Stochastik in der Schule 16 (1996) Heft 1

Testen von Hypothesen in der pharmazeutischen Praxis

Manfred Buth, Universität Hamburg

Anwendungsorientierter Mathematikunterricht von echtem Schrot und Korn orientiert sich an lebensweltlichen Zusammenhängen und sucht die Inhalte in der Praxis - immer nach der Devise, daß wir für das Leben und nicht für die Schule lernen.

In dieses Konzept fügen sich der Unterricht in Stochastik und insbesondere das Gebiet "Testen von Hypothesen" voll ein. Denn es handelt sich dabei um Unterrichtsgebiete, die vom mathematischen Standpunkt aus nicht besonders aufregend sind, sondern ihre Bedeutung für den Unterricht allein aus ihrer Nähe zu den praktischen Anwendungen erhalten. Deshalb hat der Verfasser bereits in einer früheren Veröffentlichung über die Techniken des Hypothesentestens im Bereich der Medizin berichtet. Im vorliegenden Bericht geht es um das Testen von Hypothesen bei der Arzneimittelproduktion. Dabei sind Ergebnisse herausgekommen, die man sich eigentlich schon hätte vorher ausmalen können und die bereits jetzt als Pointe vorweg genannt seien.

Wenn man im Sinne eines anwendungsorientierten Unterrichts mathematische Probleme auswählt, dann stellt sich schnell heraus, daß die mathematischen Schwierigkeiten, die damit verbunden sind, die Möglichkeiten der Schule überfordern. Beim Testen von Hypothesen bedeutet das: Man kommt um den f-Test und den t-Test nicht herum. Aber beide sind für die Schule zu schwierig.

Außerdem ergibt die Praxis kein einheitliches Bild. Das Vorgehen in der Pharmazie unterscheidet sich von demjenigen in der Medizin und das aus gutem Grund. Denn die physikalische Untersuchung einer einfachen Tablette ist viel unkomplizierter als die medizinische Behandlung und Untersuchung eines Schwerkranken, so daß beispielsweise der Umfang einer Meßreihe in der Pharmazie von untergeordneter Bedeutung bleibt, jedenfalls so lange keine chemischen Analysen erforderlich sind.

Im folgenden Abschnitt wird zunächst der mathematische Hintergrund für den f-Test und den t-Test in aller gebotenen Kürze skizziert. Anschließend wird über eine konkrete Testreihe berichtet, die sich auf die Tablettenherstellung bezieht. Die Daten stammen aus einem pharmazeutischen Praktikum der Universität Hamburg. Sie sind trotzdem sehr realitätsnah. Denn dem Verkauf der produzierten Tabletten stehen nur die strengen Bestimmungen des Arzneimittelgesetzes entgegen. Im übrigen geht es genau so zu wie in der pharmazeutischen Industrie.

Nach einem kurzen Abschnitt mit weiteren Informationen zur Praxis des Testens von Hypothesen folgt dann eine didaktische Wertung der Ergebnisse im Hinblick auf den Stochastikunterricht.

Zu besonderem Dank bin ich Herrn Prof. Dr. J. MIELCK und seiner Arbeitsgruppe vom Pharmazeutischen Institut der Universität Hamburg verpflichtet und dabei besonders Herrn Dr. A. SAKMANN und Frau K. PICKER. Sie haben mir bereitwillig die Informationen zukommen lassen, die ich im folgenden zum Nutzen der Unterrichtspraxis aufzuarbeiten versuche.

Mathematische Voraussetzungen

In diesem Abschnitt sollen diejenigen Voraussetzungen formuliert werden, die man kennen muß, um im weiteren Verlauf dieser Arbeit zu verstehen, was mathematisch gespielt wird.

Gegeben seien n Zufallsexperimente, deren Werte x_1 bis x_n reelle Zahlen sind und deren Wahrscheinlichkeitsverteilungen durch die Dichtefunktionen f_1 bis f_n beschrieben werden. Wenn die Zufallsexperimente voneinander unabhängig sind, dann wird die Wahrscheinlichkeitsdichte für die Nacheinanderausführung der Experimente durch das Produkt der Funktionen f_1 bis f_n beschrieben. Ist weiterhin eine reelle Funktion p in n Variablen gegeben, so wird durch die Gleichung

$$x = p(x_1, \dots, x_n)$$

ein neues Zufallsexperiment definiert, bei dem x genau dann als Ergebnis gelten soll, wenn die vorgegebenen Zufallsexperimente die Werte x_1 bis x_n liefern und die obige Gleichung gilt. Unter diesen Voraussetzungen lautet die Frage: Durch welche Dichtefunktion f wird das neue Zufallsexperiment beschrieben? Um das Problem zu lösen, muß man das Gebiet G(x, x+dx) im n-dimensionalen Raum ermitteln, dessen Punkte zu Werten zwischen x und x+dx führen, wenn man ihre Koordinaten in die Funktion p einsetzt. Die Antwort auf die gestellte Frage lautet dann in allgemeiner Form

$$f(x)\,dx \quad = \quad \int_{G(x,x+dx)} f_1(x_1) \cdot \ldots \cdot f_n(x_n)\, dx_1 \ldots dx_n$$

Im Fall n = 2 kann man das Problem und seine Lösung anschaulich darstellen. Der Graph der Funktion p ist dann eine dreidimensionale Fläche. Wenn man sie mit derjenigen Parallelen zur x_1-x_2-Ebene, die bei x durch die dritte Achse geht, schneidet und wenn man die entstehende Schnittlinie auf die x_1-x_2-Ebene projiziert, dann entsteht eine Linie L(x), die man in der Geographie als Höhenlinie bezeichnen würde. Sofern die beteiligten Funktionen nur einigermaßen vernünftige Eigenschaften besitzen, ist G(x, x+dx) das Gebiet zwischen den Linien L(x) und L(x+dx) (vgl. Bild 1).

Soviel zum allgemeinen Verfahren. Alles weitere besteht darin, die obige Gleichung auf konkrete Beispiele anzuwenden. Wir geben die Ergebnisse an, ohne sie jeweils im einzelnen durch eine Herleitung zu begründen.

Wenn die gegebenen Zufallsexperimente sämtlich standardnormalverteilt sind und die Funktion p durch

$$p(x_1, \ldots, x_n) \quad = \quad x_1^2 + x_2^2 + x_n^2$$

gegeben ist, dann liefert f die Dichte der Chiquadrat-Verteilung mit n Freiheitsgraden.

Falls n = 2 und das erste Zufallsexperiment standardnormalverteilt ist und das zweite chiquadratverteilt mit k Freiheitsgraden und wenn schließlich die Funktion p durch

$$p(x_1, x_2) \quad = \quad k^{1/2}\, x_1\, x_2^{-1/2}$$

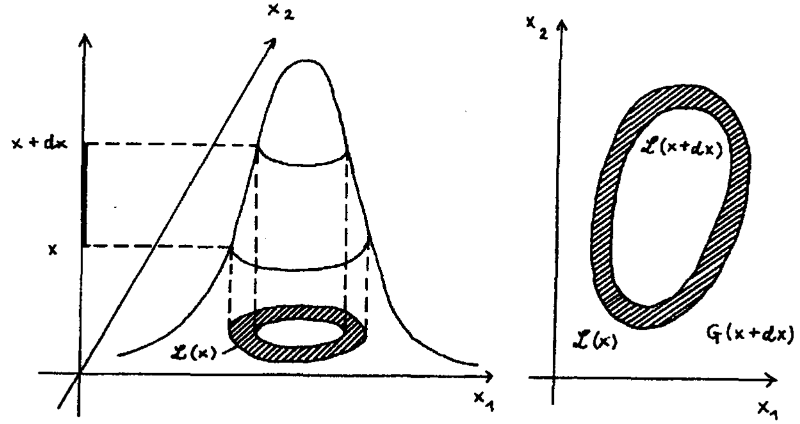

Bild 1: der Graph der Funktion p im Fall n = 2 mit Höhenlinien

gegeben ist, dann besitzt die Dichtefunktion f die t-Verteilung mit k Freiheitsgraden.

Falls aber für n = 2 beide Zufallsexperimente chiquadratverteilt sind mit k und l Freiheitsgraden und wenn p durch

$$p(x_1, x_2) \quad = \quad k^{-1}\, l\, x_1\, x_2^{-1}$$

gegeben ist, dann ist f die Dichte der f-Verteilung mit k und l Freiheitsgraden.

Diese drei Beispiele belegen zugleich, daß man den Prozeß, mit Hilfe reeller Funktionen neue Verteilungen zu konstruieren, auch iterieren kann. Aber man kann aus dem allgemeinen Verfahren noch weitere Resultate herleiten.

Falls eine Folge $x_1, x_2, \ldots, x_n$ von Ergebnissen eines Zufallsexperiments gegeben ist, dann nennen wir sie wahlweise eine Meßreihe oder eine Stichprobe. Für die Funktion p setzen wir einerseits den Stichprobenmittelwert

$$m = (1/n)(x_1 + x_2 + \dots x_n)$$

und andererseits die Stichprobenvarianz

$$s^2 = (1/(n-1))((x_1-m)^2 + (x_2-m)^2 + \dots + (x_n-m)^2)$$

Unter der Voraussetzung, daß alle Zufallsexperimente normalverteilt sind mit dem Mittelwert μ und der Varianz σ^2 kann man mit der gleichen Methode, die bisher benutzt wurde, begründen, daß

$$\gamma = \frac{m-\mu}{\sigma}\sqrt{n}$$ standardnormalverteilt ist und

$$\delta = \frac{s^2(n-1)}{\sigma^2}$$ chiquadratverteilt ist mit n-1 Freiheitsgraden.

Nun seien zwei Meßreihen gegeben. Die eine habe den Umfang n_1 und gehöre zur Normalverteilung mit dem Mittelwert μ_1 und der Varianz σ_1^2. Die andere habe den Umfang n_2 und gehöre zur Normalverteilung mit dem Mittelwert μ_2 und der Varianz σ_2^2. Dann läßt sich auch die folgende Frage mit den bisher verwendeten Methoden klären:

Unter welcher Voraussetzung kann die Hypothese aufrecht erhalten werden, daß die beiden Varianzen σ_1^2 und σ_2^2 übereinstimmen, und unter welcher Bedingung läßt sich die weitergehende Hypothese beibehalten, daß beide Mittelwerte μ_1 und μ_2 gleich sind, falls die Hypothese $\sigma_1^2 = \sigma_2^2$ aufrecht erhalten werden kann?

Um den ersten Teil der Frage zu klären, beachte man, daß in der Gleichung

$$\frac{s_1^2}{s_2^2} \cdot \frac{\sigma_1^2}{\sigma_1^2} = \frac{\dfrac{\left(\dfrac{s_1^2(n_1-1)}{\sigma_1^2}\right)}{n_1-1}}{\dfrac{\left(\dfrac{s_2^2(n_2-1)}{\sigma_2^2}\right)}{n_2-1}}$$

sowohl der Zähler als auch der Nenner auf der rechten Seite chiquadratverteilt ist mit n_1-1 bzw. n_2-1 Freiheitsgraden. Also ist die Größe auf der linken Seite f-verteilt mit n_1-1 und n_2-1 Freiheitsgraden. Um die Hypothese $\sigma_1^2 = \sigma_2^2$ zu testen, nimmt man also die Testgröße

$$u = s_1^2/s_2^2$$

und die f-Verteilung in n_1-1 und n_2-1 Freiheitsgraden und führt einen f-Test zu vorgegebenem Signifikanzniveau α durch.

Weiterhin kann man zeigen, daß die Größe

$$\rho = \frac{(m_1 - m_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

standardnormalverteilt ist. Unter den beiden Annahmen $\mu_1 = \mu_2$ und $\sigma_1^2 = \sigma_2^2$ lautet sie

$$\rho = \frac{m_1 - m_2}{\sqrt{\sigma^2}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

Als Schätzgröße für die gemeinsame Varianz σ^2 wählt man das gewichtete Mittel

$$s^2 = \frac{s_1^2(n_1-1) + s_2^2(n_2-1)}{n_1 + n_2 - 2}$$

der beiden Stichprobenvarianzen s_1^2 und s_2^2. Die Schreibweise

$$\rho = \frac{\dfrac{m_1 - m_2}{\sqrt{\sigma^2}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}}{\sqrt{\dfrac{\dfrac{s_1^2\,(n_1-1) + s_2^2\,(n_2-1)}{\sigma^2}}{n_1 + n_2 - 2}}}$$

für ρ läßt erkennen, daß eine t-Verteilung vorliegt. Also wird die Entscheidung über Annahme oder Ablehnung der Hypothese $\mu_1 = \mu_2$ unter der zusätzlichen Voraussetzung, daß $\sigma_1^2 = \sigma_2^2$ gilt, mit Hilfe der Testgröße

$$v = \frac{m_1 - m_2}{\sqrt{\dfrac{s_1^2\,(n_1-1) + s_2^2\,(n_2-1)}{n_1 - n_2 - 2}}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

und der t-Verteilung in n_1+n_2-2 Freiheitsgraden getroffen.

<u>Erläuterung an einem Beispiel</u>

In der folgenden Tabelle sind zwei Meßreihen mit 20 bzw. 10 Meßwerten gegeben.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	6	4	5	7	3	5	9	5	6	4	5	6	4	5	6	4	5	6	4
2	3	4	2	3	6	3	4	5	4										

Die Mittelwerte und die Varianzen lauten

$$m_1 = 5{,}2 \qquad m_2 = 3{,}6$$
$$s_1^2 = 1{,}75 \qquad s_2^2 = 1{,}6$$

Daraus ergeben sich die beiden Testgrößen

$$u = 1{,}092 \qquad v = 3{,}168$$

Aus einer Tabelle entnimmt man die Funktionswerte für die aufintegrierte Verteilungsfunktion:

$$f(1{,}092;\ 19;\ 9) = 0{,}532$$
$$t(3{,}168;\ 28) = 0{,}998$$

Daraus folgt, daß die Hypothese $\sigma_1^2 = \sigma_2^2$ auf dem 99-Prozent-Niveau angenommen werden kann, aber die Hypothese $\mu_1 = \mu_2$ abzulehnen ist.

Ausführung an Beispielen aus der pharmazeutischen Praxis

In der pharmazeutischen Industrie verwendet man Tablettiermaschinen, um aus pulverförmigen Substanzen Tabletten herzustellen. An derartigen Maschinen, die man sich als ziemlich robuste Einrichtungen vorzustellen hat, rotiert eine Stahlplatte mit kreisförmigen Vertiefungen um eine vertikale Achse. An einer Stelle der Scheibe wird Tablettensubstanz angehäuft, die dann in die darunter sich drehenden Löcher fällt. Anschließend wird die Substanz glattgestrichen und von einem Stempel zur Tablette zusammengedrückt. Die dabei wirkende Kraft beträgt etwa 30 kN und erzeugt somit ungefähr den gleichen Druck, als wenn vierzig erwachsene Personen auf einem Bleistiftabsatz stehen.

Aus physikalischer Sicht dürfte der folgende Zusammenhang einleuchten: Wenn sich die Scheibe schnell unter der aufgehäuften Substanz hinwegdreht, dann gelangt die Substanz nicht mehr vollständig in die darunter liegenden Löcher und die Masse der Tabletten wird im Mittel sinken. Um diesen Effekt experimentell zu prüfen, werden in zwei Meßreihen, zu denen verschiedene Umlaufgeschwindigkeiten gehören, je n Tabletten gewogen. Dann wird zunächst mit einem f-Test geprüft, ob die Stichprobenvarianzen innerhalb eines Signifikanzniveaus als übereinstimmend gelten können, und bei positivem Ausfall dieses Teils der Untersuchung weiterhin geprüft, ob die Mittelwerte sich signifikant unterscheiden oder nicht.

Ein Beispiel mag das erläutern.

In der folgenden Tabelle stehen in der ersten Zeile die Versuchsnummern von 1 bis 20. Die zweite Meßreihe enthält die Massen der Tabletten in mg bei einer Umlaufgeschwindigkeit von 17 Umdrehungen pro Minute und die dritte Zeile die entsprechenden Werte bei einer Umlaufgeschwindigkeit von 21 Umdrehungen pro Minute.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
372	364	359	372	365	368	368	368	368	375	370	372	363	367	367	371	364	364	370	371
361	365	359	366	363	364	360	358	359	362	353	358	359	364	364	358	365	357	363	363

Die Auswertung der beiden Meßreihen ergibt folgende Werte

$$m_1 = 367{,}9 \quad m_2 = 361{,}1$$
$$s_1^2 = 15{,}15 \quad s_2^2 = 11{,}42$$

Die beiden Testgrößen sind

$$u = 1{,}327 \quad v = 5{,}944$$

Der f-Test ergibt auf dem 99-Prozent-Niveau, daß die Gleichheit der Stichprobenvarianz akzeptiert wird. Denn

$$f(1{,}327;\ 19;\ 19) = 0{,}728$$

Der einseitige t-Test ergibt auf dem 99-Prozentniveau, daß die Mittelwerte signifikant voneinander abweichen. Denn

$$t(5{,}944;\ 38) > 0{,}995$$

Standards zum Testen von Hypothesen in der pharmazeutischen Industrie

Beim Überprüfen von Tabletten wird deutlich unterschieden, ob es sich um einfache physikalische oder um aufwendige chemische Analysen handelt. Deshalb setzt das Deutsche Arzneibuch die Mindestanzahl der Versuche bei der Prüfung auf Gleichförmigkeit der Masse mit n = 20 fest, beschränkt aber den minimalen Stichprobenumfang zunächst auf n = 10, wenn die Gleichförmigkeit des Gehalts überprüft wird. Das ist verständlich. Denn im ersten Fall reichen einfache physikalische Untersuchungen aus. Im zweiten Fall sind dagegen aufwendige chemische Analysen erforderlich.

Das Institut für Arzneimittel und Medizinprodukte, besser nach seinem früheren Namen als Bundesgesundheitsamt bekannt, setzt folgenden Standard für die Überprüfung pharmazeutischer Produkte fest: f-Test, t-Test, 95-Prozent-Niveau und Stichprobenumfang mindestens 6 .

Ergebnis

In der pharmazeutischen Industrie findet man realitätsnahe Beispiele für das Testen von Hypothesen. Aber sie verwenden den f-Test und den t-Test und gehen damit eindeutig über das Niveau von Schule hinaus. Daher ist zu überlegen, ob man nicht doch das exzessive Konzept eines anwendungsorientierten Mathematikunterrichts fallen läßt und statt dessen grundlegende Kenntnisse und Begriffe zur Wahrscheinlichkeitsrechnung vermittelt. Es besteht dann die Aussicht, daß alle diejenigen, die später im Beruf mit dem Testen von Hypothesen zu tun haben, sich zwar die nötigen Kenntnisse, Einsichten und Fähigkeiten erarbeiten müssen, aber dabei schneller vorankommen, weil sie über ausreichende Grundkenntnisse verfügen.

Literatur

BUTH, M.: Zum Thema 'Testen von Hypothesen': Was man aus der Forschungspraxis für die Schule lernen kann.- in: Stochstik in der Schule 13 (1993) S. 35 - 46 und in: Praxis der Mathematik 94 (1994) S. 67 - 70

KRIZ, J.: Statistik in den Sozialwissenschaften.- Reinbek: Rowohlt 1973

KREYSZIG, E.: Statistische Methoden und ihre Anwendungen.- Götingen: Vandenhoeck & Ruprecht 1968

LEHN, J.; WEGMANN, H.: Einführung in die Statistik.- Stuttgart: Teubner 1992

Prof. Dr. M. Buth
Bataverweg 35, 22455 Hamburg