Mathematik für Biologen

Prof. Dr. Rüdiger W. Braun

Heinrich-Heine-Universität Düsseldorf

16. Dezember 2010

Testverfahren

- Es sei Θ eine Menge von Parametern. Zu jedem $\theta \in \Theta$ gebe es eine Verteilung P_θ
- Die Zufallsvariablen $X_1, \ldots, X_n$ seien unabhängig und alle nach demselben P_θ verteilt. Dieses θ sei unbekannt
- Die Menge M aller möglichen Werte $(x_1, \ldots, x_n)$ von $(X_1, \ldots, X_n)$ heißt *Stichprobenraum*
- Der Parameterraum sei in zwei Mengen H_0 und H_1 zerlegt. Dabei ist H_0 die Nullhypothese und H_1 die Alternative

Testverfahren, Fortsetzung

- Ein *Test* besteht aus der Zerlegung des Stichprobenraums M in zwei Teilmengen K_0 und K_1 derart, dass die Nullhypothese akzeptiert wird, wenn $(x_1, \dots, x_n) \in K_0$, und abgelehnt wird, wenn $(x_1, \dots, x_n) \in K_1$
- Die Menge K_0 heißt *Annahmebereich*, die Menge K_1 heißt *kritischer Bereich*

	$x \in K_0$	$x \in K_1$
$\theta \in H_0$	richtige Entscheidung	Fehler 1. Art
$\theta \in H_1$	Fehler 2. Art	richtige Entscheidung

Ein Test heißt *Signifikanztest* zum Niveau α, wenn alle Fehlerwahrscheinlichkeiten erster Art $\leq \alpha$ sind

Einseitiger oberer Binomialtest zum Niveau α

- Gegeben sind unabhängige $B(1, p)$-verteilte Zufallsvariable $X_1, \ldots, X_n$ mit unbekanntem p sowie ein Signifikanzniveau α
- Getestet wird die Nullhypothese $H_0 = \{p \leq p_0\}$ gegen die Alternative $H_1 = \{p > p_0\}$
- Die Konstante c ist so zu wählen, dass

$$\sum_{k=0}^{c} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} \geq 1 - \alpha$$

$$\sum_{k=0}^{c-1} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} < 1 - \alpha$$

- c heißt *kritischer Wert*

Einseitiger oberer Binomialtest, Fortsetzung

- Der Annahmebereich ist

$$K_0 = \left\{ (x_1, \ldots, x_n) \;\middle|\; \sum_{j=1}^{n} x_j \leq c \right\}$$

- Der kritische Bereich ist

$$K_1 = \left\{ (x_1, \ldots, x_n) \;\middle|\; \sum_{j=1}^{n} x_j > c \right\}$$

- Das bedeutet: Die Nullhypothese wird abgelehnt, falls die Anzahl der Erfolge echt größer als c ist, andernfalls wird sie angenommen

Der Fehler erster Art

- Wenn c der kritische Wert ist, dann wird die Nullhypothese abgelehnt, wenn mindestens $c + 1$ Erfolge auftreten
- Wir wahrscheinlich ist das, wenn die Nullhypothese in Wirklichkeit zutrifft?
- Die Nullhypothese ist $p \leq p_0$; im Extremfall also $p = p_0$
- Fehlerwahrscheinlichkeit erster Art

$$\sum_{k=c+1}^{n} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k}$$

- Diese Größe ist $\leq \alpha$, wenn ihre Komplementärwahrscheinlichkeit $\geq 1 - \alpha$ ist, also wenn

$$\sum_{k=0}^{c} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} \geq 1 - \alpha$$

Effektive Fehlerwahrscheinlichkeit

- In der Regel wird

$$\sum_{k=0}^{c} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k}$$

nicht genau gleich $1 - \alpha$ sein

- Der Test hat also in Wirklichkeit eine etwas kleinere Fehlerwahrscheinlichkeit erster Art als gefordert
- Daher bezeichnet man

$$\alpha_0 = \sum_{k=c+1}^{n} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} = 1 - \sum_{k=0}^{c} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k}$$

als *effektive Fehlerwahrscheinlichkeit erster Art*

Effektive Fehlerwahrscheinlichkeit im Medikamentenvergleich

- Stichprobenumfang 24

$$\begin{aligned} H_0 &: \quad p \leq 0.2 \\ H_1 &: \quad p > 0.2 \end{aligned}$$

- Kritischer Wert $c = 8$
- Bei 9 oder mehr Erfolgen wird der Fehler erster Art begangen

$$\sum_{k=9}^{24} B_{24,\,0.20}(k) = 1 - \sum_{k=0}^{8} B_{24,\,0.20}(k) = 1 - 0.96383 = 0.03617$$

- Die effektive Fehlerwahrscheinlichkeit erster Art beträgt $\alpha_0 = 0.03617$
- Das ist deutlich besser als der geforderte Wert 0.05

Tabelle von $\sum_{k=0}^{r} B_{24,\,p}(k)$

r	p	0.20	0.21	0.22	0.23	0.24
0	0.	00472	00349	00257	00189	00138
1		03306	02577	01998	01541	01183
2		11452	09387	07645	06188	04978
3		26386	22662	19326	16367	13767
4		45988	41188	36622	32330	28338
5		65589	60887	56136	51401	46743
6		81107	77469	73565	69441	65149
7		91083	88804	86206	83297	80094
8		96383	95206	93782	92092	90124
9		98738	98232	97581	96763	95754
10		99621	99438	99188	98855	98421
11		99902	99846	99765	99651	99493
12		99978	99964	99942	99908	99860
13		99996	99993	99988	99979	99967
14		99999	99999	99998	99996	99993
15					99999	99999

Fehlerwahrscheinlichkeit zweiter Art

- Die Wahrscheinlichkeit eines Fehlers zweiter Art macht nur Sinn für ein konkretes p, welches zur Alternative gehört
- Beispiel $p = 0.24$
- Der Fehler wird begangen bei 8 oder weniger Erfolgen
- Die Wahrscheinlichkeit dieses Ereignisses ist

$$\sum_{k=0}^{8} \binom{n}{k} \cdot p^k \cdot (1-p)^{24-k} = 0.90124$$

- Die Wahrscheinlichkeit eines Fehlers zweiter Art beträgt 90% für $p = 0.24$

Beispiel: Medikamentenvergleich mit 200 Probanden

- Bei unverändertem Signifikanzniveau kann die Fehlerwahrscheinlichkeit 2. Art nur durch größere Stichprobenumfänge verringert werden
- Beispiel: Medikamentenvergleich mit 200 Probanden
- Es soll zum Signifikanzniveau $\alpha = 0.05$ festgestellt werden, ob die Heilerfolgswahrscheinlichkeit von Medikament 2 über 20% liegt
- Es sei p die unbekannte Heilerfolgswahrscheinlichkeit von Medikament 2
- Es sei $p_0 = 0.2$ (die Heilerfolgswahrscheinlichkeit von Medikament 1)
- Die Nullhypothese ist $\{p \leq p_0\}$, die Alternative ist $\{p > p_0\}$

Beispiel: 200 Probanden, Fortsetzung

- Bestimme c mit

$$\sum_{k=0}^{c} \binom{n}{k} \cdot p_0^k \cdot (1-p_0)^{200-k} \geq 0.95$$

$$\sum_{k=0}^{c-1} \binom{n}{k} \cdot p_0^k \cdot (1-p_0)^{200-k} < 0.95$$

- Dazu benutzen wir die Tabelle auf der nächsten Folie

Tabelle von $\sum_{k=0}^{r} B_{200,\,p}(k)$ (Ausschnitt)

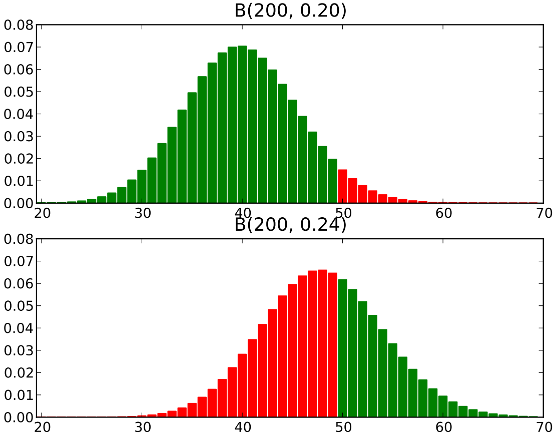
r	p	0.20	0.21	0.22	0.23	0.24
45	0.	83488	73174	60674	47254	34392
46		87375	78459	66989	53944	40722
47		90560	83062	72826	60492	47272
48		93097	86963	78073	66726	53865
49		95065	90179	82664	72503	60323
50		96550	92761	86575	77714	66482
51		97643	94780	89819	82292	72203
52		98425	96317	92441	86210	77379
53		98972	97459	94506	89479	81944
54		99343	98284	96091	92136	85868
55		99590	98867	97278	94243	89157
56		99750	99268	98145	95873	91847
57		99851	99538	98763	97103	93992
58		99913	99714	99193	98009	95663
59		99950	99827	99484	98660	96933

Beispiel: 200 Probanden, Fortsetzung

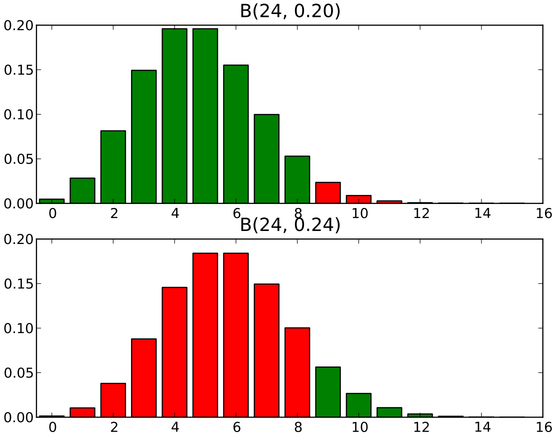
- Der kritische Wert ist $c = 49$
- Die effektive Fehlerwahrscheinlichkeit 1. Art ist 0.04935
- Der Fehler zweiter Art wird gemacht, wenn $p > p_0$, aber trotzdem nicht mehr als c Erfolge auftreten
- Für $p = 0.24$ beträgt sie

$$\sum_{k=0}^{49} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{200-k} = 0.60323$$

Fehler 1. und 2. Art bei 200 Probanden



Zum Vergleich: Fehler 1. und 2. Art bei 24 Probanden



"guaranteed non-inferiority"

- Beim Vergleich zweier Medikamente wird das neuere akzeptiert, wenn es nicht schlechter als das alte ist
- Wegen der Asymmetrie zwischen Nullhypothese und Alternative ist das eine schwierig zu erfüllende Anforderung
- Im Beispiel hat das zweite Medikament selbst bei 200 Probanden eine Chance von unter 50% auf Annahme, obwohl es deutlich besser ist als das alte