Wie man zwei Kekssorten auf Unterschiede testet

von Rhonda C. Magel, North Dakota State University, Fargo, USA.[1]

übersetzt von Klaus Krug, Bamberg

Zusammenfassung: Dieser Aufsatz stellt zwei Kursprojekte vor, die beim Testen von Hypothesen einsetzbar sind. Im ersten Projekt müssen Studenten Daten sammeln, mit dem Ziel, einen Unterschied im Mittelwert der Schokoladensplitter pro Plätzchen (Kekse) bei zwei Sorten zu überprüfen. Ferner sind Daten gesammelt (2. Projekt) worden, um einen Unterschied bei Geschmackseinstufungen zu untersuchen.

Einführung

Die Abteilung Statistik an der North Dakota State University bietet jedes Semester mehrere Kapitel eines Einführungskurses in die Statistik an. Die Voraussetzung für den Kurs sind Kenntnisse in Algebra. Es wird nicht angenommen, dass die Studenten irgendwelche Vorkenntnisse in Statistik haben. Der Zweck des Kurses ist es, die Studenten in das Gebiet der Statistik und ihrer Anwendungen auf Situationen aus dem realen Leben einzuführen.

Studenten, die diesen Kurs belegen, sind vorwiegend im zweiten oder dritten Studienjahr. Ihre Hauptfächer variieren von Biologie, über Landwirtschaft, Lebensmittelkunde, Ernährungswissenschaft, Luftfahrttechnik, Psychologie, Soziologie, Wirtschaftskunde, bis hin zum Studium kindlicher Entwicklungsphasen. Die Studenten zu motivieren, nimmt einen großen Teil der Arbeit in Anspruch.

Ich unterrichte jedes Jahr wenigstens ein Kapitel in diesem Einführungskurs in die Statistik. Während ich verschiedene Hypothesentests und Konfidenzintervalle in meinem Kurs einführe, habe ich dazu gerne einige Aktivitäten, die von den Studenten verlangen, Daten zu sammeln, statistische Tests durchzuführen, evtl. das zugehörige Konfidenzintervall zu berechnen und die Ergebnisse zu interpretieren. Ich finde, dass dies eine gute Gelegenheit ist, die Studenten aktiv in die Arbeit im Klassenzimmer einzubeziehen. Dieser Aufsatz bespricht zwei Kursprojekte, bei denen Daten gesammelt wurden, um damit statistische Tests durchzuführen.

[1] Übersetzung aus Teaching Statistics v.20(1998)3, S.81-83

Erstes Kursprojekt

Im letzten Semester habe ich nach der Einführung des t-Tests über zwei unabhängige Stichproben, des t-Tests von Stichprobenpaaren, entsprechender Konfidenzintervalle und der Behandlung von Voraussetzungen für die Durchführung der Tests, beschlossen, meine Studenten zwei Projekte zum Thema Schokoladensplitterplätzchen (Schokoladensplitterkekse) ausführen zu lassen. Im ersten Projekt wurden die Studenten aufgefordert, Daten zu sammeln, um zu testen, ob es einen offensichtlichen Anhaltspunkt dafür gab, dass der Mittelwert der Schokoladensplitter pro Plätzchen in jeder von zwei verschiedenen Sorten von Plätzchen unterschiedlich war. Vor längerer Zeit hatte ich dazu Packungen von zwei Arten von Schokoladensplitterplätzchen gekauft, die beide ungefähr dasselbe gekostet haben. Die Verpackungen der Kekse waren entfernt worden, bevor die Studenten sie sahen. Die Plätzchen waren nur als Sorte A und Sorte B gekennzeichnet. Jedem Studenten - es waren 22 – wurde ein Plätzchen von jeder Sorte gegeben und er wurde aufgefordert, die Anzahl der Splitter in jedem zu zählen. Die Plätzchen waren von annähernd gleicher Größe. Es wurde keine formelle Wägung vorgenommen, um damit herauszufinden, welches Plätzchen mehr wog oder ob ihr Gewicht annähernd gleich war. Wir haben lediglich genau getestet, ob die durchschnittliche Zahl an Schokoladensplittern pro Plätzchen bei der Herstellung gleich geblieben war oder nicht.

Das Problem, die Zahl der Schokoladensplitter in jedem Plätzchen zu zählen, wurde besprochen und Techniken, wie dies geschehen könnte, wurden erwähnt. In unserem Fall mußten wir die trockenen Plätzchen mit unseren Händen zerbrechen um die Splitter zu zählen, da wir keinen Zutritt zu einem Laboratorium hatten. Die Studenten arbeiteten beim Zählen zusammen. Dies war hilfreich, um eine größere Genauigkeit sicherzustellen. Ich erwähnte dem Kurs gegenüber, dass ich jedem Teilnehmer Splitter zum Zählen gab, damit er im Projekt einbezogen war, aber das Ergebnis sollte nicht von der Person abhängen, die damit befasst war.

Die Studenten erhielten das folgende Blatt, um darauf ihre Ergebnisse festzuhalten (Beisp.1).

Später nahmen die Studenten die Daten mit in den Computerraum und vervollständigten das Arbeitsblatt 1. Für die anschließenden Berechnungen wurde das Statistik-Programm MINITAB benutzt.

Beispiel 1:

Es sei μ_A der Mittelwert aus den Schokoladensplittern in jedem Plätzchen der Sorte A .
Es sei μ_B der Mittelwert aus den Schokoladensplittern in jedem Plätzchen der Sorte B .
Wir möchten gerne (auf dem Niveau α=0,05) einen Test durchführen, um zu sehen, ob es einen Anhaltspunkt dafür gibt, dass sich die zwei Mittelwerte unterscheiden. Eine Zufallsstichprobe der Plätzchen der Sorte A wurde gezogen und die Zahl der Schokoladensplitter in jedem Plätzchen wurde mit den folgenden Ergebnissen gezählt:

Analog ergab eine Zufallsstichprobe bei Plätzchen der Sorte B die folgenden Ergebnisse:

Anmerkung: Der Kurs erhielt die folgenden Daten:
Stichprobe A:
21, 14, 23, 15, 7, 16, 15, 15, 21, 13, 12, 15, 12, 9, 19, 7, 11, 21, 16, 16, 15, 12
Stichprobe B:
15, 21, 22, 24, 20, 14, 12, 21, 20, 9, 20, 17, 30, 23, 19, 20, 20, 20, 22, 18, 22, 25

Im ersten Teil des Arbeitsblattes 1 müssen die Studenten die Voraussetzungen für den unabhängigen doppelten t-Test überprüfen. Abgesehen davon, dass die zwei Zufallsstichproben zu den Schokoladensplitterzählungen unabhängig sein müssen, sollten die Schokoladensplitterzählungen für Plätzchen der Sorte A annähernd normalverteilt sein und auch die Schokoladensplitterzählungen für Plätzchen der Sorte B sollten annähernd normalverteilt sein. Man beachte, dass in den Abschnitten (b) und (c) die Studenten aufgefordert werden, Histogramme zu den Daten der Stichproben über die Plätzchen der Sorte A und der Sorte B zu zeichnen, um zu erkennen, ob sie spüren, dass die Voraussetzung der Normalverteilung verletzt sein wird. Den Studenten ist gesagt worden, dass es verteilungsfreie Tests gibt, falls die Normalitätsbedingung verletzt wird, aber bis jetzt sind diese nicht besprochen worden. Es zeigte sich, beide Histogramme schienen symmetrisch und glockenförmig zu sein.

Arbeitsblatt 1

$$\text{Test:}\quad \begin{matrix} H_o : \mu_A = \mu_B \\ H_1 : \mu_A \neq \mu_B \end{matrix} \qquad \alpha = 0,05$$

Geben Sie die Untersuchungsergebnisse zu A in C1 ein und die zu B in C2 ein.

a) Finden Sie für beide Plätzchensorten die Mittelwerte und die Standardabweichungen zu den Schokoladensplitterzahlen pro Plätzchen heraus.

$\overline{x}_A$ = ________ s_A = ________ $\overline{x}_B$ = ________ s_B = ________

b) Scheint die Schokoladensplitterzahl bei den Plätzchen der Sorte A annähernd normalverteilt zu sein ? Zeichnen Sie dazu ein Histogramm. Stellungnahme.

c) Scheint die Schokoladensplitterzahl bei den Plätzchen der Sorte B annähernd normalverteilt zu sein ? Zeichnen Sie dazu ein Histogramm. Stellungnahme.

d) Führen Sie den Test durch. (Es ist ein doppelter t-Test, der gleiche Varianzen voraussetzt.)

Was ist der Wert der Testgröße ?

____

Was ist der p – Wert ?

____

Sollte H_o abgelehnt werden ?

____

Das 95% Konfidenzintervall für μ_1 - μ_2 lautet:

____

Erklären Sie das Ergebnis.
Ist das Testergebnis in Einklang mit dem Ergebnis des Konfidenzintervalls ?

____

Anmerkung: (b) und (c) wurden deshalb hinzugefügt, damit die Studenten an eine der Testvoraussetzungen erinnert wurden.

Der unabhängige doppelte t-Test setzt auch gleiche Varianzen (oder Standardabweichungen) voraus. Das Arbeitsblatt verlangt, dass MINITAB dazu benutzt wird, zur Stichprobe die Standardabweichung der Schokoladensplitterzahlen zu jeder Plätzchensorte zu berechnen, um ein Gefühl dafür zu bekommen, ob diese Voraussetzung verletzt sein wird. In unserem Fall zeigte es sich, dass die Standardabweichungen der Schokoladensplitterzahlen in den Plätzchen der Sorte A und der Sorte B 4,375 bzw. 4,485 betrugen. Es bestand nicht die Ansicht, dass die Voraussetzung gleicher Varianzen verletzt wurde.

Die Nullhyothese wurde auf Grund der im Kurs gesammelten Daten abgelehnt. Ein 95% Konfidenzintervall für $\mu_A - \mu_B$ war durch (-7,65 ; -2,26) gegeben. Die Studenten waren in der Lage, eine schriftliche Erklärung für die Berechnung dieses Intervalls zu geben und zu begründen, warum die Nullhypothese, wonach die durchschnittliche Schokoladensplitterzahl in einem Plätzchen der Sorte A die gleiche war wie die durchschnittliche Schokoladensplitterzahl in einem Plätzchen der Sorte B, zu verwerfen ist.

Zweites Kursprojekt

Im 2. Projekt des Kurses sammelten die Studenten Daten zur Geschmackseinstufung beider Plätzchensorten. Jeder Student, durfte beide Plätzchensorten probieren und ihnen Rangstufen zuweisen, wenn er es wünschte. Jeder wollte die Plätzchen probieren und einordnen !! Der Grund für die Durchführung dieses Projekts war, dass ich den unabhängigen doppelten t-Test und den t-Test für Stichprobenpaare eingeführt hatte und die Studenten dazu bekommen wollte, den Unterschied zwischen zwei unabhängigen Stichproben und einem Stichprobenpaar einzusehen.

Die Tätigkeit umfasste die Umformung einer qualitativen Zufallsgröße in eine quantitative. Deshalb mussten die Studenten des Kurses annehmen, dass sie eine Zufallsstichprobe berufsmäßiger Geschmackstester darstellten. Professionelle Geschmackstester sollten im Sinn haben, welche Qualitätsansprüche sie an Schokoladenplätzchen stellen und wie viele Punkte sie in Abzug bringen, falls eine Eigenschaft nicht vorhanden ist.

Da eine Einstufung der Plätzchen in jedem der zwei Durchgänge dem gleichen Geschmackstester übertragen wurde, waren dies abhängige Einstufungen, die auf eine paarweise Datenerfassung hinausliefen. Es wurde den Studenten gegenüber erwähnt, dass ein Experiment, das Datenpaare enthält, benutzt wird, wenn man glaubt, dass die frühzeitige paarweise Anordnung deutlich die Ursache der Streuungen verringern wird. Auf jeden Fall würde jeder Geschmackstester seine eigenen Ideen darüber haben, was ein gutes Plätzchen ausmacht und wie viele Punkte für jedes fehlende Merkmal abzuziehen sind. Wenn verschiedene Geschmackstester für die Einstufung jeder Plätzchensorte eingesetzt worden wären, dann würden dadurch mehr Gründe für Streuungen hinzugefügt werden.

Das folgende Blatt wurde den Studenten zur Vervollständigung gegeben. (Beispiel 2). Die Studenten wurden aufgefordert, einander nicht um Rat zu fragen, bevor sie ihre Einstufungen aufschreiben. Der Hälfte des Kurses wurden zuerst Plätzchen der Sorte A zum Probieren gegeben und der anderen Hälfte wurde Sorte B gegeben. Rechtzeitig waren bei allen Plätzchen die Verpackungshüllen entfernt worden. Die Studenten wurden aufgefordert, einen Schluck Wasser zwischen den Geschmackstests zu trinken.

<u>Beispiel 2:</u>

Nehmen Sie an, dass die Studenten des Kurses eine Zufallsstichprobe berufsmäßiger Geschmackstester darstellen. (Für diesen Fall testen wir die Plätzchen). Jeder Student wird zwei Sorten Plätzchen, innerhalb einer Skala von "0" = grässlich bis "10" = ausgezeichnet mit "5" = geht, einordnen. Benutzen Sie nur ganze Zahlen.

(Die Daten werden in Paaren zusammengefasst. Zuerst ist die Einstufung verzeichnet, die jeder Student einem Plätzchen der Sorte A gab und an zweiter Stelle die für das Plätzchen der Sorte B.)

Anmerkung: Die tatsächlich im Kurs gesammelten Daten, lauteten wie folgt:
(5;7) (4;6) (5;8) (5;7) (6;8) (3;9) (5;10) (4;7) (5;7) (5;7) (4;7) (5;8) (6;5) (7;9) (5;3) (1;5) (4;8) (5;7) (3;7) (8;7) (4;6) (3;6)

Die Studenten nahmen die Daten später mit in den Computerraum und vervollständigten das Arbeitsblatt 2 . Berechnungen wurden unter Benutzung des Programms MINITAB ausgeführt. Im vorliegenden Fall wurde ein t-Test mit Stichprobenpaaren durchgeführt. Die Nullhypothese wurde auf Grund der Daten der Stichprobe verworfen. Als 95% Konfidenzintervall für μ_D wurde (-3,19 ; -

1,54) gefunden. Die Studenten waren erneut in der Lage, dies zu interpretieren und es in Beziehung zur Ablehnung der Nullhypothese zu setzen.

<u>Arbeitsblatt 2</u>

Es sei μ_D = Mittelwert aus der Differenz der Rangstufen zwischen den Plätzchen der Sorte A und den Plätzchen der Sorte B .

Test : H_o: $\mu_D = 0$ $\alpha = 0{,}05$

H_1: $\mu_D \neq 0$

Geben Sie die Untersuchungsergebnisse über die Plätzchen der Sorte A in C1 und entsprechende Unterlagen über die Plätzchen der Sorte B in C2 ein. Sie müssen zuerst die Differenzwerte berechnen. Geben Sie diese in C3 ein.

a) Finden Sie den Mittelwert und die Standardabweichung der Stichprobe über die Differenz der Rangstufen bei den Plätzchen der Sorte A und B heraus.

$\overline{X}_D$ = ______________________ S_D = ______________________

b) Scheinen die Differenzwerte annähernd normalverteilt zu sein ? Zeichnen Sie dazu ein Histogramm. Stellungnahme.

(Anmerkung: Im vorliegenden Fall schien das Histogramm symmetrisch und glockenförmig zu sein.)

c) Führen Sie den Test durch.

Was ist der Wert der Testgröße ? ________________________

Was ist der p – Wert ? _________________________

Schlussfolgerung:

____

d) Finden Sie ein 95% Konfidenzintervall für μ_D . Das 95% Konfidenzintervall für μ_D lautet:

____

Erklären Sie das Ergebnis. Steht es im Einklang mit dem Testergebnis ?

Schlussdiskussion

Die Studenten schienen im Kurs einen großen Teil ihrer Zeit damit zu verbringen, beide Datensätze zu den Plätzchen zu sammeln. Dies waren sinnvolle Aktivitäten, weil jeder Student mit einbezogen war, obgleich die Studenten aus einer großen Vielfalt von Ausbildungsrichtungen kamen. Die Studenten arbeiteten eifrig im Computerraum. Sie waren in der Lage, schriftliche Deutungen der Ergebnisse zu verfassen. Statt nur zu sagen "H_o wird abgelehnt", waren sie in der Lage zu schreiben "H_o wird abgelehnt, weil..." und was das bedeutete.