

Universität Wien
Institut für Mathematik
Wintersemester 2009/2010

Medizinische Statistik
Epidemiologie und χ^2 Vierfeldertest

Seminar Angewandte Mathematik
Ao. Univ. Prof. Dr. Peter Schmitt

von
Nadja Reiterer
Matrikelnr.: 0517778
csag1191@uibk.ac.at
und
Cornelia Plunger
Matrikelnr.: 0516571
csaf8340@uibk.ac.at

Inhaltsverzeichnis

1	Zusammenfassung	2
1.1	Allgemeines	2
1.2	Epidemiologie	2
1.2.1	Fachbegriffe	2
1.2.2	Lösung	3
1.2.3	Anwendungen	3
1.3	χ^2 -Vierfeldertest	4
1.3.1	Grundlagen	4
1.3.2	Durchführung des χ^2 – <i>Vierfeldertests</i> :	4
1.3.3	Schlussbemerkung	5
1.3.4	Anwendungen	5
1.4	Literaturempfehlung	6
2	Erfahrungsbericht zum Thema	7
3	Quellenangabe	8

1 Zusammenfassung

1.1 Allgemeines

Unser Vortrag setzt sich aus zwei Teilgebieten der Medizinischen Statistik zusammen:

Der erste Teil behandelt eine Problemstellung aus der Epidemiologie, der zweite Teil beschäftigt sich mit dem χ^2 Vierfeldertest.

In beiden Teilen ist ein konkretes Beispiel der Ausgangspunkt für die Herleitung einer Lösung der Problemstellung. Dazu werden alle Fachbegriffe eingeführt, die für die Lösung der Problemstellung aus dem Beispiel relevant sind. Dadurch wird die präsentierte Lösungsmethode für den Zuhörer nachvollziehbar.

Die Lösung des Problems ergibt gleichzeitig auch eine allgemeine Lösungsmethode für analoge Fragestellungen.

Den Abschluss beider Teile bildet eine kleine Diskussion über die Aussagekraft der Resultate.

Die Zuhörer sollen in die Materie eingeführt werden, zusätzlich soll es ihnen möglich sein, nach dem Vortrag, die Lösungsansätze auch auf andere Beispiele anzuwenden.

Die Vortragenden interagieren sehr stark miteinander und auch die Zuhörer werden miteinbezogen.

Eine pdf- Präsentation und die Tafel dienen als Hilfsmittel des Vortrags.

1.2 Epidemiologie

Die Epidemiologie beschäftigt sich damit, wie Krankheiten in größeren Populationen auftreten, welche Ursachen und Risikofaktoren man erkennen kann, wie sich Krankheiten in der Bevölkerung verbreiten und wie man diese prognostizieren kann.

1.2.1 Fachbegriffe

Prävalenz bezeichnet den relativen Krankenbestand, das heißt, die Wahrscheinlichkeit $P(K)$, dass eine beliebige Person aus einer Population zum Zeitpunkt t erkrankt ist.

Diagnostische Test dienen der Untersuchung des Krankheitstatus einer Person. Wir gehen im Folgenden von Tests aus, die nur zwei mögliche Ausgänge haben: Ein positives Ergebnis (Krankheit liegt vor) oder ein negatives Ergebnis (Krankheit liegt nicht vor).

Da jeder Test Mängel hat, gibt es Gütekriterien für diagnostische Tests, die Sensitivität und die Spezifität.

Sensitivität bezeichnet die bedingte Wahrscheinlichkeit, $P(T_+|K)$, dass der Test positiv reagiert, wenn die Person auch tatsächlich krank ist, mit anderen Worten, der Test reagiert richtig positiv.

Das **falsch negative Ergebnis** ist die Gegenwahrscheinlichkeit zur Sensitivität, d.h., $P(T_-|K) = 1 - P(T_+|K)$.

Spezifität bezeichnet die bedingte Wahrscheinlichkeit, $P(T_-|\bar{K})$, dass der Test negativ reagiert, wenn die Person nicht krank ist, mit anderen Worten, der Test reagiert richtig negativ.

Das **falsch positive Ergebnis** ist die Gegenwahrscheinlichkeit zur Spezifität, d.h., $P(T_+|\bar{K}) = 1 - P(T_-|\bar{K})$.

Der **positive Vorhersagewert** ist die Wahrscheinlichkeit, $P(K|T_+)$, dass die Person auch wirklich krank ist, wenn sie ein positives Testergebnis erhält. Der **negative Vorhersagewert** ist analog dazu $P(\bar{K}|T_-)$.

1.2.2 Lösung

Es gilt bekanntlich, dass $P(K|T_+) \neq P(T_+|K)$.

Aus diesem Grund kann man $P(K|T_+)$ mit Hilfe von Prävalenz, Sensitivität und Spezifität errechnen, indem man die **Bayes Formel** anwendet.

Daraus folgt, für den positive Vorhersagewert gilt:

$$P(K|T_+) = \frac{P(K) \cdot P(T_+|K)}{P(K) \cdot P(T_+|K) + P(\bar{K}) \cdot P(T_+|\bar{K})}$$

Negativer Vorhersagewert analog.

Beim Berechnen mehrerer Beispiele erkennt man leicht, dass der Vorhersagewert stark von der Prävalenz abhängig ist. Leider ist die Prävalenz für Einzelfälle nicht immer bekannt. Meist ist sie jedoch (erfreulicherweise) sehr gering, was sich aber nachteilig für den positiven Vorhersagewert auswirkt. In diesem Fall ergibt sich für den negativen Vorhersagewert einen Wert nahe 1.

In jedem Fall muss ein Arzt einen Befund ernst nehmen und weitere Tests durchführen.

Im Vortrag werden die Vorhersagewerte für einen HIV- Test hergeleitet und berechnet.

1.2.3 Anwendungen

- Andere (nicht- medizinische) Tests: Voraussetzung ist dabei, dass nur zwei Ausgänge möglich sind.
- Erkennen von Spam- Mails
- Erkennen von ähnlichen DNA- Sequenzen

1.3 χ^2 -Vierfeldertest

Der χ^2 -Vierfeldertest untersucht, ob es einen Zusammenhang zwischen zwei bestimmten Merkmalen gibt. Dabei werden jeweils Alternativmerkmale untersucht.

1.3.1 Grundlagen

Alternativmerkmale sind die für eine Untersuchung relevanten Eigenschaften, welche nur zwei gegensätzlich Ausprägungen besitzen. Beispiele dafür sind der Rhesusfaktor, alle Fragen mit Ja- Nein- Antworten,..

Statistische Tests sind von der Grundidee einem Gerichtsverfahren ähnlich: Es gilt die Unschuldsvermutung bis das Gegenteil bewiesen wurde.

Bei einem statistischen Testverfahren wird eine **Nullhypothese** erstellt, welche die Unschuld eines Verdächtigen vermutet. Gleichzeitig wird dazu eine **Alternativhypothese** aufgestellt, in der das Gegenteil behauptet wird.

In Bezug auf die Untersuchung besagt die Nullhypothese, dass es keinen Zusammenhang zwischen den beiden Merkmalen gibt, die Alternativhypothese, dass es einen Zusammenhang gibt.

In diesem Zusammenhang gibt es den **Fehler 1. Art** und den **Fehler 2. Art**. Beim Fehler 1. Art wird ein Unschuldiger verurteilt (Die Alternativhypothese wird fälschlicherweise angenommen).

Beim Fehler 2. Art wird ein Schuldiger freigesprochen (Die Nullhypothese wird fälschlicherweise angenommen). Diese Fehler sollten beim Test vermieden (bzw. möglichst klein gehalten) werden.

Demnach wird die Nullhypothese zugunsten der Alternativhypothese erst dann verworfen, wenn die Wahrscheinlichkeit einer Fehlentscheidung ziemlich klein ist.

Um den Fehler 1. Art möglichst gut zu vermeiden, hat man **Signifikanzniveaus** α eingeführt. Diese geben die Irrtumswahrscheinlichkeit an. Gemeint ist damit jene Wahrscheinlichkeit, einen Unschuldigen zu Unrecht zu verurteilen. Somit stellt α eine obere Schranke für die Wahrscheinlichkeit des Fehlers 1. Art dar, wenn die Alternativhypothese angenommen wird.

Ein Ergebnis heißt **signifikant**, wenn die Irrtumswahrscheinlichkeit höchstens 5 Prozent beträgt und **hochsignifikant**, wenn sie 1 Prozent beträgt.

1.3.2 Durchführung des χ^2 - Vierfeldertests:

1. Kontingenztafel erstellen:

Die Daten der Stichprobenerhebung werden für die Merkmale x (bzw. $\bar{x}$ und y (bzw. $\bar{y}$ in folgende Tabelle (= Kontingenztafel) eingetragen:

Merkmal	y	$\bar{y}$	gesamt
x	a	b	$a + b$
$\bar{x}$	c	d	$c + d$
gesamt	$a + c$	$b + d$	$n = a + b + c + d$

2. Prüfgröße berechnen:

$$\chi^2 = \frac{n \cdot (ad - bc)^2}{(a+b)(a+c)(c+d)(b+d)}$$

3. Testentscheidung:

Entscheidung gegen die Nullhypothese (also für Alternativhypothese), falls gilt:

$$\chi^2 > \chi_{f;1-\alpha}^2$$

Von Tabellen, die die χ^2 -Verteilung angeben kann man ablesen, wie groß $\chi_{f;1-\alpha}^2$ bei verschiedenen Signifikanzniveaus ist. Dabei hat der χ^2 -Vierfeldertest den Freiheitsgrad 1.

1.3.3 Schlussbemerkung

Abschließend kann man sagen, dass das Ergebnis signifikant ist, wenn die Alternativhypothese mit Irrtumswahrscheinlichkeit $\alpha = 0.05$ bestätigt werden kann. Jedoch ist zu beachten, dass mit Hilfe einer großen Stichprobenanzahl jeder noch so kleine Zusammenhang signifikant gemacht werden kann.

Wird die Alternativhypothese zugunsten der Nullhypothese verworfen, bedeutet das, dass es keinen signifikanten Zusammenhang zwischen den beiden Merkmalen gibt.

Dies kann zwei Ursachen haben:

1. Es gibt keinen relevanten Zusammenhang zwischen den beiden Merkmalen oder
2. Der Stichprobenumfang ist zu klein.

Im Vortrag wurde dieser Sachverhalt anhand eines Beispiels über den Zusammenhang zwischen Geschlecht und Rauchverhalten dargestellt.

1.3.4 Anwendungen

Der χ^2 -Vierfeldertest kann immer dann angewandt werden, wenn überprüft werden soll, ob zwei Alternativmerkmale in Zusammenhang stehen.

1.4 Literaturempfehlung

1. **WEISS, Christel** ⁴(2008): *Basiswissen Medizinische Statistik*. Heidelberg: Springer Medizin Verlag.
Dieses Buch gibt eine gute Einführung in das Thema.
2. **medistat GmbH (o.J.) = medistat Medizinische Statistik**. Aus: <http://www.medi-stat.de/statistik-lexikon-medizin-themen.html>.
< 23.11.2009 >
Diese Website ermöglicht es, schnell und unkompliziert Begriffe aus der medizinischen Statistik nachzuschlagen.

2 Erfahrungsbericht zum Thema

In unserem bisherigen Mathematikstudium haben wir nur sehr wenig über konkrete Anwendungen der Mathematik erfahren. Die einzigen Anwendungen waren im Bereich der Physik anzutreffen, also in einem uns relativ unbekannten und für uns (vielleicht auch deshalb) uninteressanten Bereich.

Eine Ausnahme bildete jedoch eine Vorlesung, in der wir einen kleinen Einblick in die medizinische Statistik erhalten haben. Es hat uns sehr fasziniert, zu erfahren, wie Mathematik in der Medizin angewandt wird. Deshalb sind uns die wenigen Beispiele, die wir im Rahmen dieser Lehrveranstaltung gerechnet haben, auch gut in Erinnerung geblieben.

In unserem Vortrag haben wir uns für zwei Teilgebiete der medizinischen Statistik entschieden, von denen wir glauben, dass sie relativ einfach zu verstehen sind. Deshalb könnte man diese Themen sicherlich auch Nicht-Mathematikern und Schülern näher bringen.

3 Quellenangabe

Literatur

- [1] WEISS, Christel ⁴(2008): *Basiswissen Medizinische Statistik*. Heidelberg: Springer Medizin Verlag.
- [2] BRANDSTÄTTER, Anita (2008): *Angewandte Mathematik. Schriftliche Unterlagen zur Vorlesung im Sommersemester 2008*. Innsbruck: Institut für Mathematik. Universität Innsbruck.
- [3] medistat GmbH (o.J.) = medistat Medizinische Statistik. Aus: <http://www.medi-stat.de/statistik-lexikon-medizin-themen.html>. < 23.11.2009 >
- [4] Wikipedia (2009) = Bayes Theorem. Anwendungsgebiete. Aus: <http://de.wikipedia.org/wiki/Bayestheorem>. < 23.11.2009 >