Universität Tübingen
Mathematisches Institut
Dr. Stefan Keppeler

Sommersemester 2015

Mathematik II für Biologen

Nachklausur am 05.10.2015

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 78 Punkte erreichbar, 66 Punkte $\widehat{=}$ 100% ($\widehat{=}$ Note 1,0), 50% $\widehat{=}$ 33 Punkte sind hinreichend zum Bestehen ($\widehat{=}$ Note 4,0).

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

Aufgabe 1 (4+3+3+1+1+3 = 15 Punkte)

Für eine Studie wurden in jedem Bundesland 100 Bürger gefragt, ob Sie schon einmal eine Affäre während einer festen Partnerschaft hatten. Die Anzahl derer, die die Frage bejahten (insgesamt 331), ist in der folgenden Tabelle für jedes Bundesland angegeben.

BW	BY	BE	BB	HB	HH	HE	MV	NI	NW	RP	SL	SN	ST	SH	TH
18	19	32	13	25	17	22	19	18	18	18	30	22	20	17	23

Im Bundesdurchschnitt antworteten damit 21% der Befragten mit ja. Mit Blick auf den hohen Wert für Berlin (BE) titelt eine Zeitschrift: *Berliner gehen besonders oft fremd!* Wir untersuchen mittels statistischer Tests, ob sich diese Schlagzeile mit den Daten der Studie rechtfertigen lässt.

Zunächst wählen wir die Anzahl der positiven Antworten in Berlin als Teststatistik X. Außerdem sei u der (unbekannte, wahre) Anteil der untreuen Berliner.

a) Geben Sie geeignete Hypothesen H_0 und H_A an.
Formulieren Sie diese sowohl in Worten als auch durch (Un-)Gleichungen.

b) Wie ist die Teststatistik unter H_0 verteilt?

c) Bestimmen Sie den Verwerfungsbereich zum Signifikanzniveau $\alpha = 5\%$ (laut Faustregel*).

d) Welcher Wert der Teststatistik wurde beobachtet?

e) Wie entscheidet der Test?

f) Bestimmen Sie außerdem ein 95%-Vertrauensintervall für u (laut Faustregel*).

Aufgabe 2 (2+4+2+1+1 = 10 Punkte)

Verwenden Sie weiterhin die Angaben aus Aufgabe 1.

Nun untersuchen wir mit einem χ^2-Test, ob es sein kann, dass der Anteil untreuer Bürger doch nicht vom Bundesland abhängt:

a) Geben Sie die Nullhypothese H_0 und die Alternativhypothese H_A an.

b) Bestimmen Sie den Wert der Teststatistik (wie immer mit nachvollziehbarem Rechenweg).

c) Bestimmen Sie den kritischen Wert χ^2_{krit} zum Signifikanzniveau $\alpha = 5\%$ laut Faustregel.

d) Wird die Nullhypothese verworfen?

e) Formulieren Sie das Testergebnis in einem Satz.

---

*Verwenden Sie $\Phi(1{,}64) \approx 95\%$, wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist.

Aufgabe 3 (4 Punkte)

Verwenden Sie weiterhin die Angaben aus Aufgabe 1.

Unabhängig von dem, was Sie in Aufgabe 1 & 2 herausgefunden haben, nehmen wir nun an, in Aufgabe 1 sei herausgekommen, dass der Anteil der Berliner, die die Frage bejaht haben (32%), signifikant vom Bundesdurchschnitt (21%) abweicht. Weiter nehmen wir an, in Aufgabe 2 sei herausgekommen, dass der Anteil untreuer Bürger nicht vom Bundesland abhängt. Lässt sich unter diesen Voraussetzungen die in Aufgabe 1 erwähnte Schlagzeile mithilfe der Studie rechtfertigen? Begründen Sie Ihre Antwort in maximal drei Sätzen.

Aufgabe 4 (5+2+2+2 = 11 Punkte)

Verwenden Sie weiterhin die Angaben aus Aufgabe 1.

Von den 1600 Befragten aus Aufgabe 1 werde eine Person zufällig ausgewählt (jede mit der gleichen Wahrscheinlichkeit). Wir definieren die folgenden Ereignisse:

$$J = \text{“Die Person hat die Frage bejaht.”}$$
$$B = \text{“Die Person ist aus Berlin (BE).”}$$
$$H = \text{“Die Person ist aus Hamburg (HH).”}$$

a) Bestimmen Sie $P[J]$, $P[B]$, $P[H]$, $P[J|B]$ und $P[H|J]$.

Drücken Sie die folgenden (ggf. bedingten) Wahrscheinlichkeiten in der Form $P[\ldots]$ aus, und bestimmen Sie sie.

b) Die Wahrscheinlichkeit dafür, dass die Person die Frage verneint hat.
c) Die Wahrscheinlichkeit dafür, dass die Person die Frage bejaht hat und aus Berlin ist.
d) Die Wahrscheinlichkeit dafür, dass eine Person, die die Frage bejaht hat, aus Berlin ist.

Aufgabe 5 (1+2+2+5 = 10 Punkte)

Verwenden Sie weiterhin die Angaben aus Aufgabe 1.

Betrachten Sie die 16 Zahlenwerte aus der Tabelle in Aufgabe 1 als Werte einer Stichprobe. (Sie können für diese Aufgabe vergessen, woher die Zahlen kommen.) Bestimmen Sie

a) den Median,
b) das untere und das obere Quartil
c) sowie die Medianabweichung (MAD)

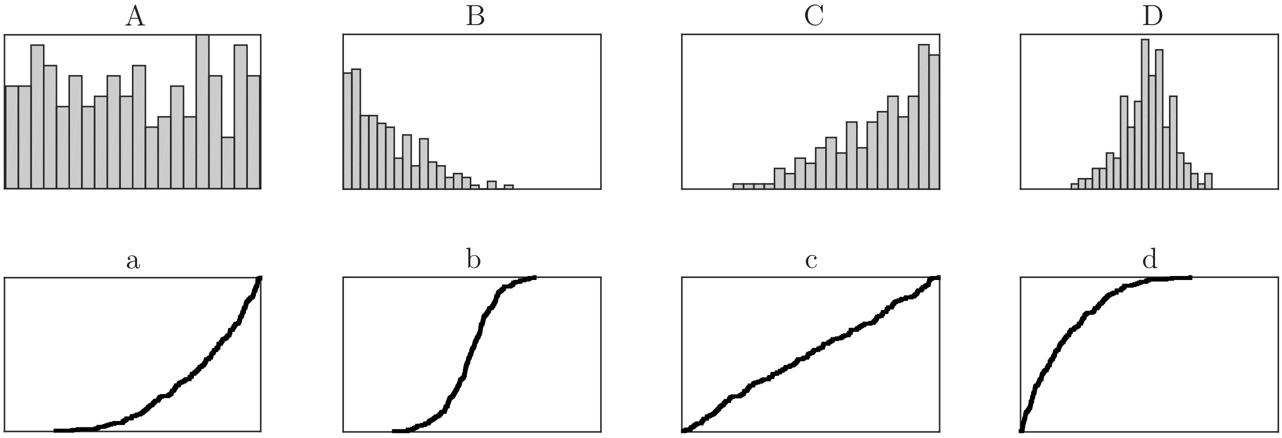
für diese Stichprobe.

d) Fertigen Sie einen Boxplot für diese Stichprobe an.
(Achse darunter oder daneben nicht vergessen!)

Aufgabe 6 (8 Punkte)

In der folgenden Abbildung ist für 4 Stichproben vom Umfang 200 jeweils ein Histogramm und die empirische Verteilungsfunktion dargestellt. Ordnen Sie die Plots einander zu, d.h. notieren Sie Paare von Groß- und Kleinbuchstaben von Plots, die von der gleichen Stichprobe stammen.

Für jede richtige Zuordnung erhalten Sie zwei Punkte, für jede falsche Zuordnung werden zwei Punkte abgezogen; nicht gemachte Zuordnungen geben null Punkte.†



Aufgabe 7 (2+6 = 8 Punkte)

Ihnen liegt eine Stichprobe mit 100 Messwerten vor. Das 95%-Vertrauensintervall (im Sinne eines beidseitigen t-Tests) für den Lageparameter μ ermitteln Sie zu

$$[5{,}4\ ,\ 6{,}2]\,, \quad \text{d.h.} \quad 5{,}4 \leq \mu \leq 6{,}2\,.$$

a) Rekonstruieren Sie den Mittelwert der Stichprobe.
b) Nun nehmen Sie weitere Messungen vor, bis der Stichprobenumfang 1000 beträgt. Lustigerweise ändert sich dabei weder der Mittelwert noch die empirische Varianz. Wie lautet jetzt das 95%-Vertrauensintervall für μ?

Die folgenden MATLAB-Schnipsel sind vielleicht hilfreich.

```
>> tinv(.95,[9 99 999])
ans =    1.8331   1.6604   1.6464

>> tinv(.975,[9 99 999])
ans =    2.2622   1.9842   1.9623
```

ZUR ERINNERUNG: Der Befehl `tinv(`q, d_f`)` liefert das q-Quantil der Studentschen t-Verteilung mit d_f Freiheitsgraden.

†Sollte sich dadurch für die Aufgabe eine negative Gesamtpunktzahl ergeben, so wird die Aufgabe mit null Punkten gewertet.

Aufgabe 8 (1+2+6+2+1 = 12 Punkte)

Eine Gnurpologin behauptet, dass Gnurpen zum Frühstück im Schnitt 100 g Xarg zu sich nehmen (hinreichendes Nahrungsangebot vorausgesetzt). Sie vermuten aber, dass der Median μ der Verteilung der aufgenommen Xarg-Menge *kleiner* ist. Um dies zu überprüfen, beobachten Sie einige Gnurpen bei der morgendlichen Nahrungsaufnahme und dokumentieren die jeweils verzehrte Xarg-Menge:

Gnurpe Nr.	1	2	3	4	5	6	7	8	9	10
Verzehrte Xarg-Menge in g	110	89	92	88	83	93	112	101	104	105

a) Bestimmen Sie den Median der Stichprobe.

Nun führen Sie einen Wilcoxontest zum Signifikanzniveau $\alpha = 5\%$ durch:

b) Geben Sie die Nullhypothese H_0 und Alternativhypothese H_A an.

c) Bestimmen Sie den Wert der Teststatistik.
Notieren Sie dabei auch alle Zwischenschritte, d.h. geben Sie insbesondere eine Tabelle an, die alle benötigten Ränge enthält.

d) Bestimmen Sie das Verwerfungskriterium.‡

e) Geben Sie die Testentscheidung an.

‡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_{krit} einseitig	3	5	8	10	13	17	21
U_{krit} zweiseitig	2	3	5	8	10	13	17