

## Mathematik II für Biologen

Klausur am 29.7.2013

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 87 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+2 = 6 Punkte)

Bei einer Schiffsexpedition wurden an 31 aufeinanderfolgenden Tagen die Anzahlen  $x_1, x_2, \dots, x_{31}$  der beobachteten Delfine aufgezeichnet. Zum Beispiel bedeutet  $x_{13} = 2$ , dass am 13. Tag genau zwei Delfine gesichtet wurden. Dabei wurde an 9 Tagen gar kein Delfin gesehen, an 11 Tagen genau einer, an 7 Tagen genau zwei, an 3 Tagen genau drei und an einem Tag sogar vier Delfine.

- Geben Sie  $x_{(13)}$  und  $x_{(26)}$  an. Dabei ist  $x_{(i)}$  wie immer der  $i$ -te Wert der geordneten Stichprobe.
- Bestimmen Sie den Median, und
- berechnen Sie den arithmetischen Mittelwert der Stichprobe  $\{x_1, x_2, \dots, x_{31}\}$ .

### Aufgabe 2 (4+2 = 6 Punkte)

Gegeben sei die Stichprobe  $\{x_i\}$  vom Umfang  $n = 5$  mit

$i$	1	2	3	4	5
$x_i$	1	2	-1	4	2

- Zeichnen Sie die empirische Verteilungsfunktion

$$F(x) = \frac{\#\{x_i : x_i \leq x\}}{n}$$

für  $-2 \leq x \leq 5$ .

- Bestimmen Sie das untere Quartil (d.h. das 25%-Quantil) der Stichprobe.

### Aufgabe 3 (2+4 = 6 Punkte)

Wir betrachten eine Stichprobe bestehend aus den Wertepaaren  $(x_i, y_i)$ ,  $i = 1, 2, 3$ , mit

$i$	1	2	3
$x_i$	3	5	7
$y_i$	2	1	6

- Bestimmen Sie  $\bar{x}$  und  $\bar{y}$ , den arithmetischen Mittelwert der  $x$ -Werte und den der  $y$ -Werte.
- Berechnen Sie den Korrelationskoeffizienten  $r_{xy}$  nach Pearson,

$$r_{xy} = \frac{\sum_{i=1}^3 (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{j=1}^3 (x_j - \bar{x})^2} \sqrt{\sum_{k=1}^3 (y_k - \bar{y})^2}}.$$

Dazu ist es hilfreich, zunächst eine Tabelle der Form

$i$	1	2	3
$x_i - \bar{x}$			
$y_i - \bar{y}$			

anzufertigen. (Ihr Rechenweg muss – wie überall – nachvollziehbar sein!)

#### Aufgabe 4

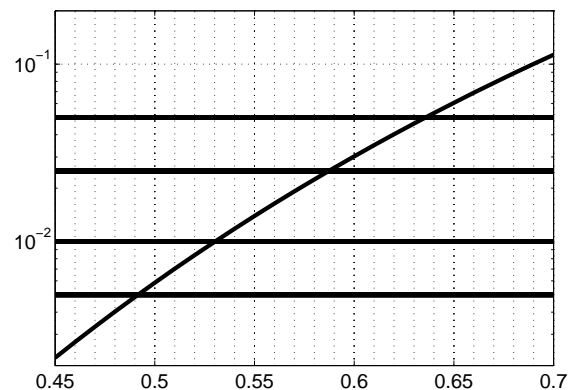
(4+2+3+1+4+1+1+4 = 20 Punkte)

In einem Experiment wird untersucht, ob Menschenaffen über ein Langzeitgedächtnis verfügen.\* Dazu werden 11 Affen in einem Raum Werkzeuge gezeigt, die sie zur Bewältigung einer Aufgabe benötigen. Danach werden die Werkzeuge in Kisten verstaut. Dabei schauen die Affen zu. Drei Jahre später führt man die Affen (einzeln) wieder in diesen Raum und fordert sie auf, die Aufgabe zu lösen. Innerhalb einer vorgegebenen Zeit suchen 10 Affen in den Kisten nach den Werkzeugen, ein Affe sieht dort nicht nach.

Wir nehmen an, dass ein Affe, der sich nicht erinnert, wo die Werkzeuge versteckt wurden, den Raum zufällig erkundet und mit Wahrscheinlichkeit  $1/2$  auch in den Kisten nachschaut. Wir untersuchen mithilfe eines Hypothesentests, ob ein Affe, der beobachtet hat, wie die Werkzeuge in den Kisten verstaut wurden, mit einer größeren Wahrscheinlichkeit dort nachsieht, und interpretieren dies ggf. als Hinweis darauf, dass sich der Affe an das drei Jahre zurückliegende Ereignis erinnert. Bezeichnen Sie mit  $p$  die Wahrscheinlichkeit dafür, dass ein Affe in den Kisten nach den Werkzeugen sucht.

- Formulieren Sie die Nullhypothese  $H_0$  und die Alternativhypothese  $H_A$ . Geben Sie jeweils sowohl eine Formulierung in Worten an, die eine Aussage über das Langzeitgedächtnis von Affen macht, als auch eine Formulierung mittels einer Gleichung oder Ungleichung, die die Wahrscheinlichkeit  $p$  enthält.
- Wählen Sie eine sinnvolle Teststatistik  $X$ , deren beobachteter Wert sich aus dem Aufgabentext bestimmen lässt; beschreiben Sie  $X$  in einem Satz.
- Wie ist  $X$  unter  $H_0$  verteilt?
- Welcher Wert von  $X$  wurde beobachtet?
- Berechnen Sie den p-Wert für diesen Test.
- Wird die Nullhypothese auf dem Signifikanzniveau  $\alpha = 5\%$  verworfen?
- Was folgern Sie aus diesem Testergebnis über das Langzeitgedächtnis von Menschenaffen (ein Satz)?
- Bestimmen Sie nun auch ein 99%-Vertrauensintervall für  $p$  im Sinne des obigen Tests. Dabei hilft der folgende MATLAB-Code und -Plot.

```
p=.4:.01:.7;
semilogy(p,1-binocdf(9,11,p))
hold on
semilogy(p,.05*ones(length(p)))
semilogy(p,.025*ones(length(p)))
semilogy(p,.01*ones(length(p)))
semilogy(p,.005*ones(length(p)))
hold off
```



\*vgl. Martin-Ordas et al., *Memory for Distant Past Events in Chimpanzees and Orangutans*, Current Biology (2013), <http://dx.doi.org/10.1016/j.cub.2013.06.017>

**Aufgabe 5**

(2+3+6+3+2+1+1 = 18 Punkte)

Gnurpen essen entweder Xarg oder Yrb, aber nie beides gleichzeitig. An einem Tag beobachten Sie morgens 47 Gnurpen, die Xarg essen und 23, die Yrb essen, mittags 14 die Xarg und 26 die Yrb essen, und abends sehen Sie nochmals 29 Gnurpen dabei zu, wie sie Xarg essen während 11 Yrb bevorzugen. Wir untersuchen mit einem  $\chi^2$ -Test, ob die Nahrungsauswahl einer Gnurpe (Xarg oder Yrb) von der Tageszeit (morgens, mittags, abends) abhängt.

- Geben Sie die Nullhypothese  $H_0$  und die Alternativhypothese  $H_A$  an.
- Erstellen Sie die Kontingenztafel der beobachteten Werte.
- Bestimmen Sie die Kontingenztafel der unter  $H_0$  erwarteten Werte.
- Berechnen Sie den Wert der Teststatistik. (Wie immer mit Rechenweg!)
- Wie lautet das Verwerfungskriterium auf dem Signifikanzniveau  $\alpha = 5\%$ ? Wie groß ist dabei die Anzahl der Freiheitsgrade?
- Wird die Nullhypothese auf dem Signifikanzniveau  $\alpha = 5\%$  verworfen?
- Hängt die Nahrungsauswahl von Gnurpen laut Testergebnis von der Tageszeit ab (ein Satz)?

**Aufgabe 6 (Verwenden Sie die Daten aus Aufgabe 5)**

(1+1+1+1+1+2 = 7 Punkte)

Von den 150 beobachteten Nahrungsaufnahmen der Gnurpen aus Aufgabe 5 wählen Sie zufällig eine aus (jede mit der gleichen Wahrscheinlichkeit). Wir definieren die folgenden Ereignisse:

$A$  := "Die ausgewählte Nahrungsaufnahme fand am Abend statt."

$X$  := "Bei der ausgewählten Nahrungsaufnahme wurde Xarg gegessen."

- Beschreiben Sie in einem Satz das Ereignis  $A^C$ .

Bestimmen Sie die folgenden (bedingten) Wahrscheinlichkeiten:

- $P[A|X]$
- $P[A|X^C]$
- $P[A \cap X]$
- $P[A]$
- $P[X|A^C]$

**Aufgabe 7**

(3+3 = 6 Punkte)

Ihnen liegt eine Stichprobe mit Stichprobenmittel  $\bar{x} = 9,7$  und Stichprobenvarianz  $s_x^2 = 9,0$  vor. Mithilfe eines zweiseitigen z-Tests auf Signifikanzniveau  $\alpha = 5\%$  möchten Sie Informationen über den Erwartungswert  $\mu$  der zugrundeliegenden Verteilung gewinnen. Wir nehmen an, dass für die Varianz  $\sigma^2$  dieser Verteilung  $\sigma^2 = s_x^2$  gilt.

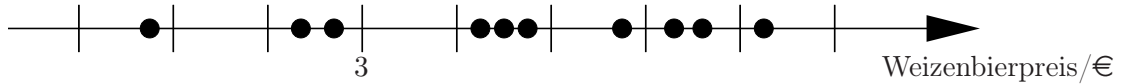
- Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit  $\mu = 9,0$  verworfen wird?
- Für welche Stichprobenumfänge  $n$  liegt  $\mu = 11,2$  im 95%-Vertrauensintervall für den Erwartungswert?

BEMERKUNG: Rechnen Sie mit  $\Phi(1,96) = 0,975$ , wobei  $\Phi$  die Verteilungsfunktion der Standardnormalverteilung bezeichnet.

### Aufgabe 8

(2+6+2+1+1+2+2+1+1 = 18 Punkte)

Eine Kommilitonin behauptet, dass ein Weizenbier in Tübinger Gaststätten im Schnitt 3€ kostet. Sie vermuten aber, dass der Median  $\mu$  der Verteilung der Weizenbierpreise größer ist. Um dies zu überprüfen, ermitteln Sie die Preise in einigen Gaststätten, stellen die Werte im folgenden (linearen, unverzerrten) Diagramm durch Punkte (•) dar,



und führen einen Wilcoxon-Test zum Signifikanzniveau  $\alpha = 5\%$  durch:

- Geben Sie die Nullhypothese  $H_0$  und Alternativhypothese  $H_A$  an.
- Bestimmen Sie den Wert der Teststatistik.  
Notieren Sie dabei als Zwischenschritt auch die Ränge aller positiven und negativen Abweichung.
- Bestimmen Sie das Verwerfungskriterium.<sup>†</sup>
- Geben Sie die Testentscheidung an.

Zum Vergleich führen Sie mit denselben Daten auch einen Vorzeichenstest (gleiche Hypothesen wie oben) durch:

- Beschreiben Sie in einem Satz die Teststatistik  $X$ .
- Wie ist die Teststatistik unter  $H_0$  verteilt?
- Bestimmen Sie den Verwerfungsbereich zum Signifikanzniveau  $\alpha = 5\%$ . Dabei hilft der folgende MATLAB-Code mit Ausgabe.

```
>> binopdf(0:10,10,.5)
ans =
Columns 1 through 6
    0.0010    0.0098    0.0439    0.1172    0.2051    0.2461
Columns 7 through 11
    0.2051    0.1172    0.0439    0.0098    0.0010

>> binocdf(0:10,10,.5)
ans =
Columns 1 through 6
    0.0010    0.0107    0.0547    0.1719    0.3770    0.6230
Columns 7 through 11
    0.8281    0.9453    0.9893    0.9990    1.0000
```

- Ermitteln Sie  $X$  aus den Daten.
- Wie entscheidet der Vorzeichen-Test?

<sup>†</sup>Der Wilcoxon-Test auf Signifikanzniveau  $\alpha = 5\%$  verwirft für Stichprobenumfang  $n$ , falls  $\min(U^+, U^-) \leq U_{\text{krit}}$  (zweiseitig) oder falls  $U^- \leq U_{\text{krit}}$  bzw.  $U^+ \leq U_{\text{krit}}$  (rechts- bzw. linksseitig). Dabei bezeichnen  $U^+$  und  $U^-$  die Rangsummen der positiven und negativen Abweichungen.

$n$	7	8	9	10	11	12	13
$U_{\text{krit}}$ einseitig	3	5	8	10	13	17	21
$U_{\text{krit}}$ zweiseitig	2	3	5	8	10	13	17