Universität Tübingen, Mathematisches Institut
Dr. Stefan Keppeler Sommersemester 2013

Mathematik II für Biologen

Übungsblatt 7 (Abgabe am 7.6.2013)

Aufgabe 24 (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 ($j = 1, \dots, n$) 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: $\lfloor \bullet\bullet\bullet \rfloor\lfloor\ \rfloor\lfloor \bullet \rfloor\lfloor \bullet\bullet \rfloor\lfloor\ \rfloor$ 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 25 (Simpsons Paradoxon) (10 Punkte)

1300 Patienten, die an ein und derselben Krankheit litten, wurden entweder mit Medikament A oder Medikament B behandelt. Die folgende Tabelle nennt die Gruppengrößen, aufgeschlüsselt nach Geschlecht, Medikament und Erfolg der Behandlung.

	Männer		Frauen	
Medikament	nicht geheilt	geheilt	nicht geheilt	geheilt
A	50	50	90	10
B	40	60	880	120

Wir wählen aus der Menge aller Patienten zufällig eine Person 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.
A = Die Person hat Medikament A bekommen.
$B = A^c$ = Die Person hat Medikament B bekommen.
G = Die Person wurde geheilt.

a) Berechnen Sie die folgenden (bedingten) Wahrscheinlichkeiten und interpretieren Sie diese in Worten. (Bsp.: $P[A|G] = 60/240 = 0{,}25$ bedeutet: "25% der Geheilten haben Medikament A bekommen.")
   (i) $P[A]$ und $P[B]$ sowie
   (ii) $P[G|A]$ und $P[G|B]$.
   Welches Medikament scheint damit eine höhere Erfolgsquote zu haben?

b) Berechnen Sie die folgenden Wahrscheinlichkeiten und interpretieren Sie diese in Worten.
   (i) $P[G|A \cap M]$ und $P[G|B \cap M]$ sowie
   (ii) $P[G|A \cap W]$ und $P[G|B \cap W]$.
   Welches Medikament scheint damit eine höhere Erfolgsquote zu haben?

c) Berechnen Sie
   (i) $P[G|M]$ und $P[G|W]$ sowie
   (ii) $P[M|B]$ und $P[M|A]$,
   um den (scheinbaren) Widerspruch zwischen (a) und (b) aufzulösen.

Aufgabe 26 (20 Zusatzpunkte)

Erreichen Sie bis spätestens 30.06.13 auf `www.khanacademy.org` *Proficiency* in den *Skills Independent probability*, *Permutations*, *Combinations*, *Permutations and combinations* und *Probability with permutations and combinations*.
Hinweise siehe Aufgabe 3 (Blatt 1).

Aufgabe 27 (χ^2-Test für Kontingenztafeln)[5] (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 25 gestellte Frage, ob eines der Medikamente A oder B besser ist als das andere, quantitativ zu beantworten. Wir gehen dabei nach Geschlechtern getrennt vor und betrachten zunächst die Frauen.

Frauen	Nicht geheilt	Geheilt	Total
A	90	10	100
B	880	120	1000
Total	970	130	1100

Wir wollen im folgenden die Nullhypothese H_0: *Der Behandlungserfolg ist unabhängig von der Wahl des Medikamentes* auf dem Signifikanz-Niveau $\alpha = 5\%$ gegen die Alternativhypothese H_A: *Der Behandlungserfolg hängt von der Wahl des Medikamentes ab* mit einem χ^2-Test testen.

Falls H_0 gelten würde, so würde man erwarten, dass

$$\frac{100}{1100} \cdot \frac{970}{1100} \cdot 1100 = 88{,}18$$

der mit A behandelten Frauen nicht geheilt würden. (In Wirklichkeit waren es 90.) Mit analogen Rechnungen für die anderen 3 Kategorien erhält man als unter H_0 erwartete Zahlen

Frauen	Nicht geheilt	Geheilt	Total
A	88,18	11,82	100
B	881,82	118,18	1000
Total	970	130	1100

Als Teststatistik nehmen wir wieder

$$\chi^2 = \sum_{\text{alle 4 Kategorien}} \frac{(\text{beobachtete Anzahl} - \text{erwartete Anzahl})^2}{\text{erwartete Anzahl}} = 0{,}349\,.$$

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 kleiner als die zum Niveau $\alpha = 5\%$ gehörige kritische Zahl $\nu + 2\sqrt{2\nu} = 3{,}83$, jenseits der H_0 verworfen wird. Also kann man aus den Zahlen nicht schließen, dass bei Frauen eines der beiden Medikamente besser wirkt als das andere.

a) Führen Sie einen entsprechenden χ^2-Test für die Männer durch.

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

Aufgabe 28 (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 $P[X \le k]$ 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 0,35 ist (d.h. die Wahrscheinlichkeit eines Fehlers 2. Art). HINWEIS: Dabei hilft `binocdf(0:16,16,.35)`

[5] 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.