Berechnung von Binomialwahrscheinlichkeiten mit Hilfe des Computers

von Helmut Wirths, Oldenburg

Zusammenfassung

Es werden Probleme diskutiert, die bei der Berechnung von Bereichwahrscheinlichkeiten im Modell der Binomialverteilung auftreten, wenn die betrachteten Bernoulli-Ketten besonders lang sind. Berichtet wird über Erfahrungen mit der expliziten und der rekursiven Gleichung zur Berechnung von Einzelwahrscheinlichkeiten und von Bereichswahrscheinlichkeiten beim Einsatz unterschiedlicher Hard- und Software. Ausführlich wird ein Algorithmus erläutert, der über die Grenzen der expliziten wie der rekursiven Gleichung hinaus Berechnungen ermöglicht. Den Abschluß bilden didaktische Reflexionen zum Computereinsatz im Stochastikunterricht, vor allem unter der didaktischen Vorgabe, motivierende und praxisnahe Fragestellungen bereits im Modell der Binomialverteilung zu behandeln.

1. Eine erste Problemstellung

In einem Grundkurs habe ich unmittelbar nach dem Erarbeiten einiger Eigenschaften der Binomialverteilung (explizite Darstellung, Erwartungswert, Histogramm) folgenden Sachverhalt (nach Kroll (1985), S. 60) vorgestellt : „Beim Bau eines Hauses werden 3000 Steine zum Verklinkern benötigt. Erfahrungsgemäß sind 1 % der Steine nicht geeignet." Im folgenden Unterrichtsgespräch entwickelte sich als Leitfrage : „Können wir einen mathematisch begründeten Rat geben, wie viele Steine bestellt werden sollen ?"

1.1 Zwei Modelle

Wie fest deterministisches Denken in den Köpfen meiner Schülerinnnen und Schüler noch verankert war, wurde mir klar, als folgende Ansätze genannt wurden : Wir suchen die Lösung von $0{,}99 \cdot n = 3000$ oder die kleinste natürliche Zahl n, die die Ungleichung $0{,}99 \cdot n \geq 3000$ löst. Als Lösung der Ungleichung wurde $n \geq 3031$ mitgenannt. Anstatt diesen Ansatz gleich zu verdammen und abzuwerten, sollte man ruhig etwas warten und darauf vertrauen, daß sich im Unterrichtsgespräch langsam eine andere Interpretation herausschält. Sind beide Interpretationen ausdiskutiert, können die Lernenden selbständig argumentieren, vergleichen und werten.

Vielleicht haben Sie auch Lernende, die nach anfänglichem Schweigen nun anfangen zu reden : „Ich stelle mir diesen großen Berg Steine vor. Ich nehme den ersten Stein in die Hand und muß mich entscheiden : Ist der Stein geeignet oder nicht ? Und so geht es Stein für Stein weiter." Schnell greifen andere diesen Impuls auf, es wird lebhaft weiter diskutiert. Nach diesen ersten Worten liegt ein Baumdiagramm zur Darstellung nahe. Aber 2^{3000} Wegstücke bei 3000 Steinen in der letzten Stufe des Baumdiagramms, eine solche Größenordnung erschreckt. Wenn wir mehr Steine bestellen, müssen wir mit entsprechend mehr Wegstücken rechnen, die zugehörigen Anzahlen wachsen rasant. Wenn wir uns die Größenordnung von 2^{3000} auch nicht mehr vorstellen können, denken können wir uns das Baumdiagramm schon. In jeder Stufe geht es um 2 Möglichkeiten : geeignet - nicht geeignet. Damit auch eine Bernoulli-Kette vorliegt, fehlt noch eine Wahrscheinlichkeit p, die in jeder Stufe gleich bleibt. Das Wort „erfahrungsgemäß" habe ich ganz bewußt in die Sachbeschreibung mit eingebaut. Es gibt Anlaß, darüber zu sprechen, wie Erfahrung zustande kommen kann. Man sollte sich an die ersten Stunden in Stochastik erinnern, als es um den Zusammenhang zwischen relativer Häufigkeit und Wahrscheinlichkeit ging.

Am Ende schält sich folgende stochastische Modellierung heraus : Es sollen n Steine ($n > 3000$) bestellt werden. Die Wahrscheinlichkeit, einen geeigneten Stein zu erhalten, beträgt für jeden Stein immer 0,99.

Stochastik in der Schule 18(1998), Heft 1, S. 43 - 54

In dieser Modellierung wird schnell deutlich, daß wir bei jeder Bestellmenge genügend Steine oder sogar mehr als wir brauchen bekommen können, aber auch das Risiko eingehen, nicht genug brauchbare Steine zu erhalten. Dieses Risiko besteht immer. Wir können es zwar dadurch kleiner machen, daß wir mehr Steine bestellen, es ganz eliminieren können wir nicht. Wesentlich zur Charakterisierung im stochastischen Modell ist also die Wahrscheinlichkeit für die Sicherheit unserer Entscheidung, die wir Sicherheitswahrscheinlichkeit nennen wollen.

Interessante Nebenfragen, die Lernende immer wieder stellen, bereichern den Unterricht : Können wir 2^{3000} als Zahl überhaupt konkret hinschreiben ? Wie viele Ziffern hat die Zahl ? Welches ist die letzte oder die vorletzte Ziffer ?

Typische Unterschiede der beiden Lösungsmodelle seien kurz zusammengestellt :

a. Im ersten, dem deterministischen Modell gilt :
   Von den gelieferten Steinen sind **immer** 1 % unbrauchbar. Kauft man 3031 Steine ein, dann hat man **immer** 3000 brauchbare.
b. Im zweitem, dem stochastischen Modell gilt :
   Von den gelieferten Steinen sind **im Mittel** 1 % unbrauchbar. Oder : Für jeden gekauften Stein beträgt **die Wahrscheinlichkeit**, unbrauchbar zu sein, 1 %. Kauft man 3031 Steine ein, dann hat man **im Mittel** 3000 brauchbare. Im Mittel bedeutet aber auch, daß es mal weniger als 3000 brauchbare Steine, ein andermal aber auch mehr sein können. Daher ist es in diesem Modell wichtig, die Wahrscheinlichkeit zu kennen, beim Einkauf von n Steinen genügend brauchbare, also mindestens 3000 zu bekommen. Diese Sicherheitswahrscheinlichkeit informiert uns, wie häufig wir bei sehr vielen Käufen unter den gegebenen Bedingungen erwarten dürfen, genügend Steine einzukaufen.

1.3 Weitere Überlegungen

Es dürfte einleuchten, daß sich Lernende auch für die Sicherheitswahrscheinlichkeit für den Einkauf von 3031 Steinen, der Lösung im deterministischen Modell, interessieren. Es gilt : $P(X \geq 3000) \approx 0{,}6$ mit $n = 3031$ und $p = 0{,}99$. In der Regel wird man diese Sicherheitswahrscheinlichkeit für zu klein halten und mehr Steine einkaufen. Auch wenn Steine im Baustoffhandel meist nur in Vielfachen von 50 oder sogar von 100 Stück verkauft werden, es finden sich immer wieder Lernende, die sich selbständig eine konkrete Sicherheitswahrscheinlichkeit vorgeben und genau die Anzahl an Steinen wissen wollen, die sie bestellen müssen, damit diese Sicherheitswahrscheinlichkeit mindestens eingehalten wird. Diese typische stochastische Fragestellung und die damit verbundene wünschenswerte Aktivität möchte ich nicht mit Hinweis auf die Usancen im Baustoffhandel unterbinden oder abwerten. Wichtig erscheint mir, daß immer darüber diskutiert wird, daß Modellbildung unter anderem auch Idealisierung oder Einengung beinhaltet, und darüber nachzudenken, worin dies in den konkreten Modellen tatsächlich besteht.

Der hier kurz angerissene Unterrichtsgang mündet in eine Phase, in der das schnelle Berechnen von Bereichswahrscheinlichkeiten wichtig ist. In den folgenden Abschnitten möchte ich über die Planung einer Unterrichtseinheit berichten, die mit der hier vorgestellten Einstiegsaufgabe beginnt. Die wesentlichen Aufgabentypen dieser Einheit werden in Wirths (1996) beschrieben und analysiert. Ich möchte darstellen, welche Probleme ich im Hinblick auf die derzeit an meiner Schule vorhandene Hard- und Software beim Einsatz eines Computers oder eines Graphik-Taschenrechners erwarten muß. Ich

selber setze einen Pentium-75-Rechner mit 8 MB Speicher und Windows 3.11 ein. Mein Computer zählt zwar nicht mehr zu den schnellsten, ist aber immer noch schneller als viele der derzeit in Schulen noch anzutreffenden Rechner.

2. Einige Erfahrungen mit der expliziten Darstellung zur Berechnung von Binomialwahrscheinlichkeiten

2.1 Die explizite Darstellung

Die Gleichung zur expliziten Berechnung der Wahrscheinlichkeiten B(n,p,k) lautet :

$B(n, p, k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$. n bedeutet hier die Anzahl aller Versuche, p die in jeder Stufe gleiche Wahrscheinlichkeit, daß das Ergebnis (z. B. „Der Stein ist brauchbar") eintritt, q die zugehörige Gegenwahrscheinlicheit (im Beispiel „Der Stein ist unbrauchbar") mit $q = 1 - p$, während k die Anzahl aller Ergebnisse zählt, die mit der Wahrscheinlichkeit p eintreten. So bedeutet im Beispiel $k = 3000$: „Es sind 3000 der bestellten n Steine brauchbar."

Ein Arbeiten mit der expliziten Darstellung führt zur direkten Übersetzung der mathematischen Terme und ist für Lernende leicht nachvollziehbar und zu verfolgen. Diesem Vorteil stehen jedoch Nachteile entgegen :

- Im Produkt $\binom{n}{k} \cdot p^k \cdot q^{n-k}$ gibt es zwei entgegengesetzte Tendenzen : Während für große n und k der Binomialkoeffizient $\binom{n}{k}$ leicht zu einem Rechner-„Overflow" führt, weil die Darstellungskapzität des Rechners überschritten wird, besteht bei den beiden Potenzen p^k sowie q^{n-k} die Gefahr, daß sie bei großen Exponenten so klein werden, daß der Rechner sie nur noch als 0 darstellen kann. Wie man beide Tendenzen geschickt kombinieren kann, wird in Abschnitt 4 beschrieben.
- Die Berechnung von Binomialkoeffizienten und von Potenzen dauert länger als das Ausführen von Grundrechenarten. Beim Berechnen von Bereichswahrscheinlichkeiten, wo eine Reihe von Einzelwahrscheinlichkeiten ermittelt und dann aufsummiert werden müssen, kann das zu einer längeren - je nach Hardwareausstattung und benutzter Software unterschiedlichen - Rechenzeit führen.

2.2 Einige Erfahrungen

Bei der Berechnung von Bereichswahrscheinlichkeiten bei der Einstiegsaufgabe unter Benutzung der expliziten Gleichung mache ich u. a. folgende Beobachtungen :

- Bei vielen Taschenrechnern kann man über „(n Cr k) * p^k * $(1 - p)^{n-k}$" Binomialwahrscheinlichkeiten direkt berechnen. Dem Einsatz des Taschenrechners werden durch die Grenzen der nCr-Taste (bei meinem SHARP EL-531 $0 \leq k \leq 69$) jedoch viel zu enge Grenzen gesetzt. Daher verzichte ich im folgenden auf die Einbeziehung des einfachen Taschenrechners in die weiteren Überlegungen.
- Exemplarisch für programmierbare Rechner soll hier der TI-82 betrachtet werden. Auch er bietet die nCr-Option. Über ein kleines Programm können einzelne Binomialwahrscheinlichkeiten und auch Bereichswahrscheinlichkeiten als Summen von Binomialwahrscheinlichkeiten berechnet werden. Auf $P(X \geq 3000)$ mit $p = 0{,}99$ oder $P(Y \leq 31)$, wobei Y die Zahl der unbrauchbaren Steine zählt, mit $p = 0{,}01$ für z. B. $n = 3031$ brauche ich nicht lange zu warten. Wenn ich $P(X < 3000)$

jedoch direkt ausrechnen und nicht aus 1 - $P(X \geq 3000)$ ermitteln lassen möchte, dann beendet die Abschaltautomatik des TI-82 den Berechnungsversuch.

- Bei Derive kann man sowohl Binomialwahrscheinlichkeiten z. B. über „B(n, p, k) := Comb(n, k) * p^k * $(1 - p)^{n-k}$" einzeln als auch kumuliert über „Summe(n, p, a, e) := $\sum_{k=a}^{e} B(n,p,k)$" für alle k aus einem zusammenhängenden Intervall [a;e] berechnen lassen. $P(X \geq 3000)$ und $P(Y \leq 31)$ erhalte ich unmittelbar nach Eingabe, während ich die direkte Berechnung von $P(X < 3000)$ nach mehr als einer Schulstunde Rechenzeit abgebrochen habe, ohne daß bis dahin ein Ergebnis vorhanden war.

Fazit : Taschenrechner mit nCr-Taste können bei Wahrscheinlichkeiten im Modell der Binomialverteilung nur bei Problemen mit nicht zu großem n und k benutzt werden. Beim TI-82 wie bei Derive zeigen sich bei der Einstiegsaufgabe deutlich Grenzen. Kann man diese Aufgabe oder eine vergleichbare in einer Klausur oder sogar im Abitur stellen, wenn der Prüfling Modellierungen wählen kann, in denen Berechnungen in angemessener Zeit nicht gelingen ? Eine Schulung der Lernenden im zumindest groben Abschätzen von zu erwartenden Laufzeiten erscheint mir unumgänglich. Außerdem ist es wichtig, die Strategie, eine Wahrscheinlichkeit über die des Gegenereignisses zu berechnen, frühzeitig in den Köpfen der Lernenden zu verankern.

Letztlich kann ich auch bei nicht zu großem n mit Hilfe der expliziten Darstellung schulrelevante Probleme bearbeiten. Die entsprechenden kurzen Programme können die Lernenden selbständig entwickeln. Für welche Aufgaben man noch Ergebnisse in akzeptablen Laufzeiten erhält, hängt unter anderem auch von der vorhandenen Hardware und der benutzten Software ab. Daher muß der Computereinsatz in jedem Einzelfall vorher in der Unterrichtsvorbereitung sorgfältig erprobt werden.

3. Einige Erfahrungen mit der rekursiven Darstellung zur Berechnung von Binomialwahrscheinlichkeiten

3.1 Die rekursive Darstellung

Es gibt noch eine andere Möglichkeit, Binomialwahrscheinlichkeiten und damit auch Bereichswahrscheinlichkeiten zu berechnen. Sie kann guten Stochastikbüchern entnommen werden. Bildet man den Quotienten $\frac{B(n, p, k)}{B(n, p, k-1)}$, wobei für B(n, p, k) und B(n, p, k - 1) der jeweils zugehörige Term der expliziten Darstellung aus Abschnitt 2 mit $1 \leq k \leq n$ eingesetzt wird, und kürzt diesen Quotienten so weit wie möglich, erhält man die rekursive Darstellung (*) : $B(n, p, k) = \frac{n+1-k}{k} \cdot \frac{p}{1-p} \cdot B(n, p, k-1)$. Startet man mit $B(n, p, 0) = q^n = (1 - p)^n$, dann kann man alle Wahrscheinlichkeiten B(n, p, k) der Reihe nach (rekursiv) mit Hilfe von Gleichung (*) berechnen. Falls die Wahrscheinlichkeit p größer als 0,5 ist, also $p > q$ gilt, lohnt es sich, mit $B(n, p, n) = p^n$ zu starten und alle Wahrscheinlichkeiten B(n, p, l) der Reihe nach zu ermitteln, in diesem Fall in umgekehrter Reihenfolge als oben beschrieben.

Beim Berechnen von Binomialwahrscheinlichkeiten mit Hilfe von Gleichung (*) werden nur Grundrechenarten benutzt, die vom Computer besonders schnell abgearbeitet werden können. Außerdem erhält man die nächste Wahrscheinlichkeit unter direkter Nutzung aller vorigen Ergebnisse. Diesen Vorteilen stehen allerdings auch Nachteile gegenüber. Man beginnt mit $P(X = 0)$ oder $P(X = n)$, also mit den kleinsten Wahrscheinlichkeiten einer Binomialverteilung, und berechnet auf dieser Basis alle weiteren Wahrscheinlichkeiten. Wenn die gewählte Startwahrscheinlichkeit so klein ist, daß sie der Rechner nur noch als „0" ausgeben kann, dann werden alle anderen Wahrscheinlichkeiten ebenfalls als „0" errechnet. Der TI-82 macht schon sehr früh dieses Problem deutlich, Turbo-Pascal-Programme und Rechenblätter bleiben davon auch nicht verschont. Man kann versuchen, das Eintreten dieses Falls etwas hinauszuzögern und mit der größeren der beiden Zahlen q^n oder p^n zu starten, jedoch hat man bei $p = q = 0{,}5$ keine Wahl. Bei Aufgaben mit $p = 0{,}5$ macht sich das Abrunden zu „0" bei größer werdendem n zuerst bemerkbar.

3.2 Einige Erfahrungen

Beim Lösen der Einstiegsaufgabe unter Benutzung der rekursiven Darstellung können wir u. a. folgende Beobachtungen machen :

- Ich habe ein kleines Programm geschrieben, mit dem der TI-82 Bereichswahrscheinlichkeiten berechnet. Er berechnet anstandslos $P(Y \leq 31)$ für $p = 0{,}01$ und $n = 3031$. Jedoch erscheint für $P(X \leq 3000)$ mit $p = 0{,}99$ und $n = 3031$ der Wert 0. Schon die erste Wahrscheinlichkeit $0{,}01^{3031}$ wird als 0 ausgegeben. Dies ist dann bereits das Endergebnis. Denn die Multiplikationen in allen weiteren Iterationen ändern diesen Wert nicht mehr. Es sieht also nach Vorteilen für die Gegenereignis-Strategie aus. Man sollte aber auch die Grenzen der Gegenereignis-Strategie vorführen. Dies ist beim TI-82 schon bei relativ kleinen Werten von n möglich. Ich formuliere eine der Kapazität des TI-82 angepaßte **Aufgabe 2** :

  „In einem Krankenhaus werden in einem Monat 343 Kinder geboren. Wie groß ist die Wahrscheinlichkeit, daß höchstens 171 davon Mädchen sind ?"

  Zur Lösung benötigen wir keinen Rechner. Auch das soll mit dieser Aufgabe demonstriert werden. Es ist aber immer wieder erstaunlich, wie selbstverständlich Lernende beginnen; zu rechnen und ein Computerprogramm einzusetzen anstatt erst nachzudenken. Verfolgen wir, was sie dann erleben. Die Zufallsgröße X zähle die Anzahl der Mädchengeburten, Y die der Jungengeburten und vereinfachend nehmen wir zunächst einmal $p = 0{,}5$ an. Ich erhalte mit dem TI-82 : $P(X \leq 171) = 0$, aber auch $P(Y \leq 171) = 0$. Ein neues Paradoxon ? Wir erwarten doch eigentlich als Summe der beiden Wahrscheinlichkeiten 1 sowie $P(X \leq 171) = P(X \geq 172) = 0{,}5$ für $p = 0{,}5$. Beim TI-82 erscheint als Ergebnis für $0{,}5^{343}$, und damit für die Startwahrscheinlichkeit, 0. Jeder Leser kann diese Aufgabe seiner Hard- oder Software anpassen.

  Man könnte Aufgabe 2 auch für das Werfen einer Laplace-Münze formulieren. Wenn aber die Aufgabe deutlich auf die Grenzen des in diesem Abschnitt betrachteten Algorithmus aufmerksam machen und außerdem die Frage nach einem neuen, besseren Algorithmus provozieren soll, dann sollte man Aufgabe 2 noch einmal stellen und in der neuen Fassung von dem aus langjährigen Beobachtungen ermittelten Mittelwert für Mädchengeburten $p = 0{,}486$ ausgehen. Denn in diesem Fall haben wir zwar die Vorstellung, daß die uns interessierenden Wahrscheinlichkeiten in der Nähe von 0,5 liegen, können sie aber mit der hier vorgestellten rekursiven Vorgehensweise mit dem TI-82 nicht ausrechnen; denn er liefert sowohl für die Anfangswahrscheinlichkeit $0{,}514^{343}$ als auch für $0{,}486^{343}$ das Ergebnis 0.
- Derive liefert eine elegante Möglichkeit, rekursives Vorgehen direkt nachzubilden. Mit „Brek(n, p, k) := IF(k = 0, $(1 - p)^n$, $\frac{(n-k+1)}{k} \cdot \frac{p}{1-p}$ · Brek(n, p, k - 1))" läßt sich übersichtlich und für Lernende nachvollziehbar die Rekursion darstellen. Es können damit Einzelwahrscheinlichkeiten berechnet werden. Leider mißlingt mir zum Bei-

spiel die Berechnung von Brek(3031, 0,99, 3000), da mein Rechner nicht genügend Speicherplatz dafür aufweist, so daß ich erst gar nicht versuche, auch Bereichswahrscheinlichkeiten über diese Möglichkeit auszurechnen. Beschränkt man sich auf Aufgaben mit kleineren Werten von n, dann gelingt diese Art rekursiver Berechnung.

- Derive bietet eine weitere Möglichkeit, die rekursive Darstellung (*) einzusetzen. Man erhält einen guten Zugang zu dieser Möglichkeit, wenn man eine Tabelle für alle Wahrscheinlichkeiten P(X = k) = B(n, p, k) erstellen will. In der ersten Spalte dieser Tabelle soll die Anzahl der Erfolge k mit $0 \le k \le n$ dargestellt werden, in der zweiten Spalte die zu k gehörige Wahrscheinlichkeit P(X = k) = B(n, p, k) und in der dritten Spalte die kumulierte Wahrscheinlichkeit $P(X \le k) = \sum_{i=0}^{k} B(n,p,i)$. Eine solche Tabelle erhält man mit der ITERATES-Funktion, die eine Matrix mit 3 Spalten und n+1 Zeilen (für die Variable k müssen nacheinander alle natürlichen Zahlen zwischen 0 und n eingesetzt werden) erzeugt. Jede einzelne Zeile dieser Matrix kann als Zeilenvektor v mit den drei Koordinaten v_1, v_2 und v_3 angesehen werden. Die erste Koordinate v_1 muß bei jedem Iterationsschritt um 1 erhöht werden. Bei der zweiten Koordinate v_2 muß der in (*) beschriebene Iterationsschritt vorgenommen werden. Mit $\frac{n-v_1}{v_1+1}\cdot\frac{p}{1-p}\cdot v_2$ wird der neue Wert von v_2 berechnet, wobei ich darauf achten muß, daß ich hier anstelle von k in (*) k + 1 einsetzen muß. Zu der dritten Koordinate v_3 muß die als zweite Koordinate v_2 berechnete Wahrscheinlichkeit hinzuaddiert werden. Nun dürfte der folgende ITERATES-Befehl verständlich sein, der eine vollständige Tabelle für die Binomialwahrscheinlichkeiten B(n, p, k) generiert.

„BinTabelle(n, p) := ITERATES([$v_1 + 1$, $\frac{n-v_1}{v_1+1}\cdot\frac{p}{1-p}\cdot v_2$, $\frac{n-v_1}{v_1+1}\cdot\frac{p}{1-p}\cdot v_2 + v_3$], v,[0, $(1-p)^n$, $(1-p)^n$], n)". Wenn ich nur den Zeilenvektor für k Erfolge erzeugen und mit seinen drei Koordinaten anzeigen lassen will, dann kann ich das mit

„BinSumme(n, p, k) := ITERATE([$v_1 + 1$, $\frac{n-v_1}{v_1+1}\cdot\frac{p}{1-p}\cdot v_2$, $v_3 + \frac{n-v_1}{v_1+1}\cdot\frac{p}{1-p}\cdot v_2$], v, [0, $(1-p)^n$, $(1-p)^n$], k)" erreichen.

Die bei der Einstiegsaufgabe interessierende Wahrscheinlichkeit $P(Y \le 31)$ kann durch Eingabe von „BinSumme(3031, 0.01, 31)" auf zwei Arten bestimmt werden : Im Exakt-Modus, bei dem Ergebnisse als ganze Zahlen oder als Brüche ausgegeben werden, oder im Näherungsmodus, bei dem Ergebnisse als Dezimalzahlen mit voreinstellbarer Genauigkeit dargestellt werden. Im Exact-Modus dauert die Berechnung für das Beispiel ca. 6 Minuten, in Näherungsmodus erhält man sofort das Ergebnis. Nach diesen beiden erfolgreichen Rechnungen interessiert mich natürlich, wie Derive dieses Mal mit der Wahrscheinlichkeit $P(X \le 2999)$ fertig wird, bei der 3000 Einzelwahrscheinlichkeiten ausgerechnet und aufaddiert werden müssen. Nach Eingabe von „BinSumme(3031, 0.99, 2999)" wird im Näherungsmodus diese Bereichswahrscheinlichkeit in ca. 26 Minuten ermittelt, was für den Unterricht vielleicht noch tolerierbar ist, in einer Prüfung jedoch immer noch zu lange dauert.

- Die rekursive Darstellung nach (*) eignet sich auch hervorragend zur Programmierung in Rechenblättern. Zuerst gibt man die Startzeile mit $q^n = (1-p)^n$ oder p^n ein, je nach Wahl der Laufrichtung, dann in der folgenden Zeile die Rekursionsvorschrift (*). Nun kann man diese Rekursionsvorschrift so oft in die folgenden Zeilen kopieren, wie Ergebnisse benötigt werden. Dies ist sehr bequem, schnell zu programmieren, von den Lernenden leicht nachzuvollziehen und auch selbständig durchzuführen. Es sollte allerdings immer vorher getestet werden, ob die Anfangswahrscheinlichkeit nicht bereits als 0 ausgegeben wird.

Fazit : Derive kann bei vielen Problemen eingesetzt werden. Mit akzeptablen Laufzeiten ist aber nur dann zu rechnen, wenn die Stichprobenanzahl n und auch die Anzahl der zu berechenden und aufzusummierenden Wahrscheinlichkeiten nicht zu groß sind. Wegen der besonders einfachen Programmierung eignet sich die rekursive Darstellung vor allem gut für Rechenblätter, wobei die schnell abrufbaren graphischen Fähigkeiten der Tabellenkalkulationsprogramme zusätzlich hervorragende Visualisierungen liefern. Alles das sind Gründe, die für den Unterrichtseinsatz von Rechenblättern sprechen. Bei sehr kleinen Werten von q^n oder p^n , die an der unteren Grenze der Darstellungsmöglichkeit für Dezimalzahlen der benutzten Software liegen, können Rundungsfehler auftreten. Daher sollte in solchen Fällen auf die Iteration verzichtet werden.

4. Einige Erfahrungen mit einem weitereren Algorithmus

4.1 Ein neuer Algorithmus

Der in Abschnitt 3 beschriebene Algorithmus beginnt mit der Berechnung der kleinsten Wahrscheinlichkeit, die bei großem n vom Rechner nur noch als „0" ausgegeben wird. Engel (1992, S. 134-136) versucht, durch Logarithmieren diesen Algorithmus zu verbessern. Hier soll der Weg von Engel nicht weiter verfolgt werden. Stattdessen wird ein anderer Algorithmus betrachtet, der direkt mit der größten Wahrscheinlichkeit der interessanten Binomialverteilung beginnt, von dieser ausgehend alle anderen Wahrscheinlichkeiten rekursiv nach beiden Richtungen hin berechnet und rechner- oder softwarebedingte Näherungsverfahren wie z. B. Logarithmieren vermeidet.

Die größte Wahrscheinlichkeit liegt in der Nähe des Erwartungswerts $\mu = n\cdot p$. Unter den Wahrscheinlichkeiten B(n, p, k) ist die für $k = [(n+1)\cdot p]$ am größten, wobei die eckige Klammer als Gauß-Klammer aufgefasst wird. Die möglichst effektive Berechnung dieser größten Wahrscheinlichkeit ist Kern eines Algorithmus, der hier mit seinen drei wesentlichen Schritten kurz erläutert werden soll. Das als Anlage abgedruckte Listing eines Programms für den TI-82 kann parallel zur Darstellung der drei Schritte gelesen werden und so eine erste Hilfe bieten. Beim ersten Lesen kann man diese Informationen überspringen und bei 4.2 fortfahren.

Schritt 1 : Die Berechnung der maximalen Wahrscheinlichkeit an der Stelle k_{max}

Will man den Binomialkoeffizienten $\binom{n}{k} = \frac{n!}{(n-k)!\cdot k!}$ kürzen, gibt es zwei Möglichkeiten :

a. Man dividiert Zähler und Nenner durch (n - k)! und erhält $\frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{k\cdot(k-1)\cdot\ldots\cdot 1}$

b. oder man kürzt durch k! und erhält $\frac{n\cdot(n-1)\cdot\ldots\cdot(k+1)}{(n-k)\cdot(n-k-1)\cdot\ldots\cdot 1}$.

Daher gibt es für B(n, p, k) folgende drei Darstellungen : $B(n,p,k) = \binom{n}{k}\cdot p^k\cdot q^{n-k} = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{k\cdot(k-1)\cdot\ldots\cdot 1}\cdot p^k\cdot q^{n-k} = \frac{n\cdot(n-1)\cdot\ldots\cdot(k+1)}{(n-k)\cdot(n-k-1)\cdot\ldots\cdot 1}\cdot p^k\cdot q^{n-k}$.

Will man einen schnellen Algorithmus erhalten, versucht man mit möglichst wenig Rechenoperationen auszukommen, die zudem möglichst auch nur Grundrechenarten sein sollten. Im Bruchterm auf der rechten Seite der Gleichung von B(n, p ,k) führt man im Nenner möglichst wenig Multiplikationen aus, wenn man das Minimum von n - k und n als größten Faktor im Nennerprodukt nimmt. Also setzen wir k_{min} = Min {k, n - k}. Zu k_{min} gehört im Zähler entweder n - k (für k_{min} = k) oder k (für k_{min} = n - k), in jedem Fall also eine Zahl z, für die z = n - k_{min} gilt. Nun kann man den Binomialkoeffizienten nach der Formel $\binom{n}{k} = \frac{z+1}{1}\cdot\frac{z+2}{2}\cdot\ldots\cdot\frac{n}{k_{min}}$ berechnen, wobei z + k_{min} = n gilt. Um dies zu sehen, muß man in den Bruchtermen der obigen Gleichung von B(n, p, k) die Zähler und die Nenner von rechts nach links lesen. Nun kommt die in Abschnitt 2 erwähnte Überlegung. Berechnet man den Binomialkoeffizienten durch Ausmultiplizieren dieses Produkts, dann besteht die Gefahr, daß es zu einem Rechner-„Overflow" kommt. Um diese Gefahr zu vermeiden, beginnt man mit dem ersten Faktor $\frac{z+1}{1}$ und multipliziert diesen so oft mit p oder q, bis das Produkt kleiner als 1 ist. Nun multipliziert man dieses mneue Produkt mit dem nächsten Faktor $\frac{z+2}{2}$ und anschließend wieder so oft mit p oder q, bis das Produkt wiederum kleiner als 1 ist. Und so fährt man mit allen weiteren Faktoren des Binomialkoeffizienten fort. Das Produkt wird schließlich immer kleiner als 1, da p und q Zahlen zwischen 0 und 1 sind und die gesuchte Wahrscheinlichkeit auch eine Zahl zwischen 0 und 1 ist. Am Ende multipliziert man das bis jetzt erhaltene Produkt noch mit p^{Restp} und q^{Restq}, wobei Restp bzw. Restq die Anzahl der an B(n, p, k) noch fehlenden und bisher im Verfahren noch nicht durch Multiplizieren „verbrauchten" Faktoren von p bzw. q sind. Restp bzw. Restq sind ganze Zahlen. Sie sind positiv, falls noch nicht alle Faktoren „verbraucht" sind, sie sind 0, falls alle Faktoren verbraucht sind und sie sind negativ, falls man zur Verkleinerung des Produkts von p mehr Faktoren als k bzw. von q mehr als n - k benutzt hat. Wer Wert auf eine möglichst exakte Rechnung legt, wird Bedenken wegen des Potenzierens äußern. Das Potenzieren muß z. B. in Turbo-Pascal unter Benutzung der ln- und der e-Funktion nachgebildet werden. Man kann dies aber umgehen. Da sowohl Restp als auch Restq ganze Zahlen sind, kann das Potenzieren in einer Schleife als wiederholte Multiplikation mit p bzw. q oder, falls Restp bzw. Restq negativ sind, mit den Kehrwerten von p bzw. q programmiert werden, so daß nur Grundrechenarten benutzt werden. Damit ist die Berechnung der Wahrscheinlichkeit B(n, p, k_{max}) abgeschlossen.

Schritt 2 : Berechnung aller Wahrscheinlichkeiten für alle k > k_{max}
Ausgang dieser Rekursion nach größer werdendem k ist die in Schritt 1 bestimmte Wahrscheinlichkeit B(n, p, k_{max}). Mit dieser Startwahrscheinlichkeit beginnt man die Rekursion mit Hilfe von Gleichung (*) aus Abschnitt 3, setzt für k zuerst k_{max} und dann der Reihe nach alle Werte größer als k_{max} ein.

Schritt 3 : Berechnung aller Wahrscheinlichkeiten für alle k < k_{max}
Ausgang dieser Rekursion nach kleiner werdendem k ist B(n, p, k_{max}) wie in Schritt 2. Mit dieser Startwahrscheinlichkeit beginnt man die Rekursion, setzt für k zuerst k_{max} und dann der Reihe nach alle Werte kleiner als k_{max} ein. Die Rekursionsgleichung muß jedoch erst noch für die Abwärtsrekursion umgeformt werden. Für die Abwärtsrekursion gilt Gleichung (**) : $B(n, p, k-1) = \frac{k}{n-k+1}\cdot\frac{1-p}{p}\cdot B(n, p, k)$.

Wer Wert auf weitere Einzelheiten legt, kann diese dazu ein informatives Struktogramm sowie ein Programm-Listing in BASIC den Seiten 80 - 82 von DIFF (1983) entnehmen.

4.2 Einige Erfahrungen

Besonders schnell wird der in 4.1 vorgestellte Algorithmus ausgeführt, wenn die Berechnung der Einzelwahrscheinlichkeiten und das Aufsummieren simultan durchgeführt wird. Das zugehörige Programm wird wesentlich aufwendiger als das für den Algorithmus in Abschnitt 3. Ich gebe es vor. Lernende können es studieren und über den Algorithmus referieren. Beim Versuch, die Aufgaben 1 und 2 unter Benutzung dieses neuen Algorithmus zu lösen, mache ich u. a. folgende Beobachtungen :

- Der TI-82 - ein Programmlisting wird als Anlage abgedruckt - benötigt für Aufgabe 2 nur ca. 30 Sekunden, um eine der Bereichswahrscheinlichkeiten $P(X \leq 171)$ oder $P(X \geq 172)$ für p = 0,5 bzw. p = 0,486 zu berechnen. Bei der Einstiegsaufgabe zum Kauf der Steine muß man etwas länger warten : Für n = 3031 ca. 95 Sekunden auf $P(X \geq 3000)$ mit p = 0,99 und $P(Y \leq 31)$ mit p = 0,01 sowie ca. 4 Minuten 40 Sekunden auf $P(X \leq 2999)$ mit p = 0,99.
- Ich habe schon häufig beobachtet, daß Lernende Hilfsmittel sinnvoll kombinieren. Hier bietet sich geradezu eine Zusammenarbeit von TI-82 und Rechenblatt an : Mit dem TI-82 wird die größte Wahrscheinlichkeit berechnet, dann werden auf dem Rechenblatt alle Wahrscheinlichkeiten des interessanten Bereichs auf der Basis der größten Wahrscheinlichkeit rekursiv generiert und graphisch dargestellt.
- Den Algorithmus setze ich in einem Turbo-Pascal-Programm um, das Code für den mathematischen Koprozessor erzeugt. Der verwendete Datentyp Extended besitzt einen Wertebereich von $1{,}9\cdot 10^{-4951}$ bis $1{,}1\cdot 10^{4932}$ und eine Genauigkeit von 19 bis 20 Dezimalstellen. Das Programm liefert bei diesen und anderen Bereichswahrscheinlichkeiten Ergebnisse unmittelbar nach Eingabe. Ich beobachte allerdings Unterschiede in den Ergebnissen, je nachdem, ob ich das Potenzieren unter Benutzung der ln- und der e-Funktion nachbilde, oder ob ich wie im Listing des TI-82-Programms dargestellt das Potenzieren als iteratives Multiplizieren ausführen lasse.

Fazit : Der TI-82 beeindruckt mit seiner Leistung. Das Turbo-Pascal-Programm mit Code für den mathematischen Koprozessor arbeitet mit extrem kurzen Laufzeiten. In den vorigen Abschnitten wurden Gründe genannt, warum eine Umsetzung in Derive nicht erprobt wird.

5. Eine besondere Testaufgabe

Nach diesen positiven Ergebnissen möchte ich die Leistungsfähigkeit des neuen Algorithmus an **Aufgabe 3** (nach ENGEL (1973) Band 1, S. 134) erproben :

„In den USA wurden 1950 bei insgesamt 3 554 119 Geburten 1 823 555 Jungen geboren. Für Jungengeburten kennt man durch langjährige Beobachtungen als relative Häufigkeit 0,514. Sind 1950 in den USA Jungengeburten deutlich seltener ?"

Interessant kann bei dieser Fragestellung zum Beispiel die Wahrscheinlichkeit dafür sein, daß 1 823 555 oder weniger Jungen geboren werden, auch wenn sich die Wahrscheinlichkeit für eine Jungengeburt nicht geändert hat. Formal interessiert uns dann $P(X \leq 1\,823\,555)$, wenn X die Anzahl der Jungengeburten zählt, die Wahrscheinlichkeit für eine Jungengeburt p = 0,514 ist und n = 3 554 119 beträgt. In Wirths (1995) habe ich $P(X \leq 1\,823\,555) \approx 2{,}7\cdot 10^{-4}$ bei einer Computeranzeige auf 4 Nachkommastellen von $2{,}6799\cdot 10^{-4}$ angegeben. Diese Lösung habe ich durch Approximation der Binomialverteilung durch die Normalverteilung ermittelt. Ich habe es lange Zeit nicht für möglich gehalten, diese Wahrscheinlichkeit auch im Modell der Binomialverteilung tatsächlich berechnen zu können. Mit dem oben dargestellten Algorithmus errechnet mein Turbo-Pascal-Programm in ca. 10 Sekunden $P(X \leq 1\,823\,555) = 2{,}6857 \cdot 10^{-4}$.

Fazit : Mit diesem Algorithmus sollte das Berechnen von Einzelwahrscheinlichkeiten und von Wahrscheinlichkeiten beliebiger Bereiche auch für sehr große Werte von n in der Größenordnung von mehreren Millionen gelingen. Das Turbo-Pascal-Programm eignet sich für eine so extreme Problemstellung wie die von Aufgabe 3 besonders gut.

6. Abschlußbemerkungen

Der Computer soll motivierende Anwendungen möglich machen, eine Herausforderung, für alle, die sich einem lebendigen Stochastikunterricht verschrieben haben. An zwei Stellen habe ich bisher jedoch durch zu lange Computerlaufzeiten jeweils stark retardierende Momente erlebt : Bei der Durchführung von Simulationen und beim Berechnen von Bereichswahrscheinlichkeiten. Ausgelöst wurde dies in beiden Fällen durch zu optimistisch angesetzte hohe Werte für die Anzahl der Wiederholungen der Simulation oder an Bernoulli-Versuchen bei Bereichswahrscheinlichkeiten. Man kann in bei Simulationen auch bei reduzierter Anzahl an Wiederholungen immer noch zu gut auswertbarem Material gelangen, beim Berechnen von Bereichswahrscheinlichkeiten muß man sich auf spezielle Aufgaben mit nicht zu großem Stichprobenumfang beschränken. Ich möchte auf Aufgaben vom Typ der hier vorgestellten in meinem Unterricht nicht verzichten, also in einigen Beispielen auch sehr große n als Stichprobenanzahl verwenden. Außerdem möchte ich unterschiedliche Gesichtspunkte, und zwar so viele wie möglich, bereits im Modell der Binomialverteilung betrachten, und erst auf einer hier gesicherten Basis von Erfahrungen und Einsichten zu stetigen Verteilungen im allgemeinen und zur Normalverteilung im besonderen übergehen.

Der Einsatz des Computers im Stochastikunterricht zwingt zum Nachdenken über Konsequenzen. Einige davon sind meiner Meinung nach :

- Für die Gestaltung von Stochastik-Curricula ergeben sich neue Möglichkeiten. Es ist nicht verwunderlich, daß analytisch einfache Verteilungen, die zur Approximation der Binomialverteilung geeignet sind, von großer Bedeutung in der Entwicklung der Stochastik waren. Denn für numerische Auswertungen sind die Gleichungen zur Berechnung von Binomialwahrscheinlichkeiten unbrauchbar, vor allem, wenn die Anzahl n der Bernoulli-Versuche groß ist. Erst umfangreiche Summen liefern die interessanten Wahrscheinlichkeiten relevanter Bereiche. So wichtig analytisch einfache Verteilungen, die zur Approximation der Binomialverteilung geeignet sind, auch heute noch sind, mit der Verfügbarkeit von Computern und entsprechend leistungsfähiger Software können wir Alternativen erproben, die vorher allenfalls nur schwer zugänglich waren, und sollten dies auch tun. Ich möchte bereits im Modell der Binomialverteilung vielfältige Erfahrungen sammeln, Einsichten vermitteln und stochastisches Denken fördern. Dies ist in einem Kurs Stochastik bereits sehr früh möglich und erfordert wenige, dazu noch elementare Begriffe und Hilfsmittel. Ein solches Vorgehen fördert die Konzentration auf Wesentliches.
- Die Probleme, die in einer Unterrichtseinheit „Prüfen und Schätzen" behandelt werden können, werden allgemein als anwendungsnah, vor allem als sehr motivierend angesehen. Aber nicht immer gelingt eine Einbettung in den eigenen Unterricht. Vor allem wenn sie erst nach einer Unterrichteinheit „Normalverteilung" eingeplant wird, fehlt meist die erforderliche Zeit. In Wirths (1996) wird über Erfahrungen berichtet, wie der Computer als Rechenknecht eingesetzt wird, dem die Aufgabe zugewiesen wird, die Wahrscheinlichkeit für Bereiche, in denen die uns interessierenden Ergebnisse liegen, zu berechnen. Wesentlich für mich ist, daß Lernende auf diesem Weg selbst Gelegenheit erhalten zu erkennen, daß Theorie gebildet werden kann, und die Aussagen dieser Theorie selbständig vorher formulieren können.

- Bei Approximationen brauche ich nicht mehr nur mit Rezepten zu arbeiten und darauf zu vertrauen, daß der errechnete Wert ein guter Näherungswert ist. Mit Computereinsatz kann ich nun sowohl im Modell der Binomialverteilung als auch in dem der Normalverteilung errechnete Werte miteinander vergleichen und die Güte der Näherung erkennen und ausdrücken.
- Die hypergeometrische Verteilung, von den Anwendungen her gesehen überaus interessant, könnte mit dem Unterrichtseinsatz einer Weiterentwicklung des hier vorgestellten Algorithmus aus ihrem Dornröschenschlaf geweckt werden.

Vor allem in den Grundkursen, in denen Lernende der Mathematik am liebsten vorzeitig aus dem Wege gehen möchten, - aber nicht nur dort ! - hat ein Stochastikunterricht mit der oben skizzierten Methodik häufig wieder Interesse für Mathematik wecken können. Diese Erfahrung kann ich nicht ignorieren, sie macht Mut zum Erproben von neuen Wegen, insbesondere auch zur Integration des Computers in den Unterricht.

Ich habe mich in diesem Beitrag auf die für meinen Unterricht derzeit einsetzbare Hard- und Software beschränkt. Es wäre interessant und wünschenswert, über den Einsatz von z. B. MathCad, Mathematica, Maple oder des TI-92 bei der Berechnung von Bereichswahrscheinlichkeiten im Modell der Binomialverteilung informiert zu werden.

6. Literatur

1. Engel, A. (1973) Wahrscheinlichkeitsrechnung und Statistik. 2 Bände.Stuttgart : Klett
2. Engel, A. (1991) Mathematisches Experimentieren mit dem PC. Stuttgart : Klett
3. Kroll, W. (1985) Elementare Stochastik im Schulunterricht - Ein Grundkurskonzept. Ergebnisse regionaler Lehrerfortbildung in Hessen. Heft 85, Fuldatal / Kassel, HILF
4. Wirths, H. (1996) Test- und Schätzverfahren. In : Mathematik in der Schule 34. Jg, 1996/11, S. 596 - 611
5. Wahrscheinlichkeitsrechnung und Statistik unter Einbeziehung von elektronischen Rechnern, Heft SR 2. DIFF 1983

Anlage Berechnen einer Bereichswahrscheinlichkeit P(U ≤ X ≤ O) mit dem TI-82

Programm Bereich

```
:  INPUT „N = ", N                        (* Anzahl Versuche *)
:  INPUT „P = ", P                        (* Erfolgswahrscheinlichkeit *)
:  INPUT „U = ", U                        (* untere Grenze Bereich *)
:  INPUT „O = ", O                        (* obere Grenze Bereich *)
:  INT((N+1)*P) → K                       (* Stelle größter Wahrscheinlichkeit *)
:  MIN(K, N - K) → A : K → C : N - A → Z  (* Beginn 1. Schritt *)
:  N - K → D : 1 → B : 1 - P → F
:  FOR(I, 1, A, 1)
:     (Z + I) / I *B → B
:     WHILE ((B > 1) AND (C > 0) : B*P → B : C - 1 → C : END
:     WHILE (B > 1) : B*F → B : D - 1 → D : END
:  END
:  FOR(I, 1, C, 1) : B*P → B : END
:  IF (D < 0) : THEN : 1/F → F : -D → D : END
:  FOR(I, 1, D, 1) : B*F → B : END        (* größte Wahrscheinlichkeit P(X = k) *)
                                          (* Ende 1. Schritt *)
:  B → W : 0 → S : 1 - P → F              (* Vorbereitung Aufwärtsrekursion *)
```

```
: IF (O > K) : THEN                         (* Beginn 2. Schritt *)
:   FOR(I, K + 1, O, 1)
:     B*P / F*(N + 1 - I) / I → B
:     IF (I ≥ U) : THEN : S + B → S : END
:   END
: END                                       (* Ende 2. Schritt *)
: W → B                                     (* Vorbereitung Abwärtsrekursion *)
: IF (U < K) : THEN                         (* Beginn 3. Schritt *)
:   K → I
:   WHILE (I > U)
:     B*I / (N - I + 1)*F / P → B
:     I - 1 → I
:     IF (I ≤ O) : THEN : S + B → S : END
:   END
: END                                       (* Ende 3. Schritt *)
: IF ((U ≤ K) AND (K ≤ O)) :THEN
:   S + W → S
: END
: DISP „S = ", S                            (* Ausgabe Bereichswahrscheinlichkeit *)
```