# Nützliche Gesetze über den Zufall - Experimente mit Excel

Manfred Borovcnik, Klagenfurt

Gesetz und Zufall sind nur scheinbare Widersprüche. Wie Gesetze über den Zufall aufgebaut sind und wie man sie nutzen kann, um sich in Situationen mit Ungewissheit besser orientieren und verhalten zu können, das soll anhand von geeigneten Analysen simulierter Daten aufgezeigt werden. EXCEL als allgemein zugängliches Tabellenkalkulationspaket bietet durchaus einige, auch graphische Möglichkeiten.

## 1. Rahmenbedingungen des Unterrichtsexperiments

### *Ziele*

Im folgenden wird eine Unterrichtsserie und deren theoretisch-didaktischer Hintergrund beschrieben. Die Deutung von Wahrscheinlichkeit als relativer Häufigkeit ist eine wesentliche Stütze für die Anwendung und das Verständnis des Begriffs. Die Gesetze der großen Zahlen, die eine eigenartige Konvergenz der relativen Häufigkeiten gegen die zugrunde liegende (oft auch unbekannte) Wahrscheinlichkeit zum Inhalt haben, sollten erarbeitet werden. Gerade diese Eigenart der ‚Konvergenz' ist verantwortlich für eine Fülle von Missverständnissen. Ihr kann man gut durch geeignete Analyse simulierter Daten Herr werden.

Daneben wird ein Verständnis für den Begriff des Restrisikos, das heute in aller Munde ist, angestrebt. Als Nebenprodukt stellt sich ein wirklich elementarer, rekursiver Zugang zur Binomialverteilung heraus, der die Vorteile einer Tabellenkalkulation bestens nutzt. Der didaktische Zugang eröffnet eine Weiterführung bis hin zu einer intuitiven Einführung in die Methode der Schätzung mittels Konfidenzintervallen, die den Begriff der Genauigkeit einer Schätzung mit dem Begriff des Restrisikos verbinden. Methodisch, aber auch im Hinblick auf das angestrebte inhaltliche Verständnis von Wahrscheinlichkeit, ist in der Planung des Unterrichts eine Auseinandersetzung mit virtuellem, auf Algorithmen basierendem Zufall, und materiellem, auf der realen Durchführung von Experimenten basierendem Zufall, erforderlich.

*Gesetze für den Zufall?*

In der klassischen Vorstellung der Naturwissenschaften geht es um die Klärung von Ursache-Wirkung-Zusammenhängen: Welche Kombination von Versuchsbedingungen muss hergestellt werden, damit dann das und das passiert. Werden in einem Experiment dann tatsächlich diese Bedingungen eingehalten, kann man aufgrund der einmal formulierten Gesetze die Folgen voraussagen. Lässt man einen Stein bestimmter Masse aus einer bestimmten Höhe fallen, so kann man dessen Geschwindigkeit nach $t_0$ Sekunden voraussagen etc., vorausgesetzt, das Gesetz gilt. Dementsprechend „einfach" ist ein naturwissenschaftliches „Gesetz" als falsch zu erkennen. Schon das Auftreten von Störgrößen und Messfehlern macht allerdings ein Überprüfen von Gesetzen schon einiges schwieriger.

Außerhalb der Naturwissenschaften kann man die Versuchsbedingungen nicht so streng kontrollieren, häufig ist man lediglich externer Beobachter. Hier versucht man, die Unregelmäßigkeiten in den Experimenten in einen stochastischen und einen mathematisch-funktionalen Teil von Zusammenhängen zu zerlegen: Der stochastische Teil beschreibt dabei sozusagen die Unvollkommenheit der Versuchsbedingungen, der funktionale Teil steht für die wesentlichen Zusammenhänge. Eine Trennung ist, wie man sich leicht denken kann, weder einfach noch eindeutig. Entsprechend vielfältig sind die Fehldeutungen in den Anwendungen (eine Methode ist die von Regression und Korrelation).

Hier sind wir auch schon bei dem Begriff von stochastischen Zusammenhängen angelangt. Stochastische Zusammenhänge, oder strenger gewendet, stochastische Gesetze sind wie auch stochastische Experimente, von ganz anderer Struktur als ihre kausalen Gegenstücke aus den Naturwissenschaften. Dies soll kurz am Münzwerfen angedeutet werden. Dieses grundlegende, am Anfang der Stochastik stehende Experiment soll wohl wiederholbar sein wie ein naturwissenschaftliches Experiment, d.h. immer unter den gleichen Bedingungen durchführbar. Das schließt mit ein, dass eine einzelne Durchführung als typisch für eine gedachte Serie von Versuchen interpretiert werden darf. Wollte man jedoch den physikalischen Experiment-Begriff auf das Münzwerfen anwenden, so wäre das Ergebnis durch die genaue Angabe, wie man zu werfen hat, eindeutig bestimmt und damit auch das Ergebnis vorweg genommen. Das aber soll bei einem idealtypischen Münzwurf nicht sein. Dieses Experiment ist geradezu so angelegt, dass es durch Wiederholung und Nicht-Kontrollieren aller verschiedensten Einzelheiten, die das Ergebnis beeinflussen könnten, zu einer (unvorhersehbaren) Fluktuation der Ergebnisse kommt, die man dann mit Wahrscheinlichkeiten beschreibt. Wenn man also von Gesetzen über den Zufall spricht, so sind das jedenfalls typischerweise nicht Gesetze im klassisch-naturwissenschaftlichen Sinne.

*Materieller und virtueller Zufall*

Es gibt keinen Zufall? Gibt es einen Zufall? Für den Laplace'schen Dämon, der alles weiß, dem auch die Zukunft nicht verschlossen ist, gibt es keinen Zufall. Es gab in der wissenschaftstheoretischen Diskussion eine lange und hitzige Debatte, dass ein solcher naiver Determinismus zu kurz greift. Vom Standpunkt der Religion könnte man den allwissenden Gott einbringen, und die persönliche Willensfreiheit – zumindest in der christlich dominierten Welt – dagegen setzen. Und wäre damit bei einem unauflöslichen Widerspruch angelangt.

Man könnte aber auch Zufall einfach als eine Denkweise interpretieren, die auch dann zu nützlichen Voraussagen, Einschätzungen, Entscheidungen etc. führt, auch wenn man den internen, möglicherweise kausalen Mechanismus der Abhängigkeiten gar nicht kennt – oder wenn es einen solchen einfach nicht gibt. Man orientiert sich dann an externen, erkennbaren Phänomenen und versucht, diese Phänomene mit statistischen ‚Gesetzen', oder besser, Regelmäßigkeiten, zu vergleichen. Wird bei diesem Vergleich kein - wesentlicher - Unterschied erkannt, so trifft man seine Voraussagen gemäß denjenigen, die sich aus den statistischen Gesetzen ergeben.

Damit ergibt sich die Notwendigkeit, Zufall zu beschreiben. Theoretisch geht das mit mathematischen Begriffen, praktisch illustrieren kann man den Zufall und damit zusammenhängende Gesetze mittels Simulation. Der Zufall wird dabei in der Wirklichkeit nachgestellt. Die Ergebnisse erhalten einen szenario-artigen Charakter. Sie zeigen auf, wie sich Ergebnisse bei 'Wirken' des reinen (fiktiven) Zufalls einstellen, wie sich Systeme entwickeln.

Dabei kann man auf zwei grundverschiedene Ansätze zurückgreifen: der materielle und der virtuelle Zufall. Beim ersteren stellt man materiell durch einen Gegenstand und Festhalten von Versuchsbedingungen wiederholbare Situationen sicher, man experimentiert wirklich und kann die Ergebnisse direkt mitverfolgen. Münzwerfen ist ein Prototyp für die materielle Realisierung von Zufall. Idealerweise hat man vorweg eine Hypothese der Gleichwahrscheinlichkeit der zwei möglichen Ergebnisse, die weitgehend akzeptiert ist. Während diese Gleichwahrscheinlichkeit nicht so wesentlich ist (man könnte auch mit unbekannten Wahrscheinlichkeiten gut umgehen und sie aus den Ergebnissen gemäß statistischen Gesetzen schätzen), ist wichtig, dass man eine statistische Unabhängigkeit der einzelnen Ergebnisse durch geeignetes Festlegen der Münzwurfbedingungen recht gut garantieren kann. Gerade diese Unabhängigkeit ist nämlich

der Angelpunkt für die Gültigkeit statistischer Gesetze.

Beim zweiten Ansatz realisiert man den Zufall fiktiv durch mathematische Algorithmen, die aus vergangenen Ergebnissen die weiteren Ergebnisse der Versuchsserie berechnen. Bestimmte Algorithmen (z.B. auf der Basis von Kongruenzen) haben nun die verblüffende Eigenschaft, dass sich sowohl die gleiche, vorgegebene Wahrscheinlichkeit für bestimmte Versuchsausgänge einstellt, als auch, und das ist wesentlich, eine Unabhängigkeit zwischen einzelnen Ergebnissen. Zumindest kann man mit Hilfe von statistischen Gesetzen einen Vergleich mit materiellem Zufall oder mit dem mathematischen Zufall anstellen, wobei dieser Vergleich zu keinem nachweisbaren Unterschied führt. Es ergibt sich hiermit die paradoxe Situation, dass man den Zufall, der idealtypischerweise für die Unmöglichkeit der präzisen Voraussage des Ergebnisses des nächsten Experiments steht, durch einen mathematisch-funktionalen Zusammenhang des nächsten Ergebnisses aus dem vorhergehenden Ergebnis (es können auch mehrere Ergebnisse sein) nachspielen kann.

### *Die Welt der Intuitionen*

Man kann davon ausgehen, dass Personen eine Fülle von individuellen Vorstellungen zu Wahrscheinlichkeit und Zufall entwickeln, unabhängig davon, ob sie eine formale Unterweisung darin erhalten. Empirische Untersuchungen geben Aufschluss sowohl über die Vielfalt als auch über gängige, sehr weit verbreitete Muster von Fehlvorstellungen. Die mathematische Einführung in die Begriffswelt soll einerseits an die bestehenden Bilder und Muster der Lernenden anbinden, diese auch tatkräftig verändern, so sie inadäquat sind, und soll zu stabilen Bildern, sogenannten sekundären Intuitionen führen. Die Kraft dieser sekundären Intuitionen liegt nach Fischbein darin, dass erst sie über eine formale Repetition der mathematischen Begriffe hinausführen; erst diese sekundären Intuitionen ermöglichen ein gutes Verstehen sowie einen raschen, sicheren Einsatz der Begriffe durch die Lernenden.

In der persönlichen Intuition vieler Menschen kommt es aus einem geradezu elementaren Bedürfnis, alles einschließlich der Zukunft unter Kontrolle zu bringen, zu einer unzulässigen Übertragung von Gesetzen auf Zufallsexperimente und Situationen, die man üblicherweise unter dem Zufall zusammenfasst. So werden einmal „erkannte“ lokale Muster einer Serie von Versuchen extrapoliert, wie ‚Kopf‘, ‚Zahl‘, ‚Kopf‘, dann muss doch ‚Zahl‘ kommen. Wie heftige Dispute zwischen Berufsspielern beim Roulette zeigen, muss über die erkannten Muster kein Konsens bestehen. So gibt es nach etwa 12-mal ‚rot‘ welche, die auf die Fortsetzung der (unheimlichen) Serie setzen, andere wiederum schwören auf den Wechsel.

Deshalb sind Gesetze über den Zufall etc. sowie der Begriff Wahrscheinlichkeit selbst vielen Fehlvorstellungen und Fehlanwendungen unterworfen. Physikalisches Gesetz und Zufall sind so ziemlich am entgegengesetzten Ende einer Skala im Hinblick auf die Voraussagbarkeit angesiedelt. Und dennoch gibt es etwas wie Gesetze, oder sagen wir einfacher, Regelmäßigkeiten in Zufallsexperimenten.

In der experimentellen Erforschung solcher Regelmäßigkeiten liegt eine Chance für den Unterricht, die Eigenheiten dieser Gesetze verständlich zu machen; wesentlich verständlicher jedenfalls als es durch eine mathematische Beschreibung möglich wird.

Aus einsichtigen Gründen wird für die Simulation von Zufall gerne der virtuelle Zufall propagiert. Man kann dabei in kurzer Zeit sehr viele Daten erzeugen, diese liegen bereits geeignet gespeichert vor und sind sofort einer graphischen sowie rechnerischen Analyse zugänglich. Alle Umstände einer Simulation mit realen Objekten, wie Münzen und Würfeln etc. fallen weg, auch die fehleranfällige und langwierige Protokollierung der Ergebnisse und deren Übertragung auf den Computer, der erst eine leistungsstarke, insbesondere auch graphische Analyse ermöglicht. Es sei jedoch davor gewarnt, allzu rasch vom virtuellen Zufall Gebrauch zu machen. Die individuellen Vorstellungen über den Zufall sind derart vielfältig, inadäquat und hartnäckig, so dass die Gefahr besteht, dass sie auf diese Weise weder berührt noch verändert werden. Was bedeutet, der Unterricht wäre gar nicht wirksam. Empirische Untersuchungen zeigen die Dominanz inadäquater Intuitionen, die eine Akzeptanz mathematischer Ergebnisse und Gesetze regelrecht verhindern.

Ein guter Kompromiss besteht darin, mit realen Objekten zu beginnen, erste Ergebnisse und Gesetzmäßigkeiten herauszuarbeiten und dann auf virtuelle Experimente überzugehen und die Ergebnisse miteinander zu vergleichen. Mit den virtuellen Experimenten erschließt sich eine weitaus größere Bandbreite für die Untersuchung des Gültigkeitsbereichs der einmal schon erkannten Gesetzmäßigkeiten. So etwa kann man die „Stabilisierung" der relativen Häufigkeiten für verschiedenste Wahrscheinlichkeiten leicht zeigen. Oder, man kann die Vergrößerung der Präzision der Schätzung von unbekannten Wahrscheinlichkeiten mit zunehmender Länge der Versuchsserie leicht demonstrieren.

## 2. Das Unterrichtsexperiment

### *Münzwerfen*

Die im folgenden geschilderte Unterrichtsserie wurde in einer 4. Klasse des BG Sankt Veit im Schuljahr 99/00 durchgeführt. Der Unterricht fand zur Gänze im PC-Labor statt. Zwanzig Lernende nahmen daran teil. Im Jahr zuvor hatten diese im Rahmen der Beschreibenden Statistik eine Einführung in EXCEL erhalten. Das Zufallsexperiment sollte eigentlich Münzwerfen sein. Die oben getroffenen Bemerkungen über virtuellen und materiellen Zufall waren Anlass, das Experiment zuerst materiell zu beginnen und erst im weiteren Verlauf auf Computer-Simulationen überzugehen.

Eine rein praktische Erwägung führte zu einer realen Modellierung des Münzwerfens mit Hilfe von Würfeln. Zum einen sollte der Lärm und das ohnehin mit der Durchführung eines Klassenexperiments verbundene Durcheinander gering gehalten werden. Zum anderen aber sind ja für die Durchführung gleiche Bedingungen und die Unabhängigkeit der Versuche voneinander zu garantieren. Das sollte mit einem Würfelbecher und fünf farblich unterschiedenen Würfeln geleistet werden. Den Farben entsprechend sollte die Reihenfolge der einzelnen Ergebnisse festgelegt werden (später wurden die Würfel einfach zusammengeschoben und das Resultat abgelesen), eine ungerade Zahl wurde als ‚Kopf', eine gerade als ‚Zahl' interpretiert und mit 1 bzw. 0 codiert. Fünf Ergebnisse sind rasch und sicher abzulesen, der Inhalt des Würfelbechers mit 5 Daten spielt auch strategisch für die spätere Auswertung eine Rolle. Die Ergebnisse der

Gruppenarbeit wurden sofort in ein Tabellenblatt am Computer protokolliert und online über die Gruppen hinweg zusammengefügt.

**Auswertung der abs/rel Häufigkeiten von g/u nach Umfang der Serie**

| Ergebnis | Codierung | Serien der Länge n *5er* | *10er* | *20er* |
|---|---|---|---|---|
| g | 0 | | | |
| g | 0 | | | |
| u | 1 | | | |
| g | 0 | | | |
| u | 1 | 2 | | |
| u | 1 | | | |
| g | 0 | | | |
| u | 1 | | | |
| g | 0 | | | |
| u | 1 | 3 | 5 | |
| g | 0 | | | |
| g | 0 | | | |
| u | 1 | | | |
| u | 1 | | | |
| g | 0 | 2 | | |
| g | 0 | | | |
| u | 1 | | | |
| g | 0 | | | |
| g | 0 | | | |
| u | 1 | 2 | 4 | 9 |
| g | 0 | | | |
| u | 1 | | | |

*Anzahl der u = Köpfe*

| *Klasse* | *Häufigkeit* |
|---|---|
| g | 311 |
| u | 289 |
| gesamt | 600 |

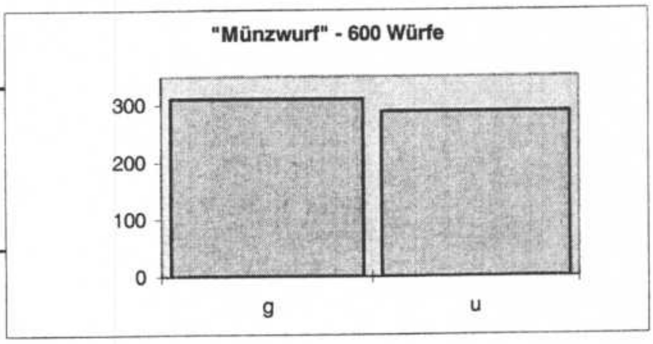


Fig. 1: Protokoll des Münzwurfexperiments.

## *Erste Analysen – der übliche Zugang*

Einen ersten Anhaltspunkt über die Wahrscheinlichkeit von ‚Kopf' gibt dessen relative Häufigkeit am Ende des gesamten Versuchs; das Ergebnis von 289 nach n=600 Versuchen gibt eine Schätzung von 0,48 und entspricht durchaus der Annahme der Gleichwahrscheinlichkeit, also einer ‚fairen' Münze. Um die Stabilisierung der relativen Häufigkeiten - das empirische Gesetz der großen Zahlen - zu belegen, ist es üblich, den Versuch nach bestimmten Zeitpunkten im Hinblick auf die Häufigkeit von ‚Kopf' auszuwerten und diese als Funktion der Zeit in einem Schaubild darzustellen. Die folgende Figur 2 zeigt deutlich die Konvergenz der relativen Häufigkeiten gegen einen Wert in der Nähe von 0,49 oder 0,50.

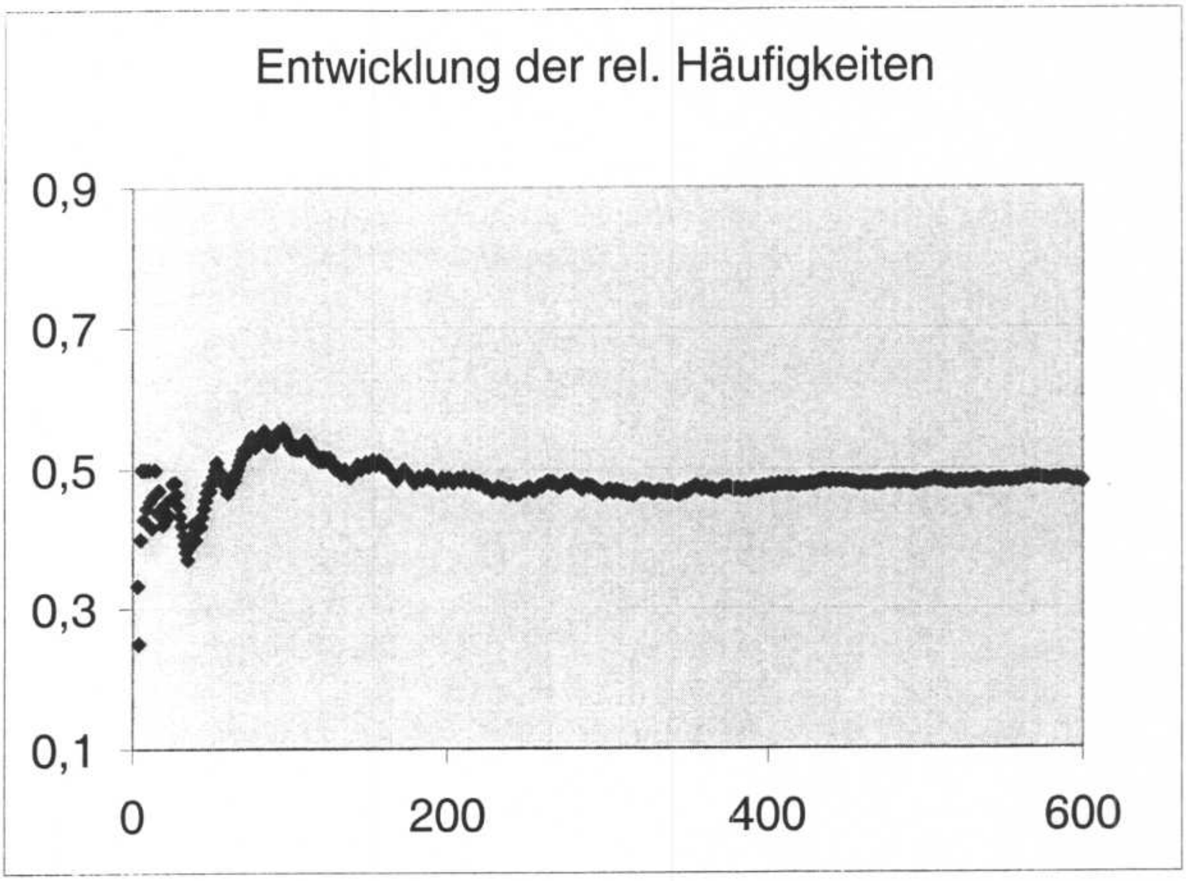


Fig. 2: Abhängigkeit der relativen Häufigkeit von Kopf vom bisherigen Verlauf der Versuchsserie.

Eine entsprechende Abbildung ziert alle Lehrbücher im einführenden Abschnitt zur Wahrscheinlichkeitsrechnung. Legion ist auch die Reihe von Einwänden gegen diese Darstellung, die nicht alle wiederholt werden sollen. Etwa suggeriert die Darstellung eine Konvergenz der relativen Häufigkeiten im Sinne der Analysis gemäß einem Bildungsgesetz, was der Unabhängigkeit der Versuche diametral widerspricht und im Sinne eines umfassenden Verständnisses des Wahrscheinlichkeitsbegriffs und seiner Deutung als relativer Häufigkeit kontraproduktiv ist.

Hier soll nur ein - später wieder aufgegriffener – Gedanke angedeutet werden: Die geringen Schwankungen der relativen Häufigkeiten im weiteren Verlauf innerhalb der einmal begonnenen Serie von Versuchen ergeben eine Illusion über die Präzision einer Schätzung der Wahrscheinlichkeit durch die relative Häufigkeit. So schwankt im Experiment der Verlauf zwischen 400 und 600 Versuchen nur unwesentlich von 0,48 auf 0,49. Die Wiederholstreuung aber, das ist die Schwankung der Zielgröße ‚relative Häufigkeit' bei einer weiteren, ganz *neu* begonnenen Serie, ist weitaus größer; eine weitere Serie könnte sich etwa bei 400 Versuchen bei 0,53 befinden und dann bis zu n=600 nur mehr auf 0,52 verändern. Während also innerhalb der einen Serie die weitere Schwankung durch die bestehenden vielen Ergebnisse eingeschränkt ist auf +/- 0,01, gibt es von Serie zu Serie eine Schwankung von etwa +/- 0,04.

Daher schränkt der Verlauf der relativen Häufigkeiten in Abhängigkeit von der Länge der Serie, wie er üblicherweise dargestellt wird, das Verständnis von und für Zufallsschwankungen unnötig ein und hinterlässt den fälschlichen Eindruck, die weiteren Ergebnisse seien von den bisherigen ‚abhängig'. Das führt im Kurzschluss u.a. zu dem berühmten ‚Gesetz' der kleinen Zahlen, wonach auch in kleineren Serien ein Ausgleich der relativen Häufigkeiten in Richtung der Wahrscheinlichkeit zu erfolgen hat.

### *Auswertung in 5-er Serien – ein neuer Zugang*

Wertet man die relative Häufigkeit von ‚Kopf' in jeder einzelnen Serie von 5 Würfen getrennt aus, berechnet man also das Ergebnis eines Wurfes mit dem Würfelbecher, so kann man zweierlei miteinander vergleichen: Einerseits hat man die Stabilisierung der relativen Häufigkeiten wie schon gehabt in Abhängigkeit von der bisherigen Länge des Gesamtversuchs, andererseits sieht man die Fluktuation der aktuellen Fünfer-Serien, die auch mit Fortschreiten des Experiments gleich groß ist wie zu Beginn. Die Berechnung ergibt sich in der Tabellenkalkulation leicht durch Einführen einer neuen Spalte und Anwenden der Summenfunktion.

Die graphische Darstellung der Berechnungen wirft schon ein differenzierteres Licht auf das eigentümliche Verhalten des Zufalls: Wenngleich sich eine Stabilisierung auf längere Sicht einstellt, sind die einzelnen neuen Versuche (nun: die fünf Ergebnisse eines Würfelbechers zusammengefasst) ungebremst dem Zufall unterworfen. Es kann keine Einschränkung und damit verbunden eine bessere Voraussage für die nächste Fünfer-Serie geben, auch wenn man das Experiment schon eine lange Zeit durchführt und die Ergebnisse notiert.

Hier ist auch eine Analogie zum Messen hilfreich. Sei unterstellt, dass die Wahrscheinlichkeit p eine physikalische Größe darstellt. Vorderhand misst man einen bekannten Wert, später werden auch unbekannte Werte p gemessen, die Messprozedur und ihre Eigenschaften bleiben unverändert. Gemessen wird der Wert p (=½) durch die relative Häufigkeit von ‚Kopf' in der jeweiligen Fünfer-Serie. In Figur 3 zeigt sich also die Variabilität der Messwerte, welche mit wenigen Ausnahmen zwischen 0,2 und 0,8 schwanken; einige Mess-Serien jedoch ergeben Werte von 0 bzw. 1 für p. Insgesamt streuen die Messwerte augenfällig, wie aus Figur 3 ersichtlich ist, um eine Achse, die durch den Wert ½ gekennzeichnet ist. In diesem Sinne scheint die Messprozedur richtig kalibriert zu sein; es wird deutlich, dass es ganz gut ist, fiktiv den zu messenden Wert schon zu kennen, damit man das Messinstrument sozusagen in seiner Eichung überprüfen kann.

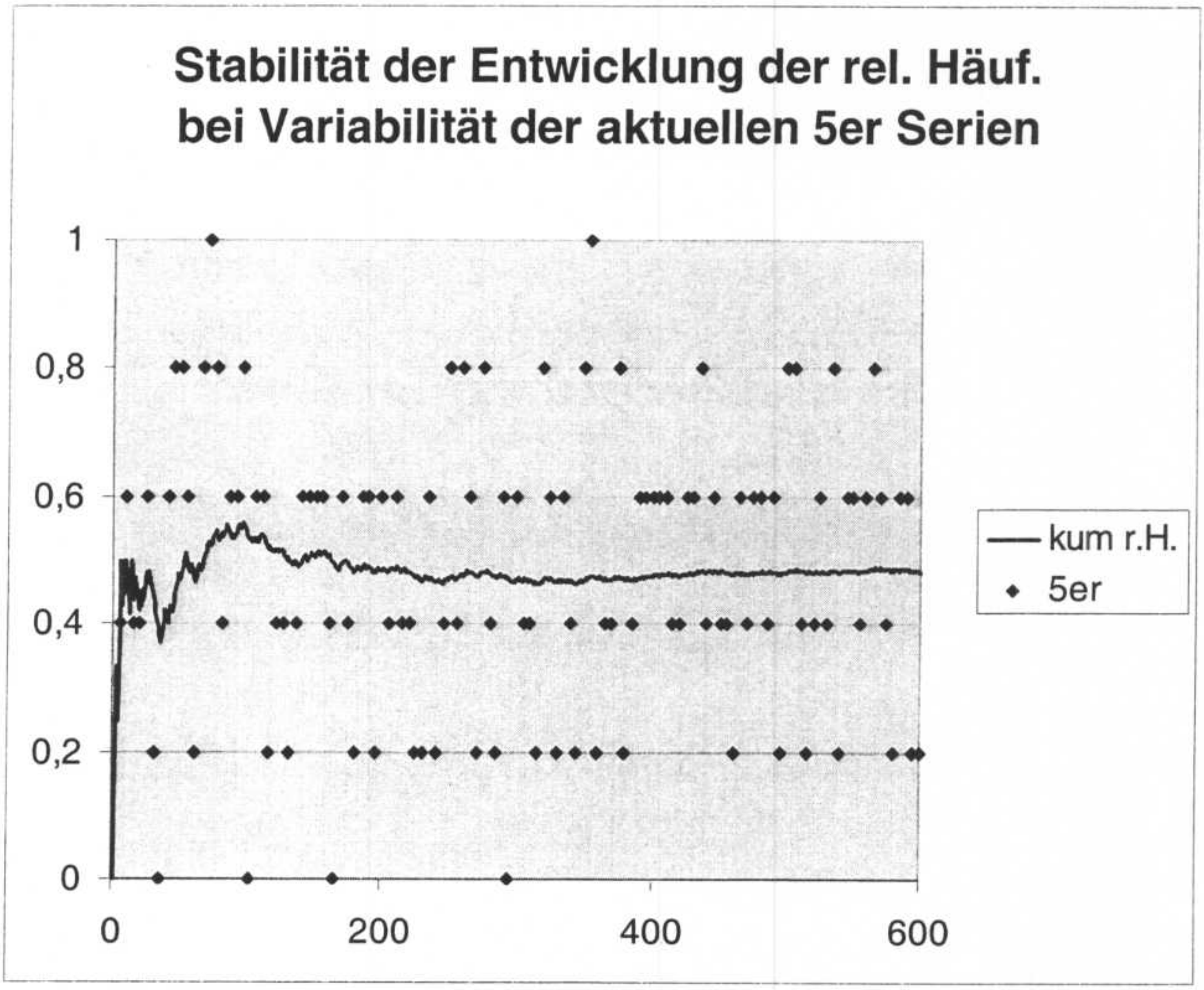


Fig. 3: Stabilität der Entwicklung der relativen Häufigkeiten bei Variabilität der aktuellen Fünfer-Serien.

Im Sinne der Messtechnik ist es aber zusätzlich ganz wesentlich, wie präzise ein Messinstrument ist, i.e. wie präzise man mit Fünfer-Serien und deren relativer Häufigkeit von ‚Kopf' die Wahrscheinlichkeit von ‚Kopf' messen kann. Die Schwankungen der Messwerte zwischen 0,2 und 0,8 haben Messfehler von +/- 0,3 zur Folge; klar ist das viel zu ungenau und stellt einen wesentlichen Nachteil der vorgeschlagenen Messprozedur dar.

Die Darstellung in Figur 4 kann man auch mit Hilfe einer Analogie zu Qualitätsregelkarten aus der industriellen Massenproduktion interpretieren. Die Einhaltung der ‚Qualität' (Abmessung, hier Wahrscheinlichkeit p=½) wird anhand von Fünfer-Serien beurteilt. Die Achse bei ½, um welche die Werte streuen, entspricht dem Sollwert, dessen Einhaltung mit der Regelkarte überprüft wird. Man ersieht aus Figur 4, dass für viele Fünfer-Serien der Sollwert einigermaßen eingehalten wird. Im Sinne der Qualitätsregelung muss man auch sogenannte Schwellenwerte benennen, bei deren Überschreitung man nicht mehr davon ausgehen kann, dass der Sollwert in der Produktion noch eingehalten wird. So etwa könnte man solche Schwellenwerte bei 0,2 und 0,8 festlegen und bei deren Überschreiten die Produktion unterbrechen und nach den Ursachen der Veränderung suchen. Erst nach Beseitigen der Störfaktoren würde man dann die Produktion weiterlaufen lassen.

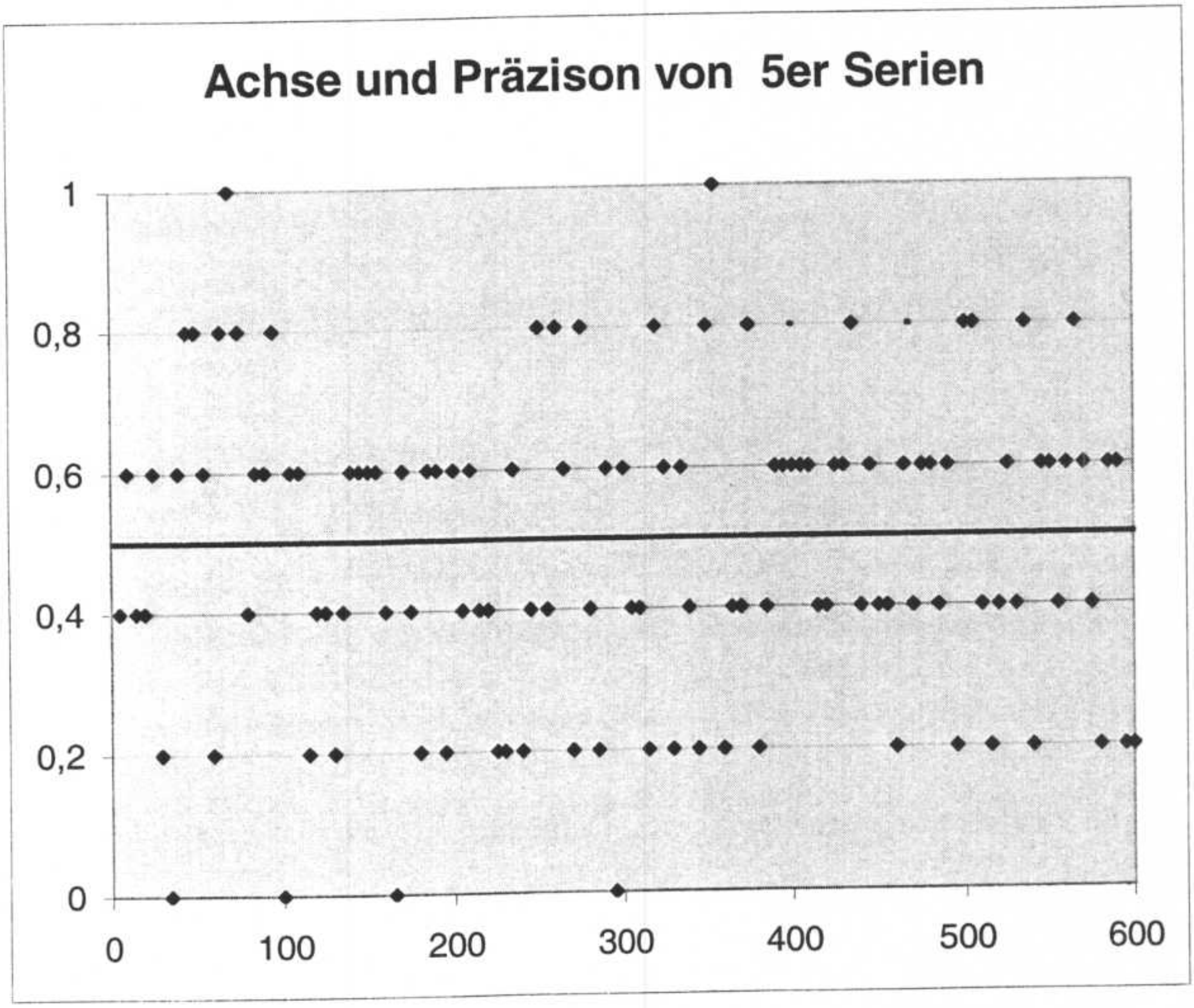


Fig. 4: Achse und Präzision der Messung von p=½ mittels relativer Häufigkeiten von Fünfer-Serien.

Aus Figur 5 ersieht man, dass die Serie Nr. 7 mit Wert 0 die untere Eingriffsgrenze unterschreitet, dann die Serie Nr. 14 mit Wert 1 die obere Eingriffsgrenze überschreitet usw. Solche Serien führen zu Abweichungen vom Sollwert, die man nicht mehr tolerieren will. Derart extreme Werte tauchen selten auf; je größer die Genauigkeitsforderungen, oder anders gesagt, je kleiner die Toleranzen, desto häufiger sind sie. Würde man die Genauigkeitsforderungen herunter setzen, so werden extreme Werte (Werte, welche diese Forderungen nicht einhalten) seltener, aber immer noch sind solche Überschreitungen möglich – speziell wenn man an die Wiederholung des gesamten Experiments denkt. Das führt zum Begriff des unvermeidbaren Restrisikos; unvermeidbar, sofern man sich nicht auf die trivialen Eingriffsgrenzen 0 und 1 zurückzieht.

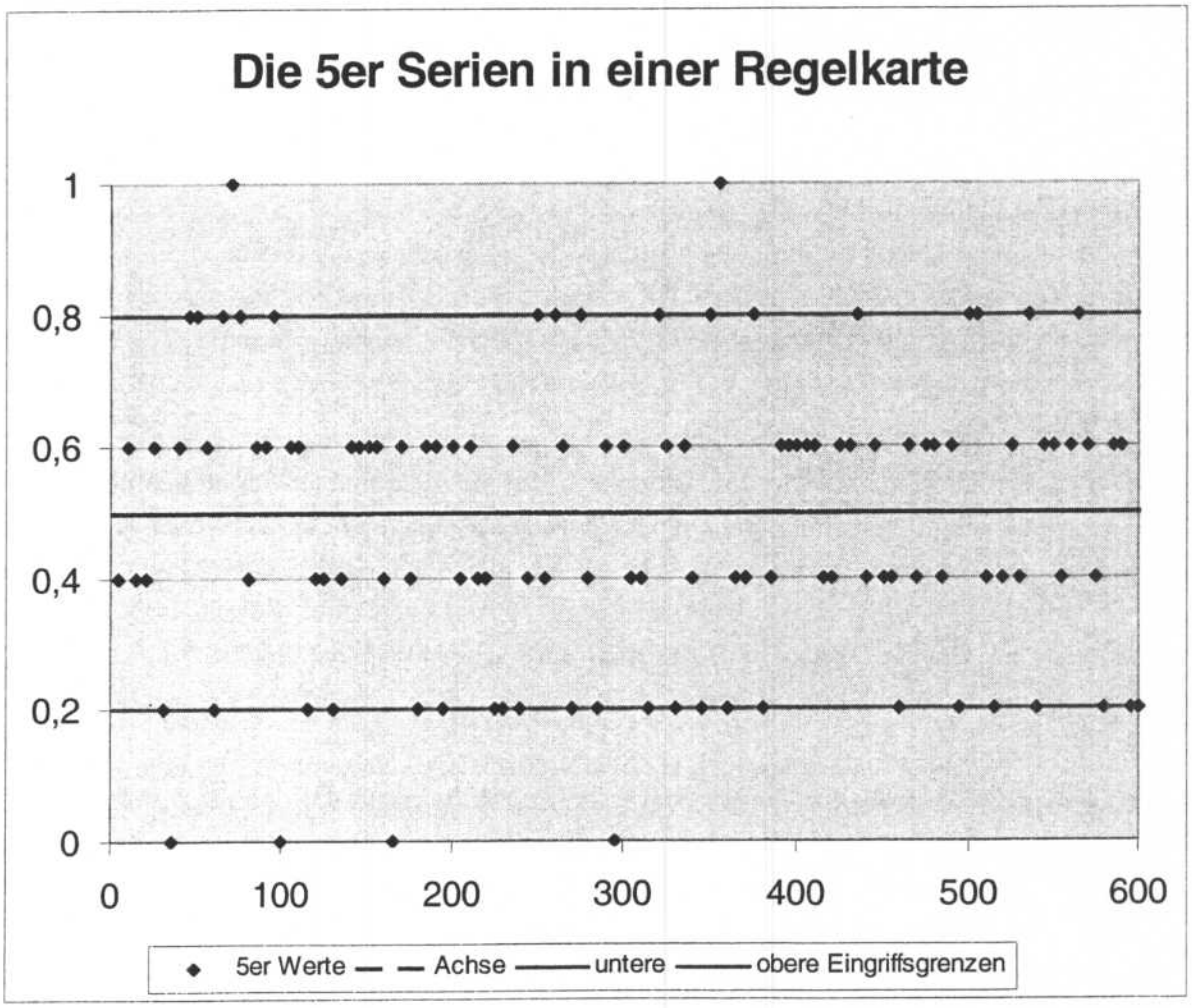


Fig. 5: Die Messwerte von Fünfer-Serien nun in einer Qualitätsregelkarte mit Sollwert bei ½ und Eingriffsgrenzen bei 0,2 sowie 0,8.

Dieser Begriff des Restrisikos zieht sich weiter durch die Statistik bis hin zu Vertrauensintervallen; immer wird es ein Risiko geben, ein Ergebnis zu erhalten, das die vorgegebenen Genauigkeitsschranken überschreitet. Es geht in den Anwendungen jedoch darum, Genauigkeitsforderungen und Restrisiko auf ein vernünftiges Maß zu reduzieren. Das meint man, wenn man sagt, den Zufall berechenbar machen. Die Berechnung gibt demnach nicht eine exakte Voraussage, was die nächste (hier) Fünfer-Serie erbringen wird, sondern bezieht sich auf die Angabe eines Restrisikos, mit dem gewisse Genauigkeitsforderungen bei (alleinigem) Wirken des Zufalls doch überschritten werden.

### *Verbesserung der Präzision*

Es wurde schon klar, dass Mess-Serien der Länge 5 wohl richtig zentriert sind, aber im Hinblick auf die Präzision der Messung eine Reihe von Wünschen offen lassen. Kann eine längere Mess-Serie Abhilfe schaffen? Werden je zwei benachbarte Fünfer-Serien zusammengefasst, so erhält man eine Zehner-Serie, schließlich erhält man rasch auch die Mess-Ergebnisse von Serien der Länge 20. In der Tabellenkalkulation fügt man einfach neue Spalten hinzu und summiert geeignet aus der Vorspalte. Die Mess-Werte stellt man sodann in gleicher Weise in Qualitätsregelkarten dar.

Aus Figur 6 wird die Zunahme der Präzision von Fünfer- auf Zwanziger-Serien überdeutlich. Die Mess-Werte liegen nun alle, mit einer einzigen Ausnahme, innerhalb der Schwellenwerte 0,2 und 0,8, ja sogar deutlich noch enger um die Achse ½. Bei gleichen Genauigkeitsforderungen hat man also ein geringeres Restrisiko, diese Forderungen zu verfehlen. Man könnte auch umgekehrt formulieren, bei gleichem Restrisiko, das man einzugehen bereit ist, erhält man eine größere Präzision – eine diesbezügliche Abbildung ist allerdings etwas schwieriger zu erstellen, weshalb sie entfällt.

Die an sich problematische Konvergenz der relativen Häufigkeiten wird gänzlich umgangen, der Effekt jedoch und damit das komplizierte Verhältnis von Konvergenz und Restrisiko zeigt sich überdeutlich: Die Mess-Prozedur ‚Messung der Wahrscheinlichkeit $p=½$ durch die relative Häufigkeit einer Serie von bestimmter Länge' scheint durch die Achse in Figur 6 tatsächlich das zu messen, was sie zu messen vorgibt. Ein fiktives Annähern an einen obskuren Grenzwert wird nicht zur Debatte gestellt, vielmehr zeigen die Regelkarten deutlich eine Achse in der Nähe von $p=½$. Das Annähern wird in der Qualität der Verbesserung der Präzision der Messung von 5 (auf 10 und) auf 20 Basisdaten in der zur Verfügung stehenden Mess-Serie sichtbar gemacht. Es bedarf keiner Konvergenzaussage in dem Sinne, was passiert, wenn die Länge der Serie noch (und über alle Maßen) erhöht wird. Der Effekt der Verbesserung darf ohne Vorbehalte von den bestehenden Daten übertragen werden, d.h. auch bei Erhöhung von 20 auf 40 Daten in der Mess-Serie kann man mit einer weiteren Verbesserung rechnen.

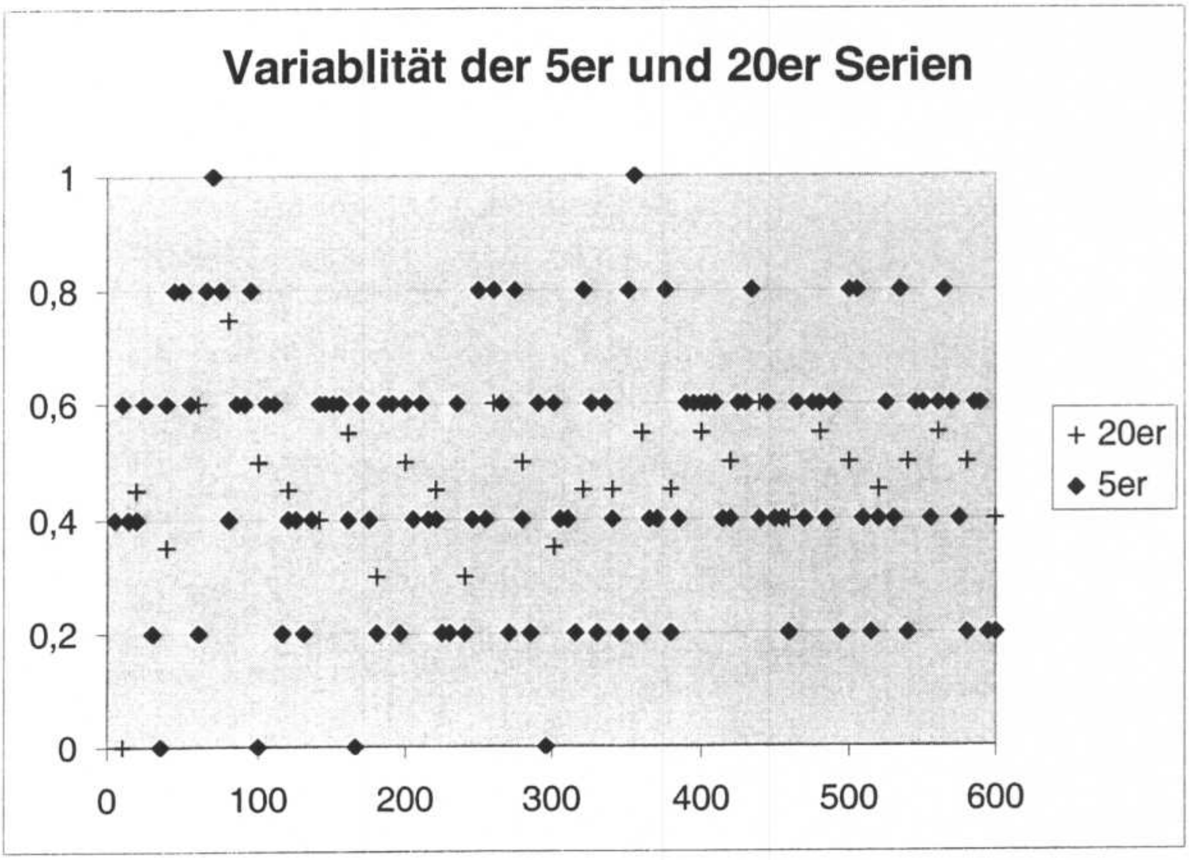


Fig. 6: Vergleich von Fünfer- und Zwanziger-Serien – die Achse bleibt gleich, die Variabilität der Mess-Werte nimmt deutlich ab.

Besonders wichtig im Hinblick auf statistisches Schließen ist jedoch der Begriff des Restrisikos; wie immer man die Länge der Mess-Serie erhöht, man hat ein gewisses Risiko zu tragen, dass noch immer Serien mit Mess-Werten außerhalb der Toleranzen auftreten, wenn schon nicht in den bestehenden Daten, so möglicherweise in einem neuen Experiment.

### 3. Untersuchung der Reichweite der erkannten Gesetze

#### *Geltungsbereich empirischer Gesetze*

Hat man bis hierher schon virtuellen Zufall über Computer-Simulationen eingesetzt, so ergeben sich u.a. zwei unterschiedliche Fragen: Aus der Sicht der Lernenden ist anzuzweifeln, ob die Ergebnisse geradezu durch den Lehrenden manipuliert worden sind, um ihnen eine Regelmäßigkeit vorzuführen. Aus allgemeinerer Sicht ist abzusichern, ob die Regelmäßigkeiten nicht Folge der für die Erzeugung des Zufalls verwendeten Algorithmen sind und daher ein reines

Artefakt darstellen. Solche Fragen stellen sich jetzt nicht, können aber bei Erweiterung der Datenerzeugung durch Computer-Simulation durch einen Vergleich mit den gegenwärtigen, materiell erzeugten Daten geprüft werden. Es kann also untersucht werden, ob die erkannten empirischen Gesetzmäßigkeiten unabhängig von der Art der Erzeugung des Zufalls gelten.

Für empirische Gesetze ist - anders als bei mathematisch-logischen und kausalen Gesetzen - besonders wichtig, ihren Geltungsbereich auszuloten; also die Frage zu untersuchen, unter welchen Bedingungen taucht das Gesetz, allenfalls in Verfeinerungen, ebenfalls auf.

- Ist das Gesetz der Achse der Messungen und der Erhöhung der Präzision der Messungen bei Verlängerung der Mess-Serie auch bei einer Wiederholung des gesamten Experiments gültig?
- Trifft es auch zu, wenn die Achse nicht bei ½ sondern bei einem anderen Wert p liegt?
- Gilt die Erhöhung der Präzision auch für andere Werte von p?
- Ist die Erhöhung der Präzision auch bei größeren Serien gültig?
- Ist die Zunahme der Präzision immer gleich groß, wenn man die Serie etwa um 100 verlängert, unabhängig von der bisherigen Länge der Serie?
- Oder ist die Zunahme der Präzision gleich groß, wenn man die Serie doppelt so lang etc. macht?
- Kann man die Erhöhung der Präzision bei gleichem Restrisiko durch eine mathematische Funktion beschreiben, welche Eigenschaften hätte allenfalls eine solche Funktion?

Usw.

Computer-simulierte Daten haben den Vorteil der raschen Erzeugung sowie der leichten Analyse in Tabellen und Graphen. Der Vergleich mit den materiell erzeugten Daten sollte nicht unterbleiben. Rasch wird man einsehen, dass die Regelmäßigkeiten dieselben sind.

### *Messung anderer Wahrscheinlichkeiten als ½*

Mit dem Würfel selbst hätte man eine schöne Möglichkeit, auf ähnliche Weise Mess-Werte für $p=1/6$ materiell zu simulieren; der Aufwand lohnt jedoch kaum. Hier sollte man bereits auf computer-erzeugte Werte ausweichen und die gewonnene Zeit auf die Analysen verwenden.

Wieder ist eine Achse der Mess-Werte zu erkennen, die in der Analogie dem Sollwert entspricht, d.h. die Achse liegt bei dem vorgegebenen p. Es liegt nahe, eine solche Achse auch dann anzunehmen, wenn man den Wert von p nicht kennt. In Figur 7 zeigt sich ein ähnliches Streuverhalten und damit eine vergleichbare Präzision wie bei $p=½$; der Streubereich lag dort bei 10 +/- 5 und ist nun etwa bei 5 +/- 5 für die Anzahl der ‚Köpfe'.

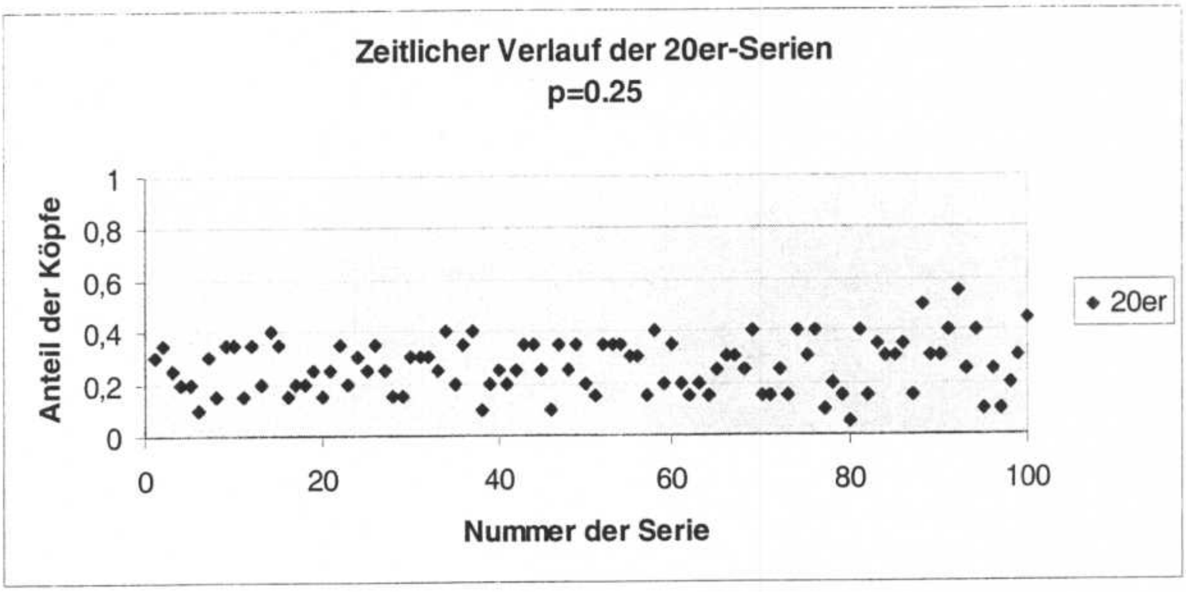


Fig. 7: Computer-erzeugte Mess-Werte für p=0,25 aus Serien der Länge 20.

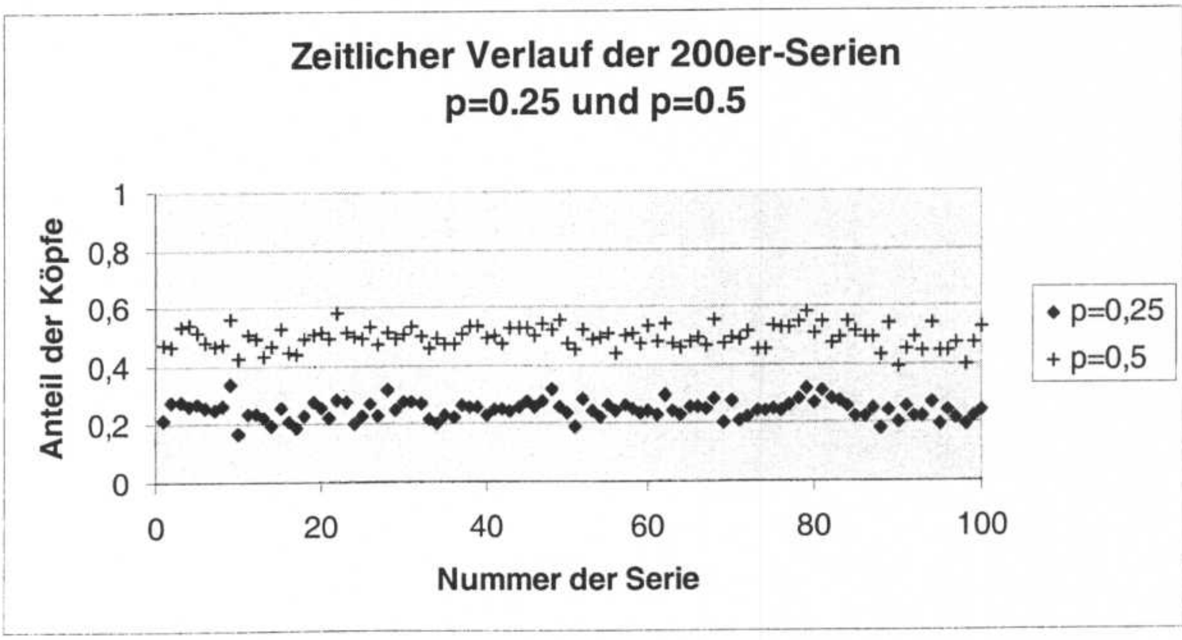


Fig. 8: Computer-erzeugte Mess-Werte für p=½ und p=0,25 aus Serien der Länge 200.

Aus Figur 8 kann man überdies noch erkennen, dass der Streubereich für p=0,25 mit +/- 12 etwas kleiner ausfällt als bei p=½ (100 +/-14) – zumindest wenn man das schon theoretisch weiß und genauer hinsieht. Die Präzision an sich scheint auch etwas mit der Größe des zu messenden Wertes zusammenzuhängen.

*Einfluss der Länge der Mess-Serie*

Die Monotonie, die Zunahme der Präzision der Messungen von 5 auf 20, tritt auch bei weiterer Verlängerung der Mess-Serie von 20 auf 200 auf, wie man leicht aus Figur 9 sieht.

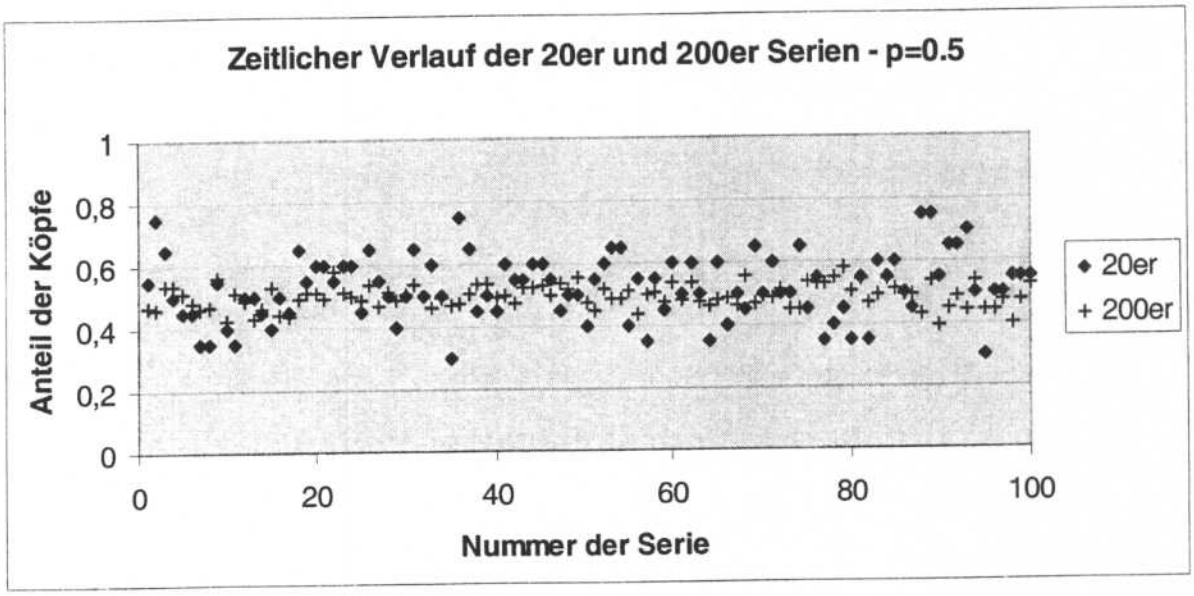


Fig. 9: Vergleich von Messungen aufgrund von 20-er- und 200-er-Serien.

Aus Figur 10 sieht man, dass das Gesetz von Stabilisierung bei Verlängerung der Serie und Variabilität jeder neuen kurzen Mess-Serie weiterhin aufrecht erhalten bleibt.

Die Fluktuation der Mess-Werte – nun von 200-er Serien – ist rein auf ein ‚Spiel des Zufalls' zurückzuführen, unterstellt man einmal, dass auch der virtuelle Zufall den Zufall repräsentieren kann. Man kann nun auch untersuchen, welche zentrierte Bandbreite etwa 95% der Mess-Werte für die Wahrscheinlichkeit umfasst. Die Schwellenwerte im Sinne der Analogie mit den Qualitätsregelkarten sind dann so angesetzt, dass eine Unterschreitung der unteren in 2,5%, eine Überschreitung der oberen ebenso in 2,5% eintritt (wiederum gibt es eine Fluktuation der tatsächlichen Prozentsätze von Experiment von Experiment). Zieht man die Achse symmetrisch in diese Bandbreite ein, so hat man die Hälfte dieser Bandbreite sozusagen als Genauigkeit, als Größenordung des Fehlers der Messung der Wahrscheinlichkeit. Entsprechende Abbildungen sind etwas mühsamer zu erstellen und seien der gebotenen Kürze wegen ausgelassen.

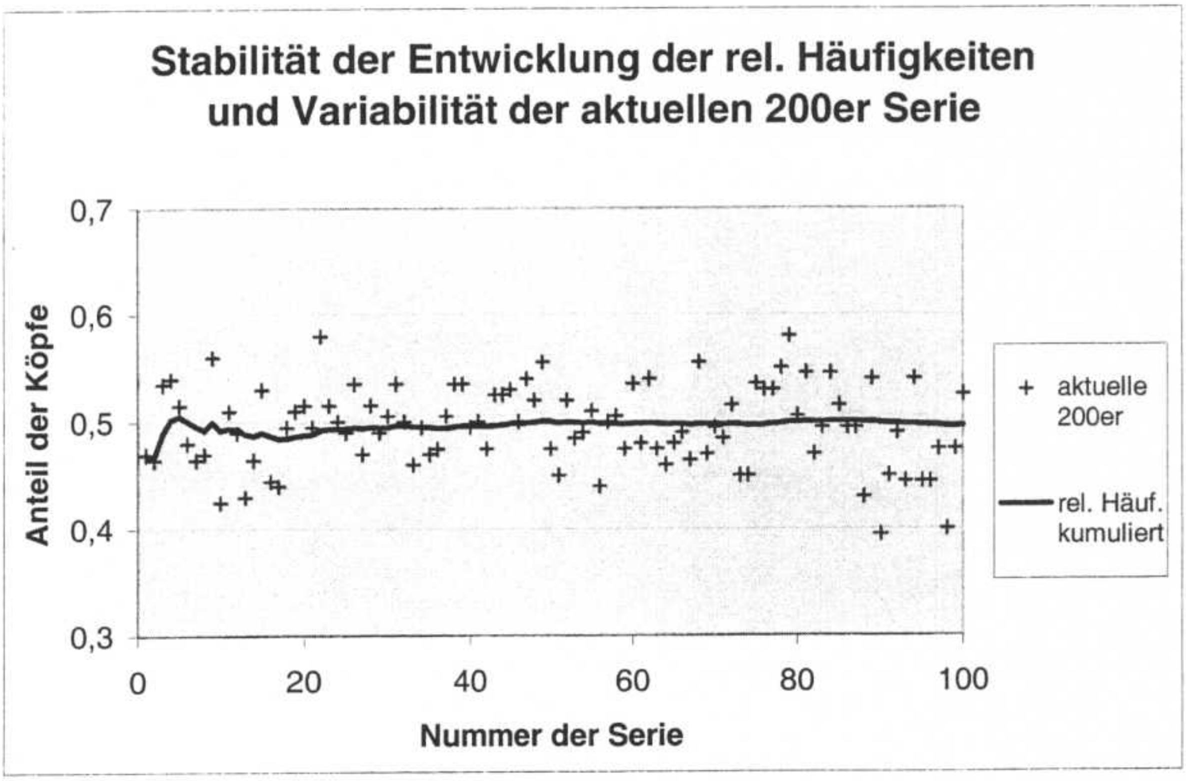


Fig. 10: Gesetz von Stabilisierung und Variabilität gilt auch für fortgesetzte 200-er-Serien.

Die Größenordnung des Spiels des Zufalls stellt sozusagen die normale Bandbreite der Variabilität der Mess-Werte dar, es besteht klarerweise ein Restrisiko – in diesem Fall 5% - der Überschreitung, also des Auftretens extremer Werte. Das Spiel des Zufalls bzw. dessen Größenordnung ist in der Analogie mit dem Messen der tolerierte Mess-Fehler; das Restrisiko ist jenes Risiko, doch einen noch größeren Mess-Fehler auf Basis der vorgegebenen Mess-Serie bestimmter Länge zu erhalten. Der Zusammenhang zwischen Größenordnung des Spiels des Zufalls und dem Restrisiko ist gegenläufig: Je größer das Spiel des Zufalls umso kleiner das Restrisiko. Beide Parameter gleichzeitig zu verbessern gelingt nur durch Verlängerung der Mess-Serie.

Der Schritt zu Vertrauensintervallen für eine unbekannte Wahrscheinlichkeit p mit Hilfe der festgestellten relativen Häufigkeit aufgrund einer Mess-Serie der Länge n ist nur ein kleiner: Wenn aufgrund etwa einer Mess-Serie der Länge 100 ein Mess-Fehler für die unbekannte Wahrscheinlichkeit von +/- 0,10 mit einem Restrisiko von 5% behaftet ist, so sollte bei Vorliegen des Mess-Werts 0,37 der ‚wahre' Wert für p nicht unter 0,37 – 0,10 = 0,27 und nicht über 0,37 + 0,10 = 0,47 liegen. (Dass die Größenordnung der Mess-Fehler neben der Länge der Mess-Serie auch noch die unbekannte Wahrscheinlichkeit als Einflussgröße hat, sei vorerst unterschlagen – der daraus resultierende Fehler ist nur bei Wahrscheinlichkeiten nahe bei 0 bzw. nahe bei 1 und sehr kleinen Mess-Serien relevant).

*Variabilität des Mess-Werts*

Es wurde schon angedeutet, das die Variabilität der Mess-Werte unterschiedlich betrachtet werden kann. Zum einen gilt es, bei Verlängerung der bestehenden Serie die Veränderlichkeit des aktuellen Mess-Werts zu beurteilen, etwa bei Erhöhung der Länge der Serie von 400 auf 500. Eine Graphik sei der Kürze wegen ausgelassen, man kann sich leicht überlegen, dass die Bandbreite des Spiels des Zufalls bei +/- 0,02 liegt, bezogen auf ein Restrisiko von 5% (und auf p-Werte, die etwa zwischen 0,3 und 0,7 liegen). Dagegen liegt die Wiederholstreuung bei einer neuen Serie, das heißt, man beginnt bei 0 und erzeugt 500 neue Daten, bei ca. +/- 0,05 – wiederum bezogen auf ein Restrisiko von 5%. Es war schon die Rede davon, dass die Analyse des Verlaufs einer Entwicklung der relativen Häufigkeiten die kleinere der beiden Wiederholstreuungen wiedergibt und damit die tatsächliche Größenordnung der Mess-Fehler verschleiert.

**4. Ausbau des Zugangs**

***Modellierung der Simulation – Ein Weg zur Binomialverteilung***

Will man etwa die Messung der Wahrscheinlichkeit ½ aufgrund einer Mess-Serie der Länge 20 simulieren, so kann man jeden einzelnen Versuch der Reihe nach gemäß dem Binomialmodell modellieren. Die Voraussetzungen im einzelnen sind: Jeder Versuch hat dieselbe ‚Erfolgswahrscheinlichkeit' (für 'Kopf') von ½ und zwar unabhängig von den Ergebnissen der bisherigen Versuche. Man kann das Modell auch schrittweise durch die Pascal'sche Rekursion beschreiben: Wenn bei 20 Versuchen z.B. 17-mal Erfolg (i.f. als 'Kopf' beim Münzwerfen angesprochen) eingetreten ist, so kann das nur daher rühren, dass bei 19 Versuchen 16mal 'Kopf' war und dann im 20. Versuch 'Kopf' aufgetreten ist, oder bei 19 Versuchen 17-mal schon 'Kopf' war und beim 20. Versuch 'Kopf' ausgeblieben ist. So kann man sich schrittweise nach oben ‚hanteln' und landet schließlich bei einem Versuch mit bekannten Wahrscheinlichkeiten.

Schließlich kann man die Rekursion schrittweise von oben nach unten – also mit zunehmender Länge der Serie – in der Tabellenkalkulation berechnen. Gewisse Erfahrung mit Tabellenkalkulation erleichtert ganz erheblich die Erstellung des Tableaus in Figur 11. Es entstehen zeilenweise die Binomialverteilungen bis n=20. Eine weitere Fortsetzung für größere n ist ohne weiteres möglich. Die zeilenweisen Binomialverteilungen, insbesondere die letzte mit n=20 können dann mit den empirischen Verteilungen der Mess-Werte aus dem Münzwurfexperiment verglichen werden. Ein tabellarischer und graphischer Vergleich zeigt eine gute Übereinstimmung, was die Güte des Binomialmodells zur Beschreibung des Münzwurfs unterstreicht.

Dieser Zugang zur Binomialverteilung verzichtet vollständig auf Kombinatorik und das umständliche Rechnen mit der Verteilungsfunktion der Binomialverteilung, die sich nun durch einfaches Summieren über Teile der Zeilen ergibt. Eine Verallgemeinerung des Tabellenblatts auf p, die von ½ verschieden sind, ist ganz einfach. Man kann p sogar variabel halten und erst bei Bedarf festlegen.

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **p=** | 0,500 | | | | | | | | | | |
| 2 | | **k=** | | | | | | | | | | |
| 3 | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 4 | Hilfszeile | *1* | *0* | *0* | *0* | *0* | *0* | *0* | *0* | *0* | *0* | *0* |
| 5 | **n = 1** | 0,500 | 0,500 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 2 | 0,250 | 0,500 | 0,250 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 3 | 0,125 | 0,375 | 0,375 | 0,125 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 4 | 0,063 | 0,250 | 0,375 | 0,250 | 0,063 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | **n = 5** | 0,031 | 0,156 | 0,313 | 0,313 | 0,156 | 0,031 | 0 | 0 | 0 | 0 | 0 |
| 10 | 6 | 0,016 | 0,094 | 0,234 | 0,313 | 0,234 | 0,094 | 0,016 | 0 | 0 | 0 | 0 |
| 11 | 7 | 0,008 | 0,055 | 0,164 | 0,273 | 0,273 | 0,164 | 0,055 | 0,008 | 0 | 0 | 0 |
| 12 | 8 | 0,004 | 0,031 | 0,109 | 0,219 | 0,273 | 0,219 | 0,109 | 0,031 | 0,004 | 0 | 0 |
| 13 | 9 | 0,002 | 0,018 | 0,070 | 0,164 | 0,246 | 0,246 | 0,164 | 0,070 | 0,018 | 0,002 | 0 |
| 14 | **n = 10** | 0,001 | 0,010 | 0,044 | 0,117 | 0,205 | 0,246 | 0,205 | 0,117 | 0,044 | 0,010 | 0,001 |
| 15 | | | | | | | | | | | | |

Fig.11: Pascal'sche Rekursion und schrittweiser Aufbau der Binomialverteilungen, aus Platzgründen nur bis n=10; p=½.

Hinsichtlich Variabilität und Verteilung lassen sich die Mess-Werte für p=½ aus den simulierten Mess-Serien der Länge 20 mit der Binomialverteilung mit n=20 und p=½, wie sie sich aus der Pascal'schen Rekursion ergibt, rasch vergleichen – die Übereinstimmung ist beeindruckend – eine entsprechende Abbildung sei hier wiederum dem Leser / der Leserin selbst überlassen.

### *Statistische Beurteilung – Ein Ausblick*

In der angewandten Statistik steht man häufig vor folgender Fragestellung: Eine Versuchsgruppe hat eine bestimmte Behandlung erhalten, die Werte der Zielgröße seien der Reihe nach mit

$$\text{VG: } x_1, x_2, \ldots, x_m$$

bezeichnet. Eine Kontrollgruppe, die sich in wesentlichen Eigenschaften von den Mitgliedern der Versuchsgruppe nicht unterscheidet, hat keine Behandlung erhalten, ihre Werte für die Zielgröße seien

$$\text{KG: } y_1, y_2, \ldots, y_n$$

Man soll aufgrund dieser Daten beurteilen, ob die Behandlung als erfolgreich angesehen werden kann. Man beobachtetet demnach einen Mittelwert-Unterschied x quer – y quer$$$ und muss nun entscheiden, ob mit gewissen Risiken klarerweise, ein solcher Unterschied ‚signifikant' von Null verschieden ist oder nicht. Ist der beobachtete Mittelwert-Unterschied signifikant größer als Null, so kann man, mit dem besagten Restrisiko, behaupten, dass die Behandlung zu größeren Werten der Zielgröße führt.

Was hat das nun mit Münzwerfen und dem Spiel des Zufalls zu tun? Zunächst sei festgehalten, dass ein festgestellter Mittelwert-Unterschied zufolge von Störgrößen von Null (entspricht keinem Unterschied in Versuchs- und Kontrollgruppe) verschieden sein kann – so etwa sind die Subjekte von Versuchs- und Kontrollgruppe doch nicht ganz gleich (gesund, stark, etc.). Ihre Eigenschaften beeinflussen daher die Zielgröße direkt. Solche Merkmale nennt man Confounder. Viele Flops in statistischen Anwendungen sind darauf zurückzuführen, dass man wesentliche Confounder übersehen hat. Schaltet man erfolgreich alle einflussreichen Confounder aus, so bleibt nur die Behandlung als einziger wesentlicher Unterschied zwischen Subjekten der Versuchs- und Kontrollgruppe übrig, sodass man die Unterschiede in der Zielgröße darauf zurückführen kann.

Es geht dann nur noch um den Vergleich des festgestellten Mittelwert-Unterschieds mit dem Spiel des reinen Zufalls. Klar ist, dass nicht alle Stichproben immer gleich ausgehen werden. Wenn nun der reine Zufall auch einen solchen Unterschied normalerweise erzeugt, so sollte man nicht von einem Erfolg der Behandlung sprechen. Erst wenn man mit der Beobachtung über dieses Spiel des Zufalls hinaus reicht, kann man die Behandlung als ‚Ursache' für die Unterschiede ausweisen; man sagt dann, die Unterschiede zwischen Versuchs- und Kontrollgruppe sind ‚statistisch signifikant'.

Was sind nun andere Stichproben, wie erhält man diese und wie kann man aus diesen Stichproben eine normale Bandbreite des Zufalls ablesen?

Schritt 1: Reiner Zufall

- Man wähle die m Daten für x aus allen kombinierten m+n Daten *zufällig*.
- Man bestimme Mittelwerte für Versuchs- und Kontrollgruppe und erhält daraus einen ersten Mess-Wert für die Differenz der Mittelwerte.

Schritt 2: Szenario von fiktiven Messungen durch Wiederholung von Schritt 1

- Die Wiederholung von Schritt 1 ergibt eine *Reihe von Mess-Werten* für die Differenz der Mittelwerte.
- Trägt man diese Mess-Werte in eine ‚Qualitätsregelkarte' ein, so kann man wie beim Münzwerfen eine Achse erkennen – diese *Basislinie* zeigt an, was zu messen ist. Wie beim Münzwerfen kann man auch eine Bandbreite einzeichnen, die 95% der Mess-Werte einfängt. Diese *95%-Linien der Schwankung* zeigen an, wie genau gemessen wird.

Schritt 3: Entscheidung

- Wenn diese 95%-Linien der Schwankung den Mittelwert-Unterschied Null nicht enthalten, so kann man sagen, die Null ist nicht durch Spiel des reinen Zufalls zu interpretieren. Mit einem Restrisiko von 5% kann man dann davon ausgehen, dass die Behandlung zu unterschiedlichen Werten in der Zielgröße führt. Typisch ist wiederum das mit bestimmten Entscheidungen verbundene Restrisiko.
- Man kann das Risiko verkleinern, indem man etwa 99%-Linien der Schwankung einzieht. Dann allerdings wird ein anderer Fehler – der sogenannte Fehler 2. Art - wahrscheinlicher, nämlich derjenige Fehler, einen bestehenden Unterschied zwischen Versuchs- und Kontrollgruppe dem Spiel des reinen Zufalls zuzuschreiben – doch das ist wieder eine andere Geschichte.

**5. Schluss**

Es sei darauf verwiesen, dass hier ein direkter Zugang zur beurteilenden Statistik möglich wird, wie er schon von Borovcnik (1985) vorgeschlagen wurde. Als Nebenprodukt ergibt sich ein Steilkurs durch die Wahrscheinlichkeitsrechnung, wobei vom einführenden Münzwurf-Experiment direkt zum Gesetz der großen Zahlen sowie zur Binomialverteilung ohne Kombinatorik geführt wird. Die Herangehensweise an das Gesetz der großen Zahlen unter Vermeidung des angesprochenen obskuren Grenzwerts der relativen Häufigkeiten geht auf eine Idee von Freudenthal (1972) zurück. Sie ist ausführlich in Borovcnik (1992) dargestellt. Die Schwierigkeiten mit physikalischen sowie stochastischen Zusammenhängen und Gesetzen ist u.a. in Borovcnik (1989) beschrieben. Zur Eigenheit von Wahrscheinlichkeit als Begriff ist einiges in Borovcnik und Peard (1996) nachzulesen. Zu intuitiven Vorstellungen von Wahrscheinlichkeit und den daraus resultierenden Schwierigkeiten für die Lernenden kann man sich in Borovcnik und Bentz (1991) orientieren.

Zur eingesetzten Tabellenkalkulation EXCEL ist zu sagen, dass sie eine weite Verbreitung hat, viele Lernende haben sie zu Hause auf ihrem PC oder dem PC der Eltern verfügbar. Das spricht für EXCEL. Diese Tabellenkalkulation hat leider auch ihre spezifischen Eigenheiten, die den Gebrauch ein wenig erschweren, wie immer wünscht sich der Anwender mehr und bessere

Führung durch die Möglichkeiten. Gute Hinweise zum effizienteren Einsatz von EXCEL findet man immer bei E. Neuwirth im Internet.

Computer-Räume an den Schulen sind weitgehend verfügbar. Mit all den Nachteilen auch, man muss das Klassenzimmer wechseln, immer wieder fallen Geräte trotz guter Wartung aus usw.

Lernende haben im Unterrichtsexperiment große Unterschiede hinsichtlich ihrer Fähigkeiten im Gebrauch der Ressourcen gezeigt, mindestens aber haben fast alle hoch motiviert am Unterricht teilgenommen.

Literatur:

Borovcnik, M.: Ein direkter Zugang zur Beurteilenden Statistik. In: *Didaktik der Mathematik* **13**(1985), 251-271.

Borovcnik, M.: Statistische Analyse von Zusammenhängen – Regression und Korrelation. In: *Schriftenreihe zur Didaktik der Mathematik d. Höheren Schule der ÖMG* **17**(1989), 1-20.

Borovcnik, M.: *Stochastik im Wechselspiel von Intuitionen und Mathematik.* Mannheim: Bibliographisches Institut 1992.

Borovcnik, M., Bentz, H.-J.: Empirical research in understanding probability. In: Kapadia, R. u. Borovcnik, M.: *Chance encounters*. Dordrecht: Kluwer 1991, 73-106.

Borovcnik, M., Peird, R.: Probability. In: Bishop, A., Clements, K., Keitel, C., Kilpatrick, J., Laborde, C.: *International Handbook of Mathematics Education*. Dordrecht: Kluwer 1996, 239-288.

Fischbein, E.: *Intuitions in science and mathematics. An educational approach*. Dordrecht: D. Reidel 1987.

Freudenthal, H.: ‚The empirical law of large numbers' or 'The stability of frequencies'. In: *Educational Studies in Mathematics* **4**(1972), 484-490.

Neuwirth, E.: http://sunsite.univie.ac.at/mailman/listinfo/Improve-excel