EINFACHE SIMULATIONSMODELLE FÜR WARTESCHLANGEN *)

von Volkmar Lindenau, Bremen

Zusammenfassung: Es wird gezeigt, daß man das Bedienungsproblem - Kundenstrom ist Poissonstrom, Bedienungszeiten sind exponentiell verteilt - bereits im MU der Klassen 8/9 durch Simulation lösen kann. Die Grundbegriffe Wartezeit, Ankunftsrate, Bedienungsrate, ... können bei der Betrachtung einfacher Sonderfälle von Bedienungsproblemen eingeführt werden. Zu gegebener Ankunftsrate und gegebener Bedienungsrate gewinnt man durch Simulation Folgen von Ankunftszeiten und Folgen von Bedienungszeiten, aus denen rekursiv die Kundenwartezeiten berechnet werden. Daraus wird die durchschnittliche Länge der Warteschlange berechnet. Durch Variation der Parameter findet man experimentell den Zusammenhang zwischen der durchschnittlichen Länge und den Parametern.

ZDM-Klassifikation: K94

1. Einführung

Warteschlangen sind uns wohlvertraut. Jeder von uns hat schon einmal in einer "Kassenschlange", z.B. im Kaufhaus, gestanden und beim Arzt, im Restaurant, vor einer Verkehrsampel, am Fahrkartenschalter, ... auf seine "Bedienung" gewartet.

In einem "Bedienungssystem" treffen laufend Kunden ein, um sich an einem Schalter bedienen zu lassen. "Kunden" und "Bedienungsschalter" haben dabei eine sehr weite Bedeutung: Autos warten vor einer Fähre auf ihre Beförderung, Werkstücke "warten" vor einer Maschine auf die Weiterbearbeitung, Ersatzteilbestellungen "warten" auf die Erledigung. Die auf ihre Bedienung wartenden Kunden bilden auf eindeutige Weise eine "Warteschlange", sobald die Reihenfolge der Bedienung festgelegt ist.

*) Überarbeitete Fassung eines auf der Tagung "Stochastik im Mathematikunterricht" (Oberwolfach, 17.-21. November 1986) gehaltenen Vortrages.

Bei der folgenden mathematischen Betrachtung von Warteschlangen wird stets vorausgesetzt, daß einige spezielle Randbedingungen erfüllt sind:

(1) Es gibt genau einen Bedienungsschalter (Hinweis: im allgemeinen Fall könnten es z.B. auch mehrere parallel arbeitende Bedienungsschalter sein).

(2) Die Warteraumgröße ist nicht begrenzt; es kommt nicht vor, daß Kunden wegen fehlender Wartemöglichkeiten oder zu langer Warteschlange unbedient weggehen.

(3) Die Kunden werden in der Reihenfolge ihres Eintreffens bedient, d.h. es gibt keine Kunden, die - etwa aufgrund von Voranmeldung - bevorzugt bedient werden u.ä.

Was soll nun die mathematische Untersuchung von Warteschlangen leisten? Unmittelbar einleuchtend ist, daß Vorhersagen über die durchschnittliche Größe der Warteschlangenlänge und über die durchschnittliche Abweichung davon, sowie über die Schalterauslastung von großem praktischen Nutzen sind, wenn es z.B. gilt, ein neu einzurichtendes Bedienungssystem baulich und personell optimal auszustatten. Damit Kunden nicht abgeschreckt werden, dürfen sie keine all zu langen Warteschlangen antreffen. Kurze Warteschlangen sind möglich durch die Einrichtung von Bedienungsschaltern mit großer Bedienungskapazität. Solche Schalter sind aber, wenn zufällig wenig Kunden kommen, schlecht ausgelastet und folglich unrentabel. Die beiden gegensätzlichen Forderungen "kurze Warteschlangen" und hohe Schalterauslastung" müssen also optimal aufeinander abgestimmt werden. Dafür sollen durch die mathematische Modellierung von Warteschlangen Entscheidungshilfen entwickelt werden.

2. Das (M,M,1)-Bedienungsmodell

Um einen Bedienungsprozess modellieren zu können, muß man Annahmen machen über die Kunden und die Art und Dauer ihrer Bedienung.

Angenommen, in einem Bedienungssystem treffen die Kunden K_1, K_2, K_3, K_4,... einzeln nacheinander zu den Zeitpunkten t_1, t_2, t_3, t_4,... ein. Die den "Kundenstrom" kennzeichnenden sog. Zwischenankunftszeiten t_2-t_1, t_3-t_2, t_4-t_3,... werden selten im voraus bekannt sein (Ausnahme: die Kunden sind vorbestellt), denn i.a. entscheidet jeder Kunde unabhängig von allen übrigen Kunden selbst über seinen Ankunftszeitpunkt im Bedienungssystem. Stellen Sie sich vor, daß Sie als Beobachter eines Bedienungsprozesses über längere Zeit, in welcher weder der Kundenstamm noch die Bedürfnisse und Gewohnheiten der Kunden sich ändern, die Zwischenankunftszeiten der Kunden registrieren. Wenn Sie nun diese Zwischenankunftszeiten in Klassen einteilen, z.B. durch Rundung auf Minuten, und sukzessive in ein Häufigkeitsdiagramm eintragen, so erhalten Sie eine Häufigkeitsverteilung, deren Form sich i.a. mit zunehmender Beobachtungsdauer stabilisiert, d.h. für das Kundenverhalten in dem vorliegenden Bedienungssystem typisch ist. Sehr oft sind die beobachteten Zwischenankunftszeiten wie in Abb. 1 verteilt:

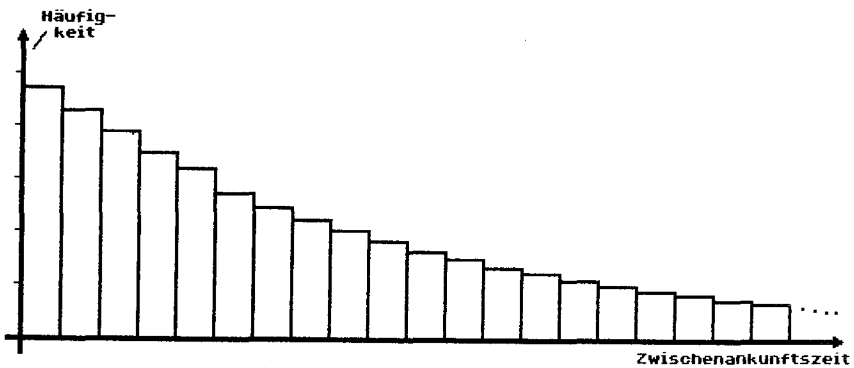


Abb.1

Für die Modellierung des Bedienungsprozesses müßte also die Wahrscheinlichkeitsverteilung der Zufallsvariablen "Zwischenankunftszeit" bekannt sein; das Gleiche gilt für die Bedienungszeiten, die für die Kunden benötigt werden. In der Praxis sind häufig die folgenden Voraussetzungen erfüllt oder wenigstens annähernd erfüllt:

(1) Die Zwischenankunftszeiten t_Z sind unabhängig und exponentialverteilt:

$$F(t)=P(t_Z\leq t)=1-e^{-\lambda t},$$

dabei ist λ die "Ankunftsrate", nämlich die durchschnittliche Zahl der pro Zeiteinheit (Abk. ZE) im Bedienungssystem eintreffenden Kunden; $^1/_\lambda$ ist dann die durchschnittliche Zwischenankunftszeit (Anmerkung: Der Graph der Dichte der Funktion F, also die Funktion $t \mapsto \lambda e^{-\lambda t}$, sieht aus wie der Graph in der Abb. 1)

(2) Die Bedienungszeiten t_B sind ebenfalls unabhängig und exponentialverteilt:

$$P(t_B\leq t)=1-e^{-\mu t};$$

hierin ist μ die Bedienungsrate, also $^1/_\mu$ die durchschnittliche Bedienungszeit. μ ist größer als λ, d.h. die im Durchschnitt benötigte Bedienungszeit $^1/_\mu$ ist kleiner als die durchschnittliche Zwischenankunftszeit $^1/_\lambda$ (anderenfalls käme es zu einem "unendlichen Anwachsen" der Warteschlange!)

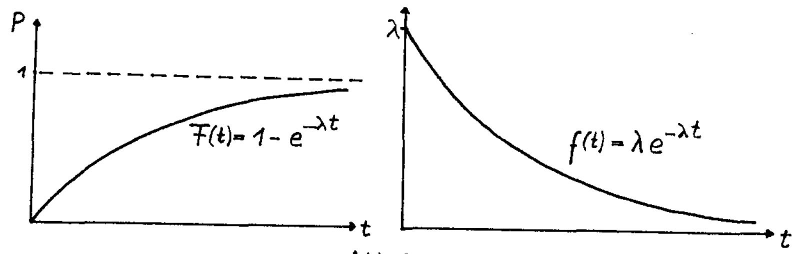


Abb. 2

Das Kürzel (M,M,1) für solch ein Bedienungssystem bringt zum Ausdruck, daß die Zwischenankunftszeiten und die Bedienungszeiten exponentialverteilt sind. Der Buchstabe M steht für "memoryless" und signalisiert die Haupteigenschaft der Exponentialverteilung: sie ist ohne Gedächtnis in dem Sinne, daß die Wartezeit auf das Eintreffen des nächsten Kunden nicht davon abhängt, wieviel Zeit man bereits ohne Erfolg gewartet hat (vgl. (3), S. 157). Die "1" deutet an, daß die Bedienung der Kunden nur an einem einzigen Bedienungsschalter erfolgt.

3. Lösung des (M,M,1)-Bedienungsproblems

Die Voraussetzungen (1) und (2) erleichtern die weitere mathematische Behandlung, weil man mit der Exponentialverteilung "sehr gut rechnen" kann. (1) impliziert, daß für die "kleine" Zeitspanne Δt (klein im Verhältnis zu $^1/_\lambda$) die Wahrscheinlichkeit dafür, daß in der Zeit Δt ein neuer Kunde eintrifft, praktisch gleich $\lambda\Delta t$ gesetzt werden kann und daß man mit $1-\lambda\Delta t$ als Wahrscheinlichkeit dafür rechnen darf, daß in der Zeit Δt kein Kunde kommt. Aus (2) ergibt sich, daß, falls ein Kunde sich am Bedienungsschalter befindet, die Bedienung dieses Kunden im Beobachtungszeitraum Δt mit der Wahrscheinlichkeit $\mu\Delta t$ beendet sein wird und daß mit der Wahrscheinlichkeit $1-\mu\Delta t$ kein Kunde weggeht.

Es sei nun $E_n(t)$ das Ereignis "Zum Zeitpunkt t befinden sich in dem Warteschlangensystem genau n Kunden", und $p_n(t)$ bezeichne die Wahrscheinlichkeit von $E_n(t)$. Man erhält eine Beziehung zwischen den Wahrscheinlichkeiten $p_n(t)$ $(n=0,1,2,...)$ und schließlich ein System von Differentialgleichungen für die $p_n(t)$, wenn man sich klar macht, auf welche Weise das Ereignis $E_n(t+\Delta t)$ von den Veränderungen im Wartesystem während des Zeitraums $[t, t+\Delta t]$ abhängt. Als Beispiel werde das Ereignis $E_2(t+\Delta t)$ betrachtet: "genau 2 Kunden im System zum Zeitpunkt $t+\Delta t$".

$E_2(t+\Delta t)$ kann bei kleinem Δt praktisch nur auf die folgenden vier Arten realisiert werden:

(1) Zum Zeitpunkt t sind 3 Kunden im System; im Zeitraum $[t, t+\Delta t]$ kommt kein neuer Kunde, aber ein Kunde geht, weil seine Bedienung gerade beendet ist; die Wahrscheinlichkeit dafür ist

$$p_3(t)\cdot(1-\lambda\Delta t)\cdot\mu\Delta t$$

(2) Zum Zeitpunkt t ist ein Kunde im System; im Zeitraum $[t, t+\Delta t]$ trifft ein neuer Kunde ein, aber kein Kunde geht; die Wahrscheinlichkeit dafür ist

$$p_1(t)\cdot\lambda\Delta t\cdot(1-\mu\Delta t)$$

(3) Zum Zeitpunkt t sind 2 Kunden im System, und während der Zeit von t bis $t+\Delta t$ passiert gar nichts; die Wahrscheinlichkeit dafür ist

$$p_2(t)\cdot(1-\lambda\Delta t)\cdot(1-\mu\Delta t)$$

(4) Zum Zeitpunkt t sind 2 Kunden im System; im Zeitraum $[t, t+\Delta t]$ kommt ein neuer Kunde, und ein Kunde geht, weil seine Bedienung gerade beendet ist; die Wahrscheinlichkeit dafür ist

$$p_2(t)\cdot\lambda\Delta t\cdot\mu\Delta t.$$

$p_2(t+\Delta t)$, also die Wahrscheinlichkeit des Ereignisses "genau 2 Kunden im Wartesystem zum Zeitpunkt $t+\Delta t$" ist gleich der Summe dieser vier Wahrscheinlichkeiten. Nach geeigneter Zusammenfassung wird

$$p_2(t+\Delta t)=p_2(t)-(\lambda+\mu)p_2(t)\Delta t+\lambda p_1(t)\Delta t+\mu p_3(t)\Delta t+(\Delta t)^2\cdot[...]$$

also

$$\frac{p_2(t+\Delta t)-p_2(t)}{\Delta t} = -(\lambda+\mu)p_2(t)+\lambda p_1(t)+\mu p_3(t)+\Delta t[...].$$

Beim Grenzübergang $\Delta t \to 0$ ergibt sich daraus die Differentialgleichung

$$p_2'(t)=-(\lambda+\mu)p_2(t)+\lambda p_1(t)+\mu p_3(t).$$

Die übrigen Differentialgleichungen erhält man auf die gleiche Weise. Insgesamt ergibt sich für die $p_n(t)$ das folgende System von Differentialgleichungen:

(I) $p_o'=-\lambda p_o+\mu p_1$

(II) $p_n'=-(\lambda+\mu)p_n+\lambda p_{n-1}+\mu p_{n+1}$ $(n=1,2,\ldots)$

Um die im Fall $\mu>\lambda$ existierende stationäre Wahrscheinlichkeitsverteilung der Zahl n der Kunden (die sich im System befinden) zu gewinnen, setzt man hierin $p_n'=0$ für $n=0,1,2,\ldots$ und erhält

(I) $0=-\lambda p_o+\mu p_1$, also $p_1=p_o\frac{\lambda}{\mu}$

(II) Für $n=1$:
$0=-(\lambda+\mu)p_1+\lambda p_o+\mu p_2$, also (wegen (I)) $p_2=p_o(\frac{\lambda}{\mu})^2$
und für beliebiges n:
$p_n=p_o(\frac{\lambda}{\mu})^n$.

Aus der Beziehung $\sum_{n=0}^{\infty}p_n=p_o\cdot\sum_{n=0}^{\infty}(\frac{\lambda}{\mu})^n=1$ folgt $p_o=1-\frac{\lambda}{\mu}$, und somit wird

$$p_n=(1-\frac{\lambda}{\mu})\cdot(\frac{\lambda}{\mu})^n,$$

d.h. daß - nach einer gewissen Anlaufphase - die Zahl n der Kunden im System geometrisch verteilt ist mit dem Parameter $\frac{\lambda}{\mu}$. Die durchschnittliche Zahl der dann im System befindlichen Kunden ist

$$\sum_{n=0}^{\infty}np_n=\sum_{n=0}^{\infty}n(1-\frac{\lambda}{\mu})\ (\frac{\lambda}{\mu})^n=\ldots=\frac{\lambda}{\mu-\lambda},$$

und die durchschnittliche Zahl der Kunden in der Warteschlange

ist

$$\sum_{n=1}^{\infty}(n-1)p_n=\ldots=\frac{\lambda^2}{\mu^2-\mu\lambda}.$$

Zahlenbeispiel:

(1) $\lambda=4$ [Kunden/Stunde], d.h. im Durchschnitt kommt alle 15 Minuten 1 Kunde.

(2) $\mu=5$ [Kunden/Stunde], d.h. im Durchschnitt beträgt die Bedienungszeit eines Kunden 12 Minuten.

Ergebnisse: $\frac{\lambda}{\mu-\lambda}=4$ (durchschnittliche Kundenzahl im System),

$\frac{\lambda^2}{\mu^2-\mu\lambda}=3.2$ (durchschnittliche Kundenzahl in der Warteschlange),

$p_o=1-\frac{\lambda}{\mu}=0.2$ (Wahrscheinlichkeit für "0 Kunden" im System, d.h. anteilige Leerlaufzeit des Bedienungsschalters).

4. Simulationsmodelle für Warteschlangen

4.1 Bemerkung zum Stochastikunterricht in der Sekundarstufe I

Im Stochastikunterricht der SI stehen die Begriffe und Methoden, die notwendig sind für die im vorangehenden Abschnitt skizzierte analytische Lösung des (M,M,1)-Bedienungsproblems, natürlich nicht zur Verfügung. Auch die Simulation des Bedienungsprozesses durch die Erzeugung von Folgen von exponentialverteilten Zufallszahlen kommt in der SI nicht in Frage (wenn die Folge (X_n) in $[0,1)$ gleichverteilt ist, so ist die Folge $(-\frac{1}{\lambda}\log(1-X_n))$ in $[0,\infty)$ exponentialverteilt mit dem Parameter λ).

Sehr fragwürdig erscheint es mir, im Stochastikunterricht der SI den auf den ersten Blick einfachen Begriff der Ankunftswahrscheinlichkeit einzuführen, um Warteschlangenprobleme behandeln zu können. Ich sehe eine Kollision mit den Zielen des Mathematikunterrichts in der SI, wenn der Lehrer wie folgt argumentiert: Eine Ankunftswahrscheinlichkeit von p=0.2 [Kunden/min] bedeutet erstens, daß im Durchschnitt alle 5 Minuten ein Kunde kommt, und zweitens, daß der Prozeß des Ankommens von Kunden simuliert werden kann durch Betrachtung einer Serie von aufeinanderfolgenden Zeitintervallen von z.B. der Länge $\Delta t=1$ Minute, wobei man in jedem dieser Zeitintervalle mit der Wahrscheinlichkeit 0.2 genau einen Kunden und mit der Wahrscheinlichkeit 0.8 gar keinen Kunden kommen läßt, oder durch Betrachtung von Zeitintervallen der Länge $\Delta t=0.5$ Minuten mit der jeweiligen Kundenankunftswahrscheinlichkeit 0.1 usw.

Daß man für "kleine" Zeitintervalle Δt mit einer zu Δt proportionalen Wahrscheinlichkeit p für die Ankunft genau eines Kunden rechnen darf - derart, daß $p/\Delta t$ die Ankunftsrate λ ist -, das ist eine Erkenntnis, die meines Erachtens das volle Verständnis der Exponentialverteilung voraussetzt und daher in der SI nicht möglich ist. Selbst der SII-Schüler wird schwer verstehen, inwiefern man für kleines Δt anstelle von

$p(t_z \leq \Delta t)=1-e^{-\lambda\Delta t}=\lambda\Delta t-\frac{(\lambda\Delta t)^2}{2!}+\ldots$ mit $\lambda\Delta t$ rechnen darf. Dem SI-Schüler wird der Lehrer geradezu verbieten, von einer Proportionalität einer Wahrscheinlichkeit zu einer anderen linear wachsenden Größe zu sprechen. Bezeichnet z.B. p_n die Wahrscheinlichkeit, beim gleichzeitigen Werfen von drei Würfeln in n solchen Würfen wenigstens einmal das Tripel (6,6,6) zu erhalten, so ist zwar $p_1=0.0046$, $p_2=0.0092$, $p_3=0.0138$, $p_4=0.0184$ (bei Rundung auf vier Nachkommastellen); dem SI-Schüler gegenüber aber wird der Lehrer betonen, daß keine Proportionalität vorliegt ($p_n=1-(1-\frac{1}{216})^n$).

4.2 Das Simulationsmodell M1

Zu den wichtigen Lernzielen des Stochastikunterrichts gehört es, stochastische Probleme durch Simulation lösen zu können. So läßt sich z.B. das Zufallsexperiment "Würfeln" leicht studieren, indem man eine Folge von Zufallszahlen $X_1, X_2, \ldots \in [0,1)$ durch die Transformation $X_i \rightarrow 1+INT(6*X_i)$ in eine Würfelzahlen-Folge verwandelt; ein Bernoulli-Experiment mit der Erfolgswahrscheinlich-

keit p wird bei Wahl einer Zufallszahl $X\in[0,1)$ durch INT(X+p) realisiert (es ist $INT(X+p)=1 \Leftrightarrow X\in[1-p,1)$, wobei das Intervall $[1-p,1)$ die Länge p hat!), und eine Bernoulli-Kette der Länge n mit der Einzelerfolgswahrscheinlichkeit p wird realisiert durch die Ausführung des Befehls

$$INT(X_1+p)+INT(X_2+p)+\ldots+INT(X_n+p),$$

wo $(X_1,\ldots,X_n)$ ein n-tupel von Zufallszahlen aus dem Intervall $[0,1)$ ist.

Da komplizierte Warteschlangenprobleme in der Praxis ohnehin durch Simulation gelöst werden, bietet es sich an, im Stochastikunterricht der SI gewisse einfache Bedienungsprozesse SI-angemessen zu simulieren.

Das im folgenden dargestellte Modell M1 soll zur Simulation all jener Bedienungsprozesse dienen, in denen jeder der potentiellen Kunden über sein Auftreten als Kunde im Bedienungssystem für sich allein entscheidet, ohne sich mit anderen Kunden abzusprechen und ohne irgendwelche Ankunftszeitpunkte zu bevorzugen, und in denen die Kunden in einem bestimmten Zeittakt gruppenweise abgefertigt werden, aber jeweils nur bis zu einer Maximalzahl m. Beispiele: PKW's/Fähre, Museumsbesucher/Führung. Die beiden wesentlichen Modellannahmen lauten deshalb wie folgt:

1. Zu dem betrachteten Bedienungssystem gehört ein fester Kundenstamm S vom Umfang N. Die Zahl N ist groß im Verhältnis zu der gerade im Bedienungssystem befindlichen Kundenzahl, und zwar so, daß N praktisch stets als die Zahl der potentiellen Kunden (außerhalb des Bedienungssystems) angesehen werden kann.

2. Man betrachte eine Folge von aufeinanderfolgenden Zeitintervallen der einheitlichen Länge 1 ZE (1 Zeiteinheit). Jede der N "Personen" aus der Menge S entschließt sich in jedem dieser Zeitintervalle mit der gleichen kleinen Wahrscheinlichkeit p, das Bedienungssystem als Kunde aufzu-

suchen (das entspricht jeweils N unabhängigen Durchführungen eines Bernoulli-Experiments mit der Trefferwahrscheinlichkeit p, so daß im Durchschnitt $N \cdot p$ Kunden pro ZE_x das Bedienungssystem aufsuchen; wegen $\binom{N}{x} p^x (1-p)^{N-x} \approx \frac{(Np)^x}{x!} e^{-Np}$ ist der Kundenstrom nahezu ein Poissonstrom wie im Bedienungssystem (M,M,1)).

3. Am Ende eines jeden der betrachteten Zeitintervalle beginnt für eine Gruppe von höchstens m Kunden die gemeinsame Bedienung; die evtl. übrigbleibenden Kunden bleiben im Warteraum in der Warteschlange. Es gilt $m > N \cdot p$.

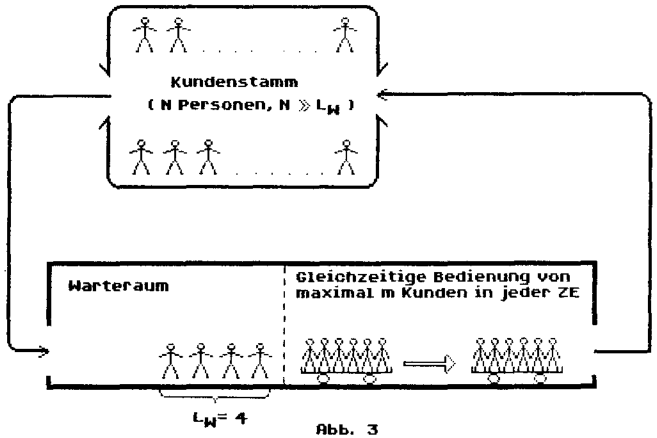


Abb. 3

Die Simulation wird für eine längere Folge von Zeitintervallen (z.B. k=1000 Intervalle) durchgeführt.

Der Simulationsalgorithmus enthält neben den Größen N (Umfang des Kundenstamms), p (Einzelerfolgswahrscheinlichkeit), m (Maximalzahl der je ZE bedienten Kunden), k (Anzahl der Simulationsintervalle) noch die Größen L_w (derzeitige Länge der Warteschlange) und SL_w (Summe aller einzelnen registrierten Warteschlangenlängen), X bezeichnet die durch die Bernoulli-Kette "ausgeloste" Zahl der neu eintreffenden Kunden.

Als Ergebnis wird SL_w/k, also die durchschnittliche Länge der Warteschlange ausgegeben. Die Wartezeit im Ankunftszeitintervall eines Kunden wird nicht mitgezählt (man könnte, wenn man wollte, diese im Durchschnitt mit einer halben Zeiteinheit berücksichtigen; das würde die Warteschlangenlänge um $\lambda/2$ erhöhen).

<u>Simulationsalgorithmus zum Modell M1:</u>

```
Eingabe: N,p,m,k                                   1)
Setze Lw: = 0, setze SLw: = 0
Für I=1 bis k wiederhole
    Setze X: = 0
    Für J=1 bis N wiederhole
        Wähle Zufallszahl Z aus [0,1)
        Setze X: = X+INT(Z+p)
    Setze Lw: = max{o;Lw+X-m}
    Setze SLw:= SLw+Lw
Ausgabe: SLw/k
```

1) Als Vorbild diente die in (2) benutzte Form der Darstellung solcher Algorithmen.

Ein Probelauf auf dem Commodore 64 mit der Eingabe $N=30$, $p=\frac{1}{6}$ (also Ankunftsrate $N \cdot p=5$), $m=6$, $k=1000$ ergab als durchschnittliche Länge der Warteschlange den Wert 1.4.

4.3 Das Simulationsmodell M2

Ein weiteres, sehr einfaches Simulationsmodell ergibt sich, wenn man die spezielle Annahme macht, daß die Gesamtzahl n der Kunden, die in dem für die Simulation ausgewählten Zeitintervall $[t_A, t_E)$ im Bedienungssystem eintreffen, im voraus bekannt ist.

Angenommen, ein Bedienungssystem wird an einem Tag innerhalb seiner Öffnungszeit $[t_A, t_E)$ insgesamt von n Kunden aufgesucht. Wenn jeder Kunde unabhängig von dem, was die übrigen Kunden tun, seine Ankunftszeit in dem Bedienungssystem wählt und wenn er darüberhinaus keine Präferenz für irgendwelche Zeitpunkte innerhalb von $[t_A, t_E)$ hat, so ist es das einfachste, wenn man sich seine Wahl eines Ankunftszeitpunktes wie das Ziehen einer Zufallszahl aus dem Zahlenintervall $[t_A, t_E)$ vorstellt. Für n Kunden sind n Zufallszahlen aus $[t_A, t_E)$ zu ziehen, und durch Ordnen dieser Zufallszahlen nach der Größe ergibt sich die Folge der Kundenankunftszeiten. Die Bedienungszeit kann konstant sein - sie ist dann sinnvollerweise kleiner als die durchschnittliche Zwischenankunftszeit $(t_E - t_A)/n$ - oder sie kann variabel gewählt werden, wobei die Folge der Bedienungszeiten nach demselben Verfahren wie die Folge der Zwischenankunftszeiten gewonnen werden kann, also als Differenzen von aufeinanderfolgenden, der Größe nach geordneten, in $[t_A, t_E)$ rechteckverteilten Zufallszahlen.

In dem folgenden Simulationsalgorithmus wird mit konstanter Bedienungszeit BZ gearbeitet. Als Intervall $[t_A, t_E)$ ist das Intervall $[0,1)$ gewählt. A(I) bezeichne die Ankunftszeit und D(I) den Endzeitpunkt der Bedienung des Kunden Nr. I. Die Wartezeit des Kunden Nr. I ist D(I)-A(I)-BZ. Der Mittelwert $\overline{T}_w$

der Wartezeit ist dann

$$\frac{1}{n} \sum_{I=1}^{n} D(I)-A(I)-BZ,$$

und $n \cdot \overline{T}_w$, also die Wartezeitensumme selbst, stimmt überein mit $\overline{L}_w \; (t_E - t_A)$, ist also im vorliegenden Fall wegen $[t_E - t_A) = [0,1)$ gleich der mittleren Warteschlangenlänge $\overline{L}_w$. Der Simulationsalgorithmus beginnt mit dem Ziehen von n Zufallszahlen aus dem Intervall $[0,1)$ (entsprechend den Ankunftszeiten der n Kunden) und dem Ordnen dieser n Zufallszahlen nach der Größe.

Der erste Kunde wird sofort bedient, also D(1)=A(1)+BZ. Der Kunde mit der Nr. I wird entweder sofort bedient (D(I)=A(I)+BZ), nämlich wenn Kunde Nr. I-1 schon weg ist (D(I-1)≤A(I)), oder seine Bedienung beginnt zum Zeitpunkt D(I-1). Die zum Endzeitpunkt evtl. noch im System befindlichen Kunden werden alle noch bedient.

Eingabe: n, BZ	
Zieh n Zufallszahlen aus [0,1) und ordne sie nach der Größe; bezeichne die geordneten Zufallszahlen der Reihe nach mit A(1), A(2),...,A(n)	
Setze D(1):=A(1)+B(Z), setze SW=0	
Für 1 von 2 bis n wiederhole	
	Wenn D(I-1)>A(I), so setze D(I)=D(I-1)+BZ, andernfalls setze D(I)=A(I)+BZ
	Setze SW:=SW+D(I)-A(I)-BZ
Ausgabe: SW	

Ein Probelauf auf dem Commodore 64 für die Eingabe n=200, BZ=0.004 ergab die mittlere Warteschlangenlänge SW=1.6. Die Zahl n übernimmt in diesem Modell die Rolle der Kundenankunftsrate λ (Kundenzahl in der Zeit 1). Die durchschnittliche Zwischenankunftszeit beträgt $\frac{1}{n}$.

5. Anmerkungen zu einer Unterrichtseinheit "Warteschlangen" im Stochastikunterricht der Sekundarstufe I

5.1 In den vorgestellten Simulationsmodellen für Bedienungsprozesse wird jeweils der mittlere Wert der zeitabhängigen Warteschlangenlänge $L_w(t)$ in einem Zeitintervall $[t_A, t_E]$, also im Grunde genommen das Integral

$$\bar{L}_w = \frac{1}{t_E - t_A} \int_{t_A}^{t_E} L_w(t)\,dt$$

bestimmt. Im Modell M1 läßt sich wegen der Betrachtung von k gleich langen Teilintervallen, in denen L_w jeweils konstant ist, $\bar{L}_w$ als gewöhnliches arithmetisches Mittel

$$\bar{L}_w = \frac{1}{k} \sum_{i=1}^{k} L_w(t_i)$$

berechnen, und im Modell M2 wird benutzt, daß $\bar{L}_w \cdot (t_E - t_A)$ als Wartezeitensumme aller in $[t_A, t_E]$ angekommenen Kunden interpretiert werden kann (stellvertretend für alle Kunden warten $\bar{L}_w$ Kunden während der ganzen Zeit). Mittelwerte zeitabhängiger Größen sind den Schülern der Sekundarstufe I wenig vertraut. Da man, wenn man eine Unterrichtseinheit "Warteschlangenmodelle" plant, vor dem Beginn der Simulation ohnehin eine Reihe von neuen Begriffen wie Ankunftsrate, Zwischenankunftszeit usw. erarbeiten muß, ist es günstig, mit der Behandlung einfacher deterministischer Warteschlangenmodelle zu beginnen, die es erlauben, neben den erwähnten Grundbegriffen auch den Begriff der mittleren Warteschlangenlänge zu erarbeiten und einzuüben.

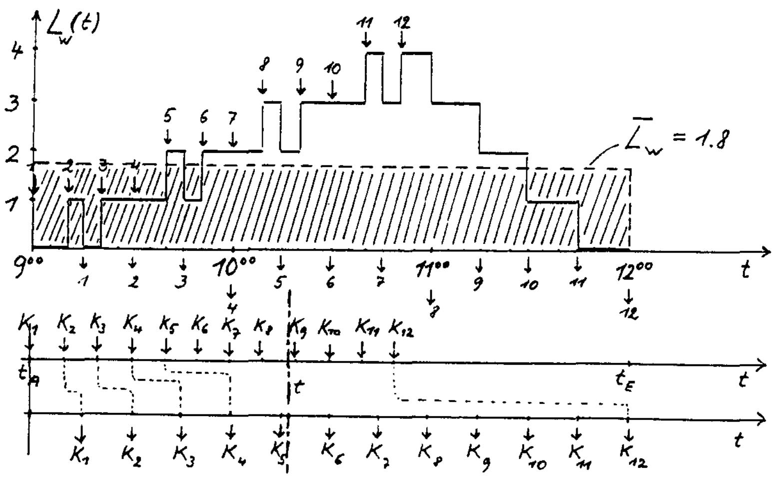
Dazu als Vorschlag die beiden folgenden deterministischen Warteschlangenmodelle:

Modell M_O:

Die Kunden kommen einzeln in konstanten Zeitabständen τ_a (Begriff "Zwischenankunftszeit"), die Bedienungszeit τ_b ist konstant.

Beispiel zum Modell M_O:

Ein Zahnarzt (Sprechzeit 9 bis 11 Uhr) hat die Patienten einzeln bestellt zu 9.00, 9.10, 9.20,...,10.50 Uhr. Die Behandlung pro Patient beträgt τ_b=15 min.



Arbeitsaufträge und Fragen an die Schüler:

1. Wann verläßt der 1. Patient, der 2. Patient,... die Praxis? (Antwort: $D(k)=t_A+k\cdot\tau_b$ mit t_A=9h, τ_b=0.25h).

2. Trage die Länge L_w der Warteschlange in Abhängigkeit von der Zeit $t\in[t_A, t_E]$ in ein Koordinatenkreuz ein und ermittle $\bar{L}_w$; gib für $\bar{L}_w$ eine geometrische Deutung in dem Graphen (Antwort: Skizze des Graphen der Treppenfunktion L_w: $[9;12] \to N_O$, $\bar{L}_w$=1.8 Personen,...)

3. Der k-te Patient wartet bis zum Beginn seiner Bedienung die Zeit $(\tau_b-\tau_a)\cdot(k-1)$. Ermittle die durchschnittliche Wartezeit aller Patienten im Zeitraum zwischen 9 und 12 Uhr. (Antwort: ....=27.5 min.).

4. Mache Dir klar, daß $\bar{L}_w\cdot(t_E-t_A)$ die Wartezeitensumme für alle Patienten im Zeitraum $[t_A, t_E]$ ist.

Modell M_o':

Gruppen von jeweils n Kunden treffen in konstanten Zeitabständen τ_a im Warteraum des Bedienungssystems ein; die Bedienungszeit ist konstant.

Beispiel zum Modell M_o':

Das Gesundheitsamt bestellt zu jeder vollen Stunde 10 Personen zur Röntgenuntersuchung. Die Personen werden einzeln nacheinander untersucht; eine Untersuchung dauert 5 Minuten. (Arbeitsaufträge und Fragen an die Schüler ähnlich wie zu dem vorangegangenen Modell M_o).

5.2 Behandlung der Simulationsmodelle M1 und M2

Wenn die Kunden nicht mehr, wie in den Modellen M_o und M_o', nach Vorbestellung, sondern nach persönlichem Entschluß in nicht vorhersagbarer Weise in dem Bedienungssystem eintreffen, so muß man zu stochastischen Warteschlangenmodellen übergehen. Daß die Kundenströme i.a. wirklich so aussehen, wie sie im Modell M1 durch ein Zufallsexperiment realisiert werden, das kann durch die Beobachtung realer Kundenströme untermauert werden (die Schüler bestimmen z.B. am Eingang des Postamts zwischen 10 und 11 Uhr die Zahl der pro Minute eintreffenden Kunden, oder der Lehrer beschafft aus Beobachtungen stammende Daten über Kundenströme). Für das "Auslosen" der in jeder Zeiteinheit im Bedienungssystem eintreffenden Kunden durch eine Bernoulli-Kette wird am besten ein PC benutzt. Gewisse Bernoulli-Ketten lassen sich aber auch leicht durch Würfelexperimente realisieren, z.B. im Fall N=30, $p=\frac{1}{6}$ durch das gleichzeitige Werfen von 30 kleinen Spielwürfeln und das Zählen der dabei erhaltenen Sechsen. Dazu gehört die Ankunftsrate $\lambda=N\cdot p=5$ (Kunden pro ZE).

Die Warteschlangenlänge L_w hat eine relativ große Streuung; so ist z.B. im (M,M,1)-Bedienungssystem die Streuung von L_w grösser als $\bar{L}_w$. Daher sind bei der Simulation lange Laufzeiten (große Intervallanzahl k) notwendig, um kurze Vertrauensintervalle für $\bar{L}_w$ zu gewinnen. Auf große Genauigkeit der Simulationsergebnisse kommt es aber im Unterricht nicht so sehr an wie dadarauf, daß das Prinzip des Lösens eines stochastischen Problems durch Simulation deutlich wird.

Das Simulationsmodell M2 greift die naive Vorstellung der Schüler vom Vorgang des zufälligen Wählens eines Ankunftszeitpunktes durch den einzelnen Kunden auf: er erfolgt wie das Wählen einer Zufallszahl aus einem Zahlenintervall. Im Unterricht liegt folglich der Schwerpunkt der Untersuchungen nicht auf der Verteilung der Zwischenankunftszeiten, sondern auf der Entwicklung des Algorithmus für das Sortieren der Kundenankunftszeitpunkte und des Algorithmus für die rekursive Berechnung der Zeitpunkte, zu denen die Kunden das System nach ihrer Bedienung verlassen.

Literatur

(1) BECKER, G. u.a.: Anwendungsorientierter Mathematikunterricht in der Sekundarstufe I. Bad Heilbrunn 1979. (S. 51-77: Unterrichtseinheit "Warteschlangen").

(2) DEUTSCHES INSTITUT FÜR FERNSTUDIEN (Hrsg.): Wahrscheinlichkeitsrechnung und Statistik unter Einbeziehung von Taschenrechnern SR3. Tübingen 1984.

(3) ENGEL, A.: Wahrscheinlichkeitsrechnung und Statistik, Band 2

(4) PANICO, J.A.: Queuing Theory. Prentice-Hall, Inc., Englewood Cliffs, N.J. 1969.