

NRW-Schulministerium verschweigt die mathematische Wahrheit bei der NRW-eigenen Erweiterung der Schwimmbadaufgabe des IQB (Mathematik-Abitur 2019)

Wenn man in der Stochastik Berechnungen und Überlegungen zu Zufallsvariablen anstellt, dann muss man mit großer Vorsicht darauf achten, ob die Objekte, die man betrachtet, „nur“ deterministische Größen oder zufallsabhängige Größen sind und ob verschiedene zufallsabhängige Größen stochastisch unabhängig voneinander sind (wie beispielsweise der Geburtsmonat und das Monateinkommen einer zufällig ausgewählten Person) oder stochastisch abhängig voneinander sind (wie beispielsweise die monatlichen Einnahmen und Ausgaben einer Person). Der Laie würde beispielsweise leicht zur intuitiven Meinung tendieren, dass der Erwartungswert des Produktes zweier Zufallsvariablen immer gleich dem Produkt ihrer Erwartungswerte sein müsse, aber dies ist bei stochastischer Abhängigkeit zwischen den beiden Zufallsvariablen keineswegs immer der Fall. Nicht alles, was der Intuition des Laien entspricht, ist in der Mathematik wahr. Und schon gar nicht im Mathematik-Gebiet Stochastik.

Wer Stochastikaufgaben entwickelt, muss selber sich vor Stochastik-Denkfehlern hüten. Der genaue Blick auf einen Aufgabenteil einer im NRW-Mathematikabitur des Jahres 2019 gestellten Aufgabe zeigt, dass die Aufgabensteller fahrlässig und zu grob mit der Stochastik umgegangen waren und deshalb irrtümlich ihre Aufgabenstellung für in der Weise lösbar hielten, wie sie es in ihrer „Modelllösung“ zeigen. Bei genauer Betrachtung wird aber klar, dass ohne eine zusätzliche, nicht im Aufgabentext vorkommende Bedingung die Vorgehensweise der Modell-„Lösung“ nicht korrekt ist. Schlimmer noch: Bei Hinzufügung der Zusatzvoraussetzung ist eine wirklich korrekte Lösung zwar möglich, aber so aufwendig und weitab von jeglicher Schulstochastik, dass man den Aufgabenteil aus diesem Grund – auch wenn er bezüglich der Voraussetzungen vollständig wäre – **leider nur als ungeeignet für Stochastik in der Schule und erst recht als ungeeignet für eine Stochastik-Abituraufgabe** bezeichnen muss.

Man fragt sich zwar, wieso denn niemand bei den Entwicklern und denjenigen, die unter fachlichen und didaktischen Aspekten die Abituraufgabenentwicklung kontrollieren sollen, das Problem erkannt und diese Teilaufgabe verworfen hat, aber hier soll es **nur** darum gehen: **Wie geht das Schulministerium damit um, wenn es nach dem Klausurtermin auf Fehler bzw. Probleme aufmerksam gemacht wird?**

Hier einige Fakten dazu:

Ich habe bereits am **14. Mai 2019** (also noch während der Korrekturphase der Klausuren) mit meiner Universitäts-Emailadresse zu dem Thema eine Email an Qua-LIS NRW (Qualitäts- und Unterstützungsagentur – Landesinstitut für Schule), nämlich an die sicher sinnvolle Adresse abitur.nrw@qua-lis.nrw.de geschrieben, und zwar mit der Betreffzeile „Kritik an zwei Teilen der Schwimmbad-Aufgabe im NRW-Mathematikabitur 2019“. Wie ich telefonisch erfahren hatte, war diese Email auch an den zuständigen Mitarbeiter des Schulministeriums weitergeleitet worden. Danach habe ich am **21. August 2019** an einen für das Zentralabitur zuständigen Mitarbeiter des Schulministeriums NRW und einen für das Zentralabitur in Mathematik zuständigen Mitarbeiter von Qua-LIS NRW eine Email mit der Betreffzeile „Genaue Betrachtung des kritischen Teils der Schwimmbadaufgabe“

geschrieben, wobei ich im Anhang in handgeschriebener Form die mathematischen Einzelheiten dargestellt hatte.

Obwohl dem Schulministerium also ausführliche Informationen zum Problem frühzeitig vorlagen und meiner Kenntnis nach **auch** von anderer Seite aus dem Wissenschaftsbereich (NRW-Universitäten) Kritik zur Aufgabe an das Schulministerium gelangt war, hat das Schulministerium im September 2019 auf die entsprechenden Internetseiten des Ministeriums, die Lehrern und Schülern per Passwort zugänglich sind, **die Aufgaben und die Modelllösungen leider ohne jeglichen erklärenden oder richtigstellenden Kommentar gebracht.**

Genauere Darstellung des Problems:

Es gibt im Aufgabenpool 2019 des IQB (Institut zur Qualitätsentwicklung im Bildungswesen) eine Stochastikaufgabe, die ich hier als „IQB-Schwimmbadaufgabe“ bezeichne. Eine Modifikation dieser Aufgabe sowohl für den Leistungskurs als auch für den Grundkurs kam in diesem Jahr 2019 bei den Stochastik-Abituraufgaben des Landes NRW vor und wird von mir kurz als „NRW-Schwimmbadaufgabe“ bezeichnet. Aus urheberrechtlichen Gründen kann ich hier den Aufgabentext nicht bringen, aber wer nicht durch Schulkontakte eine Möglichkeit zur Einsicht hat, kann in einem Abiturvorbereitungsbuch aus dem Stark-Verlag oder Klett-Verlag die Originalaufgabe sehen, allerdings nicht die Original-Modelllösung, sondern nur eine von den Verlagen unwesentlich erweiterte oder geänderte Lösung.

Es geht im Folgenden um den **Teil b der Leistungskursversion** der „Schwimmbadaufgabe“ im NRW-Mathematikabitur 2019. Da dieser Teil der LK-Version identisch mit dem Teil c der Grundkursversion dieser Aufgabe ist, gelten alle Feststellungen **auch für Teil c der Grundkursversion.**

Zu Teil b (2)

Es soll die Wahrscheinlichkeit dafür ermittelt werden, „dass der Besitzer des Kiosks an dem betrachteten Tag erwartete Einnahmen von den Jahreskartenbesitzern hat, die mindestens 1000 Euro betragen.“

Bei dieser Aufforderung würden natürlich die Lehrer ins Grübeln geraten: Es geht also um die Wahrscheinlichkeit dafür, dass „erwartete Einnahmen“ (Erwartungswert der Einnahmen also?) 1000 Euro erreichen oder übertreffen. Aber bevor die Lehrer allzu lange grübeln, schauen sie auf die Modelllösung: Die erwarteten Einnahmen werden dort in gewisser Weise mit der Anzahl X der Besucher mit Jahreskarte (im Folgenden als JK-Besucher abgekürzt) verknüpft, und die Verteilung von X ist ja gegeben. Aber die Art der Verknüpfung in der Modelllösung ist, wie noch gezeigt wird, ohne eine geeignete Zusatzvoraussetzung gar nicht mathematisch korrekt, doch eine solche Zusatzvoraussetzung fehlt leider im Aufgabentext. Das dürfte wohl der in Stochastik durchschnittlich vorgebildete Lehrer i.a. allerdings gar nicht merken, denn ihm erscheint die Modelllösung in der Tat wohl als recht plausibel.

Wenn man die Aufgabenstellung, die gemeint und letztlich nur aus der Modelllösung erschließbar ist, in der Terminologie der Stochastik formuliert, dann lautet sie: Es soll die Wahrscheinlichkeit

$P(Z \text{ größer gleich } 1000)$

ermittelt werden, wobei Z der bedingte Erwartungswert $E(A|X)$ ist und A die Gesamtkioskausgaben der das Schwimmbad an dem Tag besuchenden Jahreskartenbesitzer sind. (Übrigens liefert der Wikipediaartikel zu „Bedingter Erwartungswert“ eine durchaus sehr brauchbare Information.) Natürlich gehören bedingte Erwartungswerte in keiner Weise zum Stoff der Schulstochastik, und dem größten Teil der Lehrer dürfte der Begriff „bedingter Erwartungswert“ auch in der vielleicht schon lange zurückliegenden Stochastik-Einführungsvorlesung noch nie begegnet sein.

Kritikpunkt 1 zum Teil b (2) :

Da mancher verständige Schüler mit einem Erwartungswert natürlich die Vorstellung von einer festen, zu einer Zufallsvariablen gehörenden Zahl verbindet und sicher noch nie etwas von einem bedingten Erwartungswert gehört hat, wird er vielleicht einfach mal den Erwartungswert der Gesamtausgaben A der Jahreskartenbesitzer ausrechnen. Er kann sich überlegen, dass nach dem Satz von der totalen Wahrscheinlichkeit jeder von den 2000 Jahreskartenbesitzern mit Wahrscheinlichkeit $0,9$ mal 0 + $0,1$ mal $0,5$ Kioskausgaben von 4 Euro hat, mit Wahrscheinlichkeit $0,9$ mal 0 + $0,1$ mal $0,3$ Ausgaben von 12 Euro und mit Wahrscheinlichkeit $0,9$ mal 1 + $0,1$ mal $0,2$ keine Ausgaben hat. Das führt dazu, dass der (nicht bedingte) Erwartungswert, also der Erwartungswert schlechthin für die Kioskausgaben jedes Jahreskartenbesitzers gleich

$$(0,1 \text{ mal } 0,5) \text{ mal } 4 + (0,1 \text{ mal } 0,3) \text{ mal } 12 = 0,1 \text{ mal } 5,6 = 0,56 \text{ Euro}$$

ist und folglich

$$E(A) = 2000 \text{ mal } 0,56 = 1120 \text{ Euro}$$

gilt. Nun kann der Schüler (der hoffentlich bei seiner Interpretation der Aufgabenstellung nun nicht allzu verwirrt ist) hinschreiben:

$$P(1120 \text{ größer gleich } 1000) = 1$$

In Worten: Mit Wahrscheinlichkeit 1 kann der Kioskbesitzer Einnahmen von den Jahreskarteninhabern erwarten, die mindestens 1000 Euro betragen.

Da den Aufgabenstellern klar sein muss, dass **mangels Vertrautheit mit bedingten Erwartungswerten** diese „Lösung“ für etliche kluge Schüler, sofern sie überhaupt etwas Vernünftiges zur Lösung schaffen, **äußerst naheliegend** ist, müsste unbedingt in der Modelllösung hierauf hingewiesen werden und zugleich festgelegt werden, wie diese Art von „Lösung“ zu bewerten ist. Und zwar ist es natürlich erwägenswert, für eine solche „Lösung“ die volle Punktzahl zu vergeben. Jeglicher Hinweis auf diese Möglichkeit der „Lösung“ fehlt allerdings.

Kritikpunkt 2 zum Teil b (2):

Dieser Kritikpunkt ist der **weitaus schwerer wiegende**.

Die Modelllösung erscheint zwar auf den ersten Blick und vor allem auch für in Stochastik weniger Kundige als sehr plausibel. (Das heißt übrigens nicht, dass man beim Durchlesen des Aufgabentextes unbedingt so leicht auf den Gedankengang der Modelllösung kommt). **Aber** das Vorgehen bei der Modelllösung ist **gar nicht korrekt** und liefert gar keine immer richtige Lösung, es funktioniert nämlich nur unter einer geeigneten **Zusatzvoraussetzung**, die keineswegs aus dem Aufgabentext hervorgeht.

Entscheidend am Vorgehen der Modell-“Lösung“ ist die **vermeintliche** Beziehung:

Erwartungswert von A im Falle von n Besuchern mit Jahreskarte = 5,6 mal n

Eine solche Proportionalität lässt sich aus den Voraussetzungen im Aufgabentext ohne Zusatzvoraussetzungen gar nicht herleiten, und das kann man sich auch inhaltlich klar machen: Die Wahrscheinlichkeitsverteilung der Ausgaben eines JK-Besuchers (4 bzw. 12 bzw. 0 Euro mit Wahrscheinlichkeit 0,5 bzw. 0,3 bzw. 0,2) gibt Wahrscheinlichkeiten schlechthin an, und die können ja nach dem Satz von der totalen Wahrscheinlichkeit durch Gewichtung von unterschiedlichen bedingten Wahrscheinlichkeiten bei verschiedenen JK-Besucher-Anzahlen entstehen. Es kann ja einerseits die Tendenz bestehen, dass ein JK-Besucher bei hoher Zahl von anderen JK-Besuchern leicht auf Bekannte trifft, mit denen er dann gemeinsam beispielsweise etwas trinkt. Es kann andererseits auch die Tendenz bestehen, dass bei einer sehr hohen Zahl von JK-Besuchern entsprechende Schlangen an der Kiosk-Kasse manche JK-Besucher von einem Kauf am Kiosk abhalten. Und damit ist klar, dass die obige strikte Proportionalität zwischen dem (bedingten) Erwartungswert von A und der Anzahl n der JK-Besucher auch von der Sache her gar nicht anzunehmen ist – wenn man nicht Zusatzvoraussetzungen macht.

Diese inhaltliche Überlegung liefert schon die Idee für ein Gegenbeispiel: Man nimmt für ein Gegenbeispiel an, dass jeder JK-Besucher 4 Euro (oder auch 12 Euro) dann mit einer etwas kleineren Wahrscheinlichkeit als 0,5 (bzw. 0,3) ausgibt, wenn die Zahl der JK-Besucher außer ihm selber kleinergleich einer geeigneten Schranke c ist, und umgekehrt 4 Euro (oder auch 12 Euro) dann mit einer etwas größeren Wahrscheinlichkeit ausgibt, wenn die Anzahl der JK-Besucher außer ihm selber größer als diese Schranke c ist. Das „etwas kleiner“ und „etwas größer“ muss dabei natürlich so austariert werden, dass sich mit dem Satz von der totalen Wahrscheinlichkeit die nicht-bedingten Wahrscheinlichkeiten 0,5 bzw. 0,3 für die Ausgabe von 4 bzw. 12 Euro ergeben. Die **exakten mathematischen Einzelheiten** finden sich in der pdf-Datei **schwimmbadgegenbeispiel.pdf**

Was wäre nun eine vernünftige Zusatzvoraussetzung dafür, dass die in der Modelllösung unzulässigerweise verwendete Proportionalität

$$E(A | X = n) = 5,6 \text{ mal } n$$

gilt? Eine sinnvolle Formulierung einer solchen Bedingung kann lauten:

Die Kioskausgaben jedes Jahreskartenbesitzers sind stochastisch unabhängig von der Anzahl der im Schwimmbad erscheinenden Jahreskartenbesitzer, wobei er selber ggf. bei dieser Anzahl nicht mitzuzählen ist. Die exakte mathematische Begründung findet sich in der pdf-Datei **schwimmbadhinreichend.pdf**

Zu Teil b (1)

Die folgende Kritik bezieht sich auf die Modelllösung zum Teil b (1) und ist weit weniger wichtig als die Kritik zum Teil b (2).

Für den Erwartungswert wird in der Modelllösung zu b (1)

$$0,5 \text{ mal } 660 \text{ mal } 4 + 0,3 \text{ mal } 660 \text{ mal } 12$$

angegeben. Da den Schülern i.a. bewusst ist, dass ein Erwartungswert die Summe über alle Werte ist, die die Zufallsvariable annimmt, jeweils multipliziert mit der entsprechenden Wahrscheinlichkeit, wird ein sehr oberflächlicher Schüler u.U. zu der unsinnigen Vorstellung verleitet, dass die Gesamteinnahmen **nur** die drei Werte $660 \text{ mal } 4$, $660 \text{ mal } 12$ und $660 \text{ mal } 0$ annehmen können.

Viel sinnvoller und selbstverständlich algebraisch äquivalent wäre der Term

$$660 \text{ mal } (0,5 \text{ mal } 4 + 0,3 \text{ mal } 12)$$

oder auch der noch größeren Klarheit halber

$$660 \text{ mal } (0,5 \text{ mal } 4 + 0,3 \text{ mal } 12 + 0,2 \text{ mal } 0)$$

Denn die Gesamteinnahmen sind die Summe aus 660 Zufallsvariablen, die alle denselben Erwartungswert $0,5 \text{ mal } 4 + 0,3 \text{ mal } 12 + 0,2 \text{ mal } 0$ haben, und daraus ergibt sich der Erwartungswert durch Linearität bzw. Additivität der Erwartungswertbildung. $E(A)$ ist also die Summe aus 660 gleichen Zahlenwerten, wobei jeder Zahlenwert der Erwartungswert der Ausgaben eines Schwimmbadbesuchers ist.

Zu einem Teil, der nur in der Leistungskursversion vorkommt: Der Aufgabenteil c (4) in der Leistungskursversion der NRW-Schwimmbadaufgabe ist keine NRW-Erfindung, sondern entspricht dem Aufgabenteil 2 d der IQB-Aufgabe. Dennoch wird hier auch dieser Teil kommentiert.

Zu Teil c (4) der Leistungskursversion:

Hier wie in der IQB-Version soll eine fiktive Argumentation kritisiert werden. Ganz offensichtlich ist, dass folgende Formulierung in dieser Argumentation aber gar nicht kritisiert werden soll (und in der Modelllösung auch gar nicht kritisiert wird): „...mithilfe einer normalverteilten Zufallsgröße...., die für alle reellen Zahlen definiert ist“. Eine solche Formulierung ist nicht korrekt, sie verdreht Definitionsbereich und Wertebereich. Eine normalverteilte Zufallsgröße hat grundsätzlich als Wertebereich die Menge aller reellen Zahlen. Sie kann auch als Definitionsbereich diese Menge haben, aber das hängt ganz vom gewählten Modell ab.

Außerdem hätte selbstverständlich in die Aufgabe noch die Voraussetzung gehört, dass nie mehrere Badegäste exakt gleichzeitig eintreffen. Oder man hätte in die Aufgabenstellung die Frage danach bringen sollen, welche Voraussetzung man braucht, damit die Ermittlung des betreffenden Zeitpunkts überhaupt sinnvoll ist. Ein gründlich denkender Schüler verliert Zeit mit dem Grübeln darüber, ein weniger gründlich denkender

Schüler, der dann ggf. die gleiche Punktzahl erhält, verliert keine Zeit durch Nachdenken über das Problem.

Schlussbemerkung:

Vielleicht liegt der Entwicklung der Aufgabe das Bestreben zu Grunde, eine auch mathematisch interessante Aufgabe zu stellen, die teilweise etwas höhere Anforderungen stellt. Wenn hierbei nun eine partiell sehr problematische Aufgabe entsteht, dann mag dies vielleicht in gewissem Maße entschuldbar sein.

Wenn dann aber vom NRW-Schulministerium die nachträglich erkannten Probleme der Aufgabe bei der Bereitstellung der Originalaufgaben und Modellösungen **völlig unkommentiert bleiben und somit verschwiegen werden**, dann kann das nicht als akzeptabel bezeichnet werden.

Stand 7.10.2019 / 19.1.2020