Hierbei ist es allerdings so, dass Volker Guiard durchaus die theoretische Möglichkeit eines solchen Artefakts als gegeben sieht. Aufgrund der Schwankungen der Umlaufzeit der Lokomotive hält er ein solches Artefakt aber in der Praxis für ausgeschlossen. Diese theoretische Möglichkeit beruht auf der unterschiedlichen Umlaufzeit zwischen Außen- und Innenkreis der Modelleisenbahn (siehe Tabelle 2 seines Artikels). Jene Erkenntnis lässt sich unmittelbar in eine Verbesserung des Versuchsaufbaus umsetzen: Allerdings würde hierzu ein weiterer Sensor benötigt, was den Hard- und Softwareaufwand erhöhen würde. Im Versuchsaufbau von Hagel und Tschapke wird vor der Weiche mit dem Reedkontakt 1 die Abfrage einer Zufallszahl ausgelöst, die die Weiche steuert. Die Berechnung der Zufallszahlen läuft anschließend fortlaufend weiter, so dass sich für Innen- und Außenkreis eine unterschiedliche mittlere Anzahl berechneter Zufallszahlen bis zum nächsten Überfahren von Reedkontakt 1 ergibt.
Die Verbesserung des Versuchsaufbaus würde darin bestehen, dass nach Überfahren von Reedkontakt 1 die Berechnung der Zufallszahlen unterbrochen wird. Zusätzlich wird ein Reedkontakt 3 unmittelbar hinter der Zusammenführung von Innen- und Außenkreis benötigt: Nach Überfahren dieses zusätzlichen Kontaktes wird die Berechnung wieder aufgenommen. Dies würde bedeuten, dass die Differenz D zwischen Außen- und Innenkreis gleich 0 wäre; d.h. die von Volker Guiard identifizierte theoretische Möglichkeit eines Artefaktes würde durch einen solchen Versuchsaufbau ausgeschlossen. Zufallszahlen würden dann nur bei der Fahrt auf dem Teilstück der Eisenbahnstrecke berechnet, das Außen- und Innenkreis gemeinsam ist.

Diskussion der Zeitraffersimulation

Unabhängig von dieser wertvollen Erkenntnis aus Volker Guiards Beitrag möchte ich dennoch in Frage stellen, ob seine Analyse hinreichend ist, um durch den Pseudozufallszahlengenerator verursachte Artefakte bei der eigentlich interessierenden Messung auszuschließen. Ein Problem bei der Analyse ist (wie Volker Guiard richtig bemerkt), dass die Umlaufzeit der Lokomotive schwankt. Dies ist zum einen durch den Verschleiß (verstärkt durch den Umpolungsvorgang) sowie Erwärmung des Antriebs und Schwankungen der Spannungsversorgung während des Versuchslaufs bedingt; zum anderen ist es natürlich so, dass der PC nur zu diskreten Zeitabständen Δt eine neue Zufallszahl liefert, während auch ohne Verschleiß die Umlaufzeit der Lokomotive im allgemeinen Fall kein Vielfaches von Δt sein wird (was bereits ein gewisses „Quantisierungsrauschen“ bedingen würde). Aus diesen Gründen wäre es interessant gewesen zu untersuchen, ob der hier verwendete Pseudozufallszahlengenerator dazu neigt, mit einer gewissen Regelmäßigkeit längere Sequenzen von Nullen bzw. Einsen zu erzeugen. Eine Angabe der maximalen Länge einer Sequenz von Nullen bzw. Einsen sowie der Länge der Sequenz mit der größten Häufigkeit innerhalb einer Periode wäre hierzu hilfreich. Bei längeren Sequenzen ist die Wahrscheinlichkeit höher, auch bei nicht-idealer Synchronisation einen benachbarten, gleichen Wert abzufragen. Was die Regelmäßigkeit angeht ist diese Information zwar auch in den von Volker Guiard berechneten mittleren Autokorrelationswerten zum Teil enthalten, jedoch nicht was die Länge von Sequenzen gleicher Werte angeht. Allerdings zeigt die von Volker Guiard durchgeführte Zeitraffersimulation des Eisenbahnexperiments, dass solche denkbaren Effekte vermutlich keine Rolle spielen.

In der Zeitraffersimulation ist es Volker Guiard für bestimmte Parameterkombinationen gelungen, signifikante Abweichungen von der Erwartung zu erzeugen, die als akausale Korrelation fehlinterpretiert werden könnten. Diese Parameterkombinationen weichen jedoch von denen des Realexperiments so weit ab, dass er die von ihm untersuchten Spezifika des Pseudozufallszahlengenerators als Ursache für die im Realexperiment beobachtete akausale Korrelation ausschließt. Wenn Simulation und Realexperiment bei gleichem Pseudozufallszahlengenerator voneinander abweichende Ergebnisse liefern, heißt das aber zunächst nur, dass das Umlaufverhalten der Lokomotive in der Simulation unzureichend nachgebildet wurde. Dies ist völlig unabhängig davon, ob man der Hypothese von Hagel und Tschapke folgt, dass die Lokomotive über einen Selbsterhaltungstrieb verfügt, oder ob man vermutet, dass ein systematischer Fehler im Versuchsaufbau zur beobachteten akausalen Korrelation führt.
Für das Umlaufverhalten hat Volker Guiard zwei konstante mittlere Umlaufzeiten für Innen- und Außenkreis angenommen sowie einen konstanten Variationskoeffizienten, der für beide Strecken gleich ist. Hierfür hat er Werte verwendet, die sich auf Angaben von Hagel und Tschapke stützen, sowie Werte eingesetzt, die den von ihm untersuchten Effekt begünstigen sollen. Dies ist jedoch ein sehr stark vereinfachtes Modell. Bei der realen Lokomotive werden sich sowohl die mittlere Umlaufzeit als auch der Variationskoeffizient im Laufe des Experiments infolge von Verschleiß sowie Erwärmung des Antriebs verändern (Schwankungen der Spannungsversorgung dürften hierbei keinen Einfluss haben). In Abbildung 1 der Stellungnahme von Hagel und Tschapke im Anschluss an die Kommentare zum Originalartikel erkennt man eine gewisse Schiefe der Verteilung: Ausgehend vom Mittelwert 415000 fällt die Häufigkeit zu größeren N hin steiler ab als zu kleineren N. Eine mögliche Ursache hierfür ist die zeitliche Veränderung von mittlerer Umlaufzeit und Variationskoeffizient während der hier gefahrenen 10000 Runden.
Man könnte vermuten, dass der durch die vereinfachte Modellierung des Umlaufverhaltens der Lokomotive gemachte Fehler sich bei der Vielzahl simulierter Experimente (1000 anstelle von 10) wieder aufhebt. Gerade die Nichtübereinstimmung von Simulation und Realexperiment zeigt aber, dass das nicht der Fall ist. Ob dies darauf zurückzuführen ist, dass für die Schwankung der Umlaufzeit eine unzutreffende Verteilungsfunktion angenommen wurde oder ob die Schwankungen im Realexperiment gar nicht zufällig sind, sondern durch einen Selbsterhaltungstrieb der Lokomotive bedingt sind, ist dabei völlig unerheblich.

Messergebnisse der Realexperimente

An dieser Stelle sei nochmals daran erinnert, dass die von Hagel und Tschapke berichtete akausale Korrelation nicht regelmäßig auftritt. Für die erste Versuchsreihe am IPP Köln wurde eine (signifikante) Abweichung von 2,8 σ beobachtet. Bei einer Replikation am IPP Köln wurde bei verdoppelter Geschwindigkeit eine (nicht-signifikante) Abweichung von 1,5 σ beobachtet. In Anhang 2 wird von einer Replikation am IPP Genf mit etwas verminderter mittlerer Umlaufzeit (10 statt 11,5 Sekunden) berichtet, bei der eine (nicht-signifikante) Abweichung von 1,7 σ beobachtet wurde. Bei einer weiteren Replikation am IGPP Freiburg wurde beim vorzeitigen Abbruch des Experiments (infolge Lokomotivschadens) eine signifikante Abweichung von mehr als 3 σ beobachtet; leider liegen hier keine Angaben für die mittlere Umlaufzeit vor. Diese Informationen legen den Verdacht nahe, dass das Auftreten

der akausalen Korrelation von der mittleren Umlaufzeit der Lokomotive abhängt. Der von Hagel und Tschapke berichtete Einfluss einer unbeabsichtigten, geringfügigen Gleistrennung (die selbstverständlich auch die Umlaufzeit beeinflusst), kann ebenfalls als Indiz für diese Hypothese gewertet werden (während Hagel und Tschapke natürlich die dabei beobachtete Entwicklung von $S(n)$ im Sinne ihrer Hypothese interpretieren). Besonders auffallend sind hier die beiden Ergebnisse vom IPP Köln, bei denen eine Halbierung der mittleren Umlaufzeit auch ungefähr zu einer Halbierung der Abweichung vom Erwartungswert geführt hat. Eine Halbierung der mittleren Umlaufzeit bedeutet natürlich auch eine Halbierung der Anzahl der Pseudozufallsgeneratoraufrufe. Sofern mit dem Pseudozufallsgenerator ein systematischer Fehler verknüpft ist, wird dessen Einfluss ebenfalls halbiert.
Falls tatsächlich eine solche Abhängigkeit zwischen dem Auftreten der akausalen Korrelation und der mittleren Umlaufzeit besteht, liegt es nahe, die Ursache in der Synchronisation zwischen Lokomotive und Pseudozufallszahlengenerator zu vermuten. Verblüffend ist hierbei allerdings, dass – sofern die akausale Korrelation auftritt – die Abweichung von der Erwartung immer in die gleiche Richtung erfolgt. Der von Volker Guiard identifizierte Effekt besitzt aber gerade diese Eigenschaft, dass unabhängig von der Kodierung der Innenkreis bevorzugt würde (siehe Gleichung (10) und den folgenden Absatz in seinem Beitrag.)

Verwendung eines Pseudozufallsgenerators

Der Beitrag von Volker Guiard erhöht durch Angabe des exakten Pseudozufallszahlenalgorithmus sowie durch die Untersuchung gewisser statistischer Eigenschaften erheblich die Transparenz des von Hagel und Tschapke durchgeführten Experiments. Trotz seines Fazits möchte ich die Verwendung eines Pseudozufallszahlengenerators weiterhin kritisieren. Bei dem hier betrachteten Pseudozufallszahlengenerator ist $M=2^{24}=16777216$. In der ursprünglichen Versuchskonfiguration wurde der Pseudozufallszahlengenerator im Mittel 60000 mal pro Runde aufgerufen (die Stellungnahme von Hagel und Tschapke gibt für eine optimierte Software auf einem schnelleren PC einen Mittelwert von 415000 an). Nach 279,62 $(=M/60000)$ Runden ist also der Pseudozufallszahlenvorrat erschöpft und die gleiche Sequenz wird wieder von neuem erzeugt. Ein Versuch erstreckt sich über 10 mal 10000 Runden. Innerhalb dieser jeweils 10000 Runden wird also ungefähr 36 mal die gleiche Sequenz von Pseudozufallszahlen verwendet (bei optimierter Software sogar noch häufiger). An dieser Stelle möchte ich auch nochmals daran erinnern, wie gering die Abweichung von der Erwartung sein muss, um den im ersten Experiment erzielten Wert für $S(n) \approx 2{,}8\sigma \approx 886$ zu erreichen. Geht man von der idealen Erwartung $S(n)=0$ aus, müsste bei lediglich 443 der insgesamt 100000 Runden bzw. nur etwa 44 mal innerhalb eines Teilexperiments von 10000 Runden der Innenkreis bevorzugt werden, obwohl eigentlich der Außenkreis „dran“ ist.
Solche groben Betrachtungen sind natürlich keine exakte mathematische Analyse, zeigen aber, dass die gesamte Anordnung gegenüber systematischen Fehlern sehr empfindlich ist. Wie bereits erwähnt zeigen die vorliegenden Messergebnisse, dass zwischen Auftreten der akausalen Korrelation und mittlerer Umlaufzeit eine Abhängigkeit besteht. Dies legt nahe, dass die akausale Korrelation durch einen systematischen Fehler verursacht wird. Ich halte daher die Verwendung eines Pseudozufallszahlengenerators mit einer Periodendauer, die klein gegenüber der Anzahl benötigter Zufallszahlen ist, für außerordentlich ungeeignet. Es

ist zwar auffällig, dass innerhalb eines Teilexperiments die notwendige Bevorzugung des Innenkreises (ca. 44) in der gleichen Größenordnung liegt wie die Wiederverwendung der Zufallszahlenreihe (ca. 36), was aber nichts heißen muss. Schließlich zeigt die Replikation am IPP Genf mit nicht-signifikantem Ergebnis bei leicht verminderter Umlaufzeit (10 statt 11,5 Sekunden), dass noch weitere (unbekannte) Voraussetzungen für das Zustandekommen der akausalen Korrelation erfüllt sein müssen.

Fazit

Volker Guiard hat mit seinem Beitrag dankenswerterweise die Möglichkeiten zur Analyse des Experiments von Hagel und Tschapke wesentlich erweitert. Um das Experiment als Simulation nachzubilden, fehlen jedoch noch wesentliche Daten über das Umlaufverhalten der Lokomotive in Abhängigkeit von der Zeit. Die von Hagel und Tschapke zur Verfügung gestellte Darstellung der Häufigkeit von Pseudozufallszahlengeneratoraufrufen enthält diese Information leider nicht. Stattdessen sollte hierzu lieber eine Darstellung gewählt werden, die ähnlich wie für *S(n)* in den Abbildungen 5 und 6 ihres Beitrags die zeitliche Entwicklung der Aufrufe pro Runde im Laufe des Experiments enthält. Falls den Autoren diese Daten vorliegen, wäre es sehr erfreulich, wenn sie eine solche Darstellung nachliefern könnten bzw. für zukünftige Experimente die Versuchssoftware so erweitern, dass diese Daten zur Verfügung gestellt werden können.

Die zugrunde liegende Frage, ob es sich bei dem von Hagel und Tschapke beobachteten Effekt um eine veritable akausale Korrelation oder ein Artefakt infolge eines systematischen Fehlers im Versuchsaufbau handelt, kann nur im Realexperiment entschieden werden. Deswegen möchte ich anregen, den Versuchsaufbau wie eingangs vorgeschlagen um einen dritten Sensor zu erweitern und die Software entsprechend anzupassen. Hierdurch wird die von Volker Guiard identifizierte Möglichkeit eines Artefakts in jedem Fall ausgeschlossen. Dieser Aufwand sollte unbedingt betrieben werden, auch wenn das Auftreten dieses Effekts von Guiard als außerordentlich unwahrscheinlich angesehen wird.

Last but not least möchte ich darauf drängen den Pseudozufallszahlengenerator gegen eine andere Zufallszahlenquelle auszutauschen. Falls bei ansonsten gleichen Bedingungen der Effekt nach einem solchen Austausch erhalten bleibt, könnte der Pseudozufallszahlengenerator als Verursacher mit großer Wahrscheinlichkeit ausgeschlossen werden. Falls nicht, müsste man von einem Artefakt ausgehen, das womöglich durch eine andere als die hier diskutierte Ursache bedingt wäre.

Literatur

Hagel, J.; Tschapke, M. (2002): Zum experimentellen Nachweis akausaler Korrelationseffekte in unbelebten Systemen. *Zeitschrift für Anomalistik* 2, 6-31.

ECKHARD ETZOLD

Pseudo-RNG durch echten Zufallsgenerator ersetzen

Schon in den ersten Kommentaren zur Arbeit von Hagel und Tschapke (2002) wurde vermutet, dass der verwendete Pseudo-RNG die eigentliche Ursache der Effekte sein könnte. Guiards Aufsatz demonstriert zunächst einmal, welch hoher experimenteller und formalistischer Aufwand betrieben werden muss, um einer sich recht schnell einstellenden Vermutung von Artefakten als Ursache dieser signifikanten Effekte auf den Grund zu gehen. Guiard übernimmt zunächst unhinterfragt die Definitionen von Hagel und Tschapke:

B = -1, falls 0 < z < 0,5 (=> Innenkreis)
B = 1, falls 0,5 < z < 1 (=> Außenkreis).

So sind alle B's entweder +1 oder -1. Hagel und Tschapke verwenden als Ergebnis eine Größe S(n) (n=Anzahl der gefahrenen Runden), die die Differenz D(n) zwischen der Anzahl der auf dem Innenkreis gefahrenen Runden und der Anzahl der auf dem Außenkreis gefahrenen Runden darstellen soll. Tatsächlich definieren sie S(n) aber wie folgt:

$$S(n) = \sum_{k=1}^{n} \left[B_{innen}(k) - B_{aussen}(k) \right]$$

Daraus würde sich ergeben, dass S(n) tatsächlich das Doppelte der Differenz D(n) darstellt. Die Berechnungen für die Standardabweichung führen Hagel und Tschapke dagegen für S(n)=D(n) durch. Christian Drohm, der für mich den Java-Programmcode für eine Computersimulation des Eisenbahnversuchs erstellte, wies mich schon Anfang November 2001 auf diesen Fehler hin, und eine diesbezügliche Anfrage beantwortete Prof. Dr. Hagel in dem Sinne, dass die Definition fehlerhaft war („... es fehlt tatsächlich der Faktor 1/2 in der Definitionsformel für S"). Meiner Bitte an ihn, diesen Fehler in der *Zeitschrift für Anomalistik* noch zu berichtigen, wurde nicht entsprochen. Ein schriftliches Erratum mit einer korrekten Definition erschien jedenfalls bis jetzt nicht. Unter der hieraus resultierenden Annahme, dass die im Ausgangsartikel beschriebene Definition für S(n) die maßgebliche sei, würden sich die Ergebnisse halbieren, und die dann verbleibenden Werte zwischen 0,7 und 1,4 Standardabweichungen für die abgebildeten Kumulationskurven würden dann keiner Erklärung mehr bedürfen.

Die Sache verkompliziert sich aber noch weiter: In dem Basic-Programm, das den Bahnumlauf steuert und die Ergebnisdaten sichert, werden beim Start die Werte für den Bahninnenlauf und den Bahnaußenlauf (hier als „ioben" und „iunten" definiert) korrekt auf 0 gesetzt. Später folgt dann (jetzt allerdings korrekt berechnet):

```
if r < 0.5d0 then
  iunten=iunten+1
else
  ioben=ioben+1
end if
```

Das heißt, die Definition für B(Innenkreis) und B(Außenkreis) muss korrekt heißen:

B = 0, falls 0 < z < 0,5 (=> Innenkreis)
B = 1, falls 0,5 < z < 1 (=> Außenkreis).

Und dann kann auch die Formel $\sigma = \sqrt{n}$ gelten für die Berechnung der Standardabweichung. Unklar bleibt nun, wie die Kumulationskurven erstellt wurden. Liegen ihnen die im Artikel genannten (falschen) Formeln zugrunde oder gehen sie auf die richtig errechneten Werte des QBasic-Programms zurück? Ich möchte letzteres hier zugunsten der Autoren annehmen.
Guiards Ergebnis besagt nun, dass die zeitlichen Zufallsschwankungen in der Umlaufperiode mögliche Artefakte im Pseudo-Zufallszahlenstrom derart verwischen, dass dadurch die beobachteten signifikanten Bevorzugungen nicht erklärt werden können. Diesem möchte ich auf einem anderen Wege noch einmal nachspüren:
Da der Eisenbahnversuch bisher leider nur mit hohem mechanischen Aufwand realisiert wurde und dadurch bedingt ein Versuchslauf mit 100000 Perioden weit über eine Woche dauerte, lag die Überlegung nahe, den Versuch auf dem Computer zu simulieren, um einerseits seine Grundmechanismen genauer zu studieren, andererseits aber auch durch die hohe Rechengeschwindigkeit die Versuchsdauer erheblich auf wenige Minuten zu reduzieren. Damit war die Hoffnung verbunden, dass sich der von Hagel und Tschapke behauptete anomale Effekt auch im Computer erzeugen lasse. Für die Realisierung konnte ich Christian Drohm gewinnen, der den Programm-Quelltext in Java erstellte.[1]
Der Eisenbahnversuch von Hagel und Tschapke (2002) besteht zunächst aus einem Schwingkreis (=Bahnumlauf), dessen Periode nach jedem Zyklus durch eine Weiche in eine längere oder kürzere Schwingungsperiode umgeschaltet werden kann. Programmiertechnisch entspricht das einem Oszillator, der auf zwei verschiedene Grundfrequenzen umschaltbar ist. Der Umschalter selbst wird von einem Abfragemechanismus gesteuert, der auch als Oszillator begriffen werden kann und mit einer mehrtausendfachen Abtastrate pro Umlaufperiode einen Pseudo-RNG aufruft. Übertragen auf eine Computersimulation bedeutet das die Realisierung von zwei voneinander unabhängig laufenden Rechteck-Oszillatoren. Der erste ist ein niedrigfrequenter Oszillator (LF) mit einstellbarem Tastverhältnis, der zwischen zwei verschiedenen Grundfrequenzen („high LF" und „low LF") umgeschaltet werden kann. Der zweite ist ein höherfrequenter Oszillator (HF), der auf einer fest eingestellten Grundfrequenz läuft und zu dem Zeitpunkt, wenn LF das Ende einer Periode erreicht hat, die „Weiche" des LF entweder auf „high LF" und „low LF" umschaltet, je nachdem, ob HF sich gerade im Zustand 0 oder 1 befindet. Die Toleranz der Grundfrequenzen beider Oszillatoren ist einstellbar und kann durch ein Zufallsrauschen beeinflusst werden. Für die Zufallszahlenerzeugung wurde ein Java-Modul entwickelt, das aus dem Grundrauschen der PC-eigenen Soundkarte einen echten Zufallszahlen-Bitstrom gewinnt, ersatzweise konnte auch der saatzahlgesteuerte Java-eigene Pseudo-REG verwendet werden. Der niedrigfrequente Oszillator (LF) konnte zwischen 4,84 Hz (nLow) oder 42,51 Hz (nHigh) umgeschaltet werden, der höherfrequente Oszillator (HF) war auf 10,6 kHz eingestellt. Die Werte für die Toleranz von LF, die durch den REG bestimmt wurden, waren vorwählbar. Die folgende Tabelle zeigt die Ergebnisse jeweils für 100000 Schwingungsperioden des LF (nHigh

[1] Eine Runtime-Version dieser Simulation stelle ich auf Anfrage gerne per E-Mail zur Verfügung.

= Anzahl der Perioden, die auf 42,51 Hz entfielen):

Tabelle 1: Computersimulation des Eisenbahnversuchs

Toleranz	S(nHigh)	**z-Wert (2-seitig)**
< 0,0001	-36667	-115,95
0,0001	-37240	-117,76
0,0002	-36821	-116,43
0,0004	-23890	-75,54
0,0006	-10377	-32,81
0,0008	-2974	-9,40
0,001	-291	-0,92
0,005	-59	-0,18
0,01	-167	-0,52
0,05	-121	-0,38

Es ist klar erkennbar, dass bevorzugt die niedrigere Grundfrequenz auftrat, und die Autokorrelationen zwischen den verschiedenen Frequenzen bei geringster Toleranz hochsignifikant ansteigen. Aber schon beim Wert von 0,001 Toleranz sind die Autokorrelationen so weit verrauscht, dass keine signifikanten Bevorzugungen einer der beiden niedrigfrequenten Grundfrequenzen mehr auftreten. Damit kann ich die von Volker Guiard gefundenen Ergebnisse zunächst bestätigen. Zugleich wird aber damit auch ein weiteres Problem markiert: Nach dem von Hagel und Tschapke behaupteten „Asymmetrie-Effekt im Nulllauf" hätte sich hier auch eine signifikante Bevorzugung einer der beiden Grundfrequenzen, nämlich jener, die das System weniger beansprucht, einstellen können. Dies war auch nicht der Fall. Aus all dem lassen sich nun folgende Schlussfolgerungen ziehen:

1. Es gelingt offenbar nicht, eine Simulation des Eisenbahnversuchs im Computer zu realisieren, die ähnlich anomale Ergebnisse liefert, wie sie scheinbar im realen Eisenbahnversuch auftreten.
2. Artefakte, bedingt durch Autokorrelationen im Zusammenspiel von Zufallsgenerator und variablen Periodendauern können weitgehend ausgeschlossen werden.
3. Solange formalistische und methodische Unklarheiten bestehen, kann noch nicht von einem gesicherten anomalen Effekt gesprochen werden.
4. Obwohl es zur Zeit so aussieht als ob der Pseudo-RNG nicht die Ursache für die beobachteten Effekte ist, möchte ich trotzdem um eine Wiederholung des Eisenbahnversuchs bitten, in der der Pseudo-RNG durch einen echten Zufallsgenerator ersetzt wird. Nur so wird es m.E. neben der Arbeit, die Volker Guiard schon geleistet hat, wirklich möglich sein, Autokorrelationen des Pseudo-RNGs als Ursache definitiv auszuschließen. Sollten hier vom Zufallsverhalten abweichende Effekte auftreten, so geht die Suche nach einer Erklärung in die nächste Runde.

WILFRIED KUGEL

Traue keinem Experiment mit dieser Software!

Der Beitrag von Volker Guiard ist sicher sehr verdienstvoll, was die theoretische Analyse des Algorithmus der Microsoft-QBASIC-Funktion RND betrifft, geht aber meines Erachtens leider am – rein computertechnischen – Kern des Problems vorbei.
Anlässlich einer Anfrage von Alan Vaughan (Mind Tech Systems, Los Angeles) führte ich 1993 eine Analyse des in Microsoft-QBASIC implementierten Zufallsgenerators RND durch. Alan Vaughan hatte seinerzeit ASW-Experimente durchgeführt, bei denen seine Versuchspersonen hochsignifikante Ergebnisse erzielten. Die targets wurden auf einem MAC unter BASIC mit der Funktion RND erzeugt. Ihm kamen jedoch Zweifel an der Qualität der so entstandenen Zufallszahlenfolgen (targets).
Das Ergebnis meiner empirischen Tests (mit Hilfe von SPSS, die ich anlässlich der vorliegenden Notiz wiederholte) war, dass sich am PC (beim MAC liegt die Situation ähnlich) unter DOS mit Microsoft-Programm QBASIC und der Funktion RND systembedingt nur unter sehr beschränkten Bedingungen (eigentlich fast nie) Zufallszahlenfolgen erzeugen lassen, was allerdings lange bekannt ist! (Bei dem genannten Programm Microsoft-QBASIC 4.5 handelt es sich um ein rares QBASIC-Update von 1997, das sich in den Grundfunktionen von der ursprünglichen Version nicht unterscheidet.) Folgende Probleme treten auf:
QBASIC ist ein DOS-Programm, welches man zwar unter WINDOWS in einem DOS-Fenster laufen lassen kann, das aber dafür nicht konzipiert wurde. Läuft der PC nicht *ausschließlich* unter DOS (welches kein Multitasking kennt), so ist unter WINDOWS der Einfluss gleichzeitig ablaufender Systemprozesse im PC auf die Funktion von RND kaum abschätzbar, denn das DOS-Fenster selbst unterliegt dem Multitasking. (Es ist für viele Anwendungen ein gravierender Nachteil von WINDOWS, dass das Multitasking prinzipiell nicht ausgeschaltet werden kann.)
Die Funktion RND (Zufallszahlengenerator) von QBASIC liefert eine Dezimalzahl zwischen 0 und 1 mit der (einfachen) Genauigkeit von 16 Dezimalstellen, die man auf verschiedene Art in Zufallszahlen umformen kann. Die Funktion RND wird durch die Anweisung RANDOMIZE [Startwert] initialisiert. [Startwert] wiederum wird gewöhnlich durch die Funktion TIMER initialisiert. Die Funktion TIMER ruft die Systemuhr auf und gibt (laut QBASIC-Hilfe-Funktion) die seit Mitternacht verstrichenen Sekunden im Format von 5 Stellen und 2 Dezimalen zurück. Tatsächlich können allerdings mehr (falsche) Dezimale abgerufen werden.

Fall 1

Gibt man die Anweisung RANDOMIZE TIMER gleich am Anfang des Programms und *nur ein einziges Mal*, so kann man unter Umständen durch wiederholte (nicht zu schnell aufeinander folgende) Aufrufe von RND eine passable Zufallszahlenfolge erhalten, da nur ein Eintrittspunkt in den Zufallszahlenalgorithmus von RND generiert wird. Entscheidend ist dabei allerdings auch, wie die Werte von RND weiter verarbeitet werden, denn es kann u.U. durch Rundungsfehler zu Regelmäßigkeiten in der Verteilung der erzeugten Zufallszahlenfolge kommen. Schließlich sind bei dieser Methode alle Glieder in der für ein Experiment benötigten Zufallszahlenfolge von Anfang an festgelegt, was oft nicht erwünscht ist.

Fall 2

Gibt man die Anweisung RANDOMIZE TIMER vor *jedem* Aufruf von RND, so führt dies zu Regelmäßigkeiten bei der Initialisierung von RND und damit *stets* zu Regelmäßigkeiten in der generierten „Zufallszahlen"-Folge. Die Ursachen sind folgende:
Die Funktion TIMER ruft die Systemuhr des PC auf und gibt die seit Mitternacht verstrichenen Sekunden (nebst Dezimalen) zurück. Wie empirisch überprüft wurde, sind *alle* Stellen der von TIMER zurückgegebenen Dezimalzahl stets *nicht-zufällig*. Es treten – offenbar aus technischen Gründen – bestimmte Ziffern in den Dezimalen überhaupt nicht auf! Dies führt zu Regelmäßigkeiten bei der Initialisierung von RND und damit zu sequentiellen Abhängigkeiten in der erzeugten „Zufallszahlen"-Folge. (Der Autokorrelations-Test von SPSS zeigt starke Signifikanzen.)
Die für Fall 2 entstehende Situation ist etwa damit vergleichbar, dass man zwar eine sehr gute Zufallszahlentabelle hat (hier den Algorithmus von RND), diese Tabelle aber nur an bestimmten Stellen aufschlagen werden kann, was zu Regelmäßigkeiten in der „Zufallszahlen"-Folge führt.

Ein weiteres Problem

Diese sequentiellen Abhängigkeiten werden noch wesentlich größer, wenn man RND kurz nacheinander mehrmals aufruft, wie in dem Beispielprogramm (für einen Würfel mit 6 Alternativen) von RND, das in der QBASIC-Hilfefunktion angegeben ist:

```
RANDOMIZE TIMER
x% = INT(RND * 6) + 1
y% = INT(RND * 6) + 1
PRINT "2 Würfe mit einem Würfel: Wurf 1 ="; x%; "und Wurf 2 ="; y%
```

Nun sind stets die Übergangswahrscheinlichkeiten gestört: y% hat signifikant zu häufig den selben Wert wie das zuvor generierte x%. Die Folgen beider Werte sind also korreliert!
Dieser Fehler tritt erstaunlicherweise auch dann auf, wenn RANDOMIZE nur einmal am Anfang des Programms aufgerufen wird (Fall 1)!! Die Ursache für diese Fehlfunktion von RND beim kurz aufeinander folgenden Erzeugen zweier „Zufallszahlen" konnte ich nicht klären. Es scheint sich um einen systembedingten Fehler zu handeln.

Alle genannten Fehler von RND treten im Fall von nur 2 Alternativen (wie von mir getestet) besonders stark in Erscheinung.

Fehlerhafte Systemuhr

Der Grund für die nichtzufälligen Dezimalstellen von TIMER ist, dass die interne Systemuhr eines PC (aber auch eines MAC) seit jeher und bis heute nicht vernünftig funktioniert. Sie wird vom Prozessor nur äußerst schleppend oder auch manchmal gar nicht aktualisiert. (Deshalb geht die Systemuhr auch stets nach.) Das Problem lässt sich auch durch schnellere Prozessoren *nicht* beheben!

Abhilfe kann nur dadurch geschaffen werden, dass eine externe und schnelle Uhr benutzt wird (z.B. als Steckkarte erhältlich), von der die aktuelle Zeit (mit doppelter Genauigkeit) abgerufen wird. Dann dürfen auch nur die letzten Dezimalstellen (ab 1/1000) zur Initialisierung von RANDOMIZE benutzt werden. Dafür muss eine eigene Funktion programmiert werden. Allerdings ist dann eigentlich RND überhaupt nicht mehr erforderlich, den man kann die höheren Dezimalstellen der externen Uhr (die bezüglich des Abrufzeitpunktes zufällig sind) direkt in Zufallszahlen umformen.

Wenn man echte (und verlässliche) Zufallszahlen erhalten will, ist es auf jeden Fall besser, von vornherein einen externen Geiger-Müller-Zähler zu benutzen (über eine A/D-Karte, die man auch von QBASIC aus ansprechen kann). Dazu braucht man bei entsprechend empfindlichem Zählrohr nicht einmal ein radioaktives Präparat, sondern die Weltraumstrahlung reicht aus. Zufallszahlengeneratoren, die auf Basis einer Rauschdiode arbeiten, sind nicht zu empfehlen, da sie (unter anderem) empfindlich gegen Einstrahlungen im Radiobereich sind und deshalb aufwendig abgeschirmt werden müssen.

Nachdem ich Alan Vaughan auf das Problem mit RND aufmerksam gemacht hatte, ließ er eine zusätzliche Randomisierungs-Prozedur schreiben, welche die Regelmäßigkeiten in den Zufallszahlen (targets) zerstörte. Die Ergebnisse erneuter ASW-Tests wurden nun zufällig! Damit stellten sich die ursprünglich hochsignifikanten Resultate als Artefakte heraus – die Versuchspersonen hatten also die in den „Zufallszahlen“ enthaltenen (sequentiellen) Regelmäßigkeiten erkannt und für ihre calls verwertet.

Es wurden hier einige Fehler der Funktion RND beschrieben. Vermutlich gibt es weitere. Die Frage von Volker Guiard „Kann ein Pseudozufallszahlengenerator eine akausale Korrelation vortäuschen?“ muss also bezüglich RND, RANDOMIZE und TIMER von QBASIC unbedingt mit „Ja“ beantwortet werden. Man sollte keinem experimentellen Ergebnis trauen, das unter Benutzung dieser Software entstand.

Stellungnahme des Autors

VOLKER GUIARD

Konkrete Quelle möglicher Artefakte weiterhin unklar

Ausgehend von meinem Befund, dass die unterschiedlichen Kreisdurchlaufzeiten zumindest im Prinzip eine Ursache für ein Artefakt sein könnten, schlägt Bernd Dürrer einen neuen Versuchsaufbau vor. Dies kann ich nur begrüßen. Anschließend nennt Bernd Dürrer weitere Störgrößen, die er als Artefaktursache verdächtigt, was ich jedoch nicht immer nachvollziehen kann. Es mag sicherlich sein, dass durch Verschleiß die Geschwindigkeit sinkt. Damit driftet die analog zur Abbildung 1 von Hagel und Tschapke (2002, S. 67) vorzustellende Verteilung der Umlaufzeiten in Richtung kleinerer Werte. Für den von mir gezeigten Artefakt ist es aber nur entscheidend, dass bei den beiden Verteilungen der Umlaufzeiten (für innen und außen) mit hoher Wahrscheinlichkeit zwei spezifische Umlaufzeiten R_{innen} und $R_{außen}$ auftreten, deren Autokorrelationen sich extrem unterscheiden. Wegen der hohen Varianz der Umlaufzeiten sind die Wahrscheinlichkeiten für konkrete Einzelzeiten jedoch sehr

gering. Aber auch, wenn die Wahrscheinlichkeiten R_{innen} und $R_{außen}$ zunächst recht hoch sein sollten, so kann dieses bei starkem Verschleißeffekt nicht während der gesamten Versuchszeit der Fall sein, womit durch Verschleiß dieser Artefakt eher verhindert wird.
Die Abfrage einer neuen Zufallszahl fällt in das Intervall zwischen zwei Zufallszahlerzeugungen, wobei stets die im Speicher vorhandene letzte Zufallszahl abgerufen wird. Das hat aber nichts mit einem „Quantisierungsrauschen" zu tun. Durch die Verwendung der jeweils vorhandenen Zufallszahl wird lediglich bewirkt, dass die kontinuierliche Verteilung der Umlaufzeiten in eine diskrete Verteilung der (für unsere Belange) effektiven Umlaufzeiten überführt wird. Die effektiven Umlaufzeiten entsprechen jeweils einer ganzzahligen Anzahl von Zufallszahlaufrufen. In meiner Simulation wurde natürlich eine diskrete Verteilung effektiver Umlaufzeiten simuliert.
Einen Einfluss der unterschiedlichen Längen der Sequenzen von Nullen bzw. Einsen kann ich mir nicht vorstellen. Die Anzahl der zwischen zwei Abfragen erzeugten Zufallszahlen ist so groß, dass nicht zwei Abfragen zur selben Sequenz gehören können. Andererseits sind die Summen dieser Sequenzen von Einsen bzw. Nullen gleich lang, da in einer Periode von *M* Zufallszahlen gleich viele Einsen und Nullen auftreten.
Weiterhin betont Bernd Dürrer, dass die genaue Form der Umlaufzeitverteilung eine Rolle spielen könnte. Er stellt fest, dass in der Abbildung 1 von Hagel und Tschapke (2002, S. 67) eine leichte Links-Schiefe (also negative Schiefe) zu erkennen ist. Ein solcher Einfluss ist aber unwahrscheinlich, da nur wichtig ist, ob mit hoher Wahrscheinlichkeit bestimmte „ungünstige" Umlaufzeiten auftreten. Treten diese Umlaufzeiten bei einem Umlauf nicht auf, so ist es ziemlich belanglos, welche sonstige Umlaufzeit dann auftritt, es sei denn, es gibt noch weitere „ungünstige" Umlaufzeiten R_{innen} und $R_{außen}$ in der unmittelbaren Umgebung.
Bernd Dürrer sucht eine Ursache für diese Schiefe. Es gibt aber kein Naturgesetz, welches besagt, dass eine natürliche Verteilung immer symmetrisch sein muss.
Es ist natürlich durchaus möglich, die Simulation mit einer schiefen Verteilung zu wiederholen. Erzeugt man eine standardisiert normalverteilte Pseudozufallsvariable $\underline{u}$, dann ist $x = \exp(\underline{u} \cdot s)$ log-normalverteilt mit Erwartungswert $\mu_o = \sqrt{p}$, Varianz $\sigma_0^2 = p(p-1)$ und Schiefe $\gamma_1 = \sqrt{p-1} \cdot (p+2)$ (Kendall und Stuart 1958), falls $p = \exp(s^2)$ gesetzt wird. Die Variable $y = (\mu_0 - x)/\sigma_0$ hat dann den Erwartungswert 0, die Varianz 1 und die Schiefe $-\gamma_1$. Sind μ und σ $(\sigma = \mu \cdot \sigma\%/100)$ die gewünschten Werte des Erwartungswerts der Standardabweichung, so erhält man die gewünschte Zufallsvariable aus $y \cdot \sigma + \mu$. Soll diese Verteilung eine vorgegebene Schiefe γ_1 besitzen, so erreicht man dieses, indem man $c = \gamma_1^2/2$, $q = \sqrt[3]{1 + c + \sqrt{c(c+2)}}$, $p = q + (1/q) - 1$ und $s = \sqrt{\ln(q)}$ setzt. Wählt man die Schiefe zwischen $-0{,}1$ und $-0{,}05$, dann entspricht das Verteilungsbild – zumindest nach Augenmaß – etwa der Abbildung 1 aus Hagel und Tschapke (2002, S. 67). Die Wiederholung der Zeitraffersimulation

für das Eisenbahnexperiment $(\mu_{\text{innen}} = 60000, \mu_{\text{außen}} = 120000, \sigma\% = 3)$ ergab für die beiden z-Werte aus meiner Tabelle 3 die Beträge –0,170 und –0,463 für $\gamma_1 = -0,05$ bzw. 0,155 und 0,603 für $\gamma_1 = -0,1$. Auch diese Werte zeigen also keine besonderen Auffälligkeiten.
Bernd Dürrer erwähnt, dass bei doppelter Geschwindigkeit, also halbierter Umlaufzeit, auch der Effekt halbiert wurde und meint, dass ein eventueller systematischer Fehler des Pseudozufallszahlengenerators dann ebenfalls halbiert würde. Dies ist nicht der Fall, denn mit halbierten Umlaufzeiten erhält man zwar andere Autokorrelationen, deren Abhängigkeit von der Umlaufzeit ist aber völlig unregelmäßig. Dieser empirische Befund der Effekthalbierung ist dann doch eher mit der Hypothese von Hagel und Tschapke kompatibel, sofern man annimmt, dass mit halber Umlaufzeit auch die Anpassungsfähigkeit der Lokomotive an die Zufallszahlen geringer ist.
Die grundsätzliche Kritik gegen einen Pseudozufallszahlengenerator ist sicherlich berechtigt, die konkreten Gründe sind es jedoch nicht unbedingt. Sicherlich ist der Umfang der Periode *M* beschränkt, so dass sie bei dem Experiment mehrfach durchlaufen wird. Dabei werden aber nicht lückenlos alle Zahlen der Zufallszahlenfolge verwendet, sondern es wird in sehr großen Abständen je eine Zahl herausgepickt. Somit bilden diese herausgepickten Zahlen nicht jeweils exakte Wiederholungen bisheriger Perioden.
Ob die zur Erlangung von Signifikanz erforderliche Abweichung von $S(n) \approx 2,8\sigma \approx 886$ gering erscheint, spielt keine Rolle. Wichtig ist nur, dass der p-Wert, also die Wahrscheinlichkeit dieser und extremerer Abweichungen kleiner als ein vorgegebener Wert α ist. Sofern $S(n)$ nur am Ende des Versuches mit der Signifikanzgrenze $2 \cdot \sigma$ verglichen wird, so wäre $\alpha_{\text{einseitig}} \approx 0,023$ bzw. $\alpha_{\text{zweiseitig}} = 0,046$. Die Abbildung 5 (Hagel und Tschapke 2002, S. 15) verführt aber dazu, auch vor dem Versuchsende bereits $S(n)$ mit der durch $2 \cdot \sigma_n$ definierten Parabel zu vergleichen, womit das Risiko erster Art steigt. Es gibt nämlich Fälle, für die am Versuchsende $S(u) < 2\sigma_n$ gilt, aber irgendwann vorher $S(n) > 2\sigma_n$. Im Gegensatz zur Auswertung am Versuchsende wird bei einer laufenden (sequentiellen) Auswertung dieser Fall als signifikant betrachtet, womit also das Risiko erster Art steigt. Bei sequentieller Auswertung ist daher die Signifikanzgrenze neu zu berechnen. Auf dieses Problem bin ich aber in meinem Artikel nicht eingegangen, da hierfür eine ganze Serie von Artikeln über sequentielle Testtheorie erforderlich wäre.
Außer einer Versuchswiederholung mit einem echten Zufallsgenerator wäre auch die Verwendung eines anderen Zufallsgenerators interessant. Sollte der Generator einen Artefakt verursachen, so wäre dieser stark von dem jeweiligen Generator abhängig.
Wilfried Kugel erwähnt in seinem Kommentar einige interne Probleme des verwendeten Pseudozufallszahlengenerators. Es ist richtig, dass ein solcher Generator keine echten Zufallszahlen erzeugen kann, sondern nur extrem unregelmäßige Zahlenfolgen. Wie die Systemspezifika aber die Funktion RND störend beeinflussen sollen, ist unklar. Nun ja, das korrekte Arbeiten der Funktion RND hängt von der im jeweiligen System verwendeten

Stellenanzahl ab. Doch dieses kann man testen. Andererseits könnte die Zeit zwischen zwei Zufallszahlenerzeugungen vom System abhängen. Es ist aber nicht zu sehen, wie das auf die Häufigkeiten von +1 und –1 Einfluss haben soll.

Die Funktion RND erzeugt intern eine unregelmäßige Folge aller ganzzahligen Werte c_i von 0 bis *M*–1. Hier darf noch nicht gerundet werden, da ansonsten der Zyklus kürzer als *M* wäre. Aus diesen Werten c_i wird jeweils $z_i = c_i / M$ berechnet. Eine hierbei erfolgende Rundung hat keinen Einfluss auf den weiteren Verlauf des Generators, da nicht z_i sondern c_i die nächste Zufallszahl bestimmt. Ich konnte aber nur mit Kenntnis der Zahlen z_i vermittels $c_i = z_i \cdot M$ einen vollständigen, wiederholungsfreien (also fehlerfreien) Zyklus aller M Werte c_i rekonstruieren, was bedeutet, dass auch z_i mit so viel Stellen dargestellt wurde, so dass sich jeweils c_i eindeutig rekonstruieren ließ. Trotzdem würden sich auch Rundungen von z_i nicht verheerend auswirken, da für die letztendlich interessierende Zufallszahl B_i $(=\pm 1)$ nur entscheidend ist, ob in z_i die erste Ziffer nach dem Komma kleiner als 5 ist oder nicht.

Der von Wilfried Kugel genannte Fall 2, dass vor jedem Aufruf von RND eine Initialisierung mit RANDOMIZE TIMER erfolgt, dürfte in der Praxis wohl kaum Anwendung finden. Anderenfalls bitte ich die Autoren Hagel und Tschapke, mir zu widersprechen.

Interessant wäre nun zu erfahren, welche konkreten Probleme bei der Anwendung durch Alan Vaughan auftraten und welche zusätzliche Randomisierungprozedur er schrieb. Das von Wilfried Kugel dem Generator RND entgegen gebrachte Misstrauen wäre für alle Pseudozufallszahlengeneratoren angebracht. Denn alle Generatoren weisen in irgendeiner Beziehung mehr oder weniger große Regelmäßigkeiten auf. Bei dem Test eines Pseudozufallsgenerators wird im allgemeinen die Abhängigkeit zwischen *n* (z. B. *n*=8) direkt aufeinander folgender Zahlen untersucht. Hier mag der RND durchaus ungünstige Eigenschaften aufweisen. In dem hier betrachteten Experiment werden aber nicht direkt aufeinander folgende Zahlen der Zahlenfolge verwendet, sondern es liegen große Abstände zwischen den verwendeten Zahlen. Diese Abstände haben zusätzlich auch noch starke zufällige Schwankungen. Sollte der Generator RND aber auch für diese so ausgewählten Zufallszahlen regelmäßige Eigenschaften haben, dann hätte sich das auch in einer Zeitraffersimulation zeigen müssen.

Der Kommentar von Eckhard Etzold erinnert mich dankenswerterweise an ein Versäumnis, wodurch sicherlich einige Irritationen entstanden sein mögen. Hagel und Tschapke behaupten, dass die Varianz von *S(n)* gleich *n* sei, tatsächlich ist die aber 4*n*. Nachdem ich Johannes Hagel darauf aufmerksam machte, informierte er mich, dass hier ein Schreibfehler vorlag, richtig sollte es $S(n) = \sum_{k=1}^{n} B(k)$ heißen (E-mail vom 14.3.2002). Damit ist dann die Varianz von *S(n)* tatsächlich gleich 1.Von dieser Formel bin ich in meinem Artikel ausgegangen und habe mich dann an diese Formel so gewöhnt, dass ich vergaß, auf den Widerspruch zur Originalformel hinzuwesen. Dieses *S(n)* (mit oder ohne Korrektur) ist im

Fall von „Unfallvermeidung" (d.h. der Innenkreis, also B=–1) stets negativ. Bei den in meiner Tabelle 3 demonstrierten Extremfällen habe ich deswegen auch nur solche Fälle betrachtet, die zu negativen $\overline{B}$ bzw. z führen. In den Abbildungen von Hagel und Tschapke wurde die senkrechte Koordinate anders ausgerichtet, weswegen die Kurven in positiver Richtung verliefen. Ob hier eventuell noch weitere Irrtümer vorliegen, könnten nur die Autoren des Experiments sagen.

Nun lese ich aber bei Eckhard Etzold, dass er bereits kurz vor der Veröffentlichung Johannes Hagel auf den Fehler aufmerksam machte. Dass dieser Hinweis unberücksichtigt blieb, ist in der Tat sehr bedenklich.

Eine der Schlussfolgerungen von Etzold ist mir nicht einleuchtend. Bezeichnen wir die Originaldefinition von $S(n)$ mit $S_O(n)$ und die korrigierte Definition mit $S_K(n)$ $\left[S_O(n) = 2 \cdot S_K(n)\right]$ und akzeptieren wir die von Hagel und Tschapke gewählte Signifikanzgrenze $2n$, welche sich jedoch auf eine Variable mit der Varianz $1 \cdot n$ bezieht, also z.B. auf $S_K(n)$, dann erhält man für $S_O(n)$ die Signifikanzgrenze $2 \cdot 2n = 4n$. Die Originalwerte $S_O(n)$ wären damit erst dann auffällig, sofern sie größer als $4n$ sind.

Die weiteren Ausführungen von Eckhard Etzold zeigen, dass er über die Interna des BASIC-Programms informiert ist. Hierbei stellt sich heraus, dass der eigentliche Fehler bei Hagel und Tschapke in dem Satz unter der Formel für *S(n)* steckt. Dort wurden die vorher nicht definierten Werte B_{innen} und $B_{\text{außen}}$ mit +1 bzw. –1 identifiziert, was zur Verwechslung mit *B* führt, aber scheinbar nicht so gemeint war. Vermutlich ist folgende Wertzuordnung gemeint:

Durchlauf	B_{innen}	$B_{\text{außen}}$
Innenkreis	1	0
Außenkreis	0	1

Somit ist nach *n* Umläufen die Anzahl der Innenumläufe – in Anlehnung an die Symbole iunten und ioben – durch

$$\text{iinnen} = n\sum_{k=1}^{n} B_{\text{innen}}(k) \quad \text{gegeben. Analog sei} \quad \text{iaussen} = \sum_{k=1}^{n} B_{\text{außen}}(k)$$

zu verstehen. Somit ist die Differenz der Anzahlen der verschiedenen Kreisdurchläufe tatsächlich durch

$$S(n) = \sum_{k=1}^{n} \left[B_{\text{innen}}(k) - B_{\text{außen}}(k)\right]$$

gegeben. Setzt man nun $B = B_{\text{innen}} - B_{\text{außen}}$, dann gilt gemäß obiger Wertetabelle

$$B = \begin{cases} 1, & \text{für innen} \\ -1, & \text{für außen} \end{cases}$$

Diese Definition entspricht – bis auf den Vorzeichentausch – der Originaldefinition von *B*, womit dann die Originalfunktion *S(n)* mit der „korrigierten" Funktion (zumindest gemäß der Korrektur, wie sie mir mitgeteilt wurde) $S(n)=\sum_{k=1}^{n} B(k)$ identisch ist.

Etzold beschreibt anschließend eine komplizierte technische Schaltung, mit welcher ein zu dem Eisenbahnexperiment analoges Experiment simuliert werden soll. Ich kann hier jedoch keine Anomalie erkennen. Eine Anomalie ist – kurz gesagt – ein Ergebnis, welches jeder Erwartung widerspricht. Bei diesem Experiment müsste man erst mühsam zeigen, was eigentlich erwartet wird, nämlich die Gleichwahrscheinlichkeit zweier Zustände, was durchaus keine Selbstverständlichkeit darstellt. Dieses Experiment ist also eher als eine Qualitätsprüfung des Zufallszahlengenerators zu verstehen.

Eine Entscheidung, ob das Ergebnis auf ungleiche Wahrscheinlichkeiten beider Zustände oder – nach meinem Artefaktprinzip – auf unterschiedliche Autokorrelationen zurückzuführen ist, ließe sich feststellen, indem man die Kodierung vertauscht. Wählt man – in der Begriffswelt des Eisenbahnexperiments – im ersten Experiment „innen" im Fall *B*=–1 und „außen" im Fall *B*=+1 und vertauscht im zweiten Experiment diese Kodierung, dann müsste sich das Endergebnis umkehren, sofern es nur an der ungleichen Wahrscheinlichkeit von +1 und –1 liegt. Sollte das Ergebnis jedoch nur durch Autokorrelation bedingt sein, so fällt es bei beiden Experimenten in gleicher Richtung aus. In beiden Fällen wäre das Ergebnis aber als Qualitätsmangel des Generators zu interpretieren. Weiterhin ist unklar, was unter einem zweiseitigen z-Wert zu verstehen ist. Solch ein Begriff ist nur für p-Werte definiert.

Ich schließe mich den verschiedenen Aufrufen zur Wiederholung des Eisenbahnversuchs an. Es ist die einzige Möglichkeit, mehr Licht in die Angelegenheit zu bringen, alles andere ist nur Spekulation. Es wäre vorteilhaft, wenn dieses Experiment in zwei Varianten laufen könnte, einmal in der Originalversion und einmal mit der von Bernd Dürrer vorgeschlagenen Verbesserung. Weiterhin überlegte ich lange, ob ein echter Zufallsgenerator wirklich von Vorteil ist, da man seine Wahrscheinlichkeitseigenschaften auch erst empirisch prüfen müsste. Eine kleine Abweichung würde sich bei Langzeitversuchen allmählich summieren und damit vergrößern. Dagegen kann man bei einem Pseudozufallsgenerator theoretisch zeigen, dass im Laufe einer Periode beide Ereignisse exakt mit gleicher Häufigkeit auftreten. Auch wenn nur in großen Abständen Zahlen herausgegriffen werden, so sollte man erwarten, dass sich zeitweilige Schwankungen langfristig ausgleichen. Der eigentliche echte Zufall wird hier durch die Unterschiede der Umlaufzeiten hereingetragen. Diese Unterschiede können aber keinen systematischen Einfluss auf die Häufigkeit der beiden Ereignisse des Pseudozufallsgenerators ausüben, es sei denn, dass in dem technischen Gerät, welches die Abfrage realisiert, das „Durchlassen" (mit eventueller Zeitverzögerung) der Abfrage irgendwie von dem physikalischen Zustand des Speichers abhängt, der gerade die letzte Zufallszahl enthält.

Literatur

Kendall, M.G.; Stuart, A. (1958): The advanced theory of statistics. Griffin, London, 2. Edition, Volume I.