

Mathematik II für Biologen

Übungsblatt 10 (Abgabe am 25.06.2008)

Aufgabe 34 (Abi 1985, Ba-Wü)

(20 Zusatzpunkte)

- a) Fünf ideale Würfel werden gleichzeitig geworfen. Berechnen Sie die Wahrscheinlichkeiten der Ereignisse,
A: Die Würfel zeigen die gleiche Augenzahl,
B: Alle Würfel zeigen verschiedene Augenzahlen,
C: Die Würfel zeigen entweder die Augenzahlen 1, 2, 3, 4, 5 oder 2, 3, 4, 5, 6.
- b) Ein idealer Würfel wird 10mal geworfen. X beschreibe die Anzahl der dabei auftretenden Sechsen. Berechnen Sie den Erwartungswert $E[X]$ und die Varianz $\text{Var}(X)$. Ermitteln Sie diejenigen Werte von X , welche im Intervall $[E[X] - 3\sigma(X), E[X] + 3\sigma(X)]$ liegen. ($\sigma(X) = \sqrt{\text{Var}(X)}$)
Der Würfel wird n mal geworfen. Y beschreibe die Anzahl der dabei auftretenden Sechsen. Wie groß muß n mindestens sein, damit das Intervall $[E[X] - 3\sigma(X), E[X] + 3\sigma(X)]$ im Intervall $[0, n]$ enthalten ist?
- c) Um zu untersuchen, ob bei einem gegebenen Würfel die Augenzahl 6 mit der Wahrscheinlichkeit $\frac{1}{6}$ auftritt, wird dieser 300mal geworfen und die Anzahl der Sechsen gezählt.
(i) Wie lauten die Hypothesen?
(ii) Welche Zufallsvariable wird bei diesem Test betrachtet?
(iii) Wie ist diese Zufallsvariable bei zutreffender Nullhypothese verteilt?
(iv) Bestimmen Sie den Verwerfungsbereich für das Signifikanzniveau 5%. (Sie dürfen dazu die Verteilung der Teststatistik durch eine Normalverteilung annähern.)
(v) Wie wird aufgrund einer Stichprobe vom Umfang 300 entschieden, wenn 60 Sechsen auftreten?
(vi) Welcher Fehler kann dabei begangen werden?
(vii) Wie groß ist dieser Fehler, wenn die Wahrscheinlichkeit für Sechsen 0,2 beträgt?
- d) Ein idealer Würfel wird 450mal geworfen. Dabei beschreibe die Zufallsvariable Z die relative Häufigkeit der Augenzahl 6.
Berechnen Sie näherungsweise den zu $E[Z]$ symmetrischen Bereich, in den die relative Häufigkeit der Augenzahl 6 mit 95%iger Wahrscheinlichkeit fällt.
Wie oft muss man werfen, damit mit einer Wahrscheinlichkeit von mindestens 0,90 die relative Häufigkeit der Augenzahl 6 um höchstens 0,05 von $\frac{1}{6}$ abweicht?

Aufgabe 35 MATLAB

(10 Punkte)

Die Datei `fishy.dat` enthält in der ersten Spalte die Länge (in cm), in der zweiten Spalte das Gewicht (Einheit leider unbekannt) und in der dritten Spalte den DDT-Gehalt (in ppm) von $n = 96$ Welsen, die im Tennessee River in Alabama, USA, gefangen wurden (Quelle: Mendenhall, Sincich: *Statistics for Engineering and the Sciences*, Prentice Hall. Appendix III. Daten leicht geändert, daher das `y` im Dateinamen.).

- a) Bestimmen Sie den empirischen Median der DDT-Gehalte.
b) Ist dieser empirische Median signifikant (zum Signifikanz-Niveau $\alpha = 5\%$) von 10 verschieden? Beantworten Sie diese Frage mit einem zweiseitigen Vorzeichentest.
c) Bestimmen Sie das zum Vorzeichentest gehörige 95%-Vertrauensintervall $[a, b]$ für den "wahren, theoretischen" Median des DDT-Gehaltes eines solchen Fisches, indem Sie den Test aus Aufgabe (b) für verschiedene Werte wiederholen und dabei beobachten, ob der Test verwirft. Bestimmen Sie dabei a und b so genau, dass die erste Ziffer nach dem Komma sicher stimmt.

Unvollständiger MATLAB-Code:

```
load fishy.dat
ddt=fishy(:,3);
[p,h]=signtest(ddt,10) % help signtest
```

Aufgabe 36

(10 Punkte)

In einem Versuch ähnlich zu dem in Aufgabe 29 beschriebenen ergaben sich auf 23 Petri-Schalen folgende Anzahlen x_1, x_2, \dots, x_{23} von insgesamt $x_1 + \dots + x_{23} = 572$ Bakterienkolonien:

31 29 19 18 31 27 34 27 34 30 16 18 26 27 27 18 24 22 28 24 21 17 24

Wir nehmen an, dass die Anzahlen X_1, \dots, X_{23} von Bakterienkolonien unabhängig und poisson-verteilt mit einem gemeinsamen, aber unbekanntem Parameter λ sind. (Beachten Sie: X_1 ist eine Zufallsvariable. Sie bezeichnet sozusagen die Anzahl in der ersten Petri-Schale vor der Messung, wenn die tatsächliche Anzahl $x_1 = 31$ noch unbekannt ist.) Kann $\lambda = 22$ sein oder deuten obige Daten darauf hin, dass $\lambda \neq 22$?

- a) Führen Sie einen entsprechenden Test zum Signifikanz-Niveau $\alpha = 5\%$ für den Parameter λ durch. Geben Sie dabei an:
- (i) die Nullhypothese H_0 ,
 - (ii) die Alternativhypothese H_A ,
 - (iii) die Teststatistik X ,
 - (iv) deren Verteilung, falls H_0 gilt,
 - (v) den beobachteten Wert von X ,
 - (vi) das Verwerfungskriterium (gemäß Faustregel)
 - (vii) und die Testentscheidung sowie zusätzlich
 - (viii) ein 95%-Vertrauensintervall für λ (gemäß Faustregel).
- b) Angenommen, Sie hätten folgende Größen für obigen Test mit den angegebenen Daten korrekt bestimmt
- (A) den Verwerfungsbereich K ,
 - (B) das 95%-Vertrauensintervall für λ ,
 - (C) die Testentscheidung,
 - (D) den p-Wert,
 - (E) die Wahrscheinlichkeit eines Fehlers 1. Art,
 - (F) die Wahrscheinlichkeit eines Fehlers 2. Art bei genau spezifizierter Alternativhypothese $\lambda = \lambda_0$,
 - (G) die Verteilung der Teststatistik, falls H_0 gilt,
 - (H) den beobachteten Wert der Teststatistik,
- und wollten nun den gleichen Test mit einem anderen Datensatz der Größe 23 wiederholen, welche der Angaben (A) bis (H) bräuchten Sie dann nicht noch einmal zu bestimmen und welche müssten Sie neu bestimmen? Antworten Sie für (A) bis (H) jeweils ohne Begründung. (Richtige Antwort: 0.5 Punkte, Falsche Antwort: -0.5 Punkte, Keine Antwort: 0 Punkte. Eine ggf. negative Gesamtpunktzahl wird mit Teilaufgabe (a) verrechnet.)