

Warum ein Teil der Schwimmbad-Aufgabe im NRW-Abitur 2019 so nicht hätte gestellt werden dürfen

INGOLF TERVEER, MÜNSTER UND FRANZ-REINHOLD DIEPENBROCK, WUPPERTAL

Zusammenfassung: Was bei der Erstellung einer Aufgabe als Lösung derselben vorschwebt, ist nicht immer korrekt – auch wenn u.U. dank starker scheinbarer Plausibilität der Lösung fast niemand den Fehler darin bemerkt. Selbst umfangreiche organisatorische Qualitätsbestrebungen laufen hier manchmal ins Leere. Ein von DIEPENBROCK (2021) genanntes Beispiel hierfür liefern Teil c)(2) der Grundkurs- und in identischer Form Teil b)(2) der Leistungskursversion der „Schwimmbadaufgabe“ aus dem NRW-Mathematikabitur 2019, der vielleicht schon manche Leserin und manchen Leser dieser Zeitschrift während des damaligen Abiturs irritiert haben könnte. Hier wird nun ausführlich gezeigt: Was offensichtlich als Musterlösung gedacht ist, das würde Zusatzannahmen erfordern, die nicht aus dem Aufgabentext hervorgehen, und übrigens nicht unbedingt realistisch sind. Außerdem aber – und das macht den Aufgabenteil sowieso unbrauchbar für eine Abiturprüfung – würde eine korrekte Lösung unter solchen Zusatzannahmen einen souveränen, weit über die Schulmathematik hinausgehenden Umgang mit bedingten Erwartungswerten (nicht nur bedingten Wahrscheinlichkeiten) erfordern. Um so wichtiger wäre es, wenn die betreffenden Aufgabenteile aus dem derzeit zur Abiturvorbereitung bereitgestellten Pool alter NRW-Abituraufgaben entfernt würden, damit Lernende und Lehrende nicht eine fachlich falsche Verwendung des Stochastik-Kalküls verinnerlichen.

1 Einleitung

Die Schwimmbad-Aufgabe im NRW-Abitur 2019, die wir in Kapitel 2 vorstellen, weist in Teilaufgabe c)(2) der Grund- bzw. Teilaufgabe b)(2) der Leistungskursversion erhebliche Mängel auf. Gegenstand der Aufgabe sind die erwarteten Einnahmen des Betreibers eines Schwimmbadkiosks von den anwesenden Jahreskartenbesitzern. Gesucht ist die Wahrscheinlichkeit, dass diese mindestens 1000 € betragen. Die naheliegende Interpretation von „erwarteten Einnahmen“ als Erwartungswert der Einnahmen

führt dazu, dass diese Wahrscheinlichkeit entweder null oder eins sein muss, wie in Kapitel 4 erläutert wird.

In der genannten Lösungsskizze hingegen werden diese erwarteten Einnahmen in der Form $5,6 \cdot X$ mit einer konkreten Zufallsgröße X , nämlich der Anzahl der anwesenden Jahreskartenbesitzern angesetzt. Aus Sicht der Wahrscheinlichkeitsrechnung kann der gesuchte Ausdruck daher allenfalls ein bedingter Erwartungswert sein. Abgesehen davon, dass bedingte Erwartungswerte über Schulwissen deutlich hinausgehen, ist das genannte Zwischenergebnis ausschließlich unter speziellen, in Aufgabenstellung und Lösung nicht genannten Annahmen des zugrundeliegenden stochastischen Modells korrekt, vgl. Kapitel 5. Diese Annahmen sind aber keineswegs selbstverständlich, und wir illustrieren die Fehlerhaftigkeit der implizit getroffenen Annahme an die erwarteten Einnahmen in Kapitel 6 anhand eines „Tandemmodells“.

2 Aufgabenstellung und Musterlösung der Themensteller

Die hier besprochene Abituraufgabe findet sich als Teil c) einer GK-Aufgabe – bzw. Teil b) einer LK-Aufgabe – im NRW-Abitur 2019 mit folgendem Kontext¹:

Für ein Schwimmbad besitzen 2000 Personen eine Jahreskarte. Für einen bestimmten Tag beschreibt die Zufallsgröße X die Anzahl der Jahreskartenbesitzer, die das Schwimmbad besuchen. Vereinfachend soll davon ausgegangen werden, dass X binomialverteilt ist. Dabei beträgt die Wahrscheinlichkeit dafür, dass ein zufällig ausgewählter Jahreskartenbesitzer an diesem Tag das Schwimmbad besucht, 10%.

¹Die Aufgabe ist für Schülerinnen und Schüler, aber auch Lehrerinnen und Lehrer unter <https://www.standardsicherung.nrw.de> verfügbar, findet sich auch in den Abiturvorbereitungsmaterialien verschiedener Verlage, z.B. bei Klett (2020), S. 47f und Stark (2020), S. 2019-35. Die ministerielle Aufgabenstellung wurde von den Verlagen unverändert übernommen, in den Lösungsskizzen wurde die ministerielle Vorlage leicht überarbeitet, wobei die grundsätzliche Problematik erhalten blieb.

Das Verhalten der Jahreskartenbesitzer – im Folgenden stets mit JKB abgekürzt – am betrachteten Tag lässt sich mit Zufallsgrößen X_1, \dots, X_k beschreiben:

$$X_i = \begin{cases} 1 & \text{i-ter JKB besucht Schwimmbad} \\ 0 & \text{i-ter JKB besucht Schw.Bad nicht} \end{cases} \quad (1)$$

Die in der Aufgabenstellung genannte Zufallsgröße X ist demnach die Summe

$$X = \sum_{i=1}^k X_i. \quad (2)$$

Laut Aufgabenstellung ist X binomialverteilt mit Parametern $k = 2000$ und $p = 0,1$. Üblicherweise – und im Erfahrungsbereich der Schülerinnen und Schüler verankert – ist X das Ergebnis eines k -stufigen Bernoulli-Experiments, d.h.

$$X_1, \dots, X_k \text{ sind stochastisch unabhängig} \quad (3)$$

und besitzen alle eine Bernoulli-Verteilung mit Parameter p , d.h. für $x \in \{0,1\}$ gilt

$$P(X_i = x) = \begin{cases} p & \text{falls } x = 1 \\ 1 - p & \text{falls } x = 0 \end{cases}. \quad (4)$$

In der zu diskutierenden Teilaufgabe c) der GK-bzw. b) der LK-Version wird die Situation eines Kioskbesitzers spezifiziert²:

Auf dem Gelände des Schwimmbades wird ein Kiosk betrieben. Der Besitzer nimmt vereinfachend an, dass jeder Gast 4 €, 12 € oder gar kein Geld an seinem Kiosk ausgibt. Die Wahrscheinlichkeit, dass ein Gast 4 € ausgibt, betrage 50%, die Wahrscheinlichkeit, dass ein Gast 12 € ausgibt, betrage 30%.

- (1) *Am betrachteten Tag besuchen 660 Personen das Bad. Bestimmen Sie die Höhe der Einnahmen, mit denen der Besitzer des Kiosks rechnen kann.*
- (2) *Ermitteln Sie die Wahrscheinlichkeit dafür, dass der Besitzer des Kiosks an dem betrachteten Tag erwartete Einnahmen von den Jahreskartenbesitzern hat, die mindestens 1000 € betragen.*

Ohne auf die Operatoren-Nomenklatur³ genauer einzugehen, so bedeuten die verwendeten Operatoren

²Vgl. Klett (2020), S. 48, Stark (2020), S. 2019-35

³vgl. die bezüglich der Anforderungsbereiche dieser beiden Operatoren eindeutig formulierte Liste unter Standardsicherung (2020)

⁴vgl. Klett (2020), S.69ff bzw. Stark (2020), S. 2019-40, beide Quellen halten sich eng an die ministerielle Vorlage.

„Bestimme“ und „Ermittle“ aber insbesondere eine Aufgabenstellung im Anforderungsbereich 2, was gleichzeitig bedeutet, dass die Themensteller einen – für Schülerinnen und Schüler – angemessen einfachen Lösungsweg vor Augen haben.

Die einleitenden Sätze der Teilaufgabe stellen die Ausgaben der Gäste in den Vordergrund. Teil c)(1) bzw. b)(1) (GK bzw. LK) nimmt noch nicht die JKB in den Blick, ist aber von den Themenstellern offensichtlich auch als Vorbereitung von c)(2) bzw. b)(2) gedacht. Beide Teilaufgaben stellen die Ausgaben eines Schwimmbadbesuchers als Einnahmen des Kioskbesitzers dar. In c)(1) bzw. b)(1) seien die Zufallsgrößen Z_1, \dots, Z_{660} die Einnahmen des Kioskbesitzers, die er von den einzelnen Besuchern erzielt, mit

$$P(Z_i = y) = p_y := \begin{cases} 0,5 & \text{falls } y = 4 \\ 0,3 & \text{falls } y = 12 \\ 0,2 & \text{falls } y = 0 \end{cases} \quad (5)$$

und dem Erwartungswert

$$E(Z_i) = 0 \cdot 0,2 + 4 \cdot 0,5 + 12 \cdot 0,3 = 5,6. \quad (6)$$

Die Einnahmen, mit denen der Kioskbesitzer rechnen kann, sind natürlich als erwartete Einnahmen zu verstehen, also als $E(Z)$, wobei $Z = Z_1 + \dots + Z_{660}$. Aufgrund der Additivität des Erwartungswertes gilt

$$\begin{aligned} E(Z) &= E(Z_1) + \dots + E(Z_{660}) \\ &= 5,6 \cdot 660 = 3696. \end{aligned} \quad (7)$$

Der Kioskbesitzer kann mit Einnahmen in Höhe von 3696 € rechnen. Das Rechnen mit Erwartungswerten, etwa die Additivität gemäß (7) spielt in den zentralen Vorgaben des Landes NRW derzeit keine prominente Rolle. Ob die Aufgabe erfolgreich gelöst werden kann, hängt damit auch davon ab, wieviel die Lehrenden in der Qualifikationsphase über die Vorgaben der Kernlehrpläne hinausgegangen sind.

In Aufgabe c)(2) bzw. b)(2) (GK bzw. LK) geht es ausschließlich um die JKB. Der vom Aufgabensteller im Erwartungshorizont dargestellte Lösungsweg⁴, stellt nun in Analogie zu c)(1) bzw. b)(1) die erwarteten Einnahmen von n JKB als $5,6 \cdot n$ dar. Das für die Aufgabe erforderliche Ereignis lässt sich demnach als $\{5,6 \cdot X \geq 1000\}$ schreiben, die gesuchte Wahrscheinlichkeit beträgt also $P(X \geq 1000/5,6) = P(X \geq 179) \approx 0,947$.

Die Details der letzten Rechenschritte seien hier ausgelassen, denn nach unserer Auffassung liegt bei diesem kritischen Aufgabenteil ausgehend von dem Ausdruck „erwartete Einnahmen von den JKB“ schon im Ansatz ein schwerwiegender Irrtum vor, indem nämlich die Themensteller fälschlich annehmen, dass diese Einnahmen proportional zur Anzahl der anwesenden JKB sind. Die Ausgaben von einzelnen JKB, deren Kumulation letztlich zu den gesamten Einnahmen führt, kann stochastisch nämlich von etlichen Einflussgrößen abhängen: beispielsweise von der Anwesenheit oder Nichtanwesenheit jedes einzelnen JKB oder der Anzahl der anwesenden JKB.

Die Formulierung in der Aufgabenstellung c)(2) lässt – wenn man Konzepte der Stochastik verwendet – nur zwei mögliche Interpretationen zu:

1. Mit den „erwarteten Einnahmen von den JKB“ könnte ein unbedingter Erwartungswert, nämlich $E(Y)$ gemeint sein, wobei Y die gesamten Einnahmen des Kioskbesitzers von den JKB bezeichnet. In diesem Falle müsste die Wahrscheinlichkeit $P(E(Y) \geq 1000)$ berechnet werden. Abhängig von der deterministischen Größe $E(Y)$ ist das Ereignis für
 - $E(Y) \geq 1000$ das sichere Ereignis; dann ist die Wahrscheinlichkeit eins,
 - $E(Y) < 1000$ das unmögliche Ereignis; dann ist die Wahrscheinlichkeit null.

Wir werden den Erwartungswert in Abschnitt 4 ausrechnen und sehen, dass der erste Fall richtig ist. Aus der Musterlösung⁵ wird aber klar, dass den Themenstellern diese Interpretationsvariante nicht vorgeschwebt hat.

2. Schreibt man – wie in der Musterlösung – das interessierende Ereignis als $\{5,6X \geq 1000\}$, so ist damit angenommen, dass „die erwarteten Einnahmen von den JKB“ zufällig sind und es deshalb nur um den bedingten Erwartungswert $E(Y|X)$ handeln kann und analog c)(1) gilt

$$E(Y|X = n) = 5,6n \text{ bzw. } E(Y|X) = 5,6X . \quad (9)$$

Um diese Proportionalität des bedingten Erwartungswertes in n bzw. X verwenden zu können, sind allerdings geeignete Zusatzannahmen erforderlich, welche in der Aufgabe nicht genannt werden. Wir werden solche Zusatzannahmen formulieren, halten aber jetzt

schon fest, dass diese eher unplausibel sind. Ohne Gültigkeit der Gleichung (9) ist aber auch das von den Themenstellern gefundene Endergebnis, d.h. die Wahrscheinlichkeit 0,947 nicht belastbar.

3 Ein stochastisches Modell für die Einnahmen

Wir stellen nun ein Modell für die Ausgaben der JKB (Einnahmen des Kioskbesitzers) auf. Nach den einleitenden Sätzen der Aufgabenstellung hat insbesondere der i -te JKB, wenn er im Schwimmbad ist, d.h. unter der Bedingung $X_i = 1$ zufällige Ausgaben

$$Y_i \in \mathcal{Y} = \{0,4,12\} \quad (10)$$

mit den folgenden bedingten Wahrscheinlichkeiten $P(Y_i = y|X_i = 1) := p_y$ mit

$$p_y := \begin{cases} 0,5 & \text{falls } y = 4 \\ 0,3 & \text{falls } y = 12 \\ 0,2 & \text{falls } y = 0 \end{cases} . \quad (11)$$

Es darf angenommen werden, dass nicht anwesende JKB keine Ausgaben am Kiosk haben, d.h.

$$P(Y_i = y|X_i = 0) = \begin{cases} 0 & \text{falls } y = 4 \\ 0 & \text{falls } y = 12 \\ 1 & \text{falls } y = 0 \end{cases} . \quad (12)$$

Weitere Informationen über den Zusammenhang zwischen den Zufallsgrößen $X_1, \dots, X_k, Y_1, \dots, Y_k$ lassen sich dem Aufgabentext nicht entnehmen.

Mit diesen Notationen hat der Kioskbetreiber die Einnahmen $Y = Y_1 + \dots + Y_k$ von den JKB, wobei hier $k = 2000$ ist.

4 Lösung der Aufgabe mit dem Erwartungswert

Betrachten wir den ersten genannten Lösungsansatz, d.h. die Berechnung der Wahrscheinlichkeit

$$P(E(Y) \geq 1000) . \quad (13)$$

Hier gibt es keine Zufallsgröße mehr, anhand deren Werteverhalten eine Wahrscheinlichkeit berechnet werden kann. Die in (13) angegebene Wahrscheinlichkeit ist entweder 1 oder 0, je nachdem, ob der Erwartungswert ≥ 1000 oder < 1000 ist.

0/1-Wahrscheinlichkeiten findet man vielfach in der Stochstik, beispielsweise bei den Gesetzen großer

⁵Klett (2020), S. 70f bzw. Stark (2020), S. 2019-40

Zahlen. Das betrifft aber die Stochastik an der Hochschule; im gymnasialen Bereich sind solche Wahrscheinlichkeiten unüblich und auch nicht Gegenstand von Unterricht oder unterrichtlichen Vorgaben. In der Konsequenz werden Schülerinnen und Schüler die in (13) angegebene Wahrscheinlichkeit nur berechnen können, wenn sie diese Perspektive der Hochschulstochastik einnehmen können. Gegebenenfalls werden sie noch, ohne aus dem schulischen Kontext dafür eine mathematisch korrekte Begründung angeben zu können, den Erwartungswert $E(Y)$ bestimmen wollen, was zumindest einer Teilleistung für die korrekte Antwort entspricht. Jedoch ist auch die Bestimmung des Erwartungswertes für sich schon nicht ganz anspruchslos, wir stellen hier einen möglichen Rechenweg unter den Annahmen (4), (11) und (12) vor: Wegen der Linearität des Erwartungswertes ist

$$E(Y) = E(Y_1 + \dots + Y_k) = \sum_{i=1}^k E(Y_i). \quad (14)$$

Für jedes i gilt dabei

$$E(Y_i) = 4 \cdot P(Y_i = 4) + 12 \cdot P(Y_i = 12). \quad (15)$$

Nun gilt für $y = 4$ und $y = 12$ natürlich

$$P(Y_i = y) = P(Y_i = y, X_i = 1) \quad (16)$$

und somit nach dem Multiplikationssatz

$$\begin{aligned} P(Y_i = y) &= P(Y_i = y | X_i = 1) \cdot P(X_i = 1) \\ &= P(Y_i = y | X_i = 1) \cdot 0,1, \end{aligned}$$

also

$$E(Y_i) = 4 \cdot 0,5 \cdot 0,1 + 12 \cdot 0,3 \cdot 0,1 = 0,56. \quad (17)$$

Daraus ergibt sich dann $E(Y) = 0,56 \cdot k$. Im Kontext der Aufgabenstellung mit $k = 2000$ folgt $E(Y) = 0,56 \cdot 2000 = 1120$. Da dieser Wert größer als 1000 ist, handelt es sich bei dem durch die Aufgabenstellung angegebenen Ereignis um das sichere, die gesuchte Wahrscheinlichkeit ist eins.

Kenntnisse im Umgang mit bedingten Wahrscheinlichkeiten vorausgesetzt, könnten leistungsfähige Prüflinge diese Lösung durchaus finden - wenn auch vermutlich mit unverhältnismäßig großem Aufwand. Jedoch zeigt der seitens der Aufgabensteller veröffentlichte Erwartungshorizont, dass eine ganz andere Lösung intendiert war.

⁶Die Umformungen sind bei bedingten Erwartungswerten stets nur „fast sicher“, d.h. mit Wahrscheinlichkeit 1. Bei den hier vorliegenden diskreten Wahrscheinlichkeiten dürfen wir diesen Zusatz aber im Folgenden ignorieren

⁷Ereignisse, die gemeinsam eintreten wie hier $Y_i = y$ und $X = n$ werden im Folgenden nicht mit dem Schnittmengensymbol verknüpft, sondern durch Komma getrennt.

5 Lösung der Aufgabe mit dem Konzept des bedingten Erwartungswertes

Im zweiten Lösungsansatz ist die Wahrscheinlichkeit $P(E(Y|X) \geq 1000)$ zu berechnen. Auch wenn von den Themenstellern an keiner Stelle der Bezug zum Rechnen mit bedingten Erwartungswerten hergestellt wurde, so muss man doch annehmen, dass sie – wenigstens stillschweigend – die Aussage

$$E(Y|X = n) = 5,6 \cdot n \quad (18)$$

für die bedingten erwarteten Einnahmen Y des Kioskbesitzers für $X = n$ JKB im Schwimmbad in einem Modell mit insgesamt $k = 2000$ JKB vorausgesetzt haben. Um zu prüfen, unter welchen Voraussetzungen diese Aussage als richtig gelten kann, formen wir den bedingten Erwartungswert um. Auch bedingte Erwartungswerte sind additiv, d.h. es gilt⁶

$$E\left(\sum_{i=1}^k Y_i | X = n\right) = \sum_{i=1}^k E(Y_i | X = n). \quad (19)$$

Hier im diskreten Fall liegen elementare bedingte Wahrscheinlichkeiten vor, mit denen sich jeder Summand in der Form⁷

$$E(Y_i | X = n) = \sum_{y \in \{4,12\}} y \cdot \frac{P(Y_i = y, X = n)}{P(X = n)} \quad (20)$$

schreiben lässt. Für $y > 0$ ist die Wahrscheinlichkeit $P(Y_i = y, X_i = 0)$ null, dann lässt sich der Zähler des in (20) auftretenden Bruches wie folgt schreiben:

$$\begin{aligned} P(Y_i = y, X = n) &= P(Y_i = y, X_i = 0, X = n) + P(Y_i = y, X_i = 1, X = n) \\ &= P(Y_i = y, X_i = 1, X = n). \end{aligned}$$

Ist $n = 0$, so ist diese Wahrscheinlichkeit null, denn dann ist kein JKB im Schwimmbad, während gleichzeitig der i -te JKB im Schwimmbad sein soll. Das Ereignis ist also unmöglich. Es folgt $E(Y_i | X = 0) = 0$ und damit

$$E(Y|X = 0) = 0 = 5,6 \cdot 0. \quad (21)$$

Ist hingegen $n > 0$, so kann man

$$X = \tilde{X}_i + X_i \quad (22)$$

schreiben, wobei die Zufallsgröße $\tilde{X}_i := X - X_i$ die Anzahl der JKB im Schwimmbad mit Ausnahme des

- eventuell anwesenden - i -ten JKB bezeichnet. Dann gilt

$$\begin{aligned} & P(Y_i = y, X_i = 1, X = n) \\ &= P(Y_i = y, X_i = 1, \tilde{X}_i = n - 1) \\ &= P(Y_i = y, X_i = 1 | \tilde{X}_i = n - 1) P(\tilde{X}_i = n - 1) \end{aligned}$$

Wir treffen nun eine - entscheidende - Annahme:

Das Ankunfts- und Ausgabeverhalten (Y_i, X_i) des i -ten JKB sei stochastisch unabhängig von der Anzahl der übrigen JKB im Schwimmbad, d.h. von \tilde{X}_i . Gleichwertig dazu ist, dass für alle $\tilde{x} \in \{0, \dots, k - 1\}$ und $y \in \{0, 4, 12\}$ gilt:

$$P(Y_i = y, X_i = 1 | \tilde{X}_i = \tilde{x}) = P(Y_i = y, X_i = 1). \quad (23)$$

Es sei darauf hingewiesen, dass (23) insbesondere dann gilt, wenn folgende stärkere Voraussetzung gegeben ist: **Die Paare $(X_1, Y_1), \dots, (X_k, Y_k)$ sind stochastisch unabhängig**, d.h. das individuelle Besuchs- und Ausgabeverhalten der JKB weist keine wechselseitigen Abhängigkeiten auf.

Unter der Annahme (23) lässt sich nun die zuletzt gefundene Wahrscheinlichkeit weiter umformen:

$$\begin{aligned} & P(Y_i = y, X_i = 1, X = n) \\ &= P(Y_i = y, X_i = 1 | \tilde{X}_i = n - 1) P(\tilde{X}_i = n - 1) \\ &= P(Y_i = y, X_i = 1) P(\tilde{X}_i = n - 1) \\ &= P(Y_i = y | X_i = 1) P(X_i = 1) P(\tilde{X}_i = n - 1) \\ &= p_y \cdot p \cdot P(\tilde{X}_i = n - 1). \end{aligned}$$

Eine weitere Annahme betrifft nun das Ankunftsverhalten der restlichen $k - 1$ JKB, d.h. die Verteilung der Größe \tilde{X}_i . Da bereits X als binomialverteilt vorausgesetzt wurde, so wird dies auch auf \tilde{X}_i zutreffen, d.h. für \tilde{X}_i darf eine Binomialverteilung mit Parametern $k - 1$ und p angenommen werden. Daher gilt

$$\begin{aligned} & P(Y_i = y, X_i = 1, X = n) \\ &= p_y \cdot p \cdot \binom{k-1}{n-1} p^{n-1} (1-p)^{k-1-(n-1)} \\ &= p_y \cdot \binom{k-1}{n-1} p^n (1-p)^{k-n} \\ &= p_y \cdot \frac{n}{k} \cdot \binom{k}{n} p^n (1-p)^{k-n} = p_y \cdot \frac{n}{k} \cdot P(X = n). \end{aligned}$$

⁸Nimmt man zusätzlich an, dass es mehrere Tandems gibt, im Fall $k = 2000$ also etwa 1000 solcher Tandems, so lässt sich das Gegenbeispiel z.B. durch die Annahme unabhängig voneinander agierender Tandems auf die Situation der Aufgabenstellung übertragen.

Mit diesen Überlegungen folgt nun

$$\begin{aligned} & E(Y_i | X = n) \\ &= \sum_{y \in \{4, 12\}} y \cdot \frac{P(Y_i = y, X = n)}{P(X = n)} \\ &= \sum_{y \in \{4, 12\}} y \cdot \frac{P(Y_i = y, X_i = 1, X = n)}{P(X = n)} \\ &= \sum_{y \in \{4, 12\}} y \cdot \frac{p_y \cdot \frac{n}{k} \cdot P(X = n)}{P(X = n)} \\ &= \frac{n}{k} \cdot \sum_{y \in \{4, 12\}} y \cdot p_y = 5,6 \cdot \frac{n}{k}. \end{aligned}$$

Schließlich folgt

$$E(Y | X = n) = \sum_{i=1}^k E(Y_i | X = n) = \sum_{i=1}^k 5,6 \cdot \frac{n}{k} = 5,6 \cdot n.$$

Es ergibt sich also (18). Folglich ist die in der Musterlösung vorausgesetzte Gleichung korrekt, sofern das Besuchs- und Ausgabeverhalten eines einzelnen JKB stochastisch unabhängig zu dem Besuchsverhalten der übrigen JKB im Sinne von (23) ist. Diese Annahme wirft grundlegende Fragen nach einer realistischen Abhängigkeitsstruktur auf. Wird etwa ein JKB sein Kaufverhalten nicht abhängig von der Anzahl der weiteren Besucher - insbesondere der anwesenden JKB - anpassen? Sind viele JKB im Schwimmbad, so sind ggf. auch die Warteschlangen vor dem Kiosk lang, so dass die Wahrscheinlichkeit, dass ein einzelner JKB kein Eis kauft, unmittelbar abhängig von der Anzahl der JKB im Schwimmbad ist. Dies bedeutet aber, dass man die Grundannahme (23) nicht einfach treffen kann.

Nun könnte man auf die Idee kommen, dass die Grundannahme (23) gar nicht benötigt wird, sondern vielmehr die Gleichung (18) voraussetzungsfrei gültig ist. Im Folgenden werden wir hierzu ein Gegenbeispiel präsentieren, welches genau die Abhängigkeit eines JKB von den anderen anwesenden JKB modelliert.

6 Ein Gegenbeispiel mit zwei Jahreskartenbesitzern

Das Gegenbeispiel hat die Struktur eines Tandem-Modells, d.h. die JKB werden zu stochastisch gleichartigen Paaren zusammengefasst. Wichtig: Die zu einem Paar zusammengefassten JKB müssen nicht beide im Schwimmbad sein. Das Ausgabeverhalten innerhalb eines Paares wird mit einer Abhängigkeitsstruktur modelliert, für welche die in der Musterlösung angenommene Form des bedingten Erwartungswertes

tungswertes falsch wird. Auch wenn das Tandem-Modell nicht realistischer als beispielsweise die stochastische Unabhängigkeit sämtlicher Ausgaben ist, wird dennoch klar, dass man die Aussagen der Musterlösung nicht einfach hinnehmen kann. Hier betrachten wir zudem nur den Fall von $k = 2$ JKB, die dann ein einziges Tandem bilden⁸. Die Anzahl der JKB im Schwimmbad ist dann

$$X = X_1 + X_2. \quad (24)$$

Die Ausgaben der JKB betragen

$$Y = Y_1 + Y_2. \quad (25)$$

Es wird ein stochastisches Modell für die gemeinsame Verteilung von Y_1, Y_2, X_1, X_2 angegeben. Da weiter X_1, X_2 als stochastisch unabhängig und Bernoulli-verteilt mit Parameter $p = 0,1$ angenommen werden können, genügt es, die bedingten Wahrscheinlichkeiten

$$P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_1, X_2 = x_2) \quad (26)$$

für $y_i \in \mathcal{Y}, x_i \in \{0,1\}$ festzulegen, denn die gemeinsamen Wahrscheinlichkeiten berechnen sich durch

$$\begin{aligned} & P(Y_1 = y_1, Y_2 = y_2, X_1 = x_1, X_2 = x_2) \\ &= P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_1, X_2 = x_2) \cdot \\ & \quad \cdot P(X_1 = x_1, X_2 = x_2). \end{aligned}$$

Die bedingten Wahrscheinlichkeiten gemäß (26) werden nun so festgelegt, dass die Voraussetzungen bezüglich des Kaufverhaltens von JKB, nämlich

$$P(Y_i = y | X_i = 1) = p_y = \begin{cases} 0,5 & , y = 4 \\ 0,3 & , y = 12 \\ 0,2 & , y = 0 \end{cases} \quad (27)$$

für $i = 1,2$ erfüllt sind, während die Aussage (18) nicht mehr für alle $n \in \{0,1,2\}$ gilt. Es sei noch bemerkt, dass unter diesen Annahmen gilt:

$$E(Y_i | X_i = 1) = 4 \cdot 0,5 + 12 \cdot 0,3 = 5,6. \quad (28)$$

6.1 Hilfsgrößen

Es seien drei Werte $\beta_0, \beta_4, \beta_{12} \geq 0$ vorgegeben mit $\beta_0 + \beta_4 + \beta_{12} = 1$. Die β_y legen eine diskrete Wahrscheinlichkeitsverteilung β mit Erwartungswert

$$\mu_\beta = \sum_{y \in \mathcal{Y}} y \cdot \beta_y \quad (29)$$

fest. Weiter werde für $y \in \mathcal{Y}$ definiert

$$\alpha_y := \frac{p_y - p \cdot \beta_y}{1 - p} \quad (30)$$

Die α_y bilden eine Wahrscheinlichkeitsverteilung α :

- $p \cdot \beta_y \leq p = 0,1 \leq \min(0,5; 0,3; 0,2) \leq p_y$,

also $p_y - p \cdot \beta_y \geq 0$, also $\alpha_y \geq 0$.

- $\sum_{y \in \mathcal{Y}} \alpha_y = \sum_{y \in \mathcal{Y}} \frac{p_y - p \cdot \beta_y}{1 - p} = \frac{\sum_{y \in \mathcal{Y}} p_y - p \sum_{y \in \mathcal{Y}} \beta_y}{1 - p} = 1$.

Diese hat den Erwartungswert

$$\mu_\alpha = \sum_{y \in \mathcal{Y}} y \cdot \alpha_y = \frac{5,6 - p \cdot \mu_\beta}{1 - p}. \quad (31)$$

Es lässt sich leicht nachweisen, dass gilt

$$\mu_\alpha = 5,6 \Leftrightarrow \mu_\beta = 5,6. \quad (32)$$

Warum man α_y so und nicht anders festlegt, wird im nächsten Abschnitt im Beweis von Lemma 1, Gleichung (33) deutlich.

6.2 Das stochastische Modell im Gegenbeispiel

Es wird folgendes Modell für die bedingten Wahrscheinlichkeiten $P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_1, X_2 = x_2)$ der Ausgaben Y_1, Y_2 gegeben X_1, X_2 angenommen:

- Für $X_1 = X_2 = 1$ sind Y_1, Y_2 (bedingt) stochastisch unabhängig gemäß β verteilt.
- Für $X_1 = 1, X_2 = 0$ bzw. $X_1 = 0, X_2 = 1$ sind Y_1, Y_2 gemäß Hilfsverteilung α verteilt.
- Für $X_1 = X_2 = 0$ entstehen keine Ausgaben.

In Abb. 1 ist der Wahrscheinlichkeitsbaum zu diesem Beispiel angegeben.

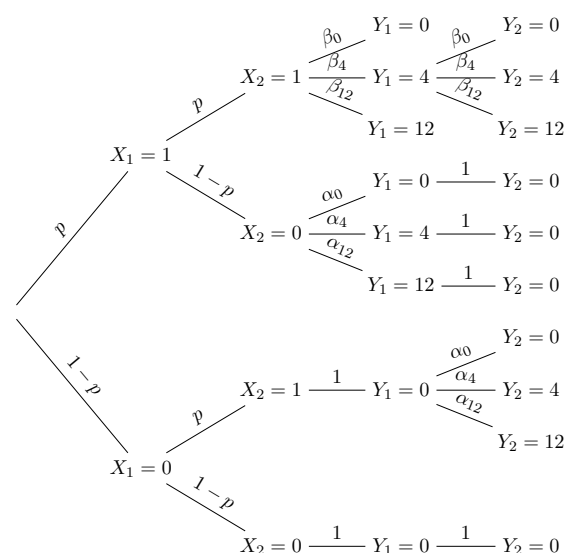


Abb. 1: Tandemmodell. Zur Vereinfachung ist im Teilbaum zu $X_1 = 1, X_2 = 1$ die letzte Stufe nur für $Y_1 = 4$ angegeben, sie ist für alle anderen Werte von Y_1 identisch.

6.3 Konformität des Modells mit der Aufgabenstellung

Lemma 1: $P(Y_i = y|X_i = 1) = p_y$ für $i = 1, 2$ und $y \in \mathcal{Y}$, d.h. das Modell aus Abschnitt 6.2 erfüllt die Annahmen der Aufgabenstellung in Aufgabenteil c)(2) der Grundkurs- bzw. b)(2) der Leistungskursversion.

Beweis: wird nur für $i = 1$ geführt, für $i = 2$ analoge Argumentation. Zunächst gilt

$$P(Y_1 = y|X_1 = 1) = \frac{P(Y_1 = y, X_1 = 1)}{P(X_1 = 1)}.$$

Der Zähler $P(Y_1 = y, X_1 = 1)$ stellt sich additiv dar aus den beiden Summanden

$$\begin{aligned} & P(Y_1 = y, X_1 = 1, X_2 = 0) \\ &= P(Y_1 = y|X_1 = 1, X_2 = 0)P(X_1 = 1, X_2 = 0) \\ &= \alpha_y \cdot p \cdot (1 - p) \end{aligned}$$

und

$$\begin{aligned} & P(Y_1 = y, X_1 = 1, X_2 = 1) \\ &= P(Y_1 = y|X_1 = 1, X_2 = 1)P(X_1 = 1, X_2 = 1) \\ &= \beta_y \cdot p^2. \end{aligned}$$

Es gilt daher nach Festlegung von α_y mit $p = 0,1$, $1 - p = 0,9$ und wenigen Umformungen:

$$\begin{aligned} P(Y_1 = y|X_1 = 1) &= \frac{\alpha_y \cdot p \cdot (1 - p) + \beta_y \cdot p^2}{p} \\ &= \alpha_y \cdot (1 - p) + \beta_y \cdot p \\ &= \frac{p_y - p \cdot \beta_y}{1 - p} (1 - p) + \beta_y \cdot p \\ &= p_y. \quad \square \end{aligned} \quad (33)$$

6.4 Bestimmung von $E(Y|X = n)$ im Gegenbeispiel

Lemma 2: Im Modell aus Abschnitt 6.2 gilt:

1. $E(Y|X = 1) = \mu_\alpha = \frac{5,6 - p \cdot \mu_\beta}{1 - p}$,
2. $E(Y|X = 2) = 2\mu_\beta$.

Weiter für $n > 0$: $E(Y|X = n) = 5,6n \Leftrightarrow \mu_\beta = 5,6$.

Beweis: Zunächst gilt

$$\begin{aligned} E(Y|X = n) &= E(Y_1 + Y_2|X = n) \\ &= E(Y_1|X = n) + E(Y_2|X = n) \\ &= 2E(Y_1|X = n), \end{aligned}$$

denn die beiden Summanden stimmen aus Symmetriegründen überein. Dann gilt:

⁹<https://www.wi.uni-muenster.de/sites/wi/files/users/terveer/sis/anhang.pdf>

Fall $n = 1$:

$$\begin{aligned} P(Y_1 = y|X = 1) &= \frac{P(Y_1 = y, X_1 + X_2 = 1)}{P(X_1 + X_2 = 1)} \\ &= \frac{P(Y_1 = y, X_1 = 1, X_2 = 0)}{P(X_1 + X_2 = 1)} \\ &\quad + \frac{P(Y_1 = y, X_1 = 0, X_2 = 1)}{P(X_1 + X_2 = 1)} \\ &= \frac{P(Y_1 = y|X_1 = 1, X_2 = 0)P(X_1 = 1, X_2 = 0)}{P(X_1 + X_2 = 1)} \\ &\quad + \frac{P(Y_1 = y|X_1 = 0, X_2 = 1)P(X_1 = 0, X_2 = 1)}{P(X_1 + X_2 = 1)} \\ &= \frac{\alpha_y p(1 - p) + \mathbf{1}_{\{0\}}(y)(1 - p)p}{2p(1 - p)} = \frac{\alpha_y + \mathbf{1}_{\{0\}}(y)}{2}. \end{aligned}$$

Hieraus folgt $E(Y_1|X = 1) = \frac{1}{2} \sum_{y \in \mathcal{Y}} y \cdot (\alpha_y + \mathbf{1}_{\{0\}}(y))$ und daraus $E(Y|X = 1) = \mu_\alpha$.

Fall $n = 2$:

$$P(Y_1 = y|X = 2) = P(Y_1 = y|X_1 = 1, X_2 = 1) = \beta_y.$$

Hieraus folgt $E(Y_1|X = 2) = \sum_{y \in \mathcal{Y}} y \cdot \beta_y = \mu_\beta$ und damit $E(Y|X = 2) = 2\mu_\beta$. \square

Aus Lemma 2 ergibt sich nun, dass die Gleichung $E(Y|X = n) = 5,6n$ nur dann für alle n gilt, wenn $\mu_\beta = 5,6$. Da aber $(\beta_0, \beta_4, \beta_{12})$ eine beliebige Wahrscheinlichkeitsverteilung auf \mathcal{Y} ist, kann diese Aussage nicht allgemein richtig sein. Die in der Musterlösung der Schwimmbadaufgabe implizit angenommene Formel für bedingte Erwartungen erfordert also in der Aufgabenstellung zusätzliche Voraussetzungen.

7 Übertragung des Gegenbeispiels auf 2000 Jahreskartenbesitzer

Das Gegenbeispiel aus dem letzten Kapitel lässt sich auf den Fall der Aufgabenstellung mit 2000 JKB in der Form übertragen, dass die JKB in 1000 stochastisch unabhängigen Tandem-Paaren auftreten. In diesem Fall ist bei geeigneter Wahl von $(\beta_0, \beta_4, \beta_{12})$ nicht nur die Annahme der Musterlösung, d.h. (18) falsch, sondern auch das angegebene Endergebnis der Aufgabe $P(5,6 \cdot X \geq 1000) = 0,947$. Die etwas mühsame Detailrechnung dazu ist digital verfügbar⁹.

8 Reflexion

Die Schwimmbadaufgabe offenbart die Schwierigkeiten bei der Erstellung adäquater Stochastik-Prüfungsaufgaben. Während Anwendungsbezug,

Originalität und das Potenzial zur Differenzierung in der Prüfungsgruppe möglichst groß sein sollen, muss eine Aufgabe mit einem zum Anwendungskontext passenden und gleichzeitig curricular verorteten stochastischen Modell lösbar sein. Die Schwimmbadaufgabe zeigt, wie letzteres bei der Suche nach einer originellen Aufgabenstellung verloren gehen kann.

Abituraufgaben des Landes NRW sind jeweils ab dem folgenden Herbst drei Jahre lang für Lehrkräfte und Schüler auf Internetseiten des NRW-Schulministeriums einsehbar. Es sollte selbstverständlich sein, dass eine so auffällig gewordene Teilaufgabe aber entfernt oder in geeigneter Weise kommentiert würde. Eine solche Gegensteuerung gegen Fehlerkonservierung ist allerdings leider unterblieben und sicher gar nicht beabsichtigt, was sich auch aus der Antwort erschließen lässt, die die Landesregierung auf eine kleine Landtagsanfrage¹⁰ gegeben hat.

Literatur

F.-R. Diepenbrock: Manche Panne beim Mathe-Abitur ist hausgemacht. F.A.Z. Bildungswelten vom 18.3.2021.

Mein Abi-Coach Mathematik Grundkurs. Klett 2020.

Antwort der Landesregierung auf die Kleine Anfrage 3880 vom 19. Juni 2020.
<https://www.landtag.nrw.de/portal/WWW/dokumentenarchiv/Dokument/MMD17-10301.pdf>.

Heruntergeladen am 07.04.2021

Abitur, Original-Prüfungsaufgaben mit Lösungen, Gymnasium-Gesamtschule NRW, Mathematik LK. Stark-Verlag 2020.

Mathematik - Standardsicherung NRW.
<https://www.standardsicherung.schulministerium.nrw.de/cms/zentralabitur-wbk/faecher/getfile.php?file=2239>.

Heruntergeladen am 28.07.2020

Anschriften der Verfasser

Dr. Ingolf Terveer

Institut für Wirtschaftsinformatik

Westfälische Wilhelms-Universität Münster

48149 Münster

terveer@uni-muenster.de

Prof.i.R. Dr. Franz-Reinhold Diepenbrock

Sadowastraße 63

42115 Wuppertal

diepenbrock@uni-wuppertal.de

¹⁰vgl. NRW (2020), Seite 2