Technische Universität Berlin
Fakultät II - Mathematik und Naturwissenschaften
Institut für Mathematik

StR.i.H.
A. Gündel-vom Hofe

WS 2012/13
14.02.2013

Modulklausur A zur „Statistik für Biologen“

Name: ....*Vorname:* ……………………….................................

Matr.-Nr.: ....*Studiengang (FU):* ....…....................................

Es sind nicht programmierbare *Taschenrechner,* das *Skript* aus der Vorlesung sowie ein handbeschriebenes zweiseitiges DIN A4-Blatt mit eigenen Notizen (*keine Aufgaben!!!*) zugelassen.

Abzugeben sind die Lösungen in *Reinschrift* mit allen *Nebenrechnungen* auf *DIN A4*-Blättern. Die Blätter sind bei Abgabe durchzunummerieren. Mit *Bleistift* oder *in Rot* geschriebene Klausuren werden *nicht gewertet*. *Lösungswege* bzw. *Lösungen* müssen nachvollzogen werden können, d.h. *sind zu dokumentieren*.

Mit 20 von 40 erreichbaren Punkten gilt die Klausur als bestanden. Zu bearbeiten sind 4 der 5 gegebenen Aufgaben. Die nicht bearbeitete Aufgabe ist auf diesem Aufgabenblatt entsprechend zu markieren, da auch nur maximal 4 Aufgaben korrigiert werden!

Unterschrift Korrektor/in: ....*Punktzahl* (*40 P*)*:* ....*Note:* ….........

---

1. Aufgabe:

Von 10 Teilnehmern einer Klausur wurden folgende Punktzahlen (von 40) erreicht:

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
$\boldsymbol{x_i}$	*23*	*32*	*35*	*21*	*31*	*3*	*13*	*32*	*40*	*35*

a) Wandeln Sie die gegebene Datenmenge zunächst in eine *geordnete Stichprobe* um und zeichnen Sie die zugehörige *empirische Verteilungsfunktion* $\hat{F}$.

b) Ermitteln Sie rechnerisch über die Indexungleichung im Skript den *empirischen Median* $x_{0.5}$, das untere *Quartil* $x_{0.25}$ und das *obere Quartil* $x_{0.75}$ sowie zusätzlich den arithmetischen Mittelwert $\bar{x}$. Zeichnen Sie mithilfe der ermittelten Werte einen *Box-Plot* zur gegebenen Stichprobe unter Kennzeichnung etwaiger Ausreißer.

c) Unter Vorgabe der Einteilung der Stichprobenwerte in die 3 Klassen $(0;20]$, $(20;30]$ und $(30;40]$ zeichne man ein *Histogramm*, bezogen auf diese Klassen.

	10,0

2. Aufgabe:

Eine Datenmessung an $n = 8$ Hühnereiern ergab bezüglich der beiden Merkmale X : *maximaler Eierdurchmesser in mm* und Y : *minimaler Eierdurchmesser in mm* folgende Werte x_i, y_i $(i = 1,\dots,10)$ sowie die zugehörigen Werte x_i^2, y_i^2 sowie $x_i y_i$:

bitte wenden!!

Max ∅ : x_i	52	55	55	56	57	58	58	60
Min ∅ : y_i	44	40	45	43	42	41	44	42
x_i^2	2704	3025	3025	3136	3249	3364	3364	3600
y_i^2	1936	1600	2025	1849	1764	1681	1936	1764
$x_i y_i$	2288	2200	2475	2408	2394	2378	2552	2520

a) Skizzieren sie die zugehörige *Punktwolke* und bestimmen Sie ihren *Schwerpunkt* $S = (\bar{x}, \bar{y})$.

b) Ermitteln die *empirische Kovarianz* cov ($\boldsymbol{x},\boldsymbol{y}$) sowie den *empirischen Korrelationskoeffizienten* r ($\boldsymbol{x},\boldsymbol{y}$) der gegebenen Punktwolke. Welchen Schluss ziehen Sie aus dem Wert?

c) Berechnen Sie die Koeffizienten für die *Ausgleichsgerade* $x = \gamma + \delta y$ *von x nach y* und zeichnen Sie diese in Ihre Punktwolkenskizze mit ein. Auf welcher Achse liegt γ ?

	10,0

3. Aufgabe:

Bei einem Multiple-Choice-Test sind 20 Fragen mit „trifft zu / trifft nicht zu“ zu beantworten.

a) Geben Sie einen geeigneten Ergebnisraum Ω in der Mengenschreibweise an, durch welchen alle möglichen Varianten an Beantwortungsmöglichkeiten für diesen Test beschrieben werden. Durch welches Urnenmodell wird das zufällige Ausfüllen eines solchen Testfragebogens simuliert, und welche Anzahl $|\Omega|$ für Ω ergibt sich dadurch?

b) Sei A das Ereignis, dass bei der zufälligen Ausfüllung dieses Tests ebenso viele „trifft zu“-Antworten auftreten wie „trifft nicht zu“-Antworten. Welches Urnenmodell kann zur Beschreibung dieses Ereignisses zugrunde gelegt werden? Was ergibt sich daraus als (Laplacesche) Wahrscheinlichkeit $P(A)$ für das Ereignis A ?

c) Sei B das Ereignis, dass bei der zufälligen Ausfüllung des Fragebogens die ersten vier Fragen mit „trifft zu“ und die letzten zwei Fragen mit „trifft nicht zu“ beantwortet sind. Beschreiben Sie wieder B als geeignete Teilmenge von Ω . Bestimmen Sie dann die (Laplaceschen) Wahrscheinlichkeiten $P(B)$ sowie $P(A \setminus B)$ für das Ereignis *„A*, aber nicht *B“* unter Anwendung einer geeigneten Formel für $|A \setminus B|$.

	10,0

4. Aufgabe:

Im Jahre 1963 wurde ein Versuch mit *23* Affen durchgeführt. Die aus *11* Affen bestehende *Versuchsgruppe* wurde einer Stresssituation ausgesetzt, indem die Affen 24 Stunden lang fleißig einen Hebel fest drücken mussten und einen Stromstoß als Bestrafung erhielten, sobald sie nachließen. Die Affen der *Kontrollgruppe* hingegen wurden in Ruhe gelassen und mussten nichts tun. Dann wurden alle Affen mit dem Polio I Virus infiziert. Das überraschende Ergebnis des Versuchs zeigt die folgende unvollständige Vierfeldertafel auf der nächsten Seite:

	Polio überlebt	**Polio nicht überlebt**	**Summe**
Versuchsgruppe	7		*11*
Kontrollgruppe	1		
Summe			*23*

a) Vervollständigen Sie die fehlenden Teile in der Tafel. Formulieren und geben Sie unter Rückgriff auf die Ereignisse L : „Affe überlebt die Polio-Infektion“ , S : „Affe wird dem Stress ausgesetzt“ sowie deren entsprechenden Gegenereignisse die bedingten Wahrscheinlichkeiten an,
   (i) dass ein Affe aus der Kontrollgruppe die Polio-Infektion nicht überlebt,
   (ii) dass ein Affe aus der Versuchsgruppe die Polio-Infektion nicht überlebt.

b) Malen Sie ein zweistufiges Baumdiagramm unter Eintragung der entsprechenden bedingten Wahrscheinlichkeiten und berechnen Sie aus diesem unter Rückgriff auf den Satz von der totalen Wahrscheinlichkeit den Wert für $P(L^C)$.

c) Formulieren und bestimmen Sie dann unter Anwendung des *Satzes von Bayes* die (bedingte) Wahrscheinlichkeit dafür, dass ein zufällig herausgegriffenes Exemplar aus der Gruppe der verstorbenen Affen gerade aus der Versuchsgruppe stammt. Vergleichen Sie Ihr Ergebnis mit dem Wert, den sie noch einmal anhand der Vierfeldertafel *direkt* – also ohne Verwendung des Satzes von Bayes – berechnen.

d) Untersuchen Sie die beiden Ereignisse L : „Affe überlebt die Polio-Infektion“ und S^C : „Affe wurde dem Stress nicht ausgesetzt“ auf stochastische Unabhängigkeit.

	10,0

<u>*5. Aufgabe:*</u>

a) Der Dozent der „Statistik für Biologen“ schreibt eine Klausur mit *120* Studenten/innen. Bei dieser Klausur benutzen *23* der Teilnehmer/innen unlautere Hilsmittel. Da der Dozent die Verwendung nicht zugelassener Hilfsmittel erahnt, führt er bei *10* zufällig ausgewählten Klausurteilnehmern eine Kontrolle durch.

   Wie groß ist die Wahrscheinlichkeit, dass der Dozent aufgrund der Zusammensetzung der zufälligen Stichprobe die Verwendung unlauterer Hilfsmittel entdeckt? Definieren Sie dazu in Worten eine geeignete Zufallsvariable X und benutzen Sie X zur Formulierung der gesuchten Wahrscheinlichkeit. Welches ist die X zugrunde liegende Wahrscheinlichkeitsverteilung und welche Parameter sind gegeben?

b) Der Dozent hat einen sich während der Klausur auffällig verhaltenen Studenten „auf dem Kieker“. Da er vermutet, dass er unzulässige Hilfsmittel benutzt, stellt er ihn nach der Klausur zur Rede. Der Student behauptet, er benötige derartige Hilfsmittel nicht, da er *80%* aller möglichen Testfragen korrekt beantworten könne. Um die Korrektheit dieser Behauptung zu überprüfen, stellt der Dozent dem Studenten nacheinander *10* zufällig ausgewählte Testfragen, die dieser beantworten soll. Da der Dozent vergesslich ist, können sich dabei Fragen auch wiederholen.

   Wie groß ist die Wahrscheinlichkeit, dass der Student, der die Wahrheit spricht, unter den an ihn gestellten Fragen tatsächlich mindestens *8* vorfindet, die er korrekt beantworten kann? Definieren Sie dazu wieder eine geeignete Zufallsvariable X und geben Sie die X zugrunde liegende Wahrscheinlichkeitsverteilung an. Welche Parameter liegen hier vor?

	10,0