Universität Tübingen
Mathematisches Institut
Dr. Stefan Keppeler

Sommersemester 09

Mathematik II für Biologen

Übungsblatt 8 (Abgabe am 19.06.2009)

Aufgabe 22 (10 Punkte)

Wir möchten zeigen, dass die Anzahl der möglichen Ergebnisse

- beim Ziehen von k Kugeln aus einer Urne mit n unterschiedlichen Kugeln
- mit Zurücklegen
- ohne Beachtung der Reihenfolge

durch $\binom{n+k-1}{k}$ gegeben ist. Dazu stellen wir n leere Behälter in einer Reihe auf. Immer wenn wir eine Kugel vom Typ j ($j = 1, \ldots, n$) aus der Urne ziehen (diese legen wir anschließend zurück in die Urne), legen wir eine (farblose) Kugel in den $j$ten Behälter. Die Anzahl der möglichen Verteilungen der k Kugeln auf n Behälter (z.B. so: $\underline{|\bullet\bullet\bullet|\;\;|\bullet|\bullet\bullet|\;\;|}$ für k=6, n=5) entspricht der gesuchten Zahl. (Warum?) Ermitteln Sie diese Zahl, indem Sie die erhaltenen Konfigurationen als eine Anordung von k Kugeln und $n-1$ Zwischenwänden betrachten.

Aufgabe 23 (Simpsons Paradoxon) (10 Punkte)

Die folgende Tabelle nennt die Zahlen von Bewerber(inne)n und davon als Student(inn)en Zugelassenen für zwei verschiedene Studiengänge in Berkeley im Jahr 1973.[6]

Studiengang	männlich		weiblich	
	Bewerber	Zugelassene	Bewerberinnen	Zugelassene
I	825	511	108	89
II	373	22	341	24

Wir wählen aus der Menge aller Bewerber(innen) eine Person zufällig aus, wobei jede Person die gleiche Wahrscheinlichkeit hat, ausgewählt zu werden. Betrachten Sie die folgenden Ereignisse:
W = Die Person ist weiblich.
$M = W^c$ = Die Person ist männlich.
I = Die Person hat sich für den Studiengang I beworben.
$II = I^c$ = Die Person hat sich für den Studiengang II beworben.
Z = Die Person wurde zum Studium zugelassen.

a) Berechnen Sie die folgenden Wahrscheinlichkeiten und interpretieren Sie diese in Worten. (Bsp.: $P[M|Z] = 0{,}8251$ bedeutet: "82,51% der Zugelassenen sind männlich.")
   (i) $P[W]$ und $P[M]$.
   (ii) $P[Z|W]$ und $P[Z|M]$.
   Wer scheint bevorzugt zugelassen zu werden, Männer oder Frauen?

b) Berechnen Sie die folgenden Wahrscheinlichkeiten und interpretieren Sie diese in Worten.
   (i) $P[Z|W \cap I]$ und $P[Z|M \cap I]$.
   (ii) $P[Z|W \cap II]$ und $P[Z|M \cap II]$.
   Wer scheint bevorzugt zugelassen zu werden, Männer oder Frauen?

c) Berechnen Sie
   (i) $P[Z|I]$ und $P[Z|II]$ sowie
   (ii) $P[I|W]$ und $P[I|M]$,
   und erklären Sie, wie sich damit der (scheinbare) Widerspruch zwischen (a) und (b) auflösen lässt.

---

[6] P. Bickel, E. A. Hammel, J. W. O'Connel, *Sex bias in graduate admissions: Data from Berkeley*, Science **187** (1975) 398–404.

Aufgabe 24 (χ^2-Test für Kontingenztafeln)[7] (10 Punkte)
Mit diesem Test lässt sich untersuchen, ob zwei Merkmale (z.B. Augenfarbe und Haarfarbe) als unabhängig voneinander angenommen werden können oder nicht.

In dieser Aufgabe wollen wir den χ^2-Test für Kontingenztafeln anwenden, um die in Aufgabe 23 gestellte Frage bezüglich einer eventuellen Diskriminierung quantitativ zu beantworten. Wir gehen dabei nach Studiengängen getrennt vor und betrachten zunächst Studiengang I.

Studiengang I	Männer	Frauen	Total
Zugelassen	511	89	600
Abgewiesen	314	19	333
Total	825	108	933

Wir wollen im Folgenden die Nullhypothese H_0: *Zulassungsentscheid und Geschlecht sind unabhängig* auf dem Signifikanz-Niveau $\alpha = 5\%$ gegen die Alternativhypothese H_A: *Zulassungsentscheid und Geschlecht sind abhängig* mit einem χ^2-Test testen.

Falls H_0 gelten würde, so würde man erwarten, dass

$$\frac{825}{933} \cdot \frac{600}{933} \cdot 933 = 530{,}55$$

der Männer zum Studiengang I zugelassen würden. (In Wirklichkeit waren es 511.) Mit analogen Rechnungen für die anderen 3 Kategorien erhält man als unter H_0 erwartete Zahlen

Studiengang I	Männer	Frauen	Total
Zugelassen	530,55	69,45	600
Abgewiesen	294,45	38,55	333
Total	825	108	933

Als Teststatistik nehmen wir wieder

$$\chi^2 = \sum_{\text{alle 4 Kategorien}} \frac{(\text{beobachtete Anzahl} - \text{erwartete Anzahl})^2}{\text{erwartete Anzahl}} = 17{,}4\,.$$

Da die Anzahl Freiheitsgrade $\nu =$ (Anzahl Zeilen -1)(Anzahl Spalten -1) $= (2-1)(2-1) = 1$ beträgt, ist der beobachtete Wert von χ^2 wesentlich größer als die zum Niveau $\alpha = 5\%$ gehörige kritische Zahl $\nu + 2\sqrt{2\nu} = 3{,}83$, jenseits der H_0 verworfen wird. Also belegen die Zahlen, dass H_A stimmt. Darüberhinaus lassen die Vorzeichen von (beobachtete Anzahl - erwartete Anzahl) die Männer benachteiligt erscheinen.

a) Führen Sie einen entsprechenden χ^2-Test für Studiengang II durch.

b) Führen Sie einen entsprechenden χ^2-Test für die Summe der beiden Matrizen zu den Studiengängen I und II durch, so dass nicht mehr zwischen den Studiengängen unterschieden wird.

Aufgabe 25 (Binomialtest) (10 Punkte)
In einer medizinischen Pilotstudie sprachen 5 von 16 Patienten auf eine *neue* Behandlung an. Sei $p \in [0,1]$ die (wahre, unbekannte) Ansprechwahrscheinlichkeit auf die *neue* Behandlung.

a) Die Ansprechwahrscheinlichkeit auf die *alte Standardbehandlung* wird mit 15% angegeben. Bestimmen Sie den Verwerfungsbereich, wenn man die Nullhypothese $p = 0{,}15$ gegen die Alternative $p > 0.15$ auf dem 5%-Niveau testet. Verwenden Sie dazu die Ausgabe des MATLAB-Befehls `binocdf(0:16,16,.15)`, die die Verteilung einer Bin(16; 0,15)-verteilten Größe X beschreibt, d.h. die Wahrscheinlichkeiten $P[X \leq k]$ für $k = 0, \ldots, 16$.
Ist die Verbesserung durch die neue Behandlung mit diesem Test signifikant?

b) Bestimmen Sie die Wahrscheinlichkeit β, dass der Test aus (a) die Nullhypothese $p = 0{,}15$ nicht verwirft, falls die wahre Ansprechwahrscheinlichkeit gleich 0,3 ist (d.h. die Wahrscheinlichkeit eines Fehlers 2. Art). HINWEIS: Dabei hilft `binocdf(0:16,16,.3)`

[7] Eine ergänzende Beschreibung finden Sie z.B. in Kapitel 7 (Seite 52–54) des Skripts von Martin Zerner, welches im Webforum im Thread *Literatur* zur Verfügung gestellt wird.