

StR.i.H.
 A. Gündel-vom Hofe

**Probeklausur zur
 „Statistik für Biologen“**

Name: Vorname:

Matr.-Nr.: Studiengang (FU):

Es sind nicht programmierbare *Taschenrechner* zugelassen. Abzugeben sind die Lösungen in *Reinschrift* mit allen *Nebenrechnungen* auf *DIN A4*-Blättern. Mit *Bleistift* oder *in Rot* geschriebene Klausuren werden *nicht gewertet*. Ein handbeschriebenes zweiseitiges *DIN A4*-Blatt mit eigenen Notizen (*keine Aufgaben!!!*) ist zugelassen. *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 entsprechend als „bearbeitet“ zu markieren sind. Es werden auch nur maximal 4 Aufgaben korrigiert !

Unterschrift Korrektor/in: Punktzahl (40 P): Note:

1. Aufgabe:

Eine Datenmessung bei 10 Objekten hat folgende Messwerte $x_i, i = 1, \dots, 10$ ergeben:

i	1	2	3	4	5	6	7	8	9	10
x_i	8,5	1,5	25	4,5	6,0	3,0	3,0	2,5	6,0	9,0

- Wandeln Sie die gegebene Datenmenge in eine *geordnete Stichprobe* um und zeichnen Sie die zugehörige *empirische Verteilungsfunktion* \hat{F} sowie die *theoretische Verteilungsfunktion* F .
- Ermitteln Sie aus der theoretischen Verteilungsfunktion den *empirischen Median* $x_{0,5}$ sowie das *untere Quartil* $x_{0,25}$ und das *obere Quartil* $x_{0,75}$. Zeichnen Sie mithilfe der ermittelten Werte einen *Box-Plot* zur gegebenen Stichprobe.
- Unter Vorgabe der Einteilung der Stichprobenwerte in die 3 Klassen $(0; 3,0]$, $(3,0; 8,0]$ und $(8,0; 30,0]$ zeichne man ein *Histogramm*, bezogen auf diese Klassen.

	10,0
--	------

2. Aufgabe:

Die folgende Tabelle beschreibt die in einer Langzeitstudie ermittelten Werte für die durchschnittliche Höhe von Sonnenblumen in Abhängigkeit von der Zeit nach dem Saattermin.

Zeit x_i	10	20	30	36	40	50
Höhe y_i	20	70	105	135	160	200

bitte wenden!!

- a) Skizzieren sie die zugehörige *Punktwolke* und bestimmen Sie ihren *Schwerpunkt* $S = (\bar{x}, \bar{y})$.
- b) Ermitteln die *empirische Kovarianz* $\text{cov}(\mathbf{x}, \mathbf{y})$ sowie den *empirischen Korrelationskoeffizienten* $r(\mathbf{x}, \mathbf{y})$ der gegebenen Punktwolke. Welchen Schluss ziehen Sie aus dem Wert?
- c) Berechnen Sie die Koeffizienten für die *Ausgleichsgerade von y nach x* und zeichnen Sie diese in Ihre Punktwolken-skizze mit ein. Welchen Wert für die Sonnenblumenhöhe würde man nach der vorgenommenen linearen Regression nach einer Zeit von $x = 39$ Tagen erwarten?

	10,0
--	------

3. Aufgabe:

Für einen in $5 \times 5 = 25$ Felder unterteilten Pfannkuchenteig stehen zur Füllung der einzelnen Felder (Pfannkuchen) zwei Varianten zur Auswahl: Pflaumenmus und Cointreau.

- a) Geben Sie einen geeigneten Ergebnisraum Ω an, wenn jedes Feld beliebig mit einer der beiden Varianten bestückt werden kann. Welches Urnenmodell steckt dahinter und wie groß ist $|\Omega|$?
- b) Sei A das Ereignis, dass zufällig 19 Felder mit Pflaumenmus und die restlichen 6 Felder mit Cointreau als Füllung versehen sind. Welches Urnenmodell legen Sie dieser zufälligen Wahl zugrunde und wie groß ist die (Laplacesche) Wahrscheinlichkeit $P(A)$ für dieses Ereignis?
- c) Sei B das Ereignis, dass bei der zufälligen Bestückung der 25 Pfannkuchenfelder die ersten 4 Felder der ersten Reihe mit einer Pflaumenmus-Füllung versehen sind, ansonsten aber jedes andere Feld beliebig mit einer der beiden Varianten bestückt ist. Beschreiben Sie B als Teilmenge von Ω und bestimmen Sie die (Laplaceschen) Wahrscheinlichkeiten $P(B)$ sowie $P(A \cup B)$.

	10,0
--	------

4. Aufgabe:

Im Jahre 1954 wurde in den USA in einem Großversuch der Salk-Impfstoff gegen Polio (Kinderlähmung) erprobt. Dazu wurden insgesamt 401.974 Kinder geimpft, darunter ein Teil – die *Versuchsgruppe* – mit dem Salk-Impfstoff, und der andere Teil – die *Kontrollgruppe* – mit einer Salzwasserlösung. Nach Ablauf einiger Jahre wurden sämtliche Kinder auf eine Polio-Infektion hin überprüft. Die Auswertung der Daten zeigte:

- (i) 115 Kinder aus der *Kontrollgruppe* von 201.229 Kindern waren an Polio erkrankt.
- (ii) 200.712 Kinder aus der *Versuchsgruppe* waren nicht an Polio erkrankt.

a) Vervollständigen Sie anhand der Daten zunächst die folgende Vierfeldertafel:

	Polio	kein Polio	Summe
Versuchsgruppe		200.712	
Kontrollgruppe	115		201.229
Summe			401.974

- b) Formulieren und berechnen Sie die bedingten Wahrscheinlichkeiten,
 - (i) dass ein Kind aus der Kontrollgruppe nicht an Polio erkrankt,
 - (ii) dass ein Kind aus der Versuchsgruppe an Polio erkrankt.

- c) Bei einem aus der Gesamtgruppe der Kinder zufällig herausgegriffenen Kind wird eine Polio-Infektion diagnostiziert. Bestimmen Sie unter Anwendung des Satzes von Bayes die (bedingte) Wahrscheinlichkeit, dass dieses Kind aus der Versuchsgruppe stammt und vergleichen Sie anschließend das Ergebnis mit dem direkt aus der Vierfeldertafel ermittelten Wert.
- d) Untersuchen Sie die beiden Ereignisse E^C : „Kind erkrankt nicht an Polio“ und K^C : „Kind stammt nicht aus der Kontrollgruppe“ auf stochastische Unabhängigkeit.

	10,0
--	------

5. Aufgabe:

- a) Das Küchenteam einer Hotelküche ordert vom Großhändler 200 Hühnereier. Es ist bekannt, dass bei diesem Händler durchschnittlich 2% der Eier ein abgelaufenes Verfallsdatum aufweisen.
- Wie groß ist die Wahrscheinlichkeit, dass in dem vom Händler zufällig zusammengestellten Sortiment von 200 Eiern sich mindestens zwei mit abgelaufenem Haltbarkeitsdatum befinden? Definieren Sie dazu eine geeignete Zufallsvariable (ZV) X mit entsprechenden Parametern und geben Sie die X zugrunde liegende Wahrscheinlichkeitsverteilung an.
- b) Der Küchenchef entnimmt der Lieferung von 200 Eiern, unter denen sich tatsächlich 4 Eier mit abgelaufenem Haltbarkeitsdatum befinden, 10 Eier für eine Überprüfung des Haltbarkeitsdatums.
- Wie groß ist die Wahrscheinlichkeit dafür, dass bei seiner Prüfung mindestens eines der Eier mit abgelaufenem Haltbarkeitsdatum entdeckt wird? Definieren Sie dazu wieder eine geeignete Zufallsvariable (ZV) X mit entsprechenden Parametern und geben Sie die X zugrunde liegende Wahrscheinlichkeitsverteilung an.

	10,0
--	------