

Mathematik II für Biologen

Klausur am 23.7.2009

Bitte schreiben Sie nicht mit Bleistift. Bitte beginnen Sie jede Aufgabe (nicht Teilaufgabe) auf einer neuen Seite. Zusätzliches Papier ist jederzeit verfügbar. **Zeigen Sie auch stets Ihren Rechenweg!** Es sind maximal 81 Punkte erreichbar, 70 Punkte $\hat{=}$ 100% ($\hat{=}$ Note 1,0), 50% $\hat{=}$ 35 Punkte sind hinreichend zum Bestehen ($\hat{=}$ Note 4,0).

Hilfsmittel: Zwei beidseitig handbeschriebene A4-Blätter, nicht internetfähiger Taschenrechner.
Bearbeitungszeit: 105 Minuten. **Viel Erfolg!**

Aufgabe 1

(2+2+4+1+4+1=14 Punkte)

Mit Hilfe von NMR-Aufnahmen versuchen Wissenschaftler vorherzusagen, für welche von zwei Optionen sich eine Testperson entscheiden wird*. Eine korrekte Vorhersage erreichen sie dabei in 36 von 50 Versuchen.

Sie möchten mit Hilfe eines statistischen Tests entscheiden, ob sie damit eine signifikant höhere Trefferquote erzielt haben, als wenn sie geraten hätten. Als Teststatistik X wählen Sie die Anzahl der richtigen Vorhersagen.

- Formulieren Sie die Nullhypothese H_0 und die Alternativhypothese H_A .
HINWEIS: Bezeichnen Sie dazu mit p die (unbekannte) Wahrscheinlichkeit für eine richtige Vorhersage. Überlegen Sie sich auch, ob Sie einen einseitigen oder einen zweiseitigen Test durchführen sollten.
- Geben Sie die Verteilung der Teststatistik unter H_0 an.
- Bestimmen Sie den Verwerfungsbereich K auf dem Signifikanzniveau $\alpha = 2\%$. Dabei hilft die Ausgabe des folgenden MATLAB-Befehls: `>> binocdf(31:50,50,.5)`

```
ans =  
Columns 1 through 10  
0.9675 0.9836 0.9923 0.9967 0.9987 0.9995 0.9998 1.0000 1.0000 1.0000  
Columns 11 through 20  
1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
```
- Was ist der beobachtete Wert der Teststatistik?
- Berechnen Sie den p-Wert für diesen Test. Dabei hilft wieder die MATLAB-Ausgabe aus Teil (c).
- Wie entscheidet der Test auf dem Signifikanzniveau $\alpha = 2\%$?

Aufgabe 2

(6+6+14+3=29 Punkte)

- Sei die Zufallsvariable X das Ergebnis eines Wurfs mit einem fairen 20seitigen Würfel, d.h. X nimmt jeden der Werte $1, 2, \dots, 20$ mit gleicher Wahrscheinlichkeit an. Berechnen Sie
 - den Erwartungswert $E[X]$,
 - die Varianz $\text{Var}(X)$ und
 - den Median med_X .



- Sie würfeln 8mal mit einem realen 20seitigen Würfel und erhalten die Ergebnisse

$$x = \{5, 20, 14, 18, 15, 12, 10, 18\}.$$

Berechnen Sie für diese Stichprobe

- den Median und
- die Quartilsdifferenz.

*vgl. J.-D. Haynes et. al., *Reading Hidden Intentions in the Human Brain*, Current Biology **17** (2007) 323-328; Datensatz dort größer.

- c) Nun würfeln Sie 100mal mit einem realen 20seitigen Würfel. Die Häufigkeiten der Würfelergbnisse X sind in der folgenden Tabelle notiert:

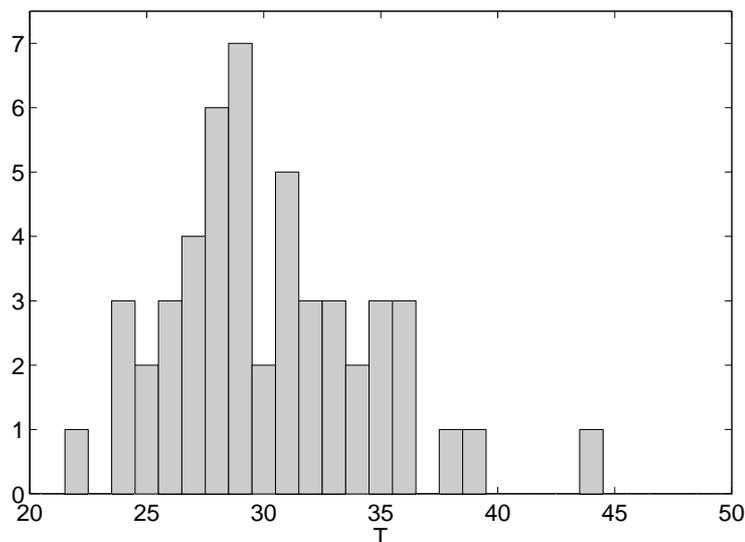
X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
#	8	2	8	4	1	9	2	2	3	7	6	4	2	8	2	9	7	1	7	8

Untersuchen Sie – mithilfe statistischer Tests auf dem Signifikanzniveau $\alpha = 5\%$ – die Frage, ob der Würfel fair ist.

Vielleicht helfen Ihnen dabei die folgenden MATLAB-Zeilen:

```
>> anzahl = [8 2 8 4 1 9 2 2 3 7 6 4 2 8 2 9 7 1 7 8]
>> sum((anzahl-5).^2)
ans = 164
>> sum(anzahl(1:10))
ans = 46
```

- (i) Wie lautet die Nullhypothese? (1 Punkt)
- (ii) Wie lautet die Alternativhypothese? (1 Punkt)
- (iii) Führen Sie dazu einen χ^2 -Test durch:
 Wie groß ist die erwartete Anzahl in jeder der 20 Klassen? (1 Punkt)
 Wie lautet die Formel für die Teststatistik? Berechnen Sie Ihren Wert. (2 Punkte)
 Bestimmen Sie das Verwerfungskriterium (gemäß Faustregel). (1 Punkt)
 Wie entscheidet der Test? (1 Punkt)
- (iv) Führen Sie nun auch einen Vorzeichen-Test durch,
 d.h. wählen Sie als Teststatistik $T := \#\{\text{Würfelergbnisse} \leq 10\}$.
 Geben Sie den beobachteten Wert der Teststatistik an. (1 Punkte)
 Bestimmen Sie den Verwerfungsbereich (gemäß Faustregel[†]). (3 Punkte)
 Wie entscheidet der Test? (1 Punkt)
- (v) Würden Sie nun dem χ^2 -Test oder dem Vorzeichen-Test glauben?
 Begründen Sie Ihre Antwort in *maximal zwei Sätzen*. (2 Punkte)
- d) Betrachten Sie nun einen Würfel, von dem Sie bereits wissen, dass er nicht fair ist. Für diesen wurde das Experiment aus Teil (c) (100maliger Wurf) 50mal durchgeführt und jeweils $T := \#\{\text{Würfelergbnisse} \leq 10\}$ notiert. Die dabei beobachteten Werte für T sind im folgenden Histogramm dargestellt.



Wie hoch schätzen Sie, darauf basierend, für diesen Würfel die Macht des Vorzeichentests aus Teil (c)?

[†]Verwenden Sie dabei $\Phi(1,96) \approx 0,975$ und *nicht* $\Phi(2) \approx 0,975$, wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist.

Aufgabe 3

(2+2=4 Punkte)

Das Verwerfungskriterium für einen beidseitigen t-Test mit 17 Freiheitsgraden auf dem Signifikanzniveau $\alpha = 4\%$ lautet

$$|T| > 2,22.$$

Das Verwerfungskriterium für einen rechtsseitigen t-Test mit 17 Freiheitsgraden auf dem Signifikanzniveau $\alpha = 4\%$ lautet

$$T > 1,86.$$

Geben Sie jeweils das Verwerfungskriterium für einen

- a) beidseitigen t-Test auf Signifikanzniveau $\alpha = 8\%$,
- b) linksseitigen t-Test auf Signifikanzniveau $\alpha = 2\%$,

mit 17 Freiheitsgraden an.

Aufgabe 4

(2+4+2+1+4=13 Punkte)

Ihr örtlicher Pizza-Service wirbt damit, dass seine durchschnittliche Lieferzeit höchstens 30min beträgt. Bei Ihren letzten 5 Bestellungen betragen die Wartezeiten bis zur Lieferung (in min)

28, 37, 29, 39 und 32.

Aufgrund dieser Daten möchten Sie die Werbeaussage des Pizza-Service' mit einem t-Test auf dem Signifikanzniveau $\alpha = 10\%$ widerlegen.

- a) Formulieren Sie die Nullhypothese H_0 und die Alternativhypothese H_A .
HINWEIS: Bezeichnen Sie dazu mit μ den (unbekannten) Erwartungswert der Lieferzeitenverteilung. Überlegen Sie sich auch, ob Sie einen einseitigen oder einen zweiseitigen Test durchführen sollten.
- b) Wie lautet die Formel für die Teststatistik? Berechnen Sie ihren Wert.
HINWEIS: Es ist sinnvoll, zunächst den Mittelwert \bar{x} und die empirische Standardabweichung s_x zu berechnen und zu notieren.
- c) Bestimmen Sie das Verwerfungskriterium.[‡]
- d) Wie entscheidet der Test?
- e) Bestimmen Sie außerdem ein 90%-Vertrauensintervall für μ (im Sinne dieses Tests).

[‡]Eine hilfreiche Tabelle finden Sie auf der letzten Seite.

Aufgabe 5

(2+1+4+2+2=11 Punkte)

Kuh Thekla und Kuh Elma waren beide an Weidefieber erkrankt. Beide wurden mit PASTOFEBRIL behandelt, und beide wurden daraufhin wieder gesund.

Wir bezeichnen mit p die (unbekannte) Wahrscheinlichkeit dafür, dass eine, an Weidefieber erkrankte und mit PASTOFEBRIL behandelte, Kuh gesund wird. Um abzuschätzen, wie groß p mindestens ist, bestimmen wir ein einseitiges 95%-Vertrauensintervall für p . Gehen Sie dazu wie folgt vor:

- Testen Sie $H_0 : p = p_0$ gegen $H_A : p > p_0$.
 - Wählen Sie als Teststatistik X die Anzahl der wieder gesund gewordenen Kühe in unserer (zugegebenermaßen kleinen) Stichprobe.
- a) Wie ist X unter H_0 verteilt?
 - b) Was ist der beobachtete Wert der Teststatistik?
 - c) Bestimmen Sie den p-Wert des Tests.
 - d) Für welche p_0 ist der p-Wert $\leq 5\%$?
 - e) Geben Sie das gesuchte 95%-Vertrauensintervall für p an.

Aufgabe 6

(2+3+5=10 Punkte)

Im einem Kino wird ein Film gezeigt, der ab 18 Jahren freigegeben ist. Bei 10% der Besucher wird das Alter kontrolliert. Die Wahrscheinlichkeit, dass ein kontrollierter Besucher unter 18 ist, beträgt 80%. Die Wahrscheinlichkeit, dass ein nicht kontrollierter Besucher unter 18 ist, beträgt 40%. Wir definieren die Ereignisse

$$U := \text{“Besucher ist unter 18”} \quad \text{und} \quad K := \text{“Besucher wird kontrolliert”} .$$

- a) Drücken Sie die Bedeutung der (bedingten) Wahrscheinlichkeit $P[U^C|K]$ in Worten aus.
- b) Drücken Sie die drei im Aufgabentext bezifferten Wahrscheinlichkeiten mithilfe der mathematischen Notation ähnlich Aufgabenteil (a) aus.
- c) Berechnen Sie die Wahrscheinlichkeit dafür, dass ein Besucher, der unter 18 Jahre alt ist, kontrolliert wird. Drücken Sie hierzu die gesuchte Wahrscheinlichkeit ebenfalls mithilfe der Notation aus den Teilen (a) und (b) aus und berechnen Sie sie.

TABELLE

Einige Quantile $t_{d_f, q}$ der Studentschen t-Verteilung für d_f Freiheitsgrade:

$d_f \setminus q$	0,90	0,95
4	1,533	2,132
5	1,476	2,015
6	1,440	1,943