aufgabenstellung doch erkannt und sich daran

möglicherweise vergeblich abgearbeitet haben? Vermutlich dürften dies doch wohl eher die (in mathematischer Hinsicht) „Intelligenteren“, „Besseren“, „Fähigeren“ gewesen sein (allerdings vielleicht auch in sozialer Hinsicht die Ungeschickteren und Unangepassteren), die nun „zu Unrecht“ Punkte verloren haben. Insofern man die Abiturprüfung nicht nur als Initiationsritual, als Gratifikation für Wohlverhalten oder als begünstigenden Verwaltungsakt versteht, sondern auch – im Sinne pädagogisch-psychologischer Diagnostik – als Leistungs- oder Fähigkeitsmessung, nicht zuletzt orientiert an Gütekriterien wie inhaltlicher und prognostischer Validität sowie Fairness, ist es ein großes Problem, wenn gerade die besonders leistungsfähigen Kandidaten schlechtere Lösungswahrscheinlichkeiten haben: Negative Trennschärfen (in der Begrifflichkeit der klassischen Testtheorie) oder nicht monotone

Item-Charakteristiken (in der Begrifflichkeit der probabilistischen Testtheorie) machen aus dieser Perspektive Testitems untauglich.

Literatur

Diepgen, R. (1987): Sequentielles Testen – auch didaktisch vielleicht eine gute Alternative. *Stochastik in der Schule* 7 (2), 9–25

Diepgen, R., Kuypers, W., Rüdiger, K. (1989): *Mathematik Sekundarstufe II Stochastik Grundkurs*. Düsseldorf: Cornelsen

Anschrift des Verfassers

Dr. Raphael Diepgen
Ruhr-Universität Bochum
Fakultät für Psychologie
Universitätsstr. 150
44780 Bochum
raphael.diepgen@rub.de

Messung von Sterndurchmessern – Ein klassisches Beispiel zur Photonenstatistik

AXEL DONGES, ISNY IM ALLGÄU

Zusammenfassung: Die experimentelle Bestimmung von Sterndurchmessern ist mit einfachen geometrisch-optischen Mitteln nicht möglich. R. Hanbury-Brown und R. Q. Twiss haben bereits 1955/56 ein Messverfahren zur Bestimmung von Sterndurchmessern entwickelt, das auf einer statistischen Analyse des Sternenlichts beruht. In diesem Beitrag werden – für den Schulunterricht stark vereinfacht – die physikalischen Grundlagen dieses Messverfahrens erläutert.

1 Einleitung

Beobachtet man mit einem Fernrohr einen weit entfernten Stern, so kann dieser mit dem Fernrohr nicht aufgelöst werden. Statt des Sterns sieht ein Beobachter das wesentlich größere Beugungsscheibchen des Sterns. Der Durchmesser des Beugungsscheibchens wird u. a. durch die Apertur des Fernrohrs festgelegt. Er hat nichts mit dem Durchmesser des Sterns zu tun. Die Bestimmung des Sterndurchmessers ist offensichtlich nicht ohne Weiteres möglich.

Dieses interessante messtechnische Problem wurde erstmals von A. A. Michelson gelöst. Er baute auf Vorschlag von A. H. L. Fizeau ein *Sterninterferometer*.

Das Sterninterferometer arbeitet – stark vereinfacht – nach folgendem Prinzip: Mit dem Licht des Sterns wird ein Doppelspalt beleuchtet und das entstehende Beugungsmuster beobachtet (siehe Abbildung 1)¹. Der Abstand $2b$ der beiden Spalte, der zunächst sehr klein gewählt wird, wird immer mehr vergrößert, bis das typische Interferenzmuster des Doppelspalts gerade verschwindet. Der Abstand der beiden Spalte ist dann gerade $2b_{\max}$, die sogenannte transversale Kohärenzlänge des Sternenlichts am Ort der Erde. Es gilt dann die einfache Beziehung (Donges 1990, S. 96)

$$2a \approx \frac{\lambda R}{2b_{\max}} \quad (1)$$

Hierbei bedeuten $2a$: Sterndurchmesser, λ : Wellenlänge des (gefilterten) Sternenlichts und R : Abstand Stern-Erde. Mit seinem Sterninterferometer hat Michelson Sterndurchmesser, genauer Sehwinkel

$$\varepsilon = \frac{2a}{R}, \quad (2)$$

bis herab zu 0,02 Bogensekunden gemessen.

¹ Tatsächlich wird das Sternenlicht mit zwei um einige Meter entfernte Spiegel in ein Fernrohr eingespiegelt und das Interferenzmuster mit dem Fernrohr beobachtet.