Manfred Borovcnik, Klagenfurt

Was man mit einer Tabellenkalkulation alles in der Ausbildung von Statistik verändern kann

Zusammenfassung:

Orientiert man sich in der Statistik an der Datenanalyse, so besteht die Möglichkeit, sich in Einzelheiten zu verlieren. Ist die Mathematik Leitlinie der Entwicklung in der Ausbildung, so nehmen die mathematischen Feinheiten überhand – auch das beeinträchtigt leicht den Überblick aber auch den Blick auf das Wesentliche. Im Folgenden werden einige Tabellenblätter, die mit EXCEL erstellt wurden, vorgestellt. Sie zeigen, dass man die Konzepte und Ideen hinter den statistischen Begriffen und Verfahren verständlich illustrieren kann.

1. Einleitung

In der Stochastik-Didaktik ist man mit dem Konzept des operativen Begriffserwerbs nicht weit gekommen. Trial-and-error-Versuche, die Begriffe richtig zu verstehen, funktionieren auch nicht. Dazu ist das Feedback einfach *zu theoretisch*. Grob gesprochen kann man mit der unsinnigsten Strategie im Lotto gewinnen und ist vielleicht dadurch für immer für ein tieferes Verständnis stochastischer Begriffe unzugänglich. Intuitionen (siehe auch Fischbein, 1987) bieten einen Ansatz; insbesondere weil sie auch unvollständige oder nicht passende intuitive Vorstellungen aufgreifen und eine Möglichkeit bieten, Veränderungen bei den Lernenden zu erreichen. Darüber hinaus sind Intuitionen affektiv „geladen"; sind sie unpassend, so sind sie auf Lerners Seite sehr hartnäckig und lassen den Erwerb neuen Wissens nicht so recht vorankommen. Deswegen ist es gut, diese Intuitionen direkt anzusprechen oder gar herauszufordern.

Wenngleich der Autor sich sehr mit den Möglichkeiten befasst hat, über die Schiene eines bewussten, reflektierten Verhältnisses zwischen Intuitionen und abstrakten Begriffen zu einem besseren Verständnis von stochastischen Konzepten in der Unterweisung zu gelangen (Borovcnik, 1991), war ihm schon früh der Stellenwert von Visualisierungen einsichtig. Visualisierungen stellen eine Form der Repräsentation abstrakter Begriffe dar. Neuerdings gibt es auch in der kognitiven Psychologie Erkenntnisse über die Relevanz geeigneter Repräsentationen für den Begriffserwerb. So hat sich die Gigerenzer-Schule verdient gemacht, den Begriff *natürlicher Häufigkeiten* zu pflegen, um damit den Typus der Bayes-Probleme direkter zugänglich zu machen (siehe Wassner e. a. 2002). Sie greift hierbei auf eine alte Sprechweise vom statistischen Dorf zurück: „Was passiert in einem statistischen Dorf mit 1000 Einwohnern?" Dann werden alle Wahrscheinlichkeiten auf den Erwartungswert für die Anzahlen in diesem Dorf herunter gebrochen. In der Repräsentation der natürlichen Häufigkeiten werden

Bayes-Probleme und bedingte Wahrscheinlichkeiten auf den Vergleich zweier Anzahlen reduziert.

Durch die zunehmende Computerisierung unserer Welt eröffnen sich überraschender Weise (überraschend, weil diese Welt selbst artifiziell ist) neue Möglichkeiten der Repräsentation von abstrakten Begriffen. Der Autor bemüht sich seit geraumer Zeit, EXCEL – stellvertretend für eine Tabellenkalkulation – zu nutzen und die damit möglichen Veranschaulichungen, Simulationen und neuen „konkreten“ Repräsentationen der Begriffe aufzuzeigen. Speziell die dynamisch veränderbare Graphik scheint gut geeignet, Lernenden die Bedeutsamkeit von Begriffen, wichtige Relationen dazu und ihre Einschränkungen interaktiv zu vermitteln. Das soll im Folgenden an zwei Themenkomplexen illustriert werden: Regression und Korrelation sowie statistisches Testen.

2. Analyse zweidimensionaler Daten im Sinne beschreibender Statistik

Zusammenhänge zwischen Merkmalen zu untersuchen ist ein Erfordernis, das in vielen Wissenschaften auftaucht. Lineare Beziehungen sind die einfachsten der Art:

$$y = a + b \cdot x,$$

wobei die Parameter a und b aus den Daten (x_i, y_i) zu schätzen sind. Allgemein geht man von einer Familie von Funktionen $f(x, \lambda_1, \lambda_1, ...)$ aus, um den Zusammenhang zwischen den beiden Merkmalen X und Y zu beschreiben. Die Suche nach den „besten“ Werten für die Parameter $\lambda_1, \lambda_1, ...$ führt auf ein Optimierungsproblem, das schon beim Modell der linearen Zusammenhänge nicht ganz einfach in Griff zu bekommen ist. Die nach der Methode der kleinsten Quadrate gewonnenen Schätzungen berechnet man in konkreten Fällen mittels Programmen wie eben einer Tabellenkalkulation. Bleibt übrig, die Aufgabe des Optimierens und ein Maß für die Güte der Lösung zu motivieren.

Interaktives Anpassen der Parameter für die beste Regressionsgerade

In der Folge sollen die Vorzüge einer Tabellenkalkulation aufgezeigt werden, auf den Einsatz von CAS und dessen Vorteile wird hier nicht eingegangen. Als Beispiel wird EXCEL dienen, das bei uns eine hohe Verbreitung hat. Die Ausführungen beziehen sich auf zwei Tabellenblätter, die in Ausschnitten in Abb. 1 und 2 gezeigt werden. Schon immer hat man die Punkwolke der Daten dargestellt und eine lineare Funktion (oder eine Parabel etc.) eingezeichnet und etwa davon geredet, dass man die Abweichungen der Punkte von der Funktion (die quadrierten Abweichungen in Richtung der Ordinate) durch Wahl der Parameter a und b möglichst klein machen will. Man kann in EXCEL eine Punktwolke mit einer willkürlichen Geraden, die visuell ganz gut passt, zeichnen und gleichzeitig diese quadratischen Abweichungen berechnen. Der Vorteil einer Tabellenkalkulation liegt darin, dass man nun einzelne Punkte oder die angepasste Gerade ändern kann. Man erhält einerseits die neue Graphik, andererseits die aktuellen Fehlerquadrate automatisch angepasst. Durch „Steuern“ der Parameter a und b

über Schieberegler erhält man eine dynamische Animation, aus der man interaktiv eine ziemlich gute Anpassung der Geraden an die Punkwolke erhält. Klar wird:

- was zu minimieren ist (die Abweichungen sind in der Darstellung ersichtlich),
- was auch bei guter Anpassung stört (einzelne Punkte, die sogar das verwendete Modell einer linearen Beziehung in Frage stellen oder die als Ausreißer einzustufen sind),
- dass Mathematik sehr hilfreich sein kann bei der Suche nach dem Optimum (es *ist* das Optimum, ich brauche dann nicht mehr weiter probieren),
- dass ein anderes Modell für die Zusammenhänge vielleicht besser passt.

Tatsächlich muss man sich auch heute noch mit lokalen Verbesserungen der Zielfunktion zufrieden geben, wenn eine Optimierung im Sinne der Analysis nicht leistbar ist; das ist u. a. dann der Fall, wenn man kompliziertere Zusammenhänge in Betracht zieht und mehrere Dimensionen für die Parameter hat. Aber man kann auch festhalten, dass die durch systematisches Probieren gefundenen Lösungen von ziemlich guter Qualität sind.

Additivität von Varianzen zur Interpretation des Bestimmtheitskoeffizienten

Ein weiteres Problem in der Anpassung von funktionalen Zusammenhängen an Daten liegt in der Beurteilung der Güte der Anpassung. Üblicherweise wird der so genannte Bestimmtheitskoeffizient (= Quadrat des Korrelationskoeffizienten) dazu herangezogen. Der Korrelationskoeffizient wird dabei als normiertes Maß der Kovariation eingeführt, ist aber keiner einfachen Deutung zugänglich, so scheint es wenigstens und das wird auch so vermittelt.

Mit einer Tabellenkalkulation wie EXCEL kann man an einem einfachen, fiktiven Datensatz mit einfachsten Methoden der deskriptiven Statistik aufzeigen, dass das Quadrat des Korrelationskoeffizienten mit den Varianzen verschiedener Daten zu tun hat:

- Die y-Daten wie sie sind – d. h., ohne Bezug zum Merkmal x,
- die aus der linearen Beziehung zu x geschätzten Daten für y, man nennt sie auch $\hat{y}$, um die Schätzung zum Ausdruck zu bringen,
- die Residuen oder Schätzfehler $y-\hat{y}$.

Aus den Berechnungen von Mittelwert und Varianz dieser Daten sieht man schnell, dass gilt:

- Mittel y = Mittel $\hat{y}$ sowie Mittel der Fehler = 0
- Varianz y = Varianz $\hat{y}$ + Varianz der Residuen

Schnell hat man für die unterschiedlichen Datensätze Interpretationen parat, die üblicherweise zur Erklärung des Bestimmtheitskoeffizienten herangezogen werden:

- Daten y mit Totaler Variation
- Daten $\hat{y}$ mit durch Regressionsbeziehung *erklärter* Variation
- Daten zu Residuen mit *nicht-erklärter* Variation

Bestimmen der Regression spielerisch

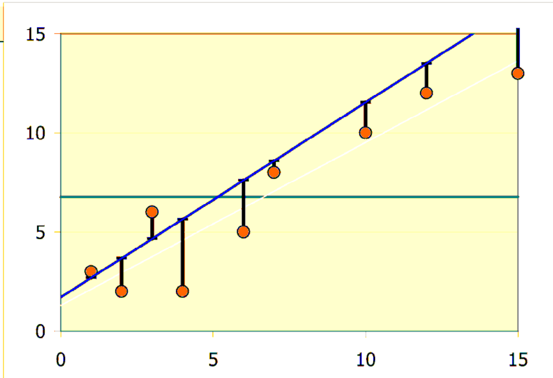
Optisches Anpassen der Parameter für die beste Regressionsgerade

Optisches Anpassen   Verbessern mit Mehrfachoperation

Modelle: für Prognosen		optimale Werte zum Vergleich		aktuell	Regler
naiv = Mittel y					
LS: y = a + b·x		1,295	a	1,709	
		0,822	b	0,983	

Anstatt die Parameter mathematisch zu optimieren, werden die Werte interaktiv angenähert. Man sieht die Auswirkungen unmittelbar.

x	y	y(a,b)	qu. Fehler (a,b)
1	3	2,69	0,10
2	2	3,67	2,80
3	6	4,66	1,80
4	2	5,64	13,25
6	5	7,61	6,79
7	8	8,59	0,35
10	10	11,54	2,36
12	12	13,50	2,25
15	13	16,45	11,90
	61	74,3	41,59
	6,78		17,089



Minimiere die Summe der quadrierten Fehler

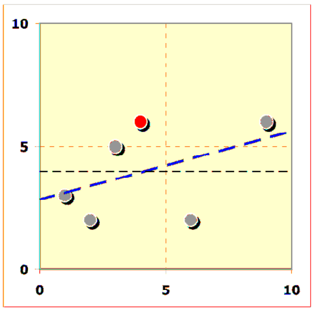
Abb. 1: Interaktives Optimieren: Die Summe der quadrierten Fehler (= die Abweichungen der Punkte von der aktuellen Geraden) ist kleiner zu machen. Steuern kann man die Steigung und den Achsenabschnitt – die optimalen Werte sind zum Vergleich ausgewiesen. Das Blatt kann auch Beweisbedürftigkeit für ein erreichtes „Optimum" erzeugen.

Additivität für Quadratsummen (Varianzen) - unabhängig von speziellen Daten

Punktwolken   Vorhersagen & Korrelation   Vergleich von naiver & linearer Vorhersage   Additivität für Quadratsummen   Inhalt

y-Daten werden aus x-Daten vorhergesagt: Vergleich von "einfachem" und linearen Modell
Die Daten werden durch den Regler interaktiv verändert.

Daten		Modell für y: Mittel		Lin. Modell für y		Resid
un-abhängig	ab-hängig	Vorhersage	Residuen	Vorhersage = Fit		
		F_1	R_1=D-F_1	F		R=D-F
x	y	m(y)	y-m(y)	$\hat{y}$	$\hat{y}$ - m(y)	y-$\hat{y}$
1	3	4,00	-1,00	3,11	-0,89	-0,11
2	2	4,00	-2,00	3,39	-0,61	-1,39
3	5	4,00	1,00	3,67	-0,33	1,33
6	2	4,00	-2,00	4,51	0,51	-2,51
4	6	4,00	2,00	3,95	-0,05	2,05
9	6	4,00	2,00	5,35	1,35	0,65



	total				Modell	Rest
Summe	24,00	24,00	0,00	24,00	0,00	0,00
Mittel	4,00	4,00	0,00	4,00	0,00	0,00
Varianz	3,60	0,00	3,60	0,67	0,67	2,93

x   y

Summe Vorhersage-Fehler

0,187 R^2

Die Herleitung erfolgt rein deskriptiv

Varianzen (später die Quadratsummen) = additiv

Abb. 2: Ein einfacher Vergleich der Varianzen für die Daten, für die vom Modell vorhergesagten Werte sowie für die Residuen zeigt eine Additivität. Die Vermutung wird erhärtet, indem man die Daten dynamisch ändert.

Man sieht rasch, dass u. A. gilt:

- $R^2 = \frac{\text{Varianz } \hat{y}}{\text{Varianz } y} = \frac{\textit{Erklärte} \text{ Varianz}}{\textit{Totale} \text{ Varianz}}$

Damit gibt es eine Reihe von Deutungen für den Korrelationskoeffizienten und dessen Quadrat.

Die „Herleitung“ entspricht eher dem Erzeugen von Vermutungen, folgt aber ganz einfachen, klar nachvollziehbaren Vorgängen. Der Vorteil einer Tabellenkalkulation liegt u. A. darin, dass man die Punktwolken dazu zeichnen sowie einzelne Punkte über Schieberegler verändern kann, sodass man einerseits die Bedeutung der einzelnen Größen „sieht“ und erkennen kann, dass es wohl für alle Datensätze gelten wird. Diese Vermutung kann bei mathematisch Interessierteren einen Bedarf für einen Nachweis erzeugen; unter Verwendung von EXCEL und dem Ansatz des Explorierens von Beziehungen kann man bei einem Lerntyp eine Beweiseinsicht, bei einem anderen Lerntyp eine Beweisnachfrage provozieren.

3. Testen von Hypothesen und die Fehler 1. und 2. Art

Beim Testen von Hypothesen kommen viele Schwierigkeiten zusammen. Es ist nicht einsichtig, was die probabilistischen Folgen verschiedener Sacheinflüsse sind:

i) Wie modelliert man das Verhalten von Probanden unter unterschiedlichen „Hypothesen“? Worin liegen die Unterschiede der Modellierungen, was sind deren empirische Konsequenzen?

ii) Was ist zu „erwarten“ in der Gruppe A wie in der Gruppe B, wenn man diese beiden Gruppen miteinander zu vergleichen hat. Wie bewertet man Abweichungen von dieser Erwartung in der Praxis?

iii) Wie kommt man zu einer Entscheidung, ob Daten aus der Gruppe A oder eben B stammen? Welche Fehler treten dabei auf?

Die angesprochenen Fehler sind auch deswegen so schwer zu verstehen und akzeptieren, weil sie fiktive Wahrscheinlichkeiten darstellen; sie gelten unter der Festsetzung, man wüsste, woher die Daten stammen. Wegen des theoretischen Gehalts der Kennziffern und wegen der globalen Betrachtung aller Möglichkeiten (die überdies nur indirekt auf die reale Situation passen, insofern sie diese gut oder schlecht modellieren), ist es von großem Vorteil, wenn man die Situation an sich und ihre Kennziffern in eine konkrete Repräsentation einbettet. Die Repräsentationen werden noch leichter lesbar, wenn man einen aussagekräftigen Kontext für das Entscheidungsproblem hat, auf den die mathematische Herangehensweise bezogen wird. Ausschnitte von zwei Tabellenblättern (Abb. 3 und 4) bieten Hilfen, sich diesen Fragen zu stellen.

Dynamische Szenarien über das Verhalten von Käfern

Beispiel: Ein Käfer wird durch ein Labyrinth mit zwei Ausgängen geschickt und kann bei einer am Ende befindlichen Geruchsquelle landen oder nicht (aus Lorenz, 1996). Hier ist es einfach, die Konsequenzen einer probabilistischen Modellierung einzusehen:

Wenn Käfer weder angezogen noch abgestoßen werden, so sollte man das Experiment als Bernoulli-Versuchsreihe modellieren können, und zwar mit $p = \frac{1}{2}$. Schickt man nun einen Käfer mehrmals durch das Labyrinth, so muss man ausschließen können, dass es Lerneffekte gibt. Das erreicht man am besten, in dem man Spekulationen über mögliche Lerneinsichten von Käfern umgeht, indem man die Geruchsquelle durch Zufall einmal in dem einen dann wieder in dem anderen Trakt platziert.

Man muss über Störeinflüsse nachdenken, welche die Käfer bei der Geruchsquelle landen lassen, die aber nichts oder nur indirekt mit dem Geruch an sich zu tun haben: Könnte etwa sein, dass sich der Käfer an der Nahrung orientiert, die mit der Geruchsquelle verbunden ist (wenn das so ist). Könnte auch sein, dass bestimmte Gerüche den Käfer anziehen, andere wieder „abstoßen". Solche *Confounder*, wie man auch sagt, muss man frühzeitig in einer „Systemanalyse" des anstehenden Problems erfassen. Ansonsten ist es später nur mehr möglich, über die Auswirkungen solcher Störgrößen zu spekulieren und die Daten werden wertlos im Sinne der Beantwortung der gestellten Fragen. Hat man jedoch Kandidaten für Störgrößen erfasst und in das Experiment integriert, so bekommt man Daten über ihre Auswirkungen und kann ihre Auswirkungen auf die Zielgröße – hier die Anzahl der Erfolge, bei der Geruchsquelle zu landen – beurteilen (siehe Borovcnik, 1995).

Will man das Experiment mit verschiedenen Käfern durchführen, fallen zwar die Lerneffekte weg, aber es mag nicht klar sein, ob alle Käfer (derselben Art) sich gleich verhalten. Fazit: Wenn sich Käfer neutral gegenüber der Geruchsquelle verhalten, so modellieren wir ihr Verhalten adäquat durch eine Binomialverteilung mit $p = \frac{1}{2}$.

In EXCEL erhält man die entsprechende Wahrscheinlichkeitsverteilung leicht und zeichnet ihre diskrete Dichte. So sieht man am besten, was die Konsequenzen unserer Modellierung sind. Schickt man einen Käfer 100mal in das Labyrinth, so wird er zwischen 40 und 60mal bei der Geruchsquelle landen. In manchen Serien weniger als 40, in manchen mehr als 60; insgesamt hat dies aber nur eine sehr kleine Wahrscheinlichkeit. Wir könnten in EXCEL auch simulieren, wie ein Experiment mit 100 Versuchen ausgeht, dann könnten wir die ganze Serie sehr oft (vielleicht 1000mal) wiederholen und daraus den Anteil an Experimenten schätzen, in denen so „extreme" Beobachtungen auftreten wie unter 40 oder mehr als 60mal bei der Geruchsquelle zu landen.

Käfer könnten aber von der Geruchsquelle angezogen oder abgestoßen werden. Dann ändert sich p auf Werte größer als $\frac{1}{2}$ bzw. kleiner als $\frac{1}{2}$. Die Voraussetzungen der Bernoulli-Versuchsreihe deuten an, wie sich Käfer – theoretisch wenigstens – verhalten müssten: immer mit derselben Erfolgswahrscheinlichkeit bei der Geruchsquelle landen, unabhängig von vergangenen Durchläufen. Wir modellieren also das Verhalten eines Käfers, der von der Geruchsquelle mäßig angezogen wird mit $p=0{,}6$. Die Folgen dieses Modells kann man wieder leicht aus dem Diagramm der entsprechenden diskreten Dichte ablesen: Große Werte der Anzahl „bei der Geruchsquelle“ werden nun um einiges wahrscheinlicher, kleine Werte sind kaum mehr wahrscheinlich. Entsprechend erwarten wir, dass sich ein Experiment mit so einem Käfer in einem Ergebnis deutlich über der 50er Marke niederschlagen wird.

Der Vorteil von EXCEL liegt wieder darin, dass man die Wahrscheinlichkeit p, bei der Geruchsquelle zu landen, dynamisch verändern kann und sofort sieht, welche probabilistischen Konsequenzen dies hat. Ergebnisse oberhalb von 60 sind nicht nur selten (haben geringe Wahrscheinlichkeit) unter der „Null“-Hypothese, sie sind überdies unter dem alternativen Wert $p=0{,}6$ ganz etwas Normales. Ein solches Ergebnis spricht demnach nicht nur *gegen* die Null-Hypothese sondern *für* eine Alternative. Klar, so hat man immer Testen von Hypothesen eingeführt, aber der Vorteil von EXCEL liegt darin, dass man die Verteilungen unter alternativen Werten „sieht“. Gleichzeitig befreit man sich von der Willkür der Wahl von $p=0{,}6$ als alternativen Wert: es kommen ja alle Werte größer oder auch kleiner als $\frac{1}{2}$ in Frage. Die Folgen einer „Vermutung“, wie sich Käfer verhalten, auf die Anzahl der Erfolge (wie oft Käfer bei der Geruchsquelle landen) können durch einen Schieberegler für p schnell erkannt und eingeschätzt werden.

Risiken einer Entscheidung für oder gegen die Nullhypothese (= Käfer sind neutral)
Entscheidet man sich bei unter 40 bzw. über 60 Erfolgen gegen die Nullhypothese, so ist klar, dass dabei Fehler möglich sind. Die Wahrscheinlichkeiten unter der Nullhypothese sind rasch visualisiert; damit sieht man den Fehler 1. Art, den so genannten α-Fehler. Jetzt sieht man aus derselben Abbildung aber auch, wie groß ein anderer Fehler wird: nämlich der Fehler 2. Art, die Nullhypothese nicht abzulehnen, obwohl eigentlich der Wert $p=0{,}6$ zutrifft – der β-Fehler. Und der ist groß! So groß hätte man ihn nicht vermutet. Klar, so hat man immer schon den Fehler 2. Art eingeführt.

Der Nachteil an statischen Zeichnungen aller Art ist aber, dass der Fehler 2. Art nur für *einen* ausgewählten Wert gezeichnet wird. Bevor ein Lerner innerlich den Widerstand gegen die konkrete Wahl von etwa $p=0{,}6$ aufgibt, dauert es lange. Dann hat man aber nur einen einzigen Wert, *ein* $\beta(p=0{,}6)$. Jetzt ergibt ein Justieren am Regler für p rasch die Einsicht, dass der β-Fehler dann sehr klein wird, wenn p sehr weit weg von $\frac{1}{2}$ liegt. So weit, dass man mit freiem Auge sieht, ob dieser Wert, etwa $p=0{,}9$ oder doch $p=\frac{1}{2}$ vorliegt. Dann braucht man aber keinen Statistiker mehr und kann selbst entscheiden, dass Käfer von der Geruchsquelle angezogen werden. Wenn hingegen sich der Wert von p der $\frac{1}{2}$ aus der Nullhypothese nähert, dann geht der β-Fehler in ungeahnte Höhen (bis zu $1-\alpha$) hinauf. Auch das kann in EXCEL dynamisch erlebt werden.

Dynamische Szenarien über das Verhalten von Käfern

Ein Käfer wird durch ein Labyrinth geschickt und kann bei
einer am Ende platzierten Geruchsquelle landen – oder nicht.
(Lorenz, Biometrie, 1988)
Szenarien: Wie verhalten sich Käfer - Wie bewährt sich die Entscheidungsregel
Mit n=100 Käfern; die Werte für p werden dynamisch mit Reglern eingestellt.

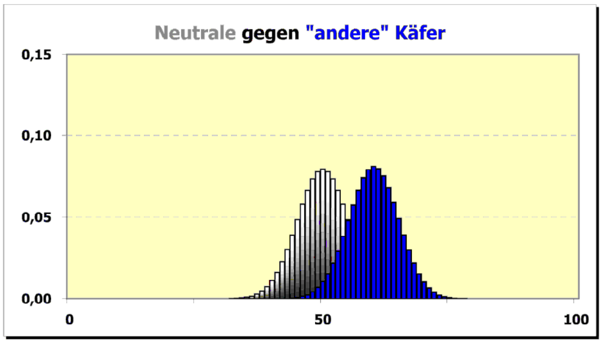
Zahl der "Erfolge" – Verteilung in Abhängigkeit vom wahren p

Zahl der "Erfolge" Risiken einer Entscheidung Gütefunktion Inhalt

Wie verhalten sich Käfer?
neutral (p=0,5)
werden angezogen (p>0,5)
werden abgestoßen (p<0,5)

p **0,60**

Ist p sehr verschieden von 0,5,
so sprechen die Daten deutlich
gegen die Nullhypothese



Die Folgen einer Vermutung, wie sich Käfer verhalten, auf die Anzahl der "Erfolge"
können durch den Regler für p schnell erkannt werden.

Abb. 3: Konsequenzen einer probabilistischen Modellierung auf die Anzahl der Erfolge: Nullverteilung und eine variable Alternative – es wird deutlich, wann man welche Möglichkeit „vorzieht" bzw. wann eine Entscheidung schwer fällt.

Risiken einer Entscheidung für oder gegen die Nullhypothese (=Käfer sind neutral)

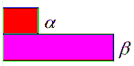
Zahl der "Erfolge" Risiken einer Entscheidung Gütefunktion Inhalt

Ergebnis	Entscheidung	Wenn H_0 "wahr"	Wenn Alternative "wahr"	
	"für" H_0	ok	β-Risiko	
	gegen H_0	α-Risiko	ok = "Macht", Alternative zu "erkennen"	

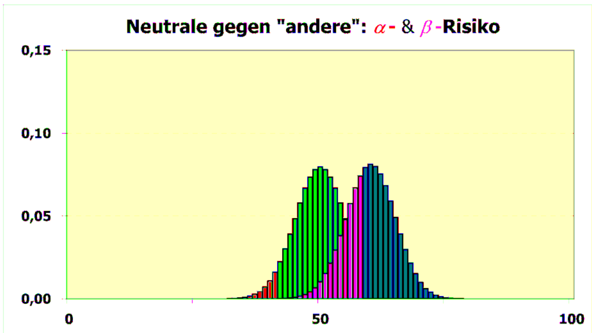
Wenn p nahe bei 0,5
dann ist β-Fehler groß

p **0,60**

α **0,12**
β ergibt sich **0,38**



α und β sind gegenläufig.



Durch Drehen an den Reglern für α und p sieht man schnell ihren Einfluss auf den β–Fehler:
Man braucht mehr Daten, will man beide Fehler verkleinern.

Abb. 4: Die Fehler 1. und 2. Art einer bestimmten Entscheidungsregel bei fester Alternative. Sowohl die Alternative als auch die Entscheidungsregel sind variabel. Bewegt sich die Alternative auf die Nullverteilung zu, wird der Fehler 2. Art immer größer. Ändert man die Entscheidung, um einen kleineren Fehler 1. Art zu haben, sieht man, wie der Fehler 2. Art gegenläufig größer wird.

Nicht nur das: die wechselseitigen Abhängigkeiten von p, α und β kann man durch Einführen weiterer Schieberegler verdeutlichen. Will man den Fehler 1. Art verkleinern, so geht automatisch der Fehler 2. Art hinauf.

Interpretation der Risiken im Kontext der Entscheidung

Entscheidet man sich für die Nullhypothese $p = \frac{1}{2}$, so mag dies bedeuten, dass wir das Verhalten der Käfer als neutral einstufen. Normalerweise wird so etwas nicht publiziert. Die „Gefahr", dennoch eine Arbeit zu veröffentlichen, obwohl die Käfer neutral sind, wird mit dem α-Fehler „gemessen". Ein kleines α bedeutet damit ein kleines Risiko, sich später zu „blamieren". Gegenläufig die Interpretation eines großen Werts für α. Also macht man das α klein.

Auch hier könnte man ohne Statistiker auskommen, wenn da nicht noch das β wäre, die Wahrscheinlichkeit eines Fehlers 2. Art: Im Kontext bedeutet es, eine wichtige Erkenntnis nicht (rechtzeitig) zu fassen und entsprechend eine wichtige Publikation zu verabsäumen. Wenn $p = 0{,}6$ eine Bahn brechende Erkenntnis über Käfer der untersuchten Art wäre, so sollte man das publizieren und weitere Forschungen dazu anregen.

Die Einsicht aus der dynamischen Animation über den gegenläufigen Zusammenhang zwischen den beiden Fehlerwahrscheinlichkeiten lassen die Problematik jedweder Forschung an neuen Erkenntnissen begreifen:

- Je lockerer man die Bedingungen für neuartige Ergebnisse macht, desto größer das Risiko von Artefakten (euphemistische Umschreibung von falschen Ergebnissen).
- Je starrer die Bedingungen für neuartige Ergebnisse, desto größer das Risiko, neue Erkenntnisse zu übersehen.

Ein anderer, für diese Zusammenhänge erhellender Kontext ist der einer Multiple-Choice-Prüfung mit je 2 (3 oder mehr) Auswahlantworten und einer Testbatterie von z. B. 100 Fragen. Hier muss man das Verhalten von jenen modellieren, die unterschiedlich gut gelernt haben und mit dem Verhalten jener kontrastieren, die sich mit reinem Raten durch die Prüfung bewegen.

Der α-Fehler entspricht dann der Entscheidung, einen Prüfling durchkommen zu lassen, obwohl er einfach rät. Beim β-Fehler geht man oft mit Vorteil zur Gegenwahrscheinlichkeit $1-\beta$, der so genannten *Macht*, über. Die Macht $1-\beta(p)$ ist dann die Chance eines Prüflings, die Prüfung positiv abzuschließen, wenn seine Wahrscheinlichkeit, ein einzelnes Item richtig zu beantworten, p beträgt. Ist α zu groß oder $1-\beta$ zu klein, so wird man den Test als nicht geeignet ansehen. Auch hier sind Abweichungen des Modells von der realen Situation zu diskutieren: Während real die Aufgaben für die Prüflinge untereinander abhängig sind, geht man in der probabilistischen Modellierung von unabhängigen Items aus. Man erhält wohl eine passende Null-Verteilung, der normale Prüfling findet sich aber kaum in den Alternativ-Verteilungen wieder. Man kann darüber spekulieren, ob das die Macht des Tests verbessert oder verschlechtert und wann welcher Effekt eintritt.

4. Zusammenfassung

Zur professionellen Datenanalyse ist EXCEL zweite Wahl, nicht jedoch für die Datenanalyse im Unterricht. Ein Vorzug von Tabellenkalkulation allgemein liegt darin, dass immer konkret gearbeitet wird: die Spalten gibt es; sie stellen die Daten bzw. die Modelldaten dar; dann ist die Differenz zu den Daten der Fehler – in einer eigenen Spalte – usw. Die Graphiken werden zu Animationen, wenn man Daten etwa über Regler steuert. Damit verlässt man die Enge von konkreten Daten und darin gefundenen Beziehungen und kann etwa erkennen, dass eine Additivität der Varianzen unabhängig von den vorliegenden Daten gilt.

Verändert man Parameter des Modells wie den Fehler 1. Art, so „sieht" man, wie sich der Fehler 2. Art dynamisch verändert. In gleicher Weise erkennt man, dass ein Fehler 2. Art von der „Distanz" zur „Null"-Verteilung abhängt: Je näher man bei dieser ist, desto schwieriger wird es, zu erkennen, dass es gerade diese Verteilung und nicht die Null-Verteilung ist. Dieser komplexe Zusammenhang wird üblicherweise mit einem statischen Bild erklärt. In der Dynamik der Verschiebung einer Verteilung hin zur Null-Verteilung liegt aber viel Potential für Einsicht.

Dynamische Veränderung von Graphiken durch gezielte Veränderung von Einflussparametern – dynamische Animation – kann viel zu einem vertieften Begriffsverständnis beitragen. Hilfreich ist, dass sich sowohl die Tabellenobjekte nicht hinter abstrakten Begriffen verstecken, dass die Verteilungen, die unterschiedlichen Gegebenheiten entsprechen, leicht gezeichnet werden und ihre Veränderungen einfach zu verfolgen sind. Man sieht die Auswirkungen – bei geeignetem graphischen Einsatz – unmittelbar.

Simple, experimentelle Mathematik wird durch EXCEL ermöglicht. Man findet neue Vermutungen und Einsichten. Manche kann man durch Beweise absichern, bei anderen erübrigt sich ein Beweis, nachdem man die Animation gesehen hat. Jedenfalls macht Arbeiten in einer Tabellenkalkulation von einigem technischen „Firlefanz" (wie dem Nachschauen von Quantilen in Tabellen) frei und regt mathematische Tätigkeiten an. Unter Tabellenkalkulationen ist EXCEL auch deswegen eine gute Wahl, weil E. Neuwirth eine Brücke zu R geschaffen hat mit seiner Software REXCEL – dieses Programm lässt von R zu EXCEL und umgekehrt wechseln und schafft für *beide* Programme eine wesentliche Verbesserung.

Alles in allem kann kaum genug betont werden, wie wichtig geeignete Repräsentationen für effizientes Lernen sind. Sie beeinflussen nicht nur Gedächtnisleistungen sondern auch tieferes Begriffsverständnis. In einer Tabellenkalkulation wird es leicht, animierte Graphiken zu erzeugen, welche als Schlüssel für das Verständnis komplizierterer Zusammenhänge angesehen werden können. Ein Vorteil liegt auch darin, dass Lernende die Graphiken und die anderen, angesprochenen Zusammenhänge selbst nachvollziehen können, sie werden aktive Mitgestalter der „Filme". Letztlich ist in EXCEL ein know how verborgen, das heute zu den Kulturtechniken zählen darf, so wie man eben ein gutes Textverarbeitungssystem oder eine Präsentationssoftware beherrschen sollte.

Literatur

Bartz, S.: Excelblatt vereinfacht Stochastik. *Stochastik in der Schule* 27 (2007) 2, 25-29.

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

Borovcnik, M., Gölles, J.: *Grundprinzipien der statistischen Datenanalyse.* Typoskript. Graz: Institut für Statistik und Systemanalyse 1995.

Borovcnik, M.: Nützliche Gesetze über den Zufall – Experimente mit Excel. Schriftenreihe Didaktik der Mathematik der Österr. Math. Ges. 32 (2001), 1-22.

Borovcnik, M.: EXCELlent Statistics – Statistik mit Unterstützung von Tabellenkalkulation. *Schriftenreihe Didaktik der Mathematik der Österr. Math. Ges.* 33 (2002).

Borovcnik, M.: Daten – Zufall – Resampling. J. Meyer (Hsg.): *Anregungen zum Stochastikunterricht* Bd. 3, Franzbecker, Hildesheim 2006b, 143-158.

Borovcnik, M.: Das Sammelbildproblem – Rosinen und Semmeln und Verwandtes: Eine rekursive Lösung mit Irrfahrten. *Stochastik in der Schule* 27 (2007) 2, 19-24.

Borovcnik, M., Neuwirth, E.: *Ein rekursiver Zugang zur Binomialverteilung.* http://www.osg.or.at/ (2006)

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

Fischbein, E.: *Intuition in Science and Mathematics. An Educational Approach.* Dordrecht: D. Reidel 1987.

Kühleitner, M.: Sammelproblem: Eine Simulation mit Excel. *Stochastik in der Schule* 27 (2007) 1, 24-26.

Lorenz, R. J.: *Grundbegriffe der Biometrie*, 4. Auflage. Stuttgart – New York: Fischer 1996.

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

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

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

Reckelkamm, B.: Der Tanz der Residuen - Erarbeitung statistischer Grundbegriffe mit Hilfe von EXCEL. *Stochastik in der Schule* 24 (2004) 3, 14-21.

Spreadsheets in Education: http://www.sie.bond.edu.au/

Wassner, C., Krauss, S., Martignon, L.: Muss der Satz von Bayes schwer verständlich sein? *Praxis der Mathematik* 44 (2002) 1, 12-16.

Whingham, D.: Normalverteilungen mit Excel. *Stochastik in der Schule* 19 (1999) 2, 28-31.

Anhang:

Inhaltsverzeichnis der Tabellenblätter zu den Themen Regression und Korrelation sowie zum Testen von Hypothesen