

## Mathematik II für Biologen

Klausur am 21.07.2014

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 86 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+1+3+2+4+1+4 = 17 Punkte)

Eine Kommilitonin behauptet, zu Boden gefallene Reißnägel würden mit Wahrscheinlichkeit  $\frac{1}{2}$  mit der Spitze nach oben liegen bleiben. Sie vermuten, dass die Wahrscheinlichkeit  $w$  für "Spitze nach oben" *geringer* ist. Um die Frage zu untersuchen, werfen Sie 200 Reißnägel auf den Boden, zählen die Anzahl der Reißnägel, die mit der Spitze nach oben liegen bleiben und führen einen Hypothesentest durch.

- Geben Sie die Nullhypothese  $H_0$  und die Alternativhypothese  $H_A$  an.
- Beschreiben Sie in einem Satz die Teststatistik  $X$ .
- Wie ist die Teststatistik unter  $H_0$  verteilt?
- Bestimmen Sie den Erwartungswert und die Varianz von  $X$  unter  $H_0$ .

Im Folgenden nehmen wir an, dass  $X$  (näherungsweise) normalverteilt ist (mit geeigneten Parametern).

- Bestimmen Sie den Verwerfungsbereich  $K$  zum Signifikanzniveau  $\alpha = 5\%$ .

In Ihrem Versuch bleiben 80 der 200 Nägel mit der Spitze nach oben liegen.

- Wie entscheidet der Test?
- Bestimmen Sie außerdem ein zweiseitiges 95%-Vertrauensintervall für die Wahrscheinlichkeit  $w$ .

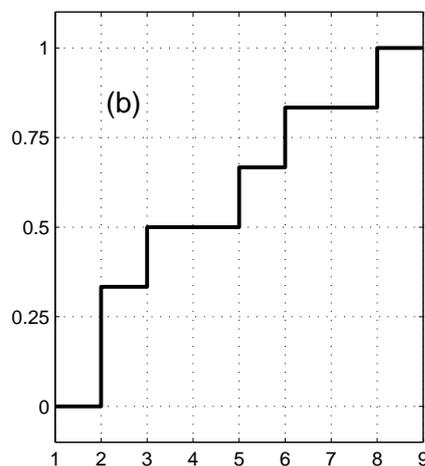
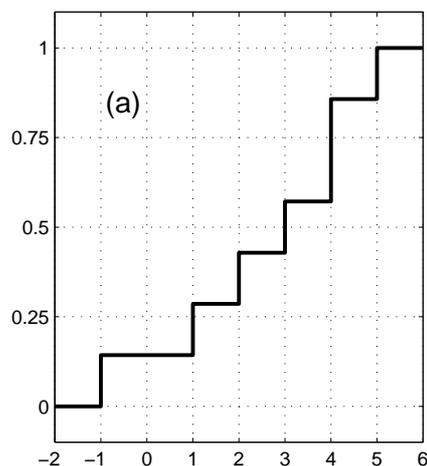
Einige Werte der Verteilungsfunktion der Standardnormalverteilung (MATLAB-Ausgabe, gerundet):

```
>> normcdf([1 1.64 1.96 3])  
ans = 0.84 0.95 0.975 0.999
```

### Aufgabe 2

(5+5 = 10 Punkte)

In den folgenden Diagrammen sind die empirischen Verteilungsfunktionen zweier Stichproben dargestellt. Bestimmen Sie jeweils den Median und die Quartilsdifferenz, und zeichnen Sie jeweils den zugehörigen Boxplot (maßstabsgetreu, mit Koordinatenachse daneben bzw. darunter).



### Aufgabe 3

(2+2+2+2+4 = 12 Punkte)

Bei einer dreiwöchigen Exkursion an die Steilküste von Nurgpirien notiert ein Forscherteam täglich die Anzahl  $x_1, x_2, \dots, x_n$  ( $n = 21$ ) der beobachteten Gnrurpen. Zum Beispiel bedeutet  $x_9 = 3$ , dass am neunten Tag genau drei Gnrurpen gesichtet wurden. Dabei wurde an 6 Tagen gar keine Gnrurpe gesehen, an 9 Tagen genau eine, an 3 Tagen genau drei, und an ebenfalls drei Tagen wurden sogar vier Gnrurpen gesehen.

- Geben Sie  $x_{(13)}$  und  $x_{(16)}$  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  $\{x_1, x_2, \dots, x_{21}\}$ .
- Berechnen Sie auch die Stichprobenvarianz.
- Zeichnen Sie die empirische Verteilungsfunktion  $F(x) = \frac{\#\{x_i : x_i \leq x\}}{n}$  für  $-1 \leq x \leq 5$ .

### Aufgabe 4

(2+1+3+2+4+2+4 = 18 Punkte)

*Verwenden Sie weiterhin die Angaben aus Aufgabe 3.*

Aufgrund vieler vorangegangener Expeditionen hatten die Forscher erwartet, im Mittel täglich 2 Gnrurpen zu sehen. Wir testen mit einem beidseitigen Vorzeichentest, ob die Beobachtungen der jüngsten Expedition mit der Annahme vereinbar sind, dass der Median der Verteilung der Anzahl der täglich gesichteten Gnrurpen 2 beträgt:

- Geben Sie die Nullhypothese  $H_0$  und die Alternativhypothese  $H_A$  an.
- Beschreiben Sie in einem Satz die Teststatistik  $X$ .
- Wie ist die Teststatistik unter  $H_0$  verteilt?
- Ermitteln Sie den Wert der Teststatistik aus den Daten.
- Bestimmen Sie den p-Wert des Tests. Dabei hilft der folgende MATLAB-Code mit Ausgabe.

```
>> binocdf(0:21,21,1/2)
ans =
Columns 1 through 8
    0.0000    0.0000    0.0001    0.0007    0.0036    0.0133    0.0392    0.0946
Columns 9 through 16
    0.1917    0.3318    0.5000    0.6682    0.8083    0.9054    0.9608    0.9867
Columns 17 through 22
    0.9964    0.9993    0.9999    1.0000    1.0000    1.0000
```

- Wie entscheidet der Vorzeichentest auf dem Signifikanzniveau  $\alpha = 5\%$ ? Formulieren Sie das Testergebnis in einem Satz.
- Bestimmen Sie nun auch ein beidseitiges 95%-Vertrauensintervall im Sinne eines t-Tests für die zu erwartende Anzahl der täglichen Gnrurpensichtungen. Verwirft ein beidseitiger t-Test auf dem Signifikanzniveau 5% die Nullhypothese, dass die zu erwartende Anzahl 2 ist? Hier hilft die folgende MATLAB-Ausgabe.

```
>> tinv(0.95,19:22)
ans =    1.7291    1.7247    1.7207    1.7171
```

```
>> tinv(0.975,19:22)
ans =    2.0930    2.0860    2.0796    2.0739
```

HINWEIS: Der Befehl `tinv( $\alpha, d$ )` berechnet das  $\alpha$ -Quantil  $q_\alpha$  der Studentischen t-Verteilung mit  $d$  Freiheitsgraden. Notieren Sie, welcher dieser Werte für Sie wichtig ist, und warum.

**Aufgabe 5**

(5+5+6 = 16 Punkte)

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

Studiengang	männlich		weiblich		gesamt	
	Bewerber	Zugelassene	Bewerberinnen	Zugelassene	Bewerb.	Zugel.
A	825	511	108	89	933	600
B	373	22	341	24	714	46
gesamt	1198	533	449	113	1647	646

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:

$\varphi$  = Die Person ist weiblich.

$\sigma$  =  $\varphi^C$  = Die Person ist männlich.

$A$  = Die Person hat sich für den Studiengang A beworben.

$B = A^C$  = Die Person hat sich für den Studiengang B beworben.

$Z$  = Die Person wurde zum Studium zugelassen.

a) Berechnen Sie die folgenden Wahrscheinlichkeiten.

(i)  $P[\varphi]$  und  $P[\sigma]$ .

(ii)  $P[Z|\varphi]$  und  $P[Z|\sigma]$ .

Wer scheint bevorzugt zugelassen zu werden, Männer oder Frauen?

b) Berechnen Sie die folgenden Wahrscheinlichkeiten.

(i)  $P[Z|\varphi \cap A]$  und  $P[Z|\sigma \cap A]$ .

(ii)  $P[Z|\varphi \cap B]$  und  $P[Z|\sigma \cap B]$ .

Wer scheint bevorzugt zugelassen zu werden, Männer oder Frauen?

c) Berechnen Sie

(i)  $P[Z|A]$  und  $P[Z|B]$  sowie

(ii)  $P[A|\varphi]$  und  $P[A|\sigma]$ ,

und erklären Sie, wie sich damit der (scheinbare) Widerspruch zwischen (a) und (b) auflösen lässt.

**Aufgabe 6**

(2+1+4+3+1+1+1 = 13 Punkte)

Verwenden Sie die Zahlen aus der Tabelle in Aufgabe 5.

Die Zahlen in Aufgabe 5 suggerieren, dass Frauen sich bevorzugt für Studiengang B bewerben, Männer dagegen lieber für Studiengang A. Wir untersuchen mit einem  $\chi^2$ -Test, ob der beobachtete Unterschied signifikant ist:

a) Formulieren Sie die Nullhypothese  $H_0$  und die Alternativhypothese  $H_A$ .

b) Geben Sie die Kontingenztafel der beobachteten Werte an.

c) Bestimmen Sie die Kontingenztafel der unter  $H_0$  erwarteten Werte. Runden Sie dabei die Einträge auf ganze Zahlen.

d) Berechnen Sie den Wert der Teststatistik. (Wie immer mit Rechenweg!)

e) Wie lautet das Verwerfungskriterium (laut Faustregel) auf dem Signifikanzniveau  $\alpha = 5\%$ ?

f) Wird die Nullhypothese auf dem Signifikanzniveau  $\alpha = 5\%$  verworfen?

g) Formulieren Sie das Testergebnis in einem Satz.

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