

All you need is luck?

Xinchi Yu und Johanna Mielke

Inhalt

1. Einleitung.....	S. 2
2. Berechnungen mit der Bernoulli-Kette.....	S. 2
2.1 Definition: Die Bernoulli Kette	S. 2
2.2 „2 von 3“: Eine Erweiterung der Bernoulli-Kette.....	S. 3
2.3 Beispiel: Der Schultest.....	S. 4
2.4 Programm zur Berechnung der Bernoulli-Ketten.....	S. 4
2.4.1 Lösung der erweiterten Bernoulli-Kette.....	S. 4
2.4.2 Berechnung der Ergebnisse.....	S. 5
2.5 Der Schulabschlusstest.....	S. 5
2.6 Programm zur Lösung der mehrstufigen Bernoulli-Kette.....	S. 6
3. Die Führerscheinprüfung	S. 6
3.1 Der Aufbau der Führerscheinprüfung.....	S. 6
3.2 Berechnung der Wahrscheinlichkeit für die richtige Beantwortung einer Frage.....	S. 6
3.3 Maximale Anzahl an falschen Antworten.....	S. 7
3.4 Programm zur Berechnung der Wahrscheinlichkeit bei der Führerscheinprüfung.....	S. 8
3.5 Auswertung.....	S. 8
3.6 Kritische Reflexion.....	S. 8
4. Der optimale Test.....	S. 9
4.1 Einleitung.....	S. 9
4.2 Verteilung von Fragen und Antworten.....	S. 9
4.2.1 Allgemeine Überlegungen.....	S. 9
4.2.2 Berechnung der optimalen Anzahl von Fragen und Antworten.....	S. 10
4.3 Programm zur Erstellung und Bewertung eines idealen Tests.....	S. 12
4.4 Bewertung eines realen Schultests.....	S. 14
4.5 Kritische Reflexion.....	S. 14
5. Erweiterungen.....	S. 14
5.1 Berechnung der Wahrscheinlichkeit pro Frage.....	S. 14
5.1.1 Idee.....	S. 14
5.1.2 Umsetzung.....	S. 14
5.2 Berechnung der Wahrscheinlichkeit unter Berücksichtigung von Minuspunkten.....	S. 14
5.2.1 Idee.....	S. 14
5.2.2 Mathematische Grundlagen.....	S. 15
5.2.3 Umsetzung.....	S. 15
5.3 Berechnung eines optimalen Tests mit der Bernoulli-Kette.....	S. 15
5.3.1 Idee.....	S. 15
5.3.2 Probleme.....	S. 15
5.3.3 Lösungsansatz.....	S. 16
5.3.4 Umsetzung.....	S. 16
6. Literaturverzeichnis.....	S. 16

1. Einleitung

„All You Need Is Luck“¹

Mit dieser Einstellung scheinen einige Personen zu versuchen, die Theorieprüfung des Führerscheins zu bestehen. Daher ist es eigentlich nicht verwunderlich, dass manche die Prüfung nicht bestehen, obwohl die Fragen vorher bekannt sind und man diese nur auswendig lernen muss.

Doch haben diese Personen wirklich nur Pech gehabt oder ist es unrealistisch, die Theorieprüfung durch Raten zu bestehen? Kann man – mit dem nötigen Glück – die Prüfung ohne Lernen schaffen? Wie hoch ist die Wahrscheinlichkeit, in der Schule einen einfach aufgebauten oder mehrstufigen Multiple-Choice-Test zu bestehen? Wie sollte ein Multiple-Choice-Test überhaupt aufgebaut werden, damit die Wahrscheinlichkeit, ihn durch Raten zu bestehen, möglichst gering ist?

Diese Fragen möchten wir mit dieser Arbeit versuchen zu beantworten.

Wir haben uns für dieses Thema entschieden, da der Führerschein in unserem Alter ein sehr aktuelles Thema ist. Das Lernen für die Prüfung ist zeitaufwendig und daher ist die Frage, ob es überhaupt notwendig ist, natürlich interessant. Multiple-Choice-Tests spielen auch in der Schule eine große Rolle. So wurde das Leseverständnis in den Fächern Deutsch und Englisch bei den Lernstandserhebungen NRW über einen Multiple-Choice-Test bewertet. Das Gleiche gilt für die französische DELF/DALF-Prüfung², die immerhin zu einem Studium in Frankreich ohne zusätzliche Sprachprüfung berechtigt. Daher ist es sehr interessant zu sehen, ob man durch diese Tests überhaupt feststellen kann, ob der Schüler den Text verstanden hat.

Für derartige Problemstellungen ist die sog. Bernoulli-Kette ein geeignetes Werkzeug. Zuerst werden wir daher die Bernoulli-Kette allgemein definieren, denn diese Definition ist die Grundlage für unsere weiteren Untersuchungen. Anschließend werden wir die Bernoulli-Kette so verändern, dass sich mit ihr einfache Multiple-Choice-Probleme behandeln lassen.

Mit dieser Formel können wir das erste einfache Anwendungsproblem untersuchen: Einen normalen Multiple-Choice-Test in der Schule. Danach weiten wir das Problem auf einen mehrstufigen Test aus, bei dem ein bestimmter Anteil von Fachgebieten zu bestehen ist. Hier erfolgt auch die erste Umsetzung mathematischer Gedanken in ein Programm, da das Rechnen mit der Bernoulli-Kette mit der Hand ziemlich umständlich ist.

Als nächstes befassen wir uns mit der Führerscheinproblematik. Bei der Frage, wie wahrscheinlich es ist, die Prüfung durch Raten zu bestehen, mussten wir zunächst den Aufbau eines Führerscheinbogens untersuchen. Mit diesen Informationen konnten wir dann die Wahrscheinlichkeit, eine Frage richtig zu beantworten, berechnen und daraus auf die Wahrscheinlichkeit, einen Bogen richtig zu beantworten, schließen. Da auch hier die Berechnung per Hand sehr unpraktisch ist, wurde auch für dieses Problem ein Programm erstellt. Anschließend versuchen wir, unsere Ergebnisse kritisch zu hinterfragen und die Übertragbarkeit der Berechnungen auf die Realität zu bewerten.

Abschließend gehen wir allgemein auf den Aufbau eines Fragebogens ein und bestimmen die optimale Verteilung von Fragen und Antworten über eine Funktion. Optimal bedeutet hier, dass die Wahrscheinlichkeit, den Test durch Raten zu bestehen, möglichst gering ist. Auch für diesen Aspekt wurde ein Programm erstellt.

2. Berechnungen mit der Bernoulli-Kette

2.1 Definition: Die Bernoulli-Kette³

Ein Zufallsexperiment, bei dem es nur zwei verschiedene Ereignisse gibt, beispielsweise „richtig“ und „falsch“ bei einem Fragebogen, nennt man Bernoulli-Versuch. Wenn man dieses Experiment n-mal wiederholt und sich dabei die Wahrscheinlichkeit, ein günstiges bzw. ungünstiges Ereignis – d.h. in unserem Falle, eine Frage richtig bzw. falsch zu beantworten – nicht ändert, die Ereignisse also unabhängig voneinander sind, so bezeichnet man dies als Bernoulli-Kette.

¹ Malcolm Higgins: „All You Need Is Luck (the Bosox Anthem)“, <http://www.amiright.com/parody/60s/thebeatles280.shtml> (Anhang 6.4)

² Stand: 2007

³ vgl. <http://www.physik-schule.de/download/pdf/Mathematik/Bernoulli-Kette.pdf> (Anhang 6.1)

Mit der Bernoulli-Kette lässt sich berechnen, wie wahrscheinlich es ist, dass genau k günstige Ereignisse entstehen. Dafür betrachtet man zuerst die Wahrscheinlichkeit, die ersten k Bernoulli-Versuche richtig zu beantworten. Die Wahrscheinlichkeit beträgt p^k , wobei p die Wahrscheinlichkeit für ein günstiges Ereignis und k die gewünschte Anzahl an günstigen Ereignissen ist.

Die Wahrscheinlichkeit, dass ein Ereignis ungünstig ist, beträgt $(1 - p)$. Da die ersten k von n Bernoulli-Versuchen günstig sein sollen, müssen die restlichen Bernoulli-Versuche ungünstig sein. Die Wahrscheinlichkeit, genau $n-k$ aufeinander folgende ungünstige Ereignisse zu erhalten, beträgt folglich $(1 - p)^{n-k}$. Die Wahrscheinlichkeit dafür, dass die günstigen und ungünstigen Ereignisse in dieser Reihenfolge auftreten, beträgt $p^k \cdot (1 - p)^{n-k}$.

Nun gibt es nicht nur eine Möglichkeit, eine bestimmte Anzahl an günstigen und ungünstigen Ereignissen zu erreichen, da es nicht entscheidend, welche Fragen richtig oder falsch beantwortet werden. Daher muss nun noch die gesamte Anzahl an günstigen Anordnungen berücksichtigt werden. Es gibt genau $\binom{n}{k}$ verschiedene Anordnungen.

Insgesamt gilt also der folgende Zusammenhang:

$$P_{k/n} = \binom{n}{k} p^k \cdot (1 - p)^{n-k}$$

2.2 „2 von 3“: Eine Erweiterung der Bernoulli-Kette

Die Bernoulli-Kette berechnet nur, wie wahrscheinlich es ist, genau eine bestimmte Anzahl an richtigen Fragen zu erzielen. Bei unserer Fragestellung muss nur eine Mindestanzahl an Fragen richtig beantwortet werden, um ein positives Testergebnis zu erhalten.

Um die Wahrscheinlichkeit für ein positives Testergebnis zu errechnen, müssen daher auch die Fälle berücksichtigt werden, bei denen mehr Fragen als die nötige Mindestanzahl korrekt beantwortet wurden.

Wir wollen diesen Aspekt an einem Beispiel verdeutlichen. Es gibt drei Fragen mit je zwei Antwortmöglichkeiten, von denen jeweils genau eine Möglichkeit richtig ist. Für ein günstiges Ergebnis (Bestehen des Tests) müssen mindestens zwei der drei Fragen richtig beantwortet werden. Mit der Bernoulli-Kette errechnet man in diesem Fall zuerst die Wahrscheinlichkeit $(p_{2/3})$, genau zwei von drei Fragen richtig zu beantworten, danach wird die Wahrscheinlichkeit $(p_{3/3})$ für genau drei von drei Fragen berechnet. Anschließend werden die beiden Wahrscheinlichkeiten addiert, sodass man die Gesamtwahrscheinlichkeit erhält (P_{ges}) .⁴

Die allgemeine Formel für genau k positive Ereignisse lautet:⁵

$$P_{k/n} = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} = \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1 - p)^{n-k}$$

Für das diskutierte Beispiel gelten die folgenden Parameter:

Zahl der Fragen: $n = 3$

Wahrscheinlichkeit eine Frage richtig zu beantworten: $p = \frac{1}{2}$

gewünschte Mindestanzahl an richtigen Fragen: $i = 2$

Berechnung:

⁴ Additionsgesetze für Wahrscheinlichkeiten, vgl. Kreyszig 1977: 53ff

⁵ vgl. 2.1

$$p_{2/3} = \frac{3!}{2!(3-2)!} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(1 - \left[\frac{1}{2}\right]\right)^{3-2}$$

$$p_{3/3} = \frac{3!}{3!(3-3)!} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(1 - \left[\frac{1}{2}\right]\right)^{3-3}$$

$$P_{ges} = p_{2/3} + p_{3/3}$$

$$P_{ges} = \frac{3!}{3!(3-3)!} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(1 - \left[\frac{1}{2}\right]\right)^{3-3} + \frac{3!}{2!(3-2)!} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(1 - \left[\frac{1}{2}\right]\right)^{3-2} = 0,5$$

Die Wahrscheinlichkeit, den Test durch Raten zu bestehen, beträgt damit 50 Prozent.

Man kann die Wahrscheinlichkeit für die korrekte Beantwortung von mindestens i von insgesamt n Fragen bei einer Wahrscheinlichkeit von p allgemein mit dem folgenden Term beschreiben:

$$P_{ges} = \sum_{k=i}^n P_{k/n} = \sum_{k=i}^n \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1-p)^{n-k}$$

2.3 Beispiel: Der Schultest

In diesem Beispiel soll ein ganz normaler Schultest untersucht werden. Der Test besteht aus zehn Fragen mit je drei Antwortmöglichkeiten, von denen immer genau eine Möglichkeit richtig ist. Um den Test zu bestehen, muss man in der Oberstufe normalerweise mehr als 40 Prozent der Punkte erreicht haben, d.h. in diesem Beispiel, dass der Schüler vier Fragen richtig beantworten muss. Wie hoch ist die Wahrscheinlichkeit, dieses Ziel zu erreichen?

Die Lösung erfolgt über die erweiterte Bernoulli-Kette:

$$P_{ges} = \sum_{k=i}^n P_{k/n} = \sum_{k=i}^n \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1-p)^{n-k}$$

$$i = 4; n = 10; p = \frac{1}{3}$$

$$\sum_{k=4}^{10} P_{k/n} = \frac{10!}{k!(10-k)!} \cdot \left(\frac{1}{3}\right)^k \cdot \left(1 - \frac{1}{3}\right)^{10-k} = 0,44$$

Der Wert liegt nur knapp unter 50 Prozent, d.h. fast jeder zweite Schüler könnte den Test bestehen, ohne zu lernen. Das Ergebnis des Tests wäre daher nicht besonders aussagekräftig.

2.4 Programm zur Bernoulli-Kette⁶

2.4.1 Lösung der erweiterten Bernoulli-Kette

Die o.g. Berechnungen lassen sich mit Hilfe unseres ersten Programms wesentlich einfacher durchführen. Dabei kann man die Wahrscheinlichkeit, eine Frage richtig zu beantworten, als Bruch in das Feld „p“ eintragen. Alternativ kann man auch jede reelle Zahl mit bis zu 42 Nachkommastellen eintragen, wenn man die reelle Zahl als Zähler und für den Nenner „1“ einträgt. Die Anzahl der benötigten richtigen Fragen trägt man im Feld „k“ und die Anzahl aller Fragen im Feld „n“ ein.⁷ Mit einem Klick auf „Berechne!“ wird die Wahrscheinlichkeit, den Test durch Raten zu bestehen, ausgerechnet.

Zusätzlich zum Endergebnis zeigt das Programm die Ergebnisse einiger Zwischenrechnungen an:

1. $n!$
2. $k!$ und $(n-k)!$
3. $k!(n-k)!$
4. p^k

⁶ vgl. Quelltext: Anhang 6.10

⁷ Überprüfung auf korrekte Eingabe siehe Anhang 6.8 und 6.9

5. $1 - p$

6. $P_{k/n}$

Die ersten drei Kontrollergebnisse beziehen sich auf den Binomialkoeffizienten $\binom{n}{k}$, der zu $\frac{n!}{k!(n-k)!}$ umgeformt wird. Die Kontrollergebnisse 4. und 5. sind weitere Teile der Gleichung $P_{k/n} = \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1-p)^{n-k}$. Da das Programm zur Berechnung des Endergebnisses die erweiterte Bernoulli-Kette anwendet, werden unter 6. die einzelnen Wahrscheinlichkeiten für die verschiedenen k angezeigt.

2.4.2 Berechnung des Ergebnisses

Der Berechnung der Wahrscheinlichkeit liegt die Gleichung $P_{k/n} = \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1-p)^{n-k}$

zugrunde. Dabei wird die gesamte Rechnung in insgesamt acht Einzelschritte zerlegt. Zuerst werden der Zähler und der Nenner des umgeformten Binomialkoeffizienten berechnet. Dazu wird eine selbst geschriebene Funktion „Fakultaet“ zur Berechnung von Fakultäten benutzt, da Delphi keine vordefinierte Funktion zur Berechnung von Fakultäten besitzt.⁸ Nachdem Zähler und Nenner berechnet wurden, wird der Zähler durch den Nenner dividiert und man erhält den Binomialkoeffizienten als Gleitkommazahl.

Im nächsten Schritt wird der Wert für p^k berechnet. Damit eine höchstmögliche Genauigkeit erzielt werden kann, werden Zähler und Nenner getrennt mit „k“ potenziert und erst dann durch Division in eine Gleitkommazahl umgewandelt. Für das Potenzieren wird die Delphi-Funktion „Power“ verwendet.

Anschließend wird $1 - p$ berechnet. Auch hier wird aus Genauigkeitsgründen zuerst mit Brüchen gerechnet und erst am Ende dividiert. Daher wird die „1“ mit dem Nenner von „p“ erweitert, der Zähler von „p“ davon subtrahiert und dann durch den Nenner von „p“ dividiert. Danach wird „(1-p)“ mit „(n-k)“ potenziert und das Endergebnis kann zusammengesetzt werden.

Da bei Multiple-Choice Tests die erweiterte Bernoulli-Kette zur Anwendung kommt, wird die eben beschriebene Rechnung in einer For-Schleife untergebracht. Diese For-Schleife wird genau „n-k+1“ mal durchlaufen, damit die Wahrscheinlichkeiten für alle „k“, die entweder gleich „n“ oder zwischen dem eingegebenen „k“ und „n“ liegen, berechnet werden.

Zum Schluss werden die Wahrscheinlichkeiten durch Addieren aller Werte zu einem Endergebnis zusammengefasst und angezeigt.

2.5 Der Schulabschlusstest

Bei einem Schulabschlusstest werden mehrere Fächer geprüft. In unserem Beispiel nehmen wir an, dass es in jedem Prüfungsfach zehn Fragen mit je drei Antwortmöglichkeiten gibt, von denen genau eine Antwortmöglichkeit richtig ist. Um ein einzelnes Fach zu bestehen, muss man mindestens fünf der zehn Fragen richtig beantwortet haben. Um den Test zu bestehen, muss man mindestens drei der fünf Prüfungsfächer bestanden haben. Das scheint nicht besonders viel zu sein, schließlich muss man pro Bogen nur 50 Prozent der Fragen richtig beantworten und auch nur in 60 Prozent der Fächer positiv abschneiden. Wie groß ist die Wahrscheinlichkeit, den Test durch Raten zu bestehen?

Lösung: Mehrstufige Kette

In diesem Beispiel gibt es zwei Bernoulli-Ketten.

Die erste Bernoulli-Kette (Einzelpfprüfung) wird mit den folgenden Werten definiert:

$$i = 5; n = 10; p = \frac{1}{3}$$

$$P_{1ges} = \sum_{k=5}^{10} P_{5/10} = \sum_{k=5}^{10} \frac{10!}{k!(10-k)!} \cdot \left(\frac{1}{3}\right)^k \cdot \left(1 - \frac{1}{3}\right)^{10-k}$$

Der Computer berechnet für diesen Term den Wert 0.2131.

⁸ vgl. Anhang 6.7

Dieses Ergebnis wird nun für die zweite Bernoulli-Kette benötigt, da man diesen Wert für p einsetzen muss. Die Parameter sind daher:

$$i = 3; n = 5; p = P_{1ges} = 0,2131$$

$$P_{2ges} = \sum_{k=3}^5 P_{3/5} = \frac{5!}{k!(5-k)!} \cdot 0,2131^k \cdot (1-0,2131)^{5-k}$$

Für diesen Term erhält man das Ergebnis 0,07, also sieben Prozent. Damit ist die Wahrscheinlichkeit, den Test durch Raten zu bestehen, eher gering. Mit anderen Testkonstruktionen lassen sich aber, wie wir im Abschnitt 4 zeigen werden, noch wesentlich niedrigere Werte erreichen.

2.6 Programm zur Lösung der mehrstufigen Bernoulli-Kette

Das zweite Programm löst die Schulabschlusstest-Problematik mit Hilfe der oben beschriebenen mehrstufigen Bernoulli-Kette. Der Grundaufbau dieses Programms wurde vom Programm für die (erweiterte) Bernoulli-Kette abgeleitet, daher unterscheidet sich der Quelltext nur geringfügig vom ersten Programm.

So wird die beim ersten Durchgang berechnete Wahrscheinlichkeit (entspricht der Wahrscheinlichkeit, in einem Fach zu bestehen) automatisch als neues „ p “ für den zweiten Durchgang eingetragen, um daraus die Wahrscheinlichkeit, in x von y Fächern zu bestehen, auszurechnen. Außerdem werden die verwendeten Werte bzw. Ergebnisse für „ p “, „ k “ und „ n “ im Verlaufsbereich mitprotokolliert. Ein Doppelklick auf eine Datenreihe zeigt diese komplett an, falls sie in der Liste nicht komplett sichtbar ist. Zudem kann man den aktuellen Verlauf in einer Datei namens „boegen-verlauf.txt“ speichern lassen, die sich im gleichen Verzeichnis wie das Programm befindet – eine eventuell vorhandene Datei gleichen Namens wird dabei nach einer Sicherheitsabfrage überschrieben. Dabei muss das Medium beschreibbar sein und der Benutzer Schreibrechte für das Verzeichnis haben.

3. Die Führerscheinprüfung

3.1 Der Aufbau einer Führerscheinprüfung⁹

Eine Theorieprüfung des Führerscheins der Klasse B (Auto) besteht aus 30 Fragen. Jede Frage hat entweder zwei, drei oder keine vorgegebene Antwort. Freitextfragen werden zu Beginn unserer Berechnungen vernachlässigt. Von den Antworten können eine, zwei oder – sofern es drei Antwortmöglichkeiten gibt – auch drei richtig sein.

Die Verteilung der Anzahl der Antwortmöglichkeiten ist nicht für alle Themengebiete gleich.

Da es für die Verteilung der Antwortmöglichkeiten keine zuverlässigen Informationen gab, wurden diese Angaben selber erhoben. Pro Themengebiet wurde eine Stichprobe von 100 Fragen betrachtet. Bei Themengebieten, die keine 100 Fragen beinhalten, wurden alle vorhandenen Fragen überprüft.¹⁰

Themengebiet	Anteil am Bogen	0 Antworten	2 Antworten	3 Antworten
Gefahrenlehre	8/30	0	11/100	89/100
Verhalten im Straßenverkehr	6/30	0	10/100	90/100
Vorfahrt	3/30	0	13/43	30/43
Verkehrszeichen	2/30	0	6/100	94/100
Umweltschutz	1/30	0	4/44	40/44
Zusatzstoff	10/30	4/100	17/100	79/100

Insgesamt gibt es 30 Fragen, bei denen man 10 Fehlerpunkte erzielen darf.

3.2 Berechnung der Wahrscheinlichkeit für die richtige Beantwortung einer Frage

Zuerst muss berücksichtigt werden, dass man auch mehrere Antwortmöglichkeiten ankreuzen kann. Man kann bei jeder Antwortmöglichkeit nur „Ja“ oder „Nein“ wählen, d.h. man kann die Antwortmöglichkeit ankreuzen oder nicht.

⁹ vgl.: <http://www.fahrschule.de/Verkehrsrecht/Pruefungsrichtlinie/> (Anhang 6.2)

¹⁰ vgl. Literaturverzeichnis

Bei einer allgemeinen Frage mit drei Antwortmöglichkeiten beträgt die Wahrscheinlichkeit, eine Antwortmöglichkeit richtigerweise anzukreuzen für jede Antwortmöglichkeit $\frac{1}{2}$. Insgesamt beträgt die Wahrscheinlichkeit, eine Frage mit drei Antwortmöglichkeiten richtig zu beantworten, $\frac{1}{2^3}$, allgemein betrachtet also $\frac{1}{2^x}$, wobei x die Anzahl der Antwortmöglichkeiten¹¹ sei. Beim Führerscheintest gibt es allerdings noch eine weitere Besonderheit: Es ist immer falsch, wenn man keine Antwortmöglichkeit ankreuzt und dies ist jedem Prüfling bekannt. Daher muss man diese Möglichkeit anschließend noch von der Zahl aller möglichen Fälle abziehen. Es gilt also: $\frac{1}{2^x - 1}$

Konkret auf die Problematik angewandt erhält folgende Wahrscheinlichkeiten:

Bei Fragen mit zwei Antwortmöglichkeiten: $\frac{1}{2^2 - 1} = \frac{1}{3}$

Bei Fragen mit drei Antwortmöglichkeiten: $\frac{1}{2^3 - 1} = \frac{1}{7}$

Die Berechnung einer Gesamtwahrscheinlichkeit unter Berücksichtigung der verschiedenen Antworttypen kann nicht über den arithmetischen Mittelwert geschehen, da die verschiedenen Fragetypen unterschiedlich oft gestellt werden. Man kann diese jedoch über den Satz von Bayes über die Totale Wahrscheinlichkeit berechnen: $p(A) = p(A/B) \cdot p(B) + p(A/\bar{B}) \cdot p(\bar{B})$.^{12,13}

Für den Bereich „Gefahrenlehre“ ergibt sich daher:¹⁴ $\frac{11}{100} \cdot \frac{1}{3} + \frac{89}{100} \cdot \frac{1}{7} = 0,16$

Analog erhält man folgende Werte für die anderen Bereiche:

Verhalten im Straßenverkehr	0,16
Vorfahrt	0,2
Verkehrzeichen	0,15
Umwelt	0,16
Zusatzstoff	0,17 ¹⁵

Nun muss man alle Ergebnisse zusammenfassen. Dazu wird erneut der Satz über die Totale Wahrscheinlichkeit verwendet. Die Wahrscheinlichkeit für die richtige Beantwortung einer beliebigen Frage beträgt

$$\frac{8}{30} \cdot 0,16 + \frac{6}{30} \cdot 0,16 + \frac{3}{30} \cdot 0,2 + \frac{2}{30} \cdot 0,15 + \frac{1}{30} \cdot 0,16 + \frac{10}{30} \cdot 0,17 = 0,17.$$

3.3 Maximale Anzahl an falschen Antworten

Jeder Bogen hat 110 Fehlerpunkte¹⁶, die auf 30 Fragen verteilt sind. Jede Frage hat also durchschnittlich 3,6 Fehlerpunkte. Insgesamt besteht man die Prüfung, wenn weniger als zehn Fehlerpunkte erreicht wurden. Die Fehlerpunkte sind über die Bereiche unterschiedlich verteilt. So findet man im Bereich Verhalten im Straßenverkehr mehr Fragen mit fünf Fehlerpunkten als im Bereich Zusatzfragen. Da die Unterschiede jedoch nicht besonders groß sind, wurde dieser Aspekt vernachlässigt. So ergibt sich:

¹¹ $x \in \mathbb{N}; x \geq 1$

¹² vgl.: Krengel 1998:24

¹³ $P(A)$: Wahrscheinlichkeit eine einzelne Frage richtig zu beantworten; A : Frage wird richtig beantwortet; B : Es gibt zwei Antwortmöglichkeiten; \bar{B} : Es gibt drei Antwortmöglichkeiten

¹⁴ Alle Werte sind der oben angegebenen Tabelle entnommen.

¹⁵ In diesem Bereich gibt es Freitextfragen. Diese Fragen werden generell als falsch gewertet, die Wahrscheinlichkeit eine solche Frage richtig zu beantworten liegt also bei 0.

¹⁶ Bei der Führerscheiprüfung werden nicht die falschen Antworten gezählt. Jede Frage ist nach der Bedeutung für das Verhalten im Straßenverkehr gewichtet.

$$10 - 3,6x > 0$$

$$x < 2,77$$

Man darf also maximal zwei Fragen falsch bzw. muss 28 von 30 Fragen korrekt beantworten.

3.4 Programm zur Berechnung der Wahrscheinlichkeit bei der Führerscheinprüfung¹⁷

Für die theoretische Führerscheinprüfung als Spezialfall der erweiterten Bernoulli-Kette bietet das im Abschnitt 2.6 beschriebene Programm einige erweiterte Funktionen. So kann man z.B. unter „Bereiche“ diejenigen Bereiche (z.B. Gefahrenlehre, Vorfahrt etc) auswählen, die man nicht beherrscht – also durch Raten schaffen möchte. Das Programm errechnet dabei die Wahrscheinlichkeit und trägt diese automatisch im Feld „p“ ein. Auf Wunsch trägt das Programm auch die Anzahl der Fragen der ausgewählten Bereiche oder die erforderliche Anzahl an Fragen ein, um die Prüfung zu bestehen. Danach errechnet das Programm, wie groß die Wahrscheinlichkeit ist, die Prüfung unter den gegebenen Bedingungen zu bestehen.

3.5 Auswertung

Die Wahrscheinlichkeit, einen Bogen durch Raten zu bestehen, liegt bei $6,47 \cdot 10^{-20}$. Zum Vergleich liegt die Wahrscheinlichkeit, „sechs Richtige“ beim Lotto zu haben, bei nur einem Versuch schon bei $7 \cdot 10^{-8}$!

Das extrem niedrige Ergebnis lässt sich durch die geringe Wahrscheinlichkeit, eine Frage richtig zu beantworten, und die hohe Anzahl an Fragen erklären.

3.6 Kritische Reflexion

Allerdings ist die „echte“ Chance, einen Bogen durch Raten zu bestehen, wesentlich höher. Einmal gibt es einige Fragen, bei denen sich Antworten logisch ausschließen. Ein Beispiel dafür ist das abgebildete Schild.¹⁹



Grafik 1

Die Antwortmöglichkeiten „Ich darf nur nach rechts abbiegen“, „Ich darf nur nach links abbiegen“ und „Ich fahre geradeaus“ schließen sich logischerweise aus. Daher ist die Wahrscheinlichkeit, eine solche Frage richtig zu beantworten, wesentlich höher als ohne diesen Ausschluss.

Außerdem zeigt der Pfeil nach rechts, er ist nicht durchgestrichen, sodass es sich logisch erschließen lässt, dass man nach rechts fahren sollte.

Diese Problematik gibt es bei auch bei anderen Fragen, wie z.B. bei der Frage „Kinder stehen auf der Straße. Was tun Sie?“ Mit gesundem Menschenverstand lässt sich erkennen, dass man wohl bremsen sollte. Da ganze Bereiche wie zum Beispiel der Bereich „Gefahrenlehre“ fast nur aus solchen Fragen bestehen, könnte man diese Bereiche aus der Berechnung ausschließen. Wenn man dann das Raten auf die technischen Fragen (Bereich Zusatzfragen) beschränkt und davon ausgeht, dass man die anderen Fragen durch Lernen oder logische Herleitung richtig beantworten kann, liegt die Wahrscheinlichkeit den Test zu bestehen nur bei $2,2 \cdot 10^{-5}$.

Ob man erkennt, dass sich Antwortmöglichkeiten ausschließen, bzw. ob man eine Frage mit dem gesunden Menschenverstand erschließen kann, hängt allerdings von der jeweiligen Person ab. Es mag auch Personen geben, die anscheinend einfache Fragen falsch beantworten. Daher lässt sich eine Wahrscheinlichkeit unter Berücksichtigung der Möglichkeit, dass sich Antworten logisch ausschließen bzw. mit dem gesunden Menschenverstand richtig beantworten lassen, nicht allgemeingültig berechnen. Deswegen haben wir diesen Aspekt nicht weiter vertieft.

Eine weitere Einschränkung ist, dass wir von der durchschnittlichen Zahl an Fehlerpunkten ausgegangen sind. Man kann die Führerscheinprüfung aber noch bestehen, wenn man fünf Fragen mit nur zwei Fehlerpunkten falsch beantwortet.

¹⁷ vgl. Quelltext Anhang 6.10

¹⁸ vgl. http://www.ohlalala.de/fun/warum_lotto.php (Anhang 6.3)

¹⁹ Quelle: http://commons.wikimedia.org/wiki/Image:Zeichen_209.svg

4. Der optimale Test

4.1 Einleitung

Nach der Betrachtung der verschiedenen Beispiele sieht man deutlich, dass man bei einem geschickten Aufbau des Tests die Wahrscheinlichkeit, den Test durch Raten zu bestehen, wesentlich beeinflussen kann. Diese Wahrscheinlichkeit wird im Folgenden als Risikowahrscheinlichkeit für den Prüfer bezeichnet. Diese Wahrscheinlichkeit sollte möglichst klein sein.

4.2 Verteilung von Fragen und Antwort

4.2.1 Allgemeine Überlegungen

Natürlich hat man, wenn man einen Test durchführt, nur eine gewisse Zeit zur Verfügung. Daher sollte man die Zahl der jeweiligen Fragen und der jeweiligen Antwortmöglichkeiten pro Frage in diesem Zeitrahmen abwägen.

Dazu wird von einer Vereinfachung ausgegangen: Alle Fragen sind gleich aufgebaut, d.h. jede Frage hat gleich viele Antwortmöglichkeiten. Außerdem wird vereinfacht angenommen, dass man um die Prüfung zu bestehen alle Fragen korrekt beantworten muss.

Zuerst wurde die Zahl der Fragen und Antwortmöglichkeiten und die Zeit zueinander im folgenden Modell in Relation gebracht:

$$ax + bxz = t \quad (I)$$

wobei x = Anzahl Fragen; z = Anzahl Antwortmöglichkeiten; t = maximale Zeit in Minuten, die für den Test verwendet werden soll; a = durchschnittliche Zeit für das Lesen und Verstehen der Frage; b = durchschnittliche Zeit, die für die Entscheidung, ob die jeweilige Antwort richtig oder falsch ist, benötigt wird (jeweils in Minuten).

Um ein realistisches Ergebnis zu erhalten, müssen einige Einschränkungen beim Definitionsbereich vorgenommen werden.

Da man mindestens eine Frage stellen muss, muss x größer/gleich als eins sein. Jeder Frage muss aus mindestens zwei Antwortmöglichkeiten bestehen.

Natürlich sind keine halben Fragen und Antworten möglich, diese Einschränkung wird jedoch erst nach der Berechnung der Nullstellen vorgenommen, da die Funktion sonst nicht stetig und damit nicht differenzierbar wäre.

a und b müssen größer als 0,5 sein, da man für das Lesen und Verstehen der Fragen sich mindestens eine halbe Minute Zeit lassen sollte. Auch hier wurde der Zahlenbereich \mathbb{R} angenommen.

Außerdem ist es notwendig, a und b im Verhältnis zu t zu beschränken.

Es ist nicht sinnvoll, einen optimalen Test zu erstellen, der dennoch so schlecht ist, dass er nach unserer Einteilung²⁰ mit „unbrauchbar“ bewertet wird. Das Ziel ist es also, a und b so zu wählen, dass keine unbrauchbaren Tests entstehen.

Ein Test, der nicht mit „unbrauchbar“ bewertet wird, muss bei einer Frage vier Antwortmöglichkeiten aufweisen.

Geht man vereinfacht davon aus, dass der obere Rand des Definitionsbereichs bei a und b gleich sein soll, so ergibt sich über Gleichung I folgender Zusammenhang:

$$a_{\max} = b_{\max}$$

$$a_{\max} + 4 \cdot a_{\max} = t \Leftrightarrow a_{\max} = \frac{t}{5} = b_{\max}$$

Wir gehen davon aus, dass a und b maximal den Wert $\frac{t}{5}$ annehmen dürfen. Es sind natürlich auch

Konstellationen möglich, bei denen beispielsweise a größer als $\frac{t}{5}$ ist und b dafür wesentlich kleiner. Bei einer solchen Konstellation erhält man allerdings kein sinnvolles Ergebnis, sodass die erarbeitete Einschränkung sinnvoll ist.

²⁰ Ein unbrauchbarer Test ist ein Test, bei dem die Wahrscheinlichkeit ihn zufällig richtig zu lösen bei $p > \frac{1}{4}$ liegt, vgl. Skala 4.4

Geht man nun von x Fragen aus, so kann x nach Gleichung I maximal den Wert $x = \frac{t}{a+2b}$ annehmen, da ansonsten z außerhalb des Definitionsbereichs liegen würde.

4.2.2 Berechnung der optimalen Anzahl von Fragen und Antworten

Für die Berechnung der Wahrscheinlichkeit, alle Fragen richtig zu beantworten, gilt folgende

Formel: $\frac{1}{z^x} = p(x)$ (II) wobei p die Wahrscheinlichkeit ist, alle Fragen richtig zu beantworten. Es wird hier außerdem davon ausgegangen, dass immer genau eine Antwortmöglichkeit richtig ist. Das Ziel sollte sein, x und z unter Berücksichtigung von I so zu wählen, dass p minimal ist. Es handelt sich also um die Berechnung eines Minimums.

Um mit der Gleichung (II) rechnen zu können, muss man sie zu einer Gleichung mit nur einer Unbekannten umformen. Das erreicht man durch das Einsetzen von I in II. Man erhält:

$$p(x) = \frac{1}{\left(\frac{t-ax}{bx}\right)^x} = e^{-x \cdot \ln\left(\frac{t-ax}{bx}\right)} \quad 21$$

Zur Berechnung des Minimums muss die Funktion abgeleitet werden. Man erhält:

$$p'(x) = e^{-x \cdot \ln\left(\frac{t-ax}{bx}\right)} \cdot \left(-\ln\left(\frac{t-ax}{bx}\right) + \frac{t}{t-ax}\right)$$

Die notwendige Bedingung für ein Minimum ist $p'(x) = 0$.

Da es keine analytische Lösung des Terms gibt, wird für die Lösung das Newton-Verfahren²² verwendet. Dieses Iterationsverfahren findet jedoch nur eine Nullstelle, auch wenn mehrere Nullstellen vorhanden sind. Daher muss gezeigt werden, dass es sich bei der gefundenen Lösung um das globale Minimum handelt.

Dafür wird zuerst über das Monotonieverhalten und die Werte an den Rändern des Definitionsbereichs gezeigt, dass es genau einen Punkt mit waagerechter Tangente gibt. Anschließend wird gezeigt, dass es sich bei dem Punkt nur um einen Tiefpunkt handeln kann.

Die schon berechnete Ableitung wird zur Bestimmung des Monotonieverhaltens verwendet.

Die für ein Minimum notwendige Bedingung lässt sich folgendermaßen umformen.

$$\begin{aligned} p'(x) &= e^{-x \cdot \ln\left(\frac{t-ax}{bx}\right)} \cdot \left(-\ln\left(\frac{t-ax}{bx}\right) + \frac{t}{t-ax}\right) = 0 \\ \Leftrightarrow \left(-\ln\left(\frac{t-ax}{bx}\right) + \frac{t}{t-ax}\right) &= 0 \quad 23 \\ \Leftrightarrow \frac{t}{t-ax} &= \ln\left(\frac{t-ax}{bx}\right) \end{aligned}$$

Wenn die Ableitung genau eine Lösung hat, dann gibt es genau einen Schnittpunkt der Funktionen

$h(x) = \frac{t}{t-ax}$ und $g(x) = \ln\left(\frac{t-ax}{bx}\right)$. Das soll im Folgenden mithilfe der Monotonie gezeigt werden.

²¹ $a, b > 0,5; z \geq 2; x \geq 1; a, b, t, x, z \in \mathbb{R}; a, b \geq \frac{t}{5}; x < \frac{t}{a+2b}; t > 5$

²² Das Newtonverfahren ist eine Möglichkeit über Iteration eine Nullstelle anzunähern, vgl.: Manfred Baum u.a. 2000:124ff, genauere Details zur Berechnung mit dem Newtonverfahren finden sich in Anhang 6.6

²³ da $e^{-x \cdot \ln\left(\frac{t-ax}{b}\right)} \neq 0$

$$g'(x) = \frac{bx}{t-ax} \cdot \left(-\frac{t}{bx^2}\right) \Rightarrow g'(x) < 0 \Rightarrow \text{Graph ist streng monoton fallend}^{24}$$

$$h'(x) = \frac{at}{(t-ax)^2} \Rightarrow h'(x) > 0 \Rightarrow \text{Graph ist streng monoton steigend}$$

Daher können sich die Graphen nur genau einmal schneiden.

Da $g(x)$ streng monoton fallend und $h(x)$ streng monoton steigend ist, folgt daraus, dass $g(x)$ einen höheren Wert als $h(x)$ am linken Rand des Definitionsbereich (d.h. $x = 1$) haben muss, wenn sich $g(x)$ und $h(x)$ schneiden.

Da a und b maximal den Wert $\frac{t}{5}$ annehmen können, ist diese Voraussetzung für maximale Werte von a und b erfüllt:

$$\left. \begin{aligned} h(1) &= \frac{t}{t - \frac{t}{5}} = \frac{5}{4} \\ g(1) &= \ln\left(\frac{t - \frac{t}{5}}{\frac{t}{5}}\right) = \ln 4 \approx 1,39 \end{aligned} \right\} g(1) > h(1)$$

Auch für kleinere Werte für a und b ist diese Ungleichung erfüllt, da $h(1)$ in diesem Falle immer kleiner würde während $g(1)$ immer größer würde. Dies gilt auch für Veränderungen beim Verhältnis von a und b , da in $h(x)$ nur a eine Rolle spielt und in $g(x)$ beide Parameter unabhängig voneinander für eine Vergrößerung des Wertes sorgen.

Es kann natürlich theoretisch auch passieren, dass der Schnittpunkt der Graphen nicht im Definitionsbereich liegt. Das ist bei unserer Funktion jedoch nicht der Fall, da $g(x)$ am rechten Rand des Definitionsbereichs $\left(x = \frac{t}{a+2b}\right)$ einen kleineren Wert als $h(x)$ hat und sich die Graphen daher schon geschnitten haben müssen:

$$g\left(\frac{t}{a+2b}\right) = \ln\left(\frac{t - \frac{at}{a+2b}}{\frac{bt}{a+2b}}\right) = \ln\left(\frac{a+2b}{b} - \frac{a(a+2b)}{b(a+2b)}\right) = \ln(2) < 1$$

$$h\left(\frac{t}{a+2b}\right) = \frac{t}{t - \frac{at}{a+2b}} = \frac{1}{\frac{a+2b-a}{a+2b}} = 1 + \frac{2b}{a} > 1$$

Es ist offensichtlich, dass diese Bedingung für alle Werte erfüllt ist.

Man kann daher sagen, dass die Funktion $p(x)$ genau eine Stelle besitzt, an der die notwendige Bedingung für ein Minimum erfüllt ist.

Abschließend muss gezeigt werden, dass es sich bei der gefundenen Extremstelle um ein Minimum handelt.

Zuerst wird ausgeschlossen, dass es sich um ein Maximum handelt. Das geht über das Verhalten des Graphen am linken Rand des Definitionsbereichs:

²⁴ Definition des Monotonieverhaltens, vgl.: Bronstein u.a. 1969:271

$$p'(x) = e^{-x \cdot \ln\left(\frac{t-ax}{bx}\right)} \cdot \left(-\ln\left(\frac{t-ax}{bx}\right) + \frac{t}{t-ax} \right)$$

$$p'(1) = e^{-\ln\left(\frac{t-a}{b}\right)} \cdot \left(-\ln\left(\frac{t-a}{b}\right) + \frac{t}{t-a} \right) \rightarrow p'(1) < 0 \rightarrow$$

negative Ableitung am linken Rand,
daher kann der Extremwert kein
Maximum sein.

Nun muss noch ausgeschlossen werden, dass es ein Sattelpunkt ist.²⁵

Dies lässt sich über einen Widerspruch bei der Monotonie der Funktion $p(x)$ zeigen.

Wir verwenden erneut die schon berechnete erste Ableitung von $p(x)$:

$$p'(x) = e^{-x \cdot \ln\left(\frac{t-ax}{bx}\right)} \cdot \left(-\ln\left(\frac{t-ax}{bx}\right) + \frac{t}{t-ax} \right)$$

Für die Betrachtung der Monotonie genügt es, die einzelnen Faktoren separat zu betrachten, da nur das Vorzeichen der Ableitung entscheidend ist.

Der erste Faktor, $e^{-x \cdot \ln\left(\frac{t-ax}{bx}\right)}$, ist aufgrund einer Eigenschaft der e-Funktion immer positiv. Beim zweiten Faktor ist das Vorzeichen nicht eindeutig.

An der Stelle $x = 1$ ist $\ln\left(\frac{t-ax}{bx}\right)$ wegen der Definition von a und b größer als $\frac{t}{t-ax}$, die Ableitung ist

also negativ. Nun ist $\ln\left(\frac{t-ax}{bx}\right)$ am rechten Rand des Definitionsbereichs wie bereits gezeigt kleiner

als $\frac{t}{t-ax}$. Ab diesem Zeitpunkt ist die Ableitung positiv.

Man kann also sagen, dass die Funktion nicht monoton ist. Wenn die Stelle x , an der der Graph $p(x)$ eine wagerechte Tangente hat, ein Sattelpunkt wäre, müsste die Funktion jedoch monoton sein.

Es konnte also gezeigt werden, dass $p(x)$ im Rahmen des Definitionsbereichs genau einen Tiefpunkt besitzt.

4.3 Programm zur Erstellung eines optimalen Tests²⁶

Dieses Programm hilft dabei, einen optimalen Ankreuzbogen zu entwerfen. Nachdem die benötigten Parameter (a , b und t) eingetragen wurden, schlägt das Programm einen Test vor, der aus x Fragen und z Antwortmöglichkeiten je Frage besteht und mit einer Wahrscheinlichkeit von p durch Raten richtig gelöst werden kann. Dabei verwendet das Programm folgende Zusammenhänge:

$$1) \quad p(x) = e^{-x \cdot \ln\left(\frac{t-ax}{bx}\right)}$$

$$2) \quad p(x) = \frac{1}{z^x}$$

$$3) \quad t = ax + bxz$$

Da das Programm zunächst die Werte für x , z und p errechnet und erst danach prüft, ob diese den Definitionsbereich aus Abschnitt 4.3.1 erfüllen, gilt neben den Bedingungen für die Parameter a , b und t aus Abschnitt 4.3.1 folgende Bedingung während der Rechnung:

$\frac{t-ax}{bx} > 0 \Leftrightarrow x < \frac{t}{a}$, da ansonsten der natürliche Logarithmus aus einer Zahl kleiner/gleich Null gebildet wird

²⁵ Ein Sattelpunkt ist ein Punkt mit waagerechter Tangente, bei dem sich jedoch das Vorzeichen der Ableitung nicht ändert. Die Funktion ist also monoton steigend bzw. fallend. Einen Sattelpunkt findet man beispielsweise bei der Funktion $f(x) = x^3$ an der Stelle $x = 0$.

²⁶ vgl. Quelltext Anhang 6.10

Damit die Wahrscheinlichkeit p möglichst klein ist, muss die Ableitung von p – wie bereits erläutert – gleich Null sein. Aus dieser Bedingung wird zunächst die Anzahl der Fragen ausgerechnet.

Dies geschieht über das sog. „Newton-Verfahren“.²⁷

Das Programm fängt bei $x = 1$ an und nähert sich darüber immer mehr der Nullstelle an. Sollte sich ein Zwischenwert für eine Nullstelle außerhalb des Definitionsbereiches D befinden, so wird x auf den Wert „ $D - 0,1$ “ gesetzt und die Nullstelle mit diesem Wert weiter angenähert.

Nachdem die Anzahl der Fragen ausgerechnet wurde, wird über den Zeitzusammenhang $t = ax + bxz$ die

entsprechende Anzahl der Antwortmöglichkeiten (z) und $p(x) = e^{-x \cdot \ln(\frac{t-ax}{bx})}$ die Wahrscheinlichkeit ausgerechnet.

Falls der Test mehr als fünf Antwortmöglichkeiten pro Frage haben sollte, bietet das Programm die Möglichkeit, einen optimierten Gegenvorschlag zu machen. Dabei wird die Anzahl der Antwortmöglichkeiten auf fünf gesetzt, danach mittels $t = ax + bxz$ die Anzahl der Fragen (x) und wie oben die Wahrscheinlichkeit p für den Gegenvorschlag ausgerechnet. Bei ungünstig gewählten Parametern kann es allerdings vorkommen, dass auch der Gegenvorschlag unrealistisch bzw. schlecht ist. Damit der Gegenvorschlag überhaupt verwendet werden kann, werden hier daher noch einige zusätzliche Konstellationen von x und z ausgeschlossen, die theoretisch noch im Definitionsbereich liegen. Es handelt sich um Tests, bei denen die Wahrscheinlichkeit p über $1/4$ liegt. In einem solchen Falle wird der Benutzer darauf hingewiesen, dass kein realistischer Gegenvorschlag möglich ist und aufgefordert, die Parameter a , b und t sinnvoller zu wählen.

Neben dem Entwerfen eines Tests kann das Programm auch einen bereits fertigen Test überprüfen. Allerdings können hier nicht die verschiedenen Parameter a , b und t berücksichtigt werden, da diese nur schwer zu ermitteln sind. Zusätzlich zur Wahrscheinlichkeit p stuft das Programm den Test noch auf einer Skala von sehr gut bis unbrauchbar ein:

$$\begin{aligned}
 p &\leq \frac{1}{100000} : \text{sehr gut} \\
 \frac{1}{10000} &\geq p \geq \frac{1}{100000} : \text{gut} \\
 \frac{1}{1000} &\geq p \geq \frac{1}{10000} : \text{akzeptabel} \\
 \frac{1}{100} &\geq p \geq \frac{1}{1000} : \text{noch akzeptabel} \\
 \frac{1}{10} &\geq p \geq \frac{1}{100} : \text{schlecht} \\
 p &\geq \frac{1}{10} : \text{sehr schlecht} \\
 p &> \frac{1}{4} : \text{unbrauchbar}
 \end{aligned}$$

Ferner bietet das Programm Tools bzw. Informationsfunktionen für die Benotung des vorgeschlagenen oder überprüften Tests an. Dabei kann man sich anzeigen lassen, wie wahrscheinlich es ist, x Fragen (und aufwärts) beim Test durch Raten richtig zu beantworten. Außerdem kann man die – nach eigenen Vorstellungen – erforderliche Anzahl an richtig beantworteten Fragen eintragen, die man für die verschiedenen (Schul)Noten („sehr gut“ bis „ungenügend“) haben muss. Das Programm zeigt dann den prozentualen Anteil der richtig zu beantwortenden Fragen am gesamten Test (für die Vergleichbarkeit mit den Richtlinien) und die Wahrscheinlichkeit, diese Note durch Raten zu erreichen, an.

Auf Wunsch kann dieses Programm auch einige Kontrollergebnisse anzeigen, es handelt sich meistens um die nicht gerundeten (Zwischen)Ergebnisse. Dazu aktiviert man den Debug-Modus.

²⁷ vgl. Anhang 6.6

4.4 Bewertung eines realen Schultests

Bei diesem Schultest, der im Fach Sozialwissenschaften geschrieben wurde²⁸, gab es elf Fragen. Jede Frage hatte vier Antwortmöglichkeiten, von denen entweder eine, zwei, drei oder vier Antworten richtig sein können.

Um die Wahrscheinlichkeit für eine Frage zu berechnen, muss man teilweise analog zum Führerschein vorgehen, da es erneut mehrere richtige Antwortmöglichkeiten geben kann. Es wird die schon hergeleitete

Formel $p = \frac{1}{2^{\text{Antwortmöglichkeiten}} - 1}$ verwendet. Die Wahrscheinlichkeit, eine Frage richtig zu beantworten

beträgt also $p = \frac{1}{2^4 - 1} = \frac{1}{15}$. Mit diesem Wert kann man die Wahrscheinlichkeit berechnen, diesen Test zu

bestehen, dass heißt mehr als die Hälfte aller Fragen richtig zu beantworten.

Diese Wahrscheinlichkeit beträgt $P_{\text{ges}} = \sum_{k=6}^{11} P_{6/11} = \frac{11!}{k!(11-k)!} \cdot \frac{1}{15^k} \cdot \left(1 - \frac{1}{15}\right)^{11-k} = 3,2 \cdot 10^{-5}$ und ist

folglich sehr gering.

Man kann also sagen, dass dieser Test sehr gut erstellt wurde, da alle wichtigen Faktoren berücksichtigt wurden. Das Bewertungsprogramm bewertet den Test mit dem Prädikat „sehr gut“.

4.5 Kritische Reflexion

Natürlich muss man auch hier bedenken, dass sich die Theorie nicht genau in die Praxis umsetzen lässt. So ist die Wahrscheinlichkeit einen Test durch Raten zu bestehen, bei unendlich vielen Antwortmöglichkeiten theoretisch unendlich niedrig. Allerdings ist es eher unwahrscheinlich, dass es unendlich viele „intelligente“²⁹ Antwortmöglichkeiten gibt. Wenn man Antwortmöglichkeiten nur ergänzt um die Risikowahrscheinlichkeit zu verringern, so wird dieses Ziel dadurch kaum erreicht, da man –selbst wenn man sich mit dem Thema nicht gut auskennt– diese Antwortmöglichkeiten logisch ausschließen kann und dadurch die Risikowahrscheinlichkeit auf einen geringeren Wert setzt. Daher sollte man, bevor man einen Test so gestaltet, dass die Risikowahrscheinlichkeit möglichst gering ist, auch überprüfen ob die Antworten realistisch sind.

Auch die Ergebnisse, die man für einen sehr hohen Wert für t und gleichzeitig niedrigem Wert für a und b erhält sind nicht zu übernehmen, da die hohe Anzahl an Antwortmöglichkeiten den Rahmen jedes sinnvollen Multiple-Choice-Test sprengen würde.

5. Erweiterungen

5.1 Berechnung der Wahrscheinlichkeit pro Frage

5.1.1 Idee

Bei unserem eingereichten Programm wird vorausgesetzt, dass die Person, die das Programm verwendet weiß, wie man die Wahrscheinlichkeit pro Frage aus einer vorgegebenen Zahl an Antwortmöglichkeiten berechnet. Da die Formel für die Wahrscheinlichkeit pro Frage nicht unbedingt bekannt ist, ist dies aber nicht vorraussetzbar. Daher wurde eine Hilfestellung ergänzt.

5.1.2 Umsetzung

Der Benutzer teilt dem Programm mit, wie viele Antwortmöglichkeiten jede Frage hat, danach berechnet das Programm die Wahrscheinlichkeit und trägt diese ein.

5.2 Berechnung der Wahrscheinlichkeit unter Berücksichtigung von Minuspunkten

5.2.1 Idee

Bei der Berechnung von Tests, die z.B. an Universitäten geschrieben werden, ist auffällig, dass es meistens für falsche Antworten Minuspunkte gibt, damit die getesteten Personen nicht versuchen, den Test durch Raten zu bestehen. Bei unseren Berechnungen wird generell davon ausgegangen, dass die Person den Test vollständig rät. Dennoch ist es interessant zu sehen, welchen Einfluss Minuspunkte auf die Wahrscheinlichkeit, den Test durch Raten zu bestehen haben.

²⁸ Zusammenschnitt von Fragen aus: Pieper, Uli u.a. 2006: S. 87-89

²⁹ Mit „intelligent“ ist in diesem Falle gemeint, dass die Antwortmöglichkeiten realistisch sind und man diese nicht ausschließen kann, weil sie zu unrealistisch erscheinen.

5.2.2 Mathematische Grundlagen

Um zu berechnen, wie wahrscheinlich es ist einen Test trotz Minuspunkten zu bestehen, muss zuerst berechnet werden, wie sich die Zahl der richtig zu beantworteten Fragen verändert. Es ist sinnvoll, im folgenden nicht mehr von richtig beantworteten Fragen sondern von erzielten Punkten zu sprechen, da man durch die Minuspunkte mehr Fragen richtig beantworten muss, um beispielsweise die Hälfte der Punkte zu erhalten.

Um die Berechnung zu verdeutlichen, die man durchführen muss um die Mindestanzahl an richtigen Fragen zu erhalten, die man benötigt um eine bestimmte Anzahl an Punkten zu erhalten, betrachten wir ein Beispiel:

In diesem Fall sollen mit zehn Fragen mindestens fünf Punkte erzielt werden. Für jede richtig beantwortete Frage erhält man einen Punkt, für jede falsch beantwortete Frage wird ein Punkt abgezogen. Ohne Minuspunkte dürfte man fünf Fragen falsch beantworten. Diese fünf Fragen müssen jetzt so aufgeteilt werden, dass man über diese Fragen gleich viele Punkte erhält und verliert; es muss also ein Gleichgewicht zwischen richtig und falsch beantworteten Fragen herrschen, damit diese Fragen keinen resultierenden Einfluss haben.

Dieses Ziel erreicht man, in dem man die Anzahl an –nach der klassischen Betrachtung– falsch beantworteten Fragen durch zwei teilt. Die mit diesem Prinzip errechneten Fragen müssen noch zu den schon richtig zu beantworteten Fragen addiert werden.

Es gilt also:

$$l = \left\lceil i + \frac{n-i}{2} \right\rceil$$

(wobei i = Mindestanzahl an Punkten; n = Anzahl an Fragen; l = neue Mindestanzahl an Fragen).

Zu beachten ist, dass l immer auf die nächste ganze Zahl gerundet werden muss, da man ansonsten unter der Mindestanzahl an Punkten landet.

5.2.3 Umsetzung

Der Benutzer kann die Option aktivieren, dass Minuspunkte vergeben werden. Dies berücksichtigt das Programm dann bei der Berechnung der Wahrscheinlichkeit und der Anzahl an richtig zu beantwortenden Fragen.

5.3 Berechnung eines optimalen Tests mit der Bernoulli-Kette

5.3.1 Idee

Unsere optimalen Tests werden momentan immer unter der Bedingung bestimmt, dass man alle Fragen richtig beantwortet. Diese Annahme wurde von uns gewählt, da die in diesem Falle zu verwendende Formel wesentlich einfacher als die Bernoulli-Kette ist. Durch Gespräche wurden uns jedoch die Vorteile deutlich, die die Berechnung mit der Bernoulli-Kette haben würde: Das Ergebnis wäre wesentlich realistischer, da man in (fast) keinem Test alle Fragen richtig beantworten muss.

5.3.2 Probleme

Es ist natürlich möglich, dass die Minima unserer Funktion und der vergleichbaren Bernoulli-Kette

$$b(x) = \frac{x!}{k!(x-k)!} \cdot \left(\frac{bx}{t-ax}\right)^x \cdot \left(1 - \frac{bx}{t-ax}\right)^{x-k} \quad \text{bei gleichen Parametern den gleichen Wert}$$

annehmen. Dies lässt sich aber nicht zeigen, da das Minimum unserer Funktion sich bei einer richtigen Antwortmöglichkeit nur annähern lässt.

Daher muss das Minimum der Bernoulli-Kette bestimmt werden. Auch dies ist numerisch nicht möglich, da die Definition der Fakultät sich im klassischen Sinne nur auf ganzzahlige x bzw. k bezieht und bei einer ganzzahligen Reihe keine Stetigkeit gegeben ist. Eine Funktion, die nicht stetig ist, lässt sich nicht differenzieren.

Es gibt zwei uns bekannte Möglichkeiten Fakultäten abzuleiten: Man kann diese exakt mit Hilfe der Gamma-Funktion³⁰ berechnen oder mithilfe der so genannten Stirling-Formel³¹ annähern. Die Gamma-Funktionen sind leider nur sehr schlecht zu handhaben und aufgrund der Tatsache, dass die Funktion sowieso schon schlecht abzuleiten ist –es muss die Produkt-, die Quotienten-, und die Kettenregel verwendet werden- und diese Klasse von Funktionen mit unseren Vorkenntnissen nicht zu verwenden ist, hatten wir uns zuerst für die nicht ganz so genaue Annäherung, die Stirling-Formel entschieden. Leider ist auch die Stirlingformel nicht zu verwenden: Die Ableitungen werden so kompliziert, dass es unmöglich ist, eine Aussage über die Funktion zu machen. Daher konnten wir hier nicht mit der Ableitung arbeiten.

5.3.3 Lösungsansatz

Da wie bereits beschrieben weder die Gamma-Funktionen noch die Stirling-Formel anwendbar ist, berechnen wir die Anzahl der Fragen über eine Minimum-Suche. Dabei werden alle möglichen Werte im Definitionsbereich ausgerechnet und danach der Wert (also die Anzahl an Fragen) mit der geringsten Wahrscheinlichkeit herausgesucht.

5.3.4 Umsetzung

Der Benutzer gibt die erforderlichen Angaben (a,b,t sowie den Anteil der richtig zu beantwortenden Fragen) an. Das Programm bestimmt daraus den Definitionsbereich, berechnet die Wahrscheinlichkeiten der darin enthaltenen Werte und sucht sich anschließend das Minimum heraus. Da x und k nur ganzzahlige Werte annehmen dürfen, werden diese Werte mathematisch gerundet.

6. Literaturverzeichnis

6.1 Printmedien

1. Baum, Manfred u.a.: LS 11, Ernest Klett Verlag, Düsseldorf
2. Blatter, Christian: Analysis 1, Springer Verlag, Berlin
3. Bronstein, Ilja u.a.: Taschenbuch der Mathematik, Verlag Harri Deutsch, Frankfurt
4. Kreyszig, Erwin: Statistische Methoden und ihre Anwendung, Vandenhoeck & Ruprecht, Göttingen
5. Pieper, Uli u.a.: Deutschland in der Wachstumskrise, Ernst Klett Schulbuchverlag, Leipzig
6. Renyi, Alfréd: Wahrscheinlichkeitsrechnung, Deutscher Verlag der Wissenschaft, Berlin
7. Fahren Lernen, Prüfbögen, Verlag Heinrich Vogel, München

6.2 Elektronische Medien

1. <http://www.fahrschule.de/Verkehrsrecht/Pruefungsrichtlinie/> (30.11.07, 16:14 Uhr)
2. <http://www.physik-schule.de/download/pdf/Mathematik/Bernoulli-Kette.pdf> (30.11.07, 16:15 Uhr)
3. http://www.ohlalala.de/fun/warum_lotto.php (01.12.07, 19:21 Uhr)
4. <http://www.amiright.com/parody/60s/thebeatles280.shtml> (01.12.07, 16:18 Uhr)
5. astragon Europa Fahrschule 2006, astragon Software GmbH, Mönchengladbach

6.3 Bildquellen

1. http://commons.wikimedia.org/wiki/Image:Zeichen_209.svg (30.11.07, 16:11 Uhr)

³⁰ Bronstein, Ilja u.a.: Taschenbuch der Mathematik, Verlag Harri Deutsch, Frankfurt

³¹ Bronstein, Ilja u.a.: Taschenbuch der Mathematik, Verlag Harri Deutsch, Frankfurt