

Sammelproblem: Eine Simulation mit Excel

MANFRED KÜHLEITNER, WIEN

Zusammenfassung: In regelmäßigen Abständen, jedenfalls zu den Welt- und Europameisterschaften im Fußball, erscheinen Sammelbilder. Auch Lebensmittelhersteller versuchen ihren Umsatz durch Beigabe von Sammelobjekten in ihren Produkten zu steigern. Dieser Artikel erklärt, wie man mit dem Tabellenkalkulationsprogramm MS ExcelTM das Sammelproblem simuliert und daraus quantitative Aussagen über die Anzahl von Käufen für eine vollständige Serie, für eine große Menge von Kunden, automatisiert erstellt.

1 Problembeschreibung

In regelmäßigen Abständen, jedenfalls zu den Welt- und Europameisterschaften im Fußball, erscheinen Sammelbilder. Auch Lebensmittelhersteller versuchen ihren Umsatz durch Beigabe von Sammelobjekten in ihren Produkten zu steigern. Wir betrachten eine Sammelbildserie, welche aus sechs verschiedenen Bildern besteht. Leider lässt sich beim Kauf einer Packung nicht feststellen, welches Bild man erwirbt. Wie viele Bilder muss man nun kaufen, um eine komplette Serie in Händen zu halten? Das ist verschieden und hängt vom Käufer ab. Ein Glückspilz kauft sechs Bilder und hat damit alle sechs verschiedenen Teile. Einem Pechvogel hingegen können auch nach 50 Käufen noch Teile fehlen. Da es in einem solchen Spiel aber nicht auf Einzelpersonen ankommt, sondern auf die große Masse, interessiert uns, wie viele Personen mit 6 Käufen auskommen, wie viele 7, 8, 9, 10 usw. Käufe benötigen.

Man kann diese Aufgabe mit Hilfe der Wahrscheinlichkeitsrechnung theoretisch lösen (Althoff, 2000). Einfacher - und vor allem flexibler in Bezug auf Änderungen des Sammelspiels - ist es, die Aufgabe mit Hilfe einer Simulation zu lösen. Stellen wir uns einen Käufer vor. Wie viele Käufe benötigt dieser? Dazu simulieren wir den Sammelvorgang durch das Werfen eines Würfels. Es wird solange gewürfelt, bis jede Augenzahl genau einmal aufgetreten ist. Die Anzahl der Würfe ist dann gleich der Anzahl der Käufe. Diese notieren wir. Anschließend wiederholen wir diese Vorgangsweise z.B. 1000-mal.

Abschließend werden die so gewonnenen Daten (Anzahl der Käufe bis zum Erlangen der vollständigen Serie von 1000 Kunden) in einer Häufig-

keitsverteilung grafisch dargestellt und deren statistische Kennzahlen (Mittelwert und Standardabweichung) ermittelt. Da die 1000 Kunden nur eine kleine Stichprobe darstellen, können wir die Genauigkeit unserer Ergebnisse prüfen, indem wir den Vorgang wiederholen und für weitere 1000 Kunden die notwendigen Käufe durch Würfeln simulieren.

Um diese sehr zeitaufwändige Arbeit des Würfeln zu automatisieren, verwenden wir das Tabellenkalkulationsmodell Excel, weil es leicht verständlich, allgemein verfügbar und in Schulen bereits häufig verwendet wird.

2 Simulation in Excel

Wie oben beschrieben, simulieren wir den Sammelvorgang durch das Werfen eines Würfels. Ein Kauf liefert daher eine zwischen 1 und 6 *gleichverteilte Zufallszahl*, wobei jede Zahl für ein Sammelbild steht. Als erstes fertigen wir in Excel ein Tabellenblatt mit 100 Würfeln nach dem Muster der Tabelle 1 an. In die erste Zeile kommt erklärender Text.

	A	B
1	Wurf Nr.	Augenzahl

Tabelle 1. Excel Tabellenblatt, Teil 1.

Das Würfeln simulieren wir in Excel mit dem Befehl ZUFALLSBEREICH(min; max). Dieser liefert gleichverteilte, ganze Zufallszahlen zwischen min und max. Ist diese Funktion nicht verfügbar, so muss man vorher das Makro „Analyse-Funktionen“ installieren.

In der Spalte A werden die einzelnen Käufe eines Kunden gezählt. Wir schreiben in die Zelle A3 die Zahl 1 und in die Zelle A4 die Formel =A3+1, die wir bis in die Zelle 102 kopieren. In der Spalte B simulieren wir das Würfeln. Dazu wird in die Zelle B3 die Formel =ZUFALLSBEREICH(1;6) geschrieben und bis in die Zelle B102 kopiert. Nun müssen wir nur mehr zählen, wie viele Käufe notwendig sind bis alle 6 Zahlen vorkommen. Durch Drücken der Taste F9 werden neue Zufallszahlen erzeugt. Diese stellen den Kaufakt für einen neuen Kunden dar. Damit ist das Würfeln automatisiert.

Nun wollen wir auch das händische Zählen zur Ermittlung der Anzahl der notwendigen Käufe automatisieren. Dazu wird die Tabelle um 6 weitere Spalten erweitert. Wir schreiben in die Zellen C2:H2 die Zahlen 1,2, ... bis 6. Zuerst stellen wir fest, welches der sechs Bilder beim ersten Kauf erworben wurde, d.h. welche Augenzahl beim ersten Wurf aufgetreten ist. Dies lässt sich in Excel mit Hilfe der Funktion WENN(Bedingung; Dann; Sonst) realisieren. Dazu schreiben wir in die Zelle C3 die Formel =WENN(\$B3=C2; 1; 0). Diese Formel kopieren wir nun bis in die Zelle H2. Anschließend schreiben wir in die Zelle C4 eine Formel, mit der wir nach jedem Kauf sagen können, wie viele Bilder des jeweiligen Typs man

bereits gekauft hat. Dazu schreiben wir in C4 die Formel =WENN(\$B4=C\$2; C3+1; C3). Diese kopieren wir nach rechts bis in die Spalte H und nach unten bis in die Zeile 102. Jetzt zählen wir in der Spalte I zeilenweise, ob noch Bilder fehlen. Dazu bestimmen wir die Anzahl der Nullen mit Hilfe des Befehls =ZÄHLENWENN(C3:H3; 0) in der Zelle I3. Auch diese Formel wird bis in die Zeile 102 kopiert. Aus der Spalte I lässt sich die Anzahl der nötigen Käufe für einen vollständigen Satz sehr leicht mit Hilfe der Formel 100-(Anzahl der Nullen in Spalte I) +1 berechnen. Wir schreiben daher in die Zelle J3 die Formel =100-ZÄHLENWENN(I3:I102; 0)+1.

	C	D	E	F	G	H	I	J
1	Anzahl #						Zähle 0	Anzahl Käufe
2	1	2	3	4	5	6		

Tabelle 2. Excel Tabellenblatt, Teil 2.

In der Zelle J3 steht somit die Anzahl der notwendigen Käufe für einen Kunden bis zur Vollständigkeit der Serie. Durch Drücken der Taste F9 erhält man neue Zufallszahlen und in der Zelle J3 automatisch die Anzahl der notwendigen Käufe dieses Kunden.

3 Wiederholte Simulation

Man kann das Drücken der Taste F9 Excel übernehmen lassen. Dazu muss man allerdings ein kleines Makro in der Programmiersprache VBA schreiben. Man ruft VBA mit der Tastenkombination Alt+F11 auf. Hier geht man auf >Einfügen/Modul. Nun erscheint ein leeres Blatt, in welches wir folgenden kurzen Code eingeben:

```
Sub Makro1()
For n = 3 To 1002
Cells(n, 11).Value
= Cells(3, 10).Value
Next n
End Sub
```

Jetzt lassen wir das Makro laufen. Es nutzt unser Tabellenblatt als „Unterprogramm“. Nach Beendigung des Makros stehen in den Zellen K3 bis K1002 die Anzahl der notwendigen Käufe bis zur Komplettierung der Serie von 1000 verschiedenen Kunden.

In freien Zellen berechnen wir mit Hilfe der Formeln =MITTELWERT(K3:K1002) bzw. =STABW(K3:K1002) den Mittelwert bzw. die

Standardabweichung dieser Zahlen. Zur Orientierung: Der Erwartungswert beträgt etwa 14,7 und die Varianz 38,9. Durch erneutes Starten des Makros erzeugen wir eine neue Stichprobe von 1000 Kunden und können so die Genauigkeit von Mittelwert und Standardabweichung beurteilen!

Um die Mehrarbeit zu nutzen, zeichnen wir noch die Häufigkeitsverteilung und die kumulative Häufigkeitsverteilung der Daten aus der Spalte K (Abbildung 1). Zur Vorbereitung ermitteln wir in einer freien Zelle mit =MAX(K2:K1002) den größten Wert der Spalte K. Wir sehen, dass unsere Werte im Intervall [6,50] liegen. Für eine Häufigkeitstabelle eignen sich daher Klassen der Breite 1. In die Zellen L3:L47 kommen die Klassengrenzen, das sind die Zahlen 6 bis 50.

	K	L
1	Daten	Klassen

Tabelle 3. Excel Tabellenblatt, Teil 3.

Jetzt gehen wir im Menü >Extras auf >Analysefunktionen und wählen >Histogramm. Als Eingabebereich wählen wir K3:K1002, als Klassenbereich I3:I47. Die Ausgabe soll in ein neues Tabellenblatt erfolgen; zusätzlich wollen wir die kumulierte Häufigkeit und eine Diagrammdarstellung.

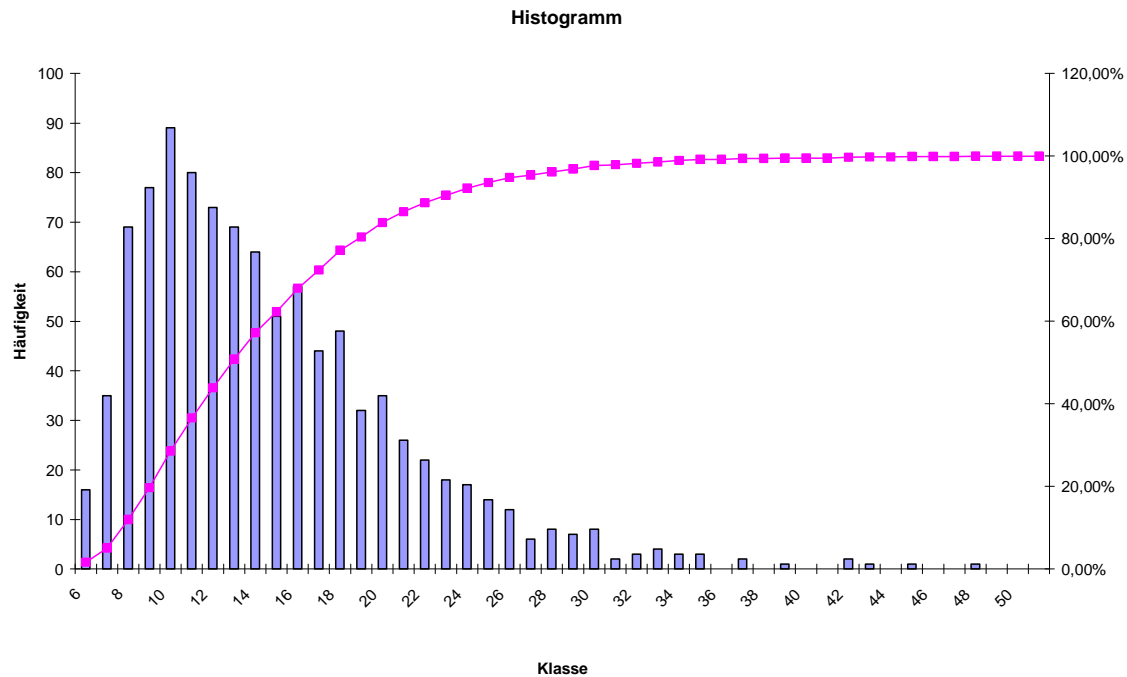


Abbildung 1: Häufigkeit (Balken) und kumulative Häufigkeit (Kurve) von 1000 simulierten Werten für die Anzahl an Käufen bis zur Vollständigkeit der Serie.

4 Schlussbemerkungen

1. Das Tabellenblatt lässt sich sehr leicht an andere Aufgabenstellungen anpassen. Einfache Variationen sind:

- Nicht 6 Symbole, sondern mehr oder weniger sind für einen kompletten Satz notwendig.
- Die Einführung von Jokern. Wer einen Joker zieht, kann jedes Symbol durch diesen ersetzen.

Welche Anzahl von Käufen ist in diesen Fällen im Mittel zu erwarten?

2. Ein fertiges Tabellenblatt wird Interessenten vom Autor auf Anfrage per Email zugesandt.

Literatur

Althoff, Heinz [2000]: Zur Berechnung der Wahrscheinlichkeiten für das Vorliegen einer vollständigen Serie (Sammelbildproblem). In: *Stochastik in der Schule* 20 (2000), Heft 1, 18-20.

Brunner, Norbert & Kühleitner Manfred [2005]: Prüfung auf Gleichwahrscheinlichkeit - eine Simulation in Excel. In: *Wissenschaftliche Nachrichten* 127 (2005), 24-28.

Brunner, Norbert & Kühleitner, Manfred [2005]: Lottofieber in Italien – eine Simulation in Excel. In: *MNU*, im Druck.

Christie, D. [2004]: Resampling mit Excel. In: *Stochastik in der Schule* 24 (2004), Heft 3, 22-27.

Meyer, J.: EXCEL-Files in der Lehrerfortbildung. Auf Anfrage beim Autor: J.M.Meyer@t-online.de

Neuwirth, E. & Arganbright, D. [2004]: *The Active Modeler: Mathematical Modeling with Microsoft Excel*. Brooks/Cole 2004.

Anschrift des Verfassers

a.o. Univ. Prof. Dr. Manfred Kühleitner,
 Institut für Mathematik, DIB, BOKU
 Gregor Mendel Straße 33
 A-1180 Wien
 E-Mail:
 manfred.kuehleitner@boku.ac.at