

DAS TESTEN VON HYPOTHESEN \*)  
ALS EIN VORGEHEN IN SECHS SCHRITTEN

L. W. JOHNSON

übersetzt von Bernd Wollring

Einen Hauptbestandteil der Einführungskurse zur Statistik im Bereich der Schule oder der Universität bildet in der Regel eine Einführung in die statistischen Prinzipien des Testens von Hypothesen über Parameter einer Grundgesamtheit (Mittelwerte, Varianzen, Regressionskoeffizienten, usw.). Oft scheinen die Studenten allerdings in den Einzelheiten steckenzubleiben, da die verschiedenen Tests bei ihrer ersten Einführung sehr unterschiedlich erscheinen. Bei dem Versuch, diese Schwierigkeiten zu überwinden, habe ich festgestellt, daß ein Zerlegen des gesamten Vorgehens beim statistischen Testen von Hypothesen in seine Grundbestandteile dem Studenten wohl einige Einsicht in die Natur dieses Vorgehens bringt. Die folgende Technik, die ich bei einem Einführungskurs zur Statistik, einem Statistikkurs in der Volkswirtschaftslehre und einem Einführungskurs zur Statistik für Studenten der Betriebswirtschaftslehre verwendet habe, ermöglicht es dem Studenten, sein Augenmerk auf diejenigen Aspekte des Testens von Hypothesen zu richten, die Überlegung erfordern. Ist dieses Grundgerüst aus sechs Schritten erst geläufig, so ist im Grunde genommen jede komplexere Problemstellung im Bereich des Testens (zum Beispiel Varianzanalyse) für die Studenten einfacher zu erfassen; denn sie können, wie einer bemerkte, "alle bis sechs zählen".

*Die Methode der sechs Schritte*

*Schritt 1: Genaues Festlegen der Hypothese*

Der erste entscheidende Schritt besteht darin, die zu testende Hypothese genau und explizit zu formulieren. Dabei soll der

\*) Originalartikel in 'TEACHING STATISTICS' (1981) Heft 2, Band 5  
'Teaching Hypothesis Testing as a Six Step Process'

Student erkennen, daß eine unkorrekt formulierte Hypothese keine Antwort auf die Frage liefern wird, die wirklich zur Debatte steht.

Insbesondere sollte die Hypothese so formuliert sein, daß sie eine geeignete Testgröße angibt, und sie sollte stets eine Nullhypothese und eine Alternativhypothese umfassen.

Ein Beispiel:

$H_0$ : Dünger A bringt denselben Ertrag wie Dünger B.

$H_A$ : Die Dünger A und B bringen unterschiedliche Erträge.

Diese Formulierung ist nicht genau genug, da man verschiedene Testgrößen benutzen kann, je nach dem, was man unter "demselben Ertrag" versteht. Zum Vergleich folgende Formulierung:

$H_0$ : Der mittlere Ertrag pro Hektar bei Verwenden des Düngers A ist derselbe wie der bei Verwenden des Düngers B.

$H_A$ : Der mittlere Ertrag pro Hektar bei Verwenden des Düngers A ist anders als der bei Verwenden des Düngers B.

Hier würde man sicher  $\bar{x}$  als Testgröße wählen.

*Schritt 2: Festlegen des Signifikanzniveaus*

Der nächste Schritt besteht im Festlegen der höchsten zulässigen Wahrscheinlichkeit, mit der man die wahre Nullhypothese zurückweist. Hier ergibt sich von selbst die Gelegenheit, Fehler erster Art und Fehler zweiter Art zu diskutieren. Wenn die Studenten die Auswirkungen dieser beiden Fehler in Anwendungssituationen nicht einschätzen können, so wird ihnen geraten, einen Wert zwischen 1 % und 5 % zu wählen.

*Schritt 3: Festlegen des Tests und Sammeln der Information aus einer Stichprobe*

Ist die Hypothese (Nullhypothese und Alternativhypothese) korrekt formuliert, so hat der Student zu entscheiden, welcher Test angemessen ist, und anschließend die zur Durchführung des Tests erforderliche Information aus einer Stichprobe zu sammeln. Betrifft der Test zum Beispiel einen einfachen Mittelwert, so wäre ein Z-Test oder ein t-Test geeignet, somit bestände die notwendige Infor-

mation (die Zutaten des Rezeptes, wenn Sie so wollen) aus dem Stichprobenmittelwert, der Standardabweichung der Grundgesamtheit (oder der Stichprobe) und dem Umfang der Stichprobe. In der Praxis erweist sich an dieser Stelle eine Diskussion der Techniken zur Stichprobengewinnung als nützlich, sofern die Daten nicht bereits verfügbar sind. Ich finde allerdings, daß man Verfahren zur Stichprobengewinnung am besten als separates Gebiet bespricht und gehe daher an dieser Stelle nur auf zufällige Stichproben ein. In jedem Fall müssen die Daten mit den Annahmen vereinbar sein, die dem Testverfahren zugrunde liegen. Wegen der unmittelbaren Schwierigkeiten, die die Studenten auf diesem Niveau haben, erhalten sie die Information aus einer Stichprobe in Form irgendeiner "eingekleideten Aufgabe".

*Schritt 4: Auffinden des kritischen Wertes für die Testgröße*

Hat man nun Schritt 2 ausgeführt und in Schritt 3 einen angemessenen Test bereitgestellt, so ist der kritische Wert der erforderlichen Testgröße (zum Beispiel Z, t, F oder Wilcoxon) nachzuschlagen. Gerade hier möchte ich betonen, daß die Kenntnis der jeweiligen Verteilung wesentlich und notwendig ist. Hier ist ein kurzer Rückblick zu Wahrscheinlichkeitsverteilungen zweckmäßig, ebenso wie die Diskussion der Tatsache, daß einige Tests parameterfrei sind.

*Schritt 5: Berechnen der beobachteten Testgröße*

Mit Hilfe der Informationen von Schritt 3 berechne man den beobachteten Wert der passenden Testgröße.

*Schritt 6: Testen und Interpretieren*

Man vergleiche den kritischen Wert der Testgröße aus Schritt 4 mit dem berechneten aus Schritt 5 und verwende eine angemessene Regel, um über das Verwerfen oder Nicht-Verwerfen (das heißt

das Annehmen) der Nullhypothese zu entscheiden. Schließlich befindet man, was dieses Ergebnis in Hinsicht auf die in Schritt 1 festgelegte Hypothese aussagt.

#### *Diskussion*

Führt man das Testen von Hypothesen mit dieser Methode der sechs Schritte ein, so erkennen die Studenten, wie ich herausgefunden habe, sehr schnell, daß die einzige echte Denkarbeit bei Schritt 1 anfällt, wo es darum geht, welcher Test benutzt werden soll, und dann noch einmal beim Interpretieren der Ergebnisse. Den Rest mag man einfach als "black box" auffassen, oder als ein Rezept, das einfach nur mechanisch abzuarbeiten ist. Gewöhnlich beginne ich mit dem Z-Test für einen einfachen Mittelwert und komme mit der Zeit zu komplizierteren Techniken. Den meisten meiner Studenten scheint es inzwischen leichter zu fallen, Hypothesen über Regressionskoeffizienten zu testen oder Varianzanalysen durchzuführen, als das in meinen früheren Klassen der Fall war; allerdings habe ich keine empirischen Daten, um diese Behauptung zu belegen.

Einige neuere Tests beschreiben einen anderen Zugang zum Testen von Hypothesen, bei dem der Begriff des Wahrscheinlichkeitswertes (p-Wert) für einen Test benutzt wird. Der Vollständigkeit wegen bespreche ich diese Methode in meinen Klassen, indem ich einfach angebe, daß man die Schritte 2 und 4 streicht, die Schritte 1, 3 und 5 dagegen beläßt und anschließend eine geänderte Form von Schritt 6 (nennen wir ihn 6a) ausführt, wobei man die Wahrscheinlichkeit dafür bestimmt, einen derart extremen Wert der Testgröße zu beobachten, vorausgesetzt, daß die Nullhypothese zutrifft. Anschließend entscheidet man, ob dieser p-Wert zu klein ist, um die Nullhypothese anzunehmen. Die Studenten erkennen schnell, daß dieser Ansatz im Prinzip nichts anderes ist als die Methode der sechs Schritte, nur daß das Festlegen des Signifikanzniveaus nun in Schritt 6a statt in Schritt 2 erfolgt.

Mich würde interessieren, ob andere Lehrer der Statistik das gesamte Vorgehen beim Testen von Hypothesen in Schritte zerlegen, die im wesentlichen Rezepte sind. Es gibt eine riesige Vielfalt von Texten zur Statistik, und in einigen, die mir geläufig sind, wird das Testen in drei oder vier Schritte zerlegt, aber eine Diskussion in Form der sechs Schritte habe ich bislang nicht gesehen. Aber irgendwo muß es eine geben, denn sie erscheinen derart einfach, und in jedem Fall sind es diese sechs Schritte, die man beim Testen von Hypothesen stets ausführt.