

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

Fakultät Agrarwissenschaften und Landschaftsarchitektur (AuL)

Klausurfragen Bio Data Science

für Pflichtmodule

im 1. & 2. Semester B.Sc./M.Sc.

(Prüfungsleistung der Wahlpflichtmodule ist eine Portfolioprüfung)

Prüfer: Prof. Dr. Jochen Kruppa-Scheetz
Fakultät für Agrarwissenschaften und Landschaftsarchitektur
j.kruppa@hs-osnabrueck.de

Wintersemester 2024/25



*„The test of a student is not how much he knows,
but how much he wants to know.“
— Alice W. Rollins*



Alex studiert im 3. Semester und wiederholt das Modul, da er im ersten Jahr andere Prioritäten für sich gesetzt hat. Das musste sein, da er sich ziemlich im Abitur verausgabt hat. Darüber hinaus war die WG auch eher auf Party angelegt. Alex hofft jetzt über Pünktlichkeit wieder in die Bahn zu kommen. Dafür steht er jetzt immer um 5 Uhr auf! Freunde von ihm beschreiben ihn eher als extrovertiert. Er kennt Paula noch aus der Schulzeit und er überlegt, ob nicht beide Mal nach Mallorca sollten.

Nach zwei Semestern Studium an der Universität Osnabrück war es dann Jessica doch viel zu theoretisch. Etwas angewandtes sollte es sein, wo sie auch eine Fähigkeit lernt, die Frau nutzen kann. Deshalb hat sich Jessica an der Hochschule eingeschrieben. Hoffentlich lernt sie etwas nützliches, wo andere für Geld geben würden. Wer nützlich ist, ist wertvoll. Ihr Traum ist ja eine Hundeschule aufzumachen. Die großen Parties hat sie immer gemieden. Sie ist lieber mit ihrer Hündin im Teutoburgerwald.



Das ist jetzt der letzte Versuch mit einem Studium. Wenn es nicht klappt dann überlegt Jonas das **Programm der IHK zu Ausbildungsvermittlung** zu nutzen. Aber eine Runde gibt er sich noch. Struktur ist eigentlich das Wichtigste und diesmal hat er sich alle Altklausuren der Fachschaft besorgt. Dann ist er auch noch gleich der Fachschaft beigetreten um mehr Informationen abzugreifen. Und er versucht nicht seine Zeit mit Alex zu verdaddeln oder in der Fachschaft bei einem Bier oder so...

Nächstes Semester geht es nach Kanada davon hat er schon auf der Berufsschule geträumt. Deshalb konzentriert er sich sehr auf die Prüfungen. Ein Schiff ist im Hafen sicher, aber dafür ist es nicht gebaut worden. Das **International Faculty Office** der Fakultät Agrarwissenschaften und Landschaftsarchitektur hat super geholfen, aber es waren einiges an Unterlagen. Jetzt hofft er, dass Tina dann doch noch mitkommt. Aber sonst macht er das eben alleine. Obwohl das eher nicht so seine Art ist. Vielleicht sollte er sich mal einen Tipp bei Tina holen, sie wirkt sehr entschlossen.



Nach der Ausbildung wollte Nilufar eigentlich gleich anfangen zu arbeiten, aber nach einem Praktikum und der Probezeit stellte sie fest, dass es ohne einen Hochschulabschluss schwer wird Führungsverantwortung zu übernehmen. Mit Menschen kann sie schon immer und dann auch eigene Projekte mit anderen verwirklichen, das ist doch was. Mit dem notwendigen Abschluss sollte der Start um so einfacher sein. Dann ist sie keine Befehlsempfängerin mehr sondern gibt die Marschrichtung vor. Schon jetzt koordiniert Nilufar das Studium von anderen.

Paula möchte die Welt zu einem besseren Ort machen. Wenn da nicht die anderen Mitmenschen wären. Paula muss das Modul nochmal wiederholen, da es dann am Ende des Semesters zu viel für sie wurde. Eine Lerngruppe hätte geholfen, aber das ist dann gar nicht so einfach eine zu finden. Zwar kennt sie schon Nilufar, aber Nilufar ist ihr manchmal zu forsch. Ihr schwant aber, dass alleine das Studium sehr schwer werden wird. Das Abitur war schon so ein Lernhorror, das möchte sie nicht nochmal. Alex sieht sie da als Vorbild.



Sommer, Sonne, Natur. Das ist es was Steffen mag. Raus in die Komune und die Natur genießen. Leider hat Steffen noch andere Bedürfnisse, die ein Einkommen benötigen. Da Studierende mehr verdienen, würde dann in Teilzeit auch mehr rausspringen. Wenn er dann privat was anbauen kann, dann spart er gleich doppelt. Leider sind viele seiner Kommilitonen total verkrampte Karrieristen. Es geht nur ums Äußere. Dabei verliert sich Steffen gerne im Prozess. Das hat auch seinen Schulabschluss etwas verzögert. Steffen lässt sich eben Zeit.

Wille war es, die es Tina hierher gebracht hat und Wille wird es sein, die Tina dann auch zum Abschluß treibt. Nach einem Rückschlag muss Tina jetzt einige Module wiederholen, damit sie dann auch fertig wird. Ab und zu ist sie schwach gewesen und das hat dann Zeit gekostet. Das Tina es dann manchmal übertreibt, weiß sie nur zu gut, aber irgendwie muss die Kontrolle ja erhalten bleiben? Insbesondere, wenn sie mal wieder die Nacht durchgefeiert hat, verachtet Tina sich. Dann baut Nilufar sie dann bei einem Tee wieder auf.



Für Yuki war es nicht einfach. Teilweise war die Krankheit sehr hinderlich, dann war es Yuki selber. Dann muss man auch wieder auf die Beine kommen und es dauert eben seine Zeit. Aber immerhin hat Yuki es jetzt den Abschluss gekriegt und hat einen Studienplatz. Jetzt heißt es in den Rhythmus kommen und schauen, was noch so passiert. Immerhin hat Yuki schon eine kleine Gruppe gefunden, in der Yuki dann Hilfe findet. Ist aber auch sehr unübersichtlich so ein Studium. Steffen ist immer super entspannt.

Erlaubte Hilfsmittel

- Normaler Taschenrechner ohne Möglichkeit der Kommunikation mit anderen Geräten! Ausdrücklich kein Handy!
- Eine DIN A4-Seite als beidseitig, selbstgeschriebene, handschriftliche Formelsammlung. Keine digitalen Ausdrucke!
- **Die Verwendung eines roten Farbstiftes ist nicht gestattet! Korrekturfarbe!**
- *You can answer the questions in English without any consequences.*

Endnote

_____ von 20 Punkten sind aus den Multiple Choice Aufgaben erreicht.
_____ von 69 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 89 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
85.0 - 89.0	1,0
80.5 - 84.5	1,3
76.5 - 80.0	1,7
72.0 - 76.0	2,0
67.5 - 71.5	2,3
63.0 - 67.0	2,7
58.5 - 62.5	3,0
54.5 - 58.0	3,3
50.0 - 54.0	3,7
44.5 - 49.5	4,0

Es ergibt sich eine Endnote von _____.

Multiple Choice Aufgaben

- Pro Multiple Choice Frage ist *genau* eine Antwort richtig.
- Übertragen Sie Ihre Kreuze in die Tabelle auf dieser Seite.

	A	B	C	D	E	✓
Aufgabe 1						
Aufgabe 2						
Aufgabe 3						
Aufgabe 4						
Aufgabe 5						
Aufgabe 6						
Aufgabe 7						
Aufgabe 8						
Aufgabe 9						
Aufgabe 10						

- Es sind ____ von 20 Punkten erreicht worden.

Rechen- und Textaufgaben

Aufgabe	11	12	13	14	15	16	17
Punkte	10	10	10	9	10	10	10

- Es sind ____ von 69 Punkten erreicht worden.

Multiple Choice Aufgaben

Die Multiple Choice Aufgaben **unterliegen dem Zufall**. Die Reihenfolge der Antworten ist zufällig. Die Fragen und Antworten sind semantisch zufällig und haben somit **verschiedene Textvarianten**. Insbesondere die reinen Textaufgaben haben verschiedene Textvarianten. Die Semeantik mag sich unterscheiden, die Inhalte sind aber gleich.

ANOVA

1. Aufgabe

(2 Punkte)

Sie führen einen Versuch mit einer Behandlung und drei Faktorleveln durch. Danach rechnen Sie eine einfaktorielle ANOVA und es ergibt sich ein $\eta^2 = 0.12$. Welche Aussage ist richtig?

- A** ☐ Das η^2 ist die Korrelation der ANOVA. Mit der Ausnahme, dass 0 der beste Wert ist.
- B** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen nicht erklärt wird. Somit der Rest an nicht erklärbarer Varianz.
- C** ☐ Die Berechnung von η^2 ist ein Wert für die Interaktion.
- D** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird. Das η^2 ist damit mit dem R^2 aus der linearen Regression zu vergleichen.
- E** ☐ Das η^2 ist ein Wert für die Güte der ANOVA. Je kleiner desto besser. Ein η^2 von 0 bedeutet ein perfektes Modell mit keiner Abweichung. Die Varianz ist null.

2. Aufgabe

(2 Punkte)

Sie führen ein Feldexperiment durch um das Gewicht von Erbsen zu steigern. Die Pflanzen wachsen unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen. Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.3$. Welche Aussage ist richtig?

- A** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird. Daher werden 30% der Varianz durch die Behandlungsgruppen erklärt.
- B** ☐ Mit dem η^2 lässt sich auf die Qualität der Randomisierung und damit der Strukturgleichheit zwischen der Grundgesamtheit und der Stichprobe schließen. Es gilt dabei die Regel, dass ein η^2 -Wert von 1 zu bevorzugen ist.
- C** ☐ Das η^2 beschreibt den Anteil der Varianz, der durch den Forschenden entsteht. Es gilt die Regel, dass ca. 70% der Varianz eines Versuches durch die Versuchsdurchführung entstehen sollen.
- D** ☐ Es werden 70% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.
- E** ☐ Es werden 30% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.

3. Aufgabe

(2 Punkte)


Eine einfaktorielle ANOVA berechnet eine Teststatistik um zu die Nullhypothese abzulehnen. Welche Aussage über die Teststatistik der ANOVA ist richtig?

- A** ☐ Wenn die F-Statistik höher ist als der kritische Wert kann die Nullhypothese nicht abgelehnt werden. Die F-Statistik ist die Differenz der MS der Behandlung durch die MS des Fehlers.
- B** ☐ Die F-Statistik wird berechnet indem die MS der Behandlung durch die MS des Fehlers geteilt werden. Wenn die F-Statistik sich kaum von der Null unterscheidet kann die Nullhypothese nicht abgelehnt werden.

- C** ☐ Die ANOVA berechnet die T-Statistik indem den Mittelwertsunterschied der Gruppen simultan durch die Standardabweichung der Gruppen teilt. Wenn die T-Statistik höher als 1.96 ist, kann die Nullhypothese abgelehnt werden.
- D** ☐ Die ANOVA berechnet die T-Statistik aus der Multiplikation der MS Behandlung mit der MS der Fehler. Wenn die F-Statistik genau 0 ist, kann die Nullhypothese nicht abgelehnt werden.
- E** ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 0 annähert kann die Nullhypothese abgelehnt werden.

4. Aufgabe

(2 Punkte)

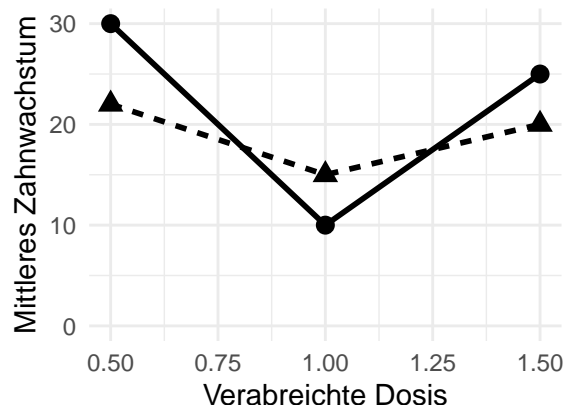
Wenn Sie mehr als zwei Gruppen als Behandlungen vorliegen haben, dann kann ein einfacher t-Test nicht für den globalen Vergleich genutzt werden. Sie entscheiden sich für eine ANOVA in . Die ANOVA analysiert dabei...

- A** ☐ ... den Unterschied zwischen der globalen Varianz und der Varianz aus verschiedenen Behandlungsguppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.
- B** ☐ ... den Unterschied zwischen mehreren Varianzen aus verschiedenen Behandlungsguppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.
- C** ☐ ... den Unterschied zwischen der Varianz in den verschiedenen Behandlungsguppen und der Varianz in einer der Behandlungsguppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in einer der Gruppen exakt zu bestimmen.
- D** ☐ ... den Unterschied zwischen der Varianz über alle Behandlungsguppen oder der Varianz aus verschiedenen Behandlungsguppen. Wenn die ANOVA signifikant ist, muss sich zwischen einem der beiden Varianzquellen entschieden werden.
- E** ☐ ... den Unterschied zwischen der F-Statistik anhand der Varianz der Gruppen. Wenn die F-Statistik exakt 0 ist, kann die Nullhypothese abgelehnt werden.

5. Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung von Vitamin D auf das Zahnwachstum bei Kanarienvögel. Der Versuch wurde an 72 Tieren durchgeführt, wobei jedes Tier eine von drei Vitamin-C-Dosen (0.5, 1 und 1.5 mg/Tag) über eine von zwei Verabreichungsmethoden erhielt. Welche Aussage ist richtig im Bezug auf eine zweifaktorielle ANOVA?



- A** ☐ Eine negative Interaktion liegt vor ($p \geq 0.5$).
- B** ☐ Keine Interaktion liegt vor ($p \leq 0.05$).
- C** ☐ Das Bestimmtheitsmaß R^2 ist groß.
- D** ☐ Die Koeffizienten sind negativ ($\beta_0 < 0; \beta_1 < 0$).
- E** ☐ Eine mittlere bis starke Interaktion liegt vor ($p \leq 0.05$)

Deskriptive Statistik & Explorative Datenanalyse

6. Aufgabe

(2 Punkte)

Wie lautet der Mittelwert und Standardabweichung von y mit 15, 4, 7, 9 und 12.

- A ☐ Es ergibt sich 8.4 +/- 9.15
- B ☐ Es ergibt sich 9.4 +/- 4.28
- C ☐ Es berechnet sich 9.4 +/- 18.3
- D ☐ Es berechnet sich 10.4 +/- 18.3
- E ☐ Sie erhalten 9.4 +/- 2.14

7. Aufgabe

(2 Punkte)

Gegeben ist y mit 14, 14, 26, 25, 22, 27 und 42. Berechnen Sie den Median, das 1st Quartile sowie das 3rd Quartile.

- A ☐ Sie erhalten 25 [14; 27]
- B ☐ Es berechnet sich 24 [15; 28]
- C ☐ Es ergibt sich 24 +/- 14
- D ☐ Es berechnet sich 26 [15; 26]
- E ☐ Sie erhalten 25 +/- 27

8. Aufgabe

(2 Punkte)

Die empfohlene Mindestanzahl an Beobachtungen für die Visualisierung mit einem Histogramm sind...

- A ☐ Die untere Grenze liegt bei einer Beobachtung.
- B ☐ 1 Beobachtung.
- C ☐ 2-5 Beobachtungen.
- D ☐ 10 Beobachtungen.
- E ☐ Wir brauchen fünf oder mehr Beobachtungen.

9. Aufgabe

(2 Punkte)

Um die Varianz zu berechnen müssen wir folgende Rechenoperationen durchführen.

- A ☐ Den Median berechnen, dann die quadratischen Abstände zum Median aufsummieren, dann die Wurzel ziehen.
- B ☐ Den Mittelwert berechnen, dann die quadratischen Abstände zum Mittelwert aufsummieren und durch die Fallzahl teilen, dann die Wurzel ziehen.
- C ☐ Wir berechnen erst den Mittelwert und dann die absoluten Abstände zu dem Mittelwert. Diese quadratischen Abstände summieren wir auf und teilen am Ende durch die Fallzahl.
- D ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend teilen wir durch die Fallzahl.
- E ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend subtrahieren wir die Fallzahl.

10. Aufgabe

(2 Punkte)

In Ihrer Abschlußarbeit wollten Sie Ihre Daten für den Ertrag in einem Barplot darstellen. Sie nutzen den Barplot auch, da der Barplot zu den meist genutzten Visualisierungen von Daten gehört. Welche statistischen Maßzahlen stellt der Barplot dar?

- A** ☐ Den Mittelwert und die Varianz.
- B** ☐ Den Median und die Standardabweichung.
- C** ☐ Durch die Abbildung des Barplot erhalten wir die Informationen über die Mittelwerte und die Standardabweichung.
- D** ☐ Der Barplot stellt den Median und die Streuung dar.
- E** ☐ Der Barplot stellt den Median und die Quartile dar.

11. Aufgabe

(2 Punkte)

Nachdem Sie in einem Feldexperiment zu Leistungssteigerung von Erdbeeren durchgeführt haben, berechnen Sie den Mittelwert und den Median. Der Mittelwert \bar{y} und der Median \tilde{y} unterscheiden sich. Welche Aussage ist richtig?

- A** ☐ Der Mittelwert und der Median sollten gleich sein, wenn keine Outlier in den Daten vorliegen.
- B** ☐ Wenn sich der Mittelwert und der Median unterscheiden, liegen vermutlich keine Outlier in den Daten vor.
- C** ☐ Der Mittelwert und der Median sollten sich unterscheiden sein, wenn Outlier in den Daten vorliegen.
- D** ☐ Wenn sich der Mittelwert und der Median nicht unterscheiden, liegen vermutlich Outlier in den Daten vor.
- E** ☐ Da sich der Mittelwert und der Median unterscheiden, ist der Datensatz nicht zu verwenden. Mittelwert und Median müssen gleich sein.

12. Aufgabe

(2 Punkte)

Ihre Betreuung der Abschlussarbeit fragt überraschend in der letzten Besprechung, ob Ihre Messwerte einer Normalverteilung genügen. Sonst könnten Sie ja gar nicht einen t-Test rechnen. Da Ihnen die Zeit wegrennt, entscheiden Sie sich für eine schnelle Visualisierung im Anhang. Welche Visualisierung nutzen Sie und welche Regel kommt zur Abschätzung einer Normalverteilung zur Anwendung?

- A** ☐ In einer explorativen Datenanalyse nutzen wir den Boxplot. Dabei sollte der Median als dicke Linie in der Mitte der Box liegen. Dann können wir von einer Normalverteilung ausgehen.
- B** ☐ Einen Barplot. Die Mittelwerte müssen alle auf einer Höhe liegen. Die Fehlerbalken haben hier keine Informationen.
- C** ☐ Einen Violinplot. Der Bauch der Violine muss hierbei einen höheren Wert annehmen als der Steg der Violine. Dann kann die Annahme einer Normalverteilung angenommen werden.
- D** ☐ Wir erstellen uns für jede Behandlung einen Dotplot und schauen, ob die Dots und damit die Varianz für jede Behandlung gleich groß sind.
- E** ☐ Nach dem Einlesen der Daten nutzen wir einen Boxplot um zu schauen, ob alle Boxen über alle Behandlungen in etwa gleich groß sind. Damit ist dann auch das IQR in allen Behandlungen in etwa gleich.

13. Aufgabe

(2 Punkte)

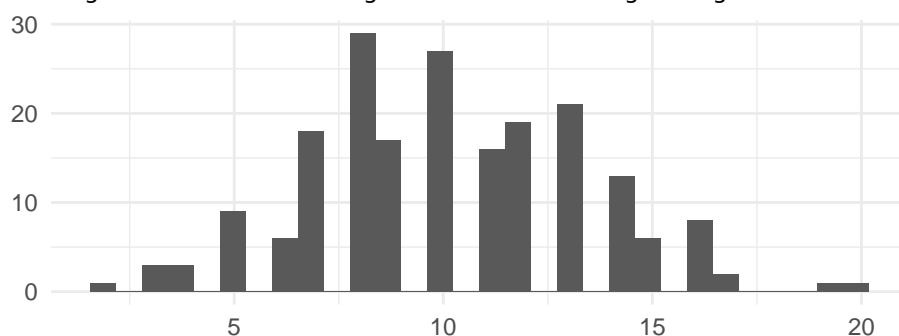
Nach der Durchführung Ihres Feldexperiments wollen Sie eine ANOVA rechnen. Dafür muss aber Ihr Messwert zumindestens approximativ einer Normalverteilung folgen. Welche der drei Abbildungen erlaubt Ihnen abzuschätzen, ob Sie eine Normalverteilung in Ihrem Endpunkt vorliegen haben?

- A ☐ Violinplot, Boxplot, Densityplot
- B ☐ Violinplot, Scatterplot, Barplot
- C ☐ Scatterplot, Mosaicplot, Boxplot
- D ☐ Histogramm, Scatterplot, Boxplot
- E ☐ Boxplot, Violinplot, Mosaicplot

14. Aufgabe

(2 Punkte)

Bevor Sie in Ihrer Abschlussarbeit einen statistischen Test rechnen, wollen Sie einmal betrachten, welcher Verteilung Ihre $n = 216$ geernteten Pflanzen folgen. Welche Verteilung ist abgebildet?



- A ☐ In dem Histogramm ist eine Ordinalverteilung dargestellt.
- B ☐ Es handelt sich um eine Poisson-Verteilung.
- C ☐ Es handelt sich um eine Binomial-Verteilung.
- D ☐ Eine multivariate Normalverteilung.
- E ☐ Es handelt sich um eine Normalverteilung.

Lineare Regression & Korrelation

15. Aufgabe

(2 Punkte)

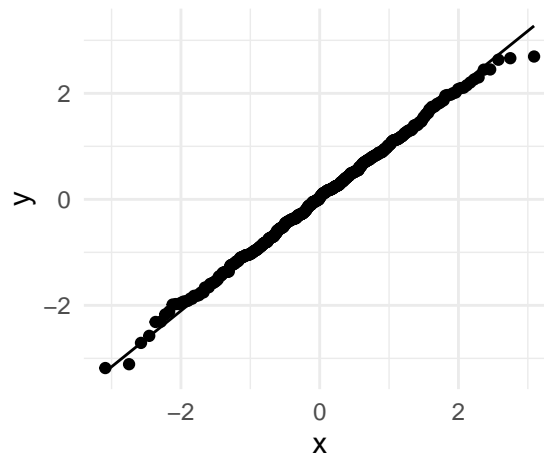
In Ihrer Abschlussarbeit wollen Sie ein prädiktives Modell rechnen. Jetzt stellt sich die Frage, was diese Entscheidung für Ihre Auswertung bedeutet. Welche Aussage ist richtig?

- A ☐ Wenn ein prädiktives Modell gerechnet werden soll, dann muss zum einen ein Trainingsdatensatz sowie ein Testdatensatz definiert werden. Dabei ist der Trainingsdatensatz meist 2/3 und der Testdatensatz 1/3 der Fallzahl groß. Der Testdatensatz dient zur Validierung.
- B ☐ Ein prädiktives Modell wird auf einem Trainingsdatensatz trainiert und anschliessend über eine explorative Datenanalyse validiert. Signifikanzen über β_i können hier nicht festgestellt werden.
- C ☐ Es wird ein Trainingsdatensatz zum Modellieren des Trainingsmodells benötigt. Der Testdatensatz dient rein zur Visualisierung. Dies gilt vor allem für ein prädiktives Modell.
- D ☐ Wenn ein prädiktives Modell gerechnet werden soll, dann muss zum einen ein Trainingsdatensatz sowie ein Testdatensatz definiert werden. Dabei ist der Trainingsdatensatz meist 1/10 und der Testdatensatz 1/3 der Fallzahl groß. Der Testdatensatz dient zur Validierung.
- E ☐ Ein prädiktives Modell möchte die Zusammenhänge von X auf Y modellieren. Hierbei geht es um die Effekte von X auf Y. Man sagt, wenn x_1 um 1 ansteigt ändert sich Y um einen Betrag β_1 .

16. Aufgabe

(2 Punkte)

Nach der Modellierung einer Regression stellt sich die Frage, ob die Residuen approximativ einer Normalverteilung folgen. Sie können einen QQ-Plot für die visuelle Überprüfung der Annahme an die Residuen nutzen. Welche Aussage ist richtig?

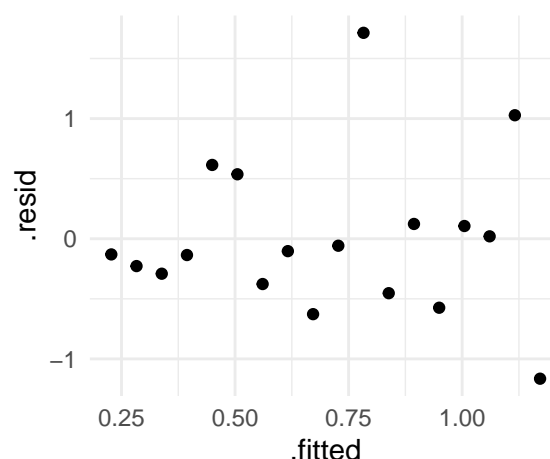


- A** ☐ Wir betrachten insbesondere die beiden Enden der Gerade. Der Rest ist mehr oder minder egal, dann ist die Annahme an die Normalverteilung der Residuen erfüllt.
- B** ☐ Wir betrachten die Punkte. Wenn die Punkte einigermaßen gleichmäßig verteilt liegen, dann gehen wir von normalen Residuen aus.
- C** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.
- D** ☐ Wir betrachten die Gerade und dabei insbesondere die beiden Enden der Gerade. Hier sollten die Punkte auf der Geraden liegen, dann ist die Annahme an die Normalverteilung der Residuen erfüllt.
- E** ☐ Wir betrachten die Gerade und dabei insbesondere die beiden Enden der Gerade. Hier sollten die Punkte auf der Geraden liegen, dann ist die Annahme an die Normalverteilung der Residuen erfüllt. Diese Annahme ist nicht erfüllt.

17. Aufgabe

(2 Punkte)

Nach der Modellierung einer Regression stellt sich die Frage, ob die Residuen (`. resid`) gleichmäßig um die gefitte Gerade liegen. Sie können folgende Abbildung für die visuelle Überprüfung der Residuen nutzen. Welche Aussage ist richtig?



- A** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Es ist ein Muster zu erkennen und wir können damit auf die Signifikanz von x_1, \dots, x_p schließen.

- B** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Diagonalen. Damit ist das Modell erfolgreich geschätzt worden.
- C** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Ein klares Muster ist zu erkennen und/oder einige Outlier sind zu beobachten.
- D** ☐ Wenn die Punkte gleichmäßig in dem positiven wie auch negativen Bereich ohne ein klares Muster liegen, dann hat unsere Modellierung geklappt. Wir können mit dem Modell weitermachen.
- E** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Vereinzelt Punkte liegen oberhalb bzw. unterhalb der Geraden um die 0 Linie weiter entfernt. Ein klares Muster ist zu erkennen.

18. Aufgabe

(2 Punkte)

Welche Aussage über den Korrelationskoeffizienten ρ ist richtig?

- A** ☐ Der Korrelationskoeffizienten ρ liegt zwischen -1 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ als standardisierte Steigung zu verstehen, wenn eine Standardisierung durchgeführt wurde. Diese Adjustierung nach Fischer muss am Anschluß der Berechnung der Korrelation durchgeführt werden.
- B** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen 0 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos. Eine Signifikanz kann nicht nachgewiesen werden.
- C** ☐ Der Korrelationskoeffizienten ρ wird wie das η^2 aus der ANOVA interpretiert. Der Korrelationskoeffizienten ρ beschreibt den Anteil an erklärter Varianz durch die Regression. Dabei gibt er jedoch eine Richtung an und kann auch negativ werden.
- D** ☐ Der Korrelationskoeffizienten ρ ist eine veraltete Darstellungsform von Effekten in der linearen Regression und wird wie das η^2 aus der ANOVA interpretiert. Der Korrelationskoeffizienten ρ beschreibt den Anteil an erklärter Varianz durch die Regression.
- E** ☐ Der Korrelationskoeffizienten ρ liegt zwischen -1 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als standardisierte Steigung verstanden werden.

19. Aufgabe


(2 Punkte)



Nach einer simplen linearen Regression zur Untersuchung vom Einfluss der Fe_3O_4 -Konzentration in $[\mu g]$ im Wasser auf den absoluten Proteingehalt von Wasserlinsen in $[kg]$ erhalten Sie einen $\beta_{Fe_3O_4}$ Koeffizienten von 6.9×10^{-7} und einen hoch signifikanten p -Wert mit 0.00032. Warum sehen Sie so einen kleinen Effekt bei einer so deutlichen Signifikanz?

- A** ☐ Die Einheit der Fe_3O_4 -Konzentration ist zu klein gewählt. Dadurch sehen wir den sehr kleinen p -Wert. Der p -Wert und die Einheit von der Fe_3O_4 -Konzentration hängen antiproportional zusammen.
- B** ☐ Das Gewicht und die Fe_3O_4 -Konzentration korrelieren sehr stark, deshalb wird der $\beta_{Fe_3O_4}$ Koeffizient sehr klein. Mit einer ANOVA kann für die Korrelation korrigiert werden und der Effektschätzer passt dann zum p -Wert.
- C** ☐ Manchmal ist die Einheit der Einflussvariable X zu groß gewählt, so dass der Anstieg von 1 Einheit in X zu einer zu großen Änderung in y führt. Daher kann der Effekt $\beta_{Fe_3O_4}$ sehr klein wirken, da der p -Wert wird auf einer einheitslosen Teststatistik bestimmt wird.
- D** ☐ Die Fallzahl ist zu hoch angesetzt. Je höher die Fallzahl ist, desto kleiner ist die Teststatistik und damit ist dann auch der p -Wert sehr klein. Es sollte über eine Reduzierung der Fallzahl nachgedacht werden. Dann sollte der Effekt zum p -Wert passen.
- E** ☐ Manchmal ist die Einheit der Einflussvariable X zu klein gewählt, so dass der Anstieg von 1 Einheit in X zu einer zu kleinen Änderung in y führt. Daher kann der Effekt $\beta_{Fe_3O_4}$ sehr klein wirken, aber auf einer anderen Einheit sehr viel größer sein. Der p -Wert wird auf einer einheitslosen Teststatistik bestimmt.

20. Aufgabe

(2 Punkte)




Neben der klassischen Regression kann die Funktion `lm()` in  auch für welche andere Art von Anwendung genutzt werden?

- ☐ **A** Neben der klassischen Verwendung der Funktion `lm()` in der linearen Regression kann auch ein Gruppenvergleich gerechnet werden. Dafür müssen aber alle Faktoren aus den Daten entfernt und numerisch umgewandelt werden. Dann kann das R Paket `{emmeans}` genutzt werden um die Korrelation zu berechnen. Eine Adjustierung ist dann nicht mehr notwendig.
- ☐ **B** Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich.
- ☐ **C** Die Funktion `lm()` berechnet die Varianzstruktur für eine ANOVA. Danach kann dann über eine explorative Datenanalyse nochmal eine Signifikanz berechnet werden. Sollte vor der Verwendung der Funktion `lm()` schon eine EDA gerechnet worden sein, so ist die Analyse wertlos.
- ☐ **D** Die Funktion `lm()` in  wird klassischerweise für die nicht-lineare Regression genutzt. Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt.
- ☐ **E** Ist die Einflussvariable X ein Faktor so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich. Dennoch muss zuerst ein lineares Modell mit der Funktion `lm()` in  gerechnet werden.

21. Aufgabe

(2 Punkte)

Wenn Ihr gemessener Endpunkt nicht einer Normalverteilung folgt, so können Sie dennoch Ihre Daten modellieren. Hierzu nutzen Sie dann das *generalisierte lineare Modell (GLM)*. Welche Aussage ist richtig?

- ☐ **A** Das GLM ist ein faktisch maschineller Lernalgorithmus, der selbstständig die Verteilungsfamilie für Y wählt.
- ☐ **B** In  ist mit dem *generalisierten linearen Modell (GLM)* eine Modellierung implementiert, die die Poissonverteilung für Zähldaten oder die Binomialverteilung für 0/1-Daten modellieren kann. Weitere Modellierungen sind in  auch mit zusätzlich geladenen Paketen nicht möglich.
- ☐ **C** Das GLM erlaubt auch nicht normalverteilte Residuen in der Schätzung der Regressionsgrade.
- ☐ **D** In  ist mit dem *generalisierten linearen Modell (GLM)* eine Modellierung implementiert, die neben der klassischen Normalverteilung auch die Poissonverteilung für Zähldaten oder die Binomialverteilung für 0/1-Daten modellieren kann.
- ☐ **E** Dank dem *generalisierten linearen Modell (GLM)* können auch andere Verteilungsfamilien – außer die Normalverteilung – mit einer linearen Regression modelliert werden. Dafür werden alle Verteilungen in eine Normalverteilung überführt und anschließend standardisiert.

Vermischte Themen

22. Aufgabe


(2 Punkte)








Die Randomisierung von Beobachtungen zu den Versuchseinheiten ist bedeutend in der Versuchsplanung. Welche der folgenden Aussagen ist richtig?

- ☐ **A** Randomisierung sorgt für Strukturgleichheit und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- ☐ **B** Strukturgleichheit ist durch Randomisierung gegeben. Leider hilft die Randomisierung noch nicht um von der Stichprobe auf die Grundgesamtheit zu schließen. Deshalb wurde das Falsifikationsprinzip entwickelt.
- ☐ **C** Randomisierung bringt starke Unstrukturiertheit in das Experiment und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- ☐ **D** Randomisierung erlaubt erst die Mittelwerte zu schätzen. Ohne Randomisierung keine Mittelwerte. Ohne Mittelwerte keine Varianz und somit auch kein statistischer Test.
- ☐ **E** Durch eine Randomisierung können wir nicht von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.

23. Aufgabe


(2 Punkte)


Viele wissenschaftliche Originalquellen sind in Englisch verfasst. Jetzt finden Sie heraus, dass auch  nur in englischer Sprache funktioniert. Warum ist das so?

- A** ☐ Programmiersprachen haben Probleme mit Umlauten und Sonderzeichen der deutschen Sprache. Daher ist die Nutzung in Deutsch in den AGBs von  untersagt.
- B** ☐ Alle Funktionen und auch Anwendungen sind in  in englischer Sprache. Die Nutzung von deutschen Wörtern ist nicht schick und das ist zu vermeiden.
- C** ☐ Die Spracherkennung von  ist nicht in der Lage Deutsch zu verstehen.
- D** ☐ Die  Pakete sind nur in englischer Sprache verfasst. Das ist aber nicht der Hauptgrund, denn  hat wie alle Programmiersprachen Probelem mit Umlauten und Sonderzeichen.
- E** ☐  Pakete sind nur in englischer Sprache verfasst. Es macht keinen Sinn  daher in Deutsch zu bedienen.

24. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit wollen Sie zu Beginn eine explorativen Datenanalyse (EDA) in  rechnen. Dafür gibt es eine generelle Abfolge von Prozessschritten. Welche ist hierbei die richtige Reihenfolge?

- A** ☐ Wir transformieren die Spalten über `mutate()` in ein `tibble` und können dann über `ggplot()` uns die Abbildungen erstellen lassen. Dabei beachten wir das wir keine Faktoren in den Daten haben.
- B** ☐ Wir lesen als erstes die Daten über `read_excel()` ein, transformieren die Spalten über `mutate()` in die richtige Form und können dann über `ggplot()` uns die Abbildungen erstellen lassen. Wichtig ist, dass wir keine Faktoren sondern nur numerische Variablen vorliegen haben.
- C** ☐ Wir lesen die Daten ein und mutieren die Daten. Dabei ist wichtig, dass wir nicht das Paket `tidyverse` nutzen, da dieses Paket veraltet ist. über die Funktion `library(tidyverse)` entfernen wir das Paket von der Analyse.
- D** ☐ Wir lesen als erstes die Daten über `read_excel()` ein, transformieren die Spalten über `mutate()` in die richtige Form und können dann über `ggplot()` uns die Abbildungen erstellen lassen.
- E** ☐ Für eine explorativen Datenanalyse (EDA) in  müssen wir als erstes die Daten über `read_excel()` einlesen. Danach müssen wir schauen, dass wir die Zeilen richtig über `mutate()` transformiert haben. Insbesondere müssen Variablen mit kontinuierlichen Werten in einen Faktor umgewandelt werden. Am Ende nutzen wir die Funktion `ggplot()` für die eigentlich EDA.

25. Aufgabe

(2 Punkte)

Gegeben ist das Modell $Y \sim X$. Welche Aussage über $n_1 = n_2$ ist richtig?

- A** ☐ Es handelt sich um ein balanciertes Design.
- B** ☐ Es handelt sich um abhängige Beobachtungen.
- C** ☐ Es liegt Varianzhomogenität vor.
- D** ☐ Es liegt Varianzhetrogenität vor.
- E** ☐ Es handelt sich um ein unbalanciertes Design.

26. Aufgabe

(2 Punkte)

Im Rahmen Ihrer Abschlussarbeit werten Sie ein Experiment mit Ferkel aus. Es geht um die Leistungssteigerung der Ferkelproduktion. Sie messen jeweils die Gewichtszunahme der Ferkel. Die Ferkel einer Muttersau sind dabei...

- A** ☐ Untereinander stark korreliert. Die Ferkel sind von einer Mutter und somit miteinander korreliert. Dies wird in der Statistik jedoch meist nicht modelliert.

- B** ☐ Je nach Stallanlage kommt eine andere Analyse in Betracht. Eine allgemeine Aussage über Ferkel und Sauen lässt sich statistisch nicht treffen.
- C** ☐ Die Ferkel stammen von der gleichen Sau und sind somit untereinander unabhängig.
- D** ☐ Untereinander abhängig. Die Ferkel stammen von einem Muttertier und haben vermutlich eine ähnliche Varianzstruktur.
- E** ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.

27. Aufgabe

(2 Punkte)

Neben der Mittelwertsdifferenz als Effektschätzer bei normalverteilten Endpunkten wird auch häufig der Effektschätzer Risk ratio bei binären Endpunkten verwendet. Welche Aussage über den Effektschätzer Risk ratio ist im folgenden Beispiel zur Behandlung von Klaueninfektionen bei Kühen richtig? Dabei sind 3 Tiere krank und 12 Tiere sind gesund.

- A** ☐ Das Verhältnis der Chancen Risk ratio ergibt ein Chancenverhältnis von 0.2. Wir sind an der Chance krank zu sein interessiert.
- B** ☐ Der Anteil der Kranken wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Risk ratio von 0.2.
- C** ☐ Da es sich um ein Chancenverhältnis handelt ergibt sich ein Risk ratio von 0.25.
- D** ☐ Es ergibt sich ein Risk ratio von 0.2, da es sich um ein Chancenverhältnis handelt.
- E** ☐ Es ergibt sich ein Risk ratio von 4, da es sich um ein Anteil handelt.

28. Aufgabe

(2 Punkte)

In der Bio Data Science wird häufig mit sehr großen Datensätzen gerechnet. Historisch ergibt sich nun ein Problem bei der Auswertung der Daten und deren Bewertung hinsichtlich der Signifikanz. Welche Aussage ist richtig?

- A** ☐ Mehr Fallzahl in Datensätzen bedeutet mehr signifikante Ergebnisse, da in mehr Daten auch mehr Informationen beinhaltet sind. Deshalb lohnen sich riesige Datensätze, die durch die vielen signifikanten Ergebnisse auch eine Menge an relevanten Erkenntnissen liefern.
- B** ☐ Riesige Datensätze haben mehr Fallzahl was zur α -Inflation führt. Durch eine Adjustierung kann dem Problem entgegengewirkt werden.
- C** ☐ Aktuell werden zu große Datensätze für die gängige Statistik gemessen. Daher wendet man maschinelle Lernverfahren für kausale Modelle an. Hier ist die Relevanz gleich Signifikanz.
- D** ☐ Big Data ist ein Problem der parametrischen Statistik. Parameter lassen sich nur auf kleinen Datensätzen berechnen, da es sich sonst nicht mehr um eine Stichprobe im engen Sinne der Statistik handelt.
- E** ☐ Eine erhöhte Fallzahl führt automatisch zu mehr signifikanten Ergebnissen auch wenn der Effekt klein ist und damit nicht relevant. Dadurch sind die Informationen zur Signifikanz in riesigen Datensätzen schwer zu verwerten, da fast alle Vergleiche signifikant sind.

Multiple Gruppenvergleiche

29. Aufgabe

(2 Punkte)

Sie haben folgende unadjustierten p-Werte gegeben: 0.03, 0.02, 0.001, 0.21 und 0.34. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.15, 0.1, 0.005, 1.05 und 1.7. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.006, 0.004, 2e-04, 0.042 und 0.068. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.

- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.006, 0.004, 2e-04, 0.042 und 0.068. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.15, 0.1, 0.005, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.15, 0.1, 0.005, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.

30. Aufgabe

(2 Punkte)

Sie rechnen einen PostHoc-Test. Nun sollen Sie ein CLD erstellen. Was bedeutet dieser Fachbegriff und welche folgende Beschreibung der Interpretation ist korrekt?

- A** ☐ Contrast letter display. Unterschiede in den Behandlungen werden durch den gleichen Buchstaben oder Symbol dargestellt. Die Interpretation des CLD führt häufig in die Irre.
- B** ☐ Compact letter display. Gleiche Buchstaben bedeuten, dass sich die Behandlungen unterscheiden. Daher ist das CLD sehr unintuitiv. Es wäre besser, wenn gleiche Buchstaben Gleichheit anzeigen würden. Dies ist aber leider in der statistischen Testtheorie nicht möglich.
- C** ☐ Compact letter detection. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt.
- D** ☐ Compact letter display. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt. Teilweise ist die Interpretation des CLD herausfordernd, da wir ja nach dem Unterschied suchen.
- E** ☐ Compact letter display. Gleiche Buchstaben zeigen Gleichheit in den Behandlungen. Die Interpretation ist deshalb sehr intuitiv und einfach. Darüber hinaus ist damit das CLD auch auf einer Linie mit der Testtheorie, da wir ja auch dort die Gültigkeit der Nullhypothese nachweisen. Wir suchen ja Gleichheit.

31. Aufgabe

(2 Punkte)

Sie haben eine zweifaktorielle ANOVA gerechnet und wollen nach einem signifikanten Ergebnis in dem Gruppenfaktor einen Posthoc-Test rechnen. Welches R Paket nutzen Sie dafür und welche Eigenschaften des Paktes sind korrekt?

- A** ☐ Das R Paket {ggplot}. Wir erhalten hier sofort eine Visualisierung der Daten. Anhand der Visualisierung lässt sich eine explorative Datenanalyse durchführen, die gleichwertig zu einem Posthoc-Test ist.
- B** ☐ Das R Paket {hmisc} erlaubt die Durchführung eines multiplen Gruppenvergleichs aus verschiedenen Modellen heraus. Aus einem hmisc Objekt lässt sich recht einfach das CLD erstellen und so über Barplots eine schnelle Interpretation der statistischen Auswertung durchführen.
- C** ☐ Das R Paket {emmeans} erlaubt die Durchführung eines multiplen Gruppenvergleichs. Aus einem emmeans Objekt lässt sich leider kein CLD erstellen. Dennoch ist das Paket einfach zu bedienen und wird deshalb genutzt. Die Interpretation der statistischen Auswertung wird über einen Barplot abgebildet.
- D** ☐ Das R Paket {emmeans} erlaubt die Durchführung eines multiplen Gruppenvergleichs. Aus einem {emmeans} Objekt lässt sich recht einfach das CLD erstellen und so über Barplots eine schnelle Interpretation der statistischen Auswertung durchführen.
- E** ☐ Das R Paket {lm}. Das Paket {lm} erstellt selbstständig Konfidenzintervalle und entsprechende p-Werte. Da wir in dem Paket nicht adjustieren müssen, ist es bei Anwendern sehr beliebt.

32. Aufgabe

(2 Punkte)

In den Humanwissenschaften werden multiple Vergleiche häufig anders behandelt als in den Agrarwissenschaften. In beiden Bereichen tritt jedoch das gleiche Phänomen bei multiplen Testen auf. Wie muss mit dem Phänomen umgegangen werden und wie ist es benannt?

- A** ☐ Beim multiplen Testen kann es zu einer α -Deflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist. Die p-Werte werden durch die Anzahl an Vergleichen geteilt

- B** ☐ Beim multiplen Testen kann es zu einer α -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern sehr viel höher. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist.
- C** ☐ Beim multiplen Testen kann es zu Varianzheterogenität kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5%. Daher müssen die p-Werte entsprechend adjustiert werden. Das Verfahren nach Welch, bekannt aus dem t-Test, ist hier häufig anzuwenden.
- D** ☐ Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern sehr viel niedriger, bei ca. 1%. Es kommt zu einer α -Hyperinflation. Dagegen kann mit der Adjustierung der p-Werte nach Bonferroni vorgegangen werden.
- E** ☐ Beim multiplen Testen kann es zu einer β -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 20%. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist.

33. Aufgabe

(2 Punkte)

In einem Feldversuch haben Sie einen Behandlungsfaktor mit mehreren Levels vorliegen. Sie rechnen einen multiplen Vergleich. Vorher hatten Sie eine einfaktorielle ANOVA mit einem signifikanten Ergebnis vorliegen. Welche Aussage ist richtig?

- A** ☐ Beim multiplen Testen kann es zu einer Effektüberschätzung (Δ -Inflation) kommen. Daher müssen die Effekte angepasst werden. Dies geschieht nicht händisch sondern intern in den angewendeten Algorithmen.
- B** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ nach Bonferroni adjustiert werden. Dafür wird der Effekt mit der Anzahl an Vergleichen k multipliziert. Dies geschieht analog zu den p-Werten.
- C** ☐ Beim multiplen Testen kann es zu einer Δ -Inflation kommen. Das globale Effektniveau liegt nicht mehr bei 20%. Daher müssen die Effekte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der Effekte nach Bonferroni das bekannteste Verfahren ist.
- D** ☐ Beim multiplen Testen muss der Effekt, wie der Mittelwertsunterschied Δ aus einem t-Test, nicht adjustiert werden.
- E** ☐ Beim multiplen Testen kann es zu einer Δ -Deflation kommen. Das globale Relevanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die Δ -Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der Δ -Werte nach Bonferroni das bekannteste Verfahren ist. Die Δ -Werte werden durch die Anzahl an Vergleichen geteilt.

Statistische Testtheorie

34. Aufgabe

(2 Punkte)

Geben ist $Pr(D|H_0)$ als mathematischer Ausdruck, welche Aussage ist richtig?

- A** ☐ Die Wahrscheinlichkeit für die Nullhypothese, wenn die Daten wahr sind.
- B** ☐ $Pr(D|H_0)$ beschreibt die Wahrscheinlichkeit die Teststatistik T_D aus den Daten D zu beobachten, wenn die Nullhypothese wahr ist.
- C** ☐ Die Inverse der Wahrscheinlichkeit unter der die Nullhypothese nicht mehr die Alternativhypothese überdeckt.
- D** ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit nicht die Daten D zu beobachten sondern die Nullhypothese, wenn diese wahr ist.
- E** ☐ Die Wahrscheinlichkeit der Daten unter der Nullhypothese in der Grundgesamtheit.

35. Aufgabe

(2 Punkte)

Die Testtheorie hat mehrere Säulen. Einer der Säulen ist das Falsifikationsprinzip. Das Falsifikationsprinzip besagt,

- A** ☐ ... dass Fehlerterme in statistischen Modellen nicht verifiziert werden können.
- B** ☐ ... dass ein schlechtes Modell durch ein schlechteres Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.
- C** ☐ ... dass ein schlechtes Modell durch ein weniger schlechtes Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.
- D** ☐ ... dass ein schlechtes Modell durch das Falsifikationsprinzip durch ein noch schlechteres Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.
- E** ☐ ... dass in der Wissenschaft immer etwas falsch sein muss. Sonst gebe es keinen Fortschritt.

36. Aufgabe

(2 Punkte)

Der Fehler 1. Art oder auch Signifikanzniveau α genannt, liegt bei 5%. Welcher der folgenden Gründe für diese Festlegung auf 5% als Signifikanzschwelle ist richtig?

- A** ☐ Der Wert ergab sich aus einer Auswertung von 1042 wissenschaftlichen Veröffentlichungen zwischen 1914 und 1948. Der Wert 5% wurde in 28% der Veröffentlichungen genutzt. Daher legte man sich auf diese Zahl fest.
- B** ☐ Da Wissenschaftler eine Schwelle für die statistische Testentscheidung benötigen wurde α in einer großen Konferenz 1945 gewählt. Damit ist $\alpha = 5\%$ eine Kulturkonstante mit einem Rang einer Naturkonstante.
- C** ☐ Da Wissenschaftler eine Schwelle für die statistische Testentscheidung benötigen wurde α historisch gewählt. Damit ist $\alpha = 5\%$ eine Kulturkonstante.
- D** ☐ Der Begründer der modernen Statistik, R. Fischer, hat die Grenze simuliert und berechnet. Dadurch ergibt sich dieser optimale Cut-Off.
- E** ☐ Als Kulturkonstante hat $\alpha = 5\%$ den Rang einer Naturkonstante und wurde nach langer Diskussion in der UN im Jahre 1983 festgesetzt. Damals auch schon mit der Zustimmung der UdSSR.

37. Aufgabe

(2 Punkte)

Betrachten wir die Teststatistik aus einem abstrakteren Blickwinkel. Beim statistischen Testen wird das „*signal*“ mit dem „*noise*“ aus den Daten D zu einer Teststatistik T_D verrechnet. Welche der Formel berechnet korrekt die Teststatistik T_D ?

- A** ☐ Es gilt $T_D = \text{signal} \cdot \text{noise}$
- B** ☐ Es gilt $T_D = \frac{\text{signal}}{\text{noise}}$
- C** ☐ Es gilt $T_D = (\text{signal} \cdot \text{noise})^2$
- D** ☐ Es gilt $T_D = \frac{\text{signal}}{\text{noise}^2}$
- E** ☐ Es gilt $T_D = \frac{\text{noise}}{\text{signal}}$

38. Aufgabe

(2 Punkte)

Eine Analogie kann helfen einen Sachverhalt besser zu verstehen. Wie kann folgende Aussage richtig in die Analogie der statistischen Testtheorie gesetzt werden?

H_0 beibehalten obwohl die H_0 falsch ist

- A ☐ In die Analogie eines Rauchmelders: *Alarm without fire*, dem α -Fehler.
- B ☐ Dem β -Fehler mit der Analogie eines Rauchmelders: *Fire without alarm*.
- C ☐ In die Analogie eines brennenden Hauses ohne Rauchmelder: *House without noise*.
- D ☐ In die Analogie eines Feuerwehrautos: *Car without noise*.
- E ☐ *Alarm with fire*, dem α -Fehler in der Analogie von Feuer.

39. Aufgabe

(2 Punkte)

Welche statistische Maßzahl erlaubt es Relevanz mit Signifikanz zu verbinden? Welche Aussage ist richtig?

- A ☐ Einem Konfidenzintervall. Das Konfidenzintervall bringt durch eine Visualisierung und zwei Intervallgrenzen die Möglichkeit mit, eine Relevanzschwelle neben der definierten Signifikanzschwelle zu definieren.
- B ☐ Über das Konfidenzintervall. Das Konfidenzintervall inkludiert eine Entscheidung über die Relevanz und zusätzlich kann über die Visualisierung des Konfidenzintervalls eine Signifikanzschwelle vom Forschenden definiert werden.
- C ☐ Das OR. Als Chancenverhältnis gibt es das Verhältnis von Relevanz und Signifikanz wieder.
- D ☐ Einem Konfidenzintervall. Das Konfidenzintervall bringt durch eine Visualisierung und drei Intervallgrenzen die Möglichkeit mit, eine Relevanzschwelle neben der Signifikanzschwelle und der α -Schwelle zu definieren.
- E ☐ Die Teststatistik. Durch den Vergleich von T_c zu T_k ist es möglich die H_0 abzulehnen. Die Relevanz ergibt sich aus der Fläche rechts vom dem T_c -Wert.

40. Aufgabe

(2 Punkte)

Welche Aussage über den p -Wert und dem Signifikanzniveau α gleich 5% ist richtig?

- A ☐ Wir schauen, ob der p -Wert größer ist als das Signifikanzniveau α und vergleichen somit Wahrscheinlichkeiten. Die Wahrscheinlichkeiten werden als Flächen unter der Kurve der Teststatistik dargestellt, wenn die H_A gilt.
- B ☐ Wir vergleichen mit dem p -Wert und dem Signifikanzniveau α absolute Werte auf einem Zahlenstrahl und damit den Unterschied der Teststatistiken, wenn die H_0 gilt.
- C ☐ Wir vergleichen mit dem p -Wert und dem Signifikanzniveau α Wahrscheinlichkeiten und damit die absoluten Werte auf einem Zahlenstrahl, wenn die H_0 gilt.
- D ☐ Wir vergleichen die Effekte des p -Wertes mit den Effekten der Signifikanzschwelle unter der Annahme der Nullhypothese. Dabei gilt, dass wir die Nullhypothese nur ablehnen können anhand des Falsifikationsprinzips.
- E ☐ Wir schauen, ob der p -Wert kleiner ist als das Signifikanzniveau α und vergleichen somit Wahrscheinlichkeiten. Die Wahrscheinlichkeiten werden als Flächen unter der Kurve der Teststatistik dargestellt, wenn die H_0 gilt.

41. Aufgabe

(2 Punkte)

Um die Testtheorie besser zu verstehen, mag es manchmal sinnvoll sein ein Beispiel aus dem Alltag zu wählen. Die Ergebnisse der Analyse durch einen statistischen Test können auch in grobe Analogie zur Wettervorhersage gebracht werden. Welche Aussage trifft am ehesten zu?

- A** ☐ In der Analogie der Durchschnittstemperatur: Wie oft tritt ein Effekt durchschnittlich ein? Wir erhalten eine Wahrscheinlichkeit für die Effekte. Zum Beispiel, wie hoch ist die Wahrscheinlichkeit für einen Mittelwert als Durchschnitt.
- B** ☐ In der Analogie der Regenwahrscheinlichkeit: ein statistischer Test gibt die Wahrscheinlichkeit für das Auftreten eines Ereignisses wieder. Die Stärke des Effektes wird nicht wiedergeben.
- C** ☐ In der Analogie der Maximaltemperatur: Was ist der maximale Unterschied zwischen zwei Gruppen. Wir erhalten hier eine Aussage über die Spannweite und den maximalen Effekt.
- D** ☐ In der Analogie der Regenwahrscheinlichkeit in einem bestimmten Gebiet: ein statistischer Test gibt die Wahrscheinlichkeit für ein Ereignis in einem Experiment mit den Daten D wieder und lässt sich kaum verallgemeinern.
- E** ☐ In der Analogie der Sonnenscheindauer: Wie lange kann mit einem entsprechenden Effekt gerechnet werden? Die Wahrscheinlichkeit für den Effekt gibt der statistische Test wieder.

42. Aufgabe

(2 Punkte)

In Ihrer Forschungsarbeit wollen Sie eine Aussage über ein untersuchtes Individuum treffen. Dazu nutzen Sie eine ANOVA als statistischen Test. Erhalten Sie eine valide Aussage aus einem statistischen Test?

- A** ☐ Ja, wir erhalten eine Aussage. Müssen aber das Individuum im Kontext der Population adjustieren.
- B** ☐ Ja, wir können ein untersuchtes Individuum nicht mit einem t-Test auswerten. Wir erhalten keine Aussage zum Individuum. Wir können aber den Effekt als Quelle der Relevanz nutzen.
- C** ☐ Ja, wir können ein untersuchtes Individuum mit einer ANOVA auswerten. Wir erhalten eine Aussage zum Individuum.
- D** ☐ Ja, wir erhalten nur eine Aussage zu zwei Individuen. Ein statistischer Test liefert Informationen zu einem Individuum im Vergleich zu einem anderen Individuum.
- E** ☐ Nein, wir können ein untersuchtes Individuum nicht mit einer ANOVA auswerten. Wir erhalten keine Aussage zum Individuum.

43. Aufgabe

(2 Punkte)

In der statistischen Testtheorie gibt es den Begriff *Power*. Was sagt der statistische Begriff *Power* aus?

- A** ☐ Die Power $1 - \beta$ wird auf 80% gesetzt. Damit liegt die Wahrscheinlichkeit für die H_0 bei 20%.
- B** ☐ Es gilt $\alpha + \beta = 1$ und somit liegt β meist bei 95%.
- C** ☐ Die Power beschreibt die Wahrscheinlichkeit die H_A abzulehnen. Wir testen die Power jedoch nicht.
- D** ☐ Die Power wird nicht berechnet sondern ist eine Eigenschaft des Tests. Die Power wird auf 80% gesetzt und beschreibt mit welcher Wahrscheinlichkeit H_A *bewiesen wird*
- E** ☐ Die Power ist nicht in der aktuellen Testthorie mehr vertreten. Wir rechnen nur noch mit dem Fehler 1. Art.

44. Aufgabe

(2 Punkte)

Sie rechnen einen statistischen Test und erhalten neben dem p-Wert noch einen Effekt wiedergegeben. Welche Aussage zum Effekt ist richtig?

- A** ☐ Der Forschende muss am Anfang wissen, ob das Ergebnis eines Experiments relevant für seine Forschung ist. Dafür kann der Effekt eines statistischen Tests genutzt werden oder auch der Prähoc-Test. Damit beschreibt der Effekt den biologischen interpretierbaren Teil eines Experiments vor der Durchführung. Zum Beispiel der Unterschied zwischen zwei Mittelwerten.
- B** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Zum Beispiel den mittleren Unterschied zwischen zwei Gruppen aus einem t-Test. Damit ist der Effekt direkt mit dem Begriff der Relevanz verbunden. Die Entscheidung über die Relevanz trifft der Forschende unabhängig von der Signifikanz eines statistischen Tests.
- C** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Modernen Algorithmen liefern keine Effekte mehr sondern nur noch bedingte Wahrscheinlichkeiten. Der Effekt spielt in der modernen Statistik keine Rollen mehr.
- D** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Damit ist der Effekt direkt mit dem Begriff der Signifikanz verbunden. Die Entscheidung über die Signifikanz trifft der Forschende unabhängig von der Relevanz eines statistischen Tests.
- E** ☐ Der Effekt eines statistischen Tests beschreibt den Output oder die Wiedergabe eines Tests in einem Computer.

45. Aufgabe

(2 Punkte)

Welche Aussage über die Entscheidung anhand des 95%-Konfidenzintervalls gegen die Nullhypothese ist richtig?

- A** ☐ Anhand des 95%-Konfidenzintervalls lässt sich wie folgt eine Entscheidung treffen. Liegt der Wert über oder gleich dem Signifikanzniveau α dann kann die Nullhypothese abgelehnt werden.
- B** ☐ Anhand des 95%-Konfidenzintervalls lässt sich wie folgt eine Entscheidung treffen. Liegt der Wert in dem Signifikanzniveauintervall α dann kann die Nullhypothese abgelehnt werden.
- C** ☐ Ist T_D höher als der kritische Wert $T_{\alpha=5\%}$ dann wird die Nullhypothese H_0 abgelehnt.
- D** ☐ Ist $Pr(D|H_0)$ kleiner als das Signifikanzniveau α gleich 5% dann wird die Nullhypothese H_0 abgelehnt.
- E** ☐ Ist in dem 95%-Konfidenzintervall nicht die Null enthalten dann wird die Nullhypothese H_0 abgelehnt.

46. Aufgabe

(2 Punkte)

Wenn Sie im Allgemeinen einen statistischen Test rechnen, dann kommen Sie um eine statistische Hypothese H nicht herum. Welche Aussage über statistische Hypothesen ist richtig?

- A** ☐ Es gibt ein statistisches Hypothesenpaar mit der Hypothese für und gegen die wissenschaftliche Fragestellung. Die Hypothesen werden H_{pro} und H_{contra} bezeichnet.
- B** ☐ Mit der Nullhypothese H_A und der Alternativhypothese H_0 gibt es zwei Hypothesen, die aber selten genutzt werden.
- C** ☐ Es gibt ein Hypothesenset bestehend aus k Hypothesen. Meistens wird die Nullhypothese H_0 und die Alternativhypothese H_A verwendet. Wegen des Falsifikationsprinzips ist es wichtig, die bekannte falsche und unbekannte richtige Hypothese mit in das Set zu nehmen.
- D** ☐ Die Hypothesen H_0 und H_A sind rein prosarischer Natur und bilden keinen mathematischen Hintergrund ab. In der Statistik wird die wissenschaftliche Fragestellung getestet. Daher stehen auch die verständlichen Hypothesen im Mittelpunkt der biologischen Interpretation.
- E** ☐ Mit der Nullhypothese H_0 und der Alternativhypothese H_A oder H_1 gibt es zwei Hypothesen.

Statistische Tests für Gruppenvergleiche

47. Aufgabe

(2 Punkte)

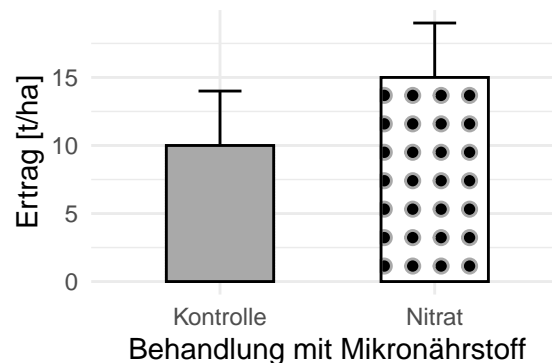
In Ihrer Abschlussarbeit rechnen Sie einen Student t-Test. Welche Aussage ist auch für den Welch t-Test richtig?

- A** ☐ Der t-Test ist ein Vortest der ANOVA und basiert daher auf dem Vergleich von Streuungsparametern
- B** ☐ Der t-Test vergleicht zwei Gruppen indem die Mittelwerte miteinander verglichen werden.
- C** ☐ Der t-Test testet generell zu einem erhöhten α -Niveau von 20%.
- D** ☐ Der t-Test berechnet die Differenz von zwei Mittelwerten als Effekt und gibt eine Entscheidung, ob sich die beiden Mittelwerte *jeweils* von Null unterscheiden.
- E** ☐ Der t-Test vergleicht die Mittelwerte von zwei Gruppen unter der strikten Annahme von Varianzhomogenität. Sollte keine Varianzhomogenität vorliegen, so gibt es keine Möglichkeit den t-Test in einer Variante anzuwenden.

48. Aufgabe

(2 Punkte)

In einer Studie zur Bewertung der Wirkung des Mikronährstoff Nitrat auf den Ertrag in t/ha von Weizen im Vergleich zu einer Kontrolle entstand folgende Abbildung. Der Versuch wurde in 11 Parzellen pro Gruppe durchgeführt. Welche Aussage ist im Bezug auf einen t-Test ist richtig?



- A** ☐ Der Test deutet auf einen signifikanten Unterschied hin. Der Effekt liegt vermutlich bei 5.
- B** ☐ Die Barplots deuten auf keinen signifikanten Unterschied. Der Effekt liegt vermutlich bei 5 unter einer groben Abschätzung. Wir müssen aber eine ANOVA rechnen um den Effekt wirklich bestimmen zu können.
- C** ☐ Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt kann nicht bei einem t-Test aus Barplots bestimmt werden.
- D** ☐ Die Barplots deuten auf keinen signifikanten Unterschied. Der Effekt liegt vermutlich bei 5 unter einer groben Abschätzung.
- E** ☐ Der Effekt und die Signifikanz lassen sich nicht aus Barplots abschätzen. Höchstens der Effekt als relativer Unterschied zwischen der Höhe der Barplots. Standard ist der mediane Unterschied aus Boxplots.

49. Aufgabe

(2 Punkte)

Welche Aussage über den gepaarten t-Test für verbundene Stichproben ist richtig?

- A** ☐ Abhängige Beobachtungen müssen gesondert in einem gepaarten t-Test modelliert werden. Wenn wiederholt an dem gleichen Tier oder Pflanze gemessen wird, dann bilden wir den Quotienten zwischen den beiden Zeitpunkten. Auf den Quotienten rechnen wir den gepaarten t-Test.
- B** ☐ Der gepaarte t-Test wird genutzt, wenn die Differenzen der Beobachtungen verbunden sind und wir dadurch die Unabhängigkeit nicht mehr vorliegen haben.

- C** ☐ Wenn die Beobachtungen nicht unabhängig voneinander sind, rechnen wir einen gepaarten t-Test. Messen wir wiederholt an dem gleichen Tier oder Pflanze dann bilden wir die Differenz zwischen den zwei Messpunkten.
- D** ☐ Beim gepaarten t-Test kombinieren wir die Vorteile des Student t-Test für Varianzhomogenität mit den Vorteilen des Welch t-Test für Varianzheterogenität. Wir bilden dafür die Differenz der Einzelbeobachtungen.
- E** ☐ Wenn die Beobachtungen unabhängig voneinander sind, rechnen wir einen gepaarten t-Test. Messen wir wiederholt an dem gleichen Tier oder Pflanze dann bilden wir das Produkt zwischen den zwei Messpunkten.

50. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit passen die Ergebnisse einer ANOVA und eines multiplen Vergleiches nicht zusammen. Nach einem Experiment mit fünf Maissorten ergibt eine ANOVA ($p = 0.045$). Sie führen anschließend die paarweisen t-Tests für alle Vergleiche durch. Nach der Adjustierung für multiples Testen ist kein p-Wert unter der α -Schwelle. Sie schauen sich auch die rohen, unadjustierten p-Werte an und finden hier als niedrigsten p-Wert $p_{3-2} = 0.052$. Welche Aussage ist richtig?

- A** ☐ Das Beispiel kann so nicht auftreten, da die ANOVA und die t-Tests algorithmisch miteinander verschränkt sind.
- B** ☐ Hier kommt der Effekt der steigenden Fallzahl auf die Anzahl an signifikante Ergebnisse zu tragen. Da die ANOVA auf mehr Fallzahl testet als die einzelnen paarweisen t-Tests, kann die ANOVA leichter einen signifikanten Unterscheid nachweisen. Die p-Werte sind immer etwas kleiner als bei den t-Tests.
- C** ☐ Das ist kein Wunder. Die ANOVA testet nicht auf der gesamten Fallzahl und die paarweisen t-Tests gewinnen immer eine oder mehr Gruppen als Fallzahl dazu. Mit steigender Fallzahl sind mehr signifikante Unterschiede zu erwarten. Die p-Werte unterscheiden sich numerisch auch kaum.
- D** ☐ Die ANOVA testet auf der gesamten Fallzahl. Es wäre besser die ANOVA auf der gleichen Fallzahl wie die einzelnen t-Tests zu rechnen.
- E** ☐ Der Fehler liegt in den t-Tests. Wenn eine ANOVA signifikant ist, dann muss zwangsweise auch ein t-Test signifikant sein.

Teil I.

Deskriptive Statistik & Explorative Datenanalyse

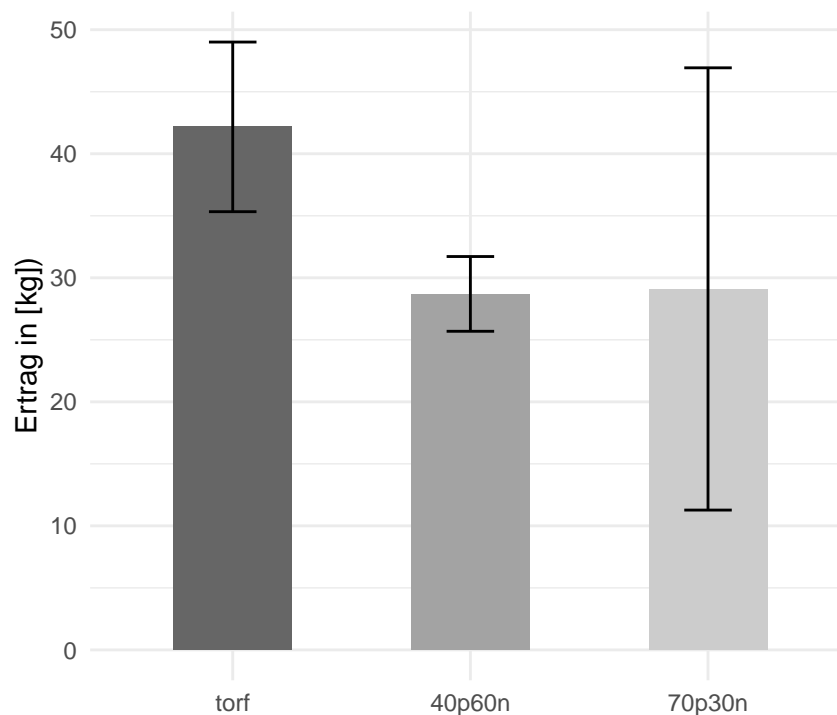
51. Aufgabe

(8 Punkte)


Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Zerforschen des Barplots Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Jessica ist es eine Möglichkeit schneller ans Ziel zu gelangen. Jessica soll in ihrer Abschlussarbeit Brokoli untersuchen. Die Behandlung in ihrer Abschlussarbeit werden verschiedene Substrattypen (*torf*, *40p60n* und *70p30n*) sein. Erheben wird Jessica als Outcome (Y) *Ertrag* benannt als *yield* in ihrer Exceldatei. Von ihrer Betreuerin erhält sie nun folgende Abbildung von Barplots, die sie erstmal zur Übung nachbauen soll, bevor sie mit dem eigentlichen Versuch beginnt.



Leider kennt sich Jessica mit der Erstellung von Barplots in  nicht aus. Deshalb braucht sie bei der Visualisierung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Erstellen Sie eine Tabelle mit den statistischen Maßzahlen der drei Barplots! *Beachten Sie die korrekte Darstellungsform der statistischen Maßzahlen!* **(3 Punkte)**
3. Erstellen Sie einen beispielhaften Datensatz im  üblichen Format, aus dem die drei Barplots *möglicherweise* erstellt wurden! **(2 Punkte)**
4. Kann Jessica einen Unterschied zwischen den Behandlungen erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

52. Aufgabe

(8 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Barplots Jonas steht vor einem ersten Problem, denn wenn es nach seiner Betreuerin geht, soll er in einem Feldexperiment Kartoffeln auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Stricken. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Jonas denkt gerne über Stricken nach. Die Behandlung waren verschiedene Bewässerungstypen (*low*, *mid* und *high*). In seiner Exceldatei hat er den Messwert (*Y*) *Ertrag* als *yield* aufgenommen. Nun soll Jonas die Daten einmal als Barplots in einer Präsentation visualisieren, damit seiner Betreuerin wieder klar wird, was er eigentlich nochmal gemacht hat und was für ein Ergebnis in einem statistischen Test zu erwarten wäre. Wäre da nicht noch etwas. Jonas und die Erschöpfung, eine unendliche Geschichte mit kniffligen Wendungen. Aber egal. Einfach mal raus um zu Schwimmen. Ohne Ziel und Uhr. Das ist was für Jonas.

treatment	yield
low	42.1
mid	34.3
mid	41.2
high	15.6
low	35.0
high	20.0
mid	34.1
low	32.9
mid	36.5
low	47.6
high	27.4
low	24.4

Leider kennt sich Jonas mit der Erstellung von Barplots nicht aus. Deshalb braucht er bei der Visualisierung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Zeichnen Sie in *einer* Abbildung die Barplots für die Behandlung von Kartoffeln! Beschriften Sie die Achsen entsprechend! **(4 Punkte)**
3. Beschriften Sie *einen* Barplot mit den gängigen statistischen Maßzahlen! **(2 Punkte)**
4. Wenn Jonas *keinen Effekt* zwischen den Behandlungen von Kartoffeln erwarten würde, wie sehen dann die Barplots aus? Antworten Sie mit einer Skizze der Barplots! **(1 Punkt)**

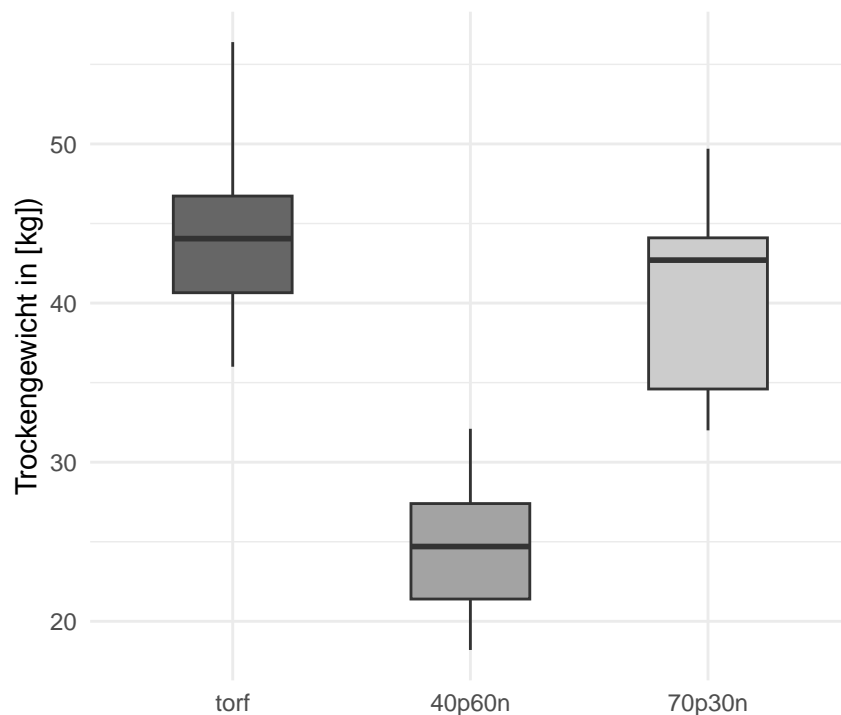
53. Aufgabe


(9 Punkte)


Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Zerforschen des Boxplots Wenn der Mangel nicht wäre, ja dann wäre wohl vieles möglich für Jessica! Aber so.. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Jessica ist es eine Möglichkeit schneller ans Ziel zu gelangen. Jessica soll in ihrem Projektbericht Erdbeeren untersuchen. Die Behandlung in ihrem Projektbericht werden verschiedene Substrattypen (*torf*, *40p60n* und *70p30n*) sein. Erheben wird Jessica als Outcome (Y) *Trockengewicht* benannt als *drymatter* in ihrer Exceldatei. Von ihrer Betreuerin erhält sie nun folgende Abbildung von Boxplots, die sie erstmal zur Übung nachbauen soll, bevor sie mit dem eigentlichen Versuch beginnt. Aber nur in passender Atmosphäre! Schon dutzende Male gesehen: Herr der Ringe. Aber immer noch großartig zusammen mit Schokobons.



Leider kennt sich Jessica mit der Erstellung von Boxplots in  nicht aus. Deshalb braucht sie bei der Visualisierung Ihre Hilfe!

1. Erstellen Sie eine Tabelle mit den statistischen Maßzahlen aus der obigen Abbildung der drei Boxplots! Beachten Sie die korrekte Darstellungsform der statistischen Maßzahlen! **(3 Punkte)**
2. Beschriften Sie *einen* der Boxplots mit den gängigen statistischen Maßzahlen! **(2 Punkte)**
3. Erstellen Sie einen beispielhaften Datensatz, aus dem die drei Boxplots *möglicherweise* erstellt wurden, im  üblichen Format! **(2 Punkte)**
4. Kann Jessica einen Unterschied zwischen den Behandlungen erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

54. Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Boxplots Tina steht vor einem ersten Problem, denn wenn es nach ihrem Betreuer geht, soll sie in einem Feldexperiment Lauch auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Am Ende dann doch besser Astronomie. Wunderbar. Eine echte Ablenkung für Tina. Die Behandlung waren verschiedene Lüftungssystemen und Folientunneln (*ctrl* und *tornado*). In ihrer Exceldatei hat sie den Outcome (Y) *Proteingehalt* als *protein* aufgenommen. Nun soll Tina die Daten einmal als Boxplots in einer Präsentation visualisieren, damit ihrem Betreuer wieder klar wird, was sie eigentlich nochmal gemacht hat und was für ein Ergebnis in einem statistischen Test zu erwarten wäre. Anhand von Boxplots lässt sich eine Aussage über die Varianzhomogenität über die Behandlungsgruppen treffen. Wäre da nicht noch etwas. Wenn die Wut nicht wäre, ja dann wäre wohl vieles möglich für Tina! Aber so.. Aber egal. Um zu Boxen geht Tina dann später nochmal raus. Echte Entspannung.

treatment	drymatter
tornado	34.7
ctrl	29.3
tornado	37.2
ctrl	23.1
ctrl	19.9
tornado	31.4
ctrl	32.9
ctrl	11.2
tornado	39.1
tornado	27.4
ctrl	18.1
tornado	33.5
tornado	42.3
ctrl	30.0
tornado	31.3
ctrl	18.5
ctrl	19.6
tornado	25.2
tornado	36.9

Leider kennt sich Tina mit der Erstellung von Boxplots nicht aus. Deshalb braucht sie bei der Visualisierung Ihre Hilfe!

1. Zeichnen Sie in *einer* Abbildung die beiden Boxplots für die zwei Behandlungen von Lauch! Beschriften Sie die Achsen entsprechend! **(5 Punkte)**
2. Wie ist Ihr Vorgehen, wenn Sie eine *gerade* Anzahl an Beobachtungen pro Gruppe haben? **(1 Punkt)**
3. Beschriften Sie *einen* der beiden Boxplots mit den gängigen statistischen Maßzahlen! **(2 Punkte)**
4. Wenn Sie *keinen Effekt* zwischen den Behandlungen von Lauch erwarten würden, wie sehen dann die beiden Boxplots aus? Antworten Sie mit einer Skizze der Boxplots! **(1 Punkt)**

55. Aufgabe

(8 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Histogramm für kategoriale Daten In einem Gespräch mit ihrer Betreuerin wird Jessica gebeten seine Daten aus einem Kreuzungsexperiment mit Lamas in einem Histogramm darzustellen. Jessica schmeißt noch eine Handvoll Schokobons in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von David Bowie. In ihrem Experiment hat er die Anzahl an weißen Blutkörperchen erst fotografiert und dann ausgezählt. Laut ihrer Betreuerin soll das Histogramm helfen, die Verteilung der die Anzahl an weißen Blutkörperchen zu bestimmen. Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für sie war schon immer der Mangel gewesen. Ein leidiges Lied. Wenn David Bowie ertönt, dann sucht die Hündin schleunigst Schutz unter dem Sofa. Jessica schüttelt den Kopf.

Die Anzahl an weißen Blutkörperchen: 5, 2, 5, 2, 1, 4, 3, 6, 3, 4, 4, 4, 2, 6, 4, 3, 6, 3, 4, 2, 6, 6, 2, 8, 7, 5, 6, 1, 3, 3, 8, 7, 2, 6, 2, 3

Leider kennt sich Jessica mit der Erstellung von Histogrammen überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Zeichnen Sie ein Histogramm um die Verteilung der Daten zu visualisieren! **(3 Punkte)**
2. Beschriften Sie die Achsen der Abbildung! **(2 Punkte)**
3. Ergänzen Sie die absoluten und relativen Häufigkeiten in der Abbildung! **(1 Punkt)**
4. Berechnen Sie aus den Daten die *Wahrscheinlichkeit* gleich oder mehr als die Anzahl 6 zu beobachten! **(1 Punkt)**
5. Berechnen Sie aus den Daten die *Chance* gleich oder mehr als die Anzahl 6 zu beobachten! **(1 Punkt)**

56. Aufgabe

(8 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Histogramm für kontinuierliche Daten In ihrer Hausarbeit möchte Yuki gerne die Daten aus einem Stallexperiment mit Hühnern in einem Histogramm darstellen. Das Histogramm erlaubt ihr dabei Rückschlüsse auf die Verteilung über den Endpunkt (Y) zu treffen Yuki schmeißt noch eine Handvoll Reese's Peanut Butter Cups in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von London Grammar. In seinem Experiment hat Yuki die mittleren dunklen Pigmentstörungen gezählt. Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für sie war schon immer die Faulheit gewesen. Ein leidiges Lied. Wenn London Grammar ertönt, dann sucht das Minischwein schleunigst Schutz unter dem Sofa. Yuki schüttelt den Kopf.

Die mittleren dunklen Pigmentstörungen: 10.3, 8.8, 9, 8.5, 11.2, 8.3, 10.9, 6.6, 10.7, 9.2, 11.8, 10.1, 11.9, 11.2, 13.9, 10.8, 11.2, 8.6, 5.9, 12.3, 11.8, 11.8

Leider kennt sich Yuki mit der Erstellung von Histogrammen überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Zeichnen Sie ein Histogramm um die Verteilung der Daten zu visualisieren! **(3 Punkte)**
2. Erläutern Sie Ihr Vorgehen um ein Histogramm für kontinuierliche Daten zu zeichnen! **(2 Punkte)**
3. Beschriften Sie die Achsen der Abbildung! **(2 Punkte)**
4. Ergänzen Sie die relativen Häufigkeiten in der Abbildung! **(1 Punkt)**

57. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Scatterplots Steffen liest laut: 'Wenn zwei kontinuierliche Variablen vorliegen, können diese in einer explorativen Datenanalyse...'. Steffen stoppt. Aus den Boxen wummert Taylor Swift und sein Mund ist verklebt von Oreos. 'Herrlich', denkt Steffen. Was waren noch gleich kontinuierliche Variablen? In seiner Hausarbeit hatte er einen Versuch in einer Klimakammer im Wendland durchgeführt. Dabei ging es um den Zusammenhang zwischen Proteingehalt [g/kg] und durchschnittlichen Anteil an Ton [%/l] im groben Kontext von Erdbeeren. Nun stellt sich die Frage für ihn, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Dafür war eine explorative Datenanalyse gut! Eine echte Herausforderung für ihn war schon immer die Romantik gewesen. Ein leidiges Lied. Dann was anderes. Wenn Harry Potter läuft, dann ist die Schlange nicht mehr da. Aber jetzt braucht er mal Entspannung!

Durchschnittlichen Anteil an Ton [%/l]	Proteingehalt [g/kg]
31.3	42.4
26.8	36.6
18.9	24.0
33.4	43.8
28.8	42.1
31.0	41.8
28.9	36.8
20.5	31.3
24.2	35.5
21.2	27.5
18.7	26.8
28.3	39.3

Leider kennt sich Steffen mit der Erstellung einer explorativen Datenanalyse für kontinuierliche Daten überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Erstellen Sie eine Visualisierung für die Datentabelle. Beschriften Sie die Achsen entsprechend! **(4 Punkte)**
2. Schätzen Sie eine Gerade durch die Punkte! **(1 Punkt)**
3. Beschriften Sie die Gerade mit den gängigen statistischen Maßzahlen! Geben Sie die numerischen Zahlenwerte mit an! **(3 Punkte)**
4. Wenn *ein* Effekt von x auf y vorhanden wäre, wie würde die Gerade verlaufen und welche Werte würden die statistischen Maßzahlen annehmen? **(2 Punkt)**

58. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Mosaicplots Zwei kategoriale Variablen darzustellen ist nicht so einfach. Steffen hatte erst über einen Mittelwert nachgedacht, dann aber die Idee verworfen. Wäre da nicht noch was anderes. Eine echte Herausforderung für ihn war schon immer die Romantik gewesen. Ein leidiges Lied. Dabei hatte er sich in ein Gewächshausexperiment im Emsland zum einen die Behandlung KI-gesteuert [ja/nein] und zum anderen die Messung Chlorophyllgehalt unter Zielwert [ja/nein] im Kontext von Erbsen angeschaut. Jetzt möchte seine Betreuerin erstmal die langen Tabellen mit ja/nein in einer explorativen Datenanalyse zusammengefasst bekommen. Sonst geht es bei seiner Hausarbeit nicht weiter. Was super nervig ist. Einfach mal raus um zu Ringen. Ohne Ziel und Uhr. Das ist was für Steffen.

Chlorophyllgehalt KI-gesteuert unter Zielwert		Chlorophyllgehalt KI-gesteuert unter Zielwert	
ja	nein	nein	nein
nein	ja	nein	nein
nein	ja	ja	ja
nein	ja	ja	nein
nein	nein	ja	ja
ja	nein	ja	ja
ja	nein	ja	ja
nein	ja	ja	ja
ja	ja	ja	nein
nein	ja	ja	ja
ja	ja	ja	ja
nein	ja	ja	ja
ja	nein	nein	ja
nein	nein	nein	ja
nein	nein	nein	ja
ja	ja	ja	ja
nein	ja	nein	ja
ja	nein	ja	ja
nein	ja	nein	nein

Leider kennt sich Steffen mit der Erstellung einer explorativen Datenanalyse für kategoriale Daten überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Stellen Sie den Zusammenhang zwischen den beiden kategorialen Variablen in einer zusammenfassenden Tabelle dar! **(3 Punkte)**
2. Visualisieren Sie den Zusammenhang zwischen den beiden kategorialen Variablen! **(3 Punkte)**
3. Berechnen Sie die Verhältnisse in der Visualisierung! Welche Annahme haben Sie getroffen? **(2 Punkte)**
4. Wenn *ein* Effekt von der Behandlung vorliegen würde, wie würde die Tabelle und die Visualisierung aussehen? **(2 Punkt)**

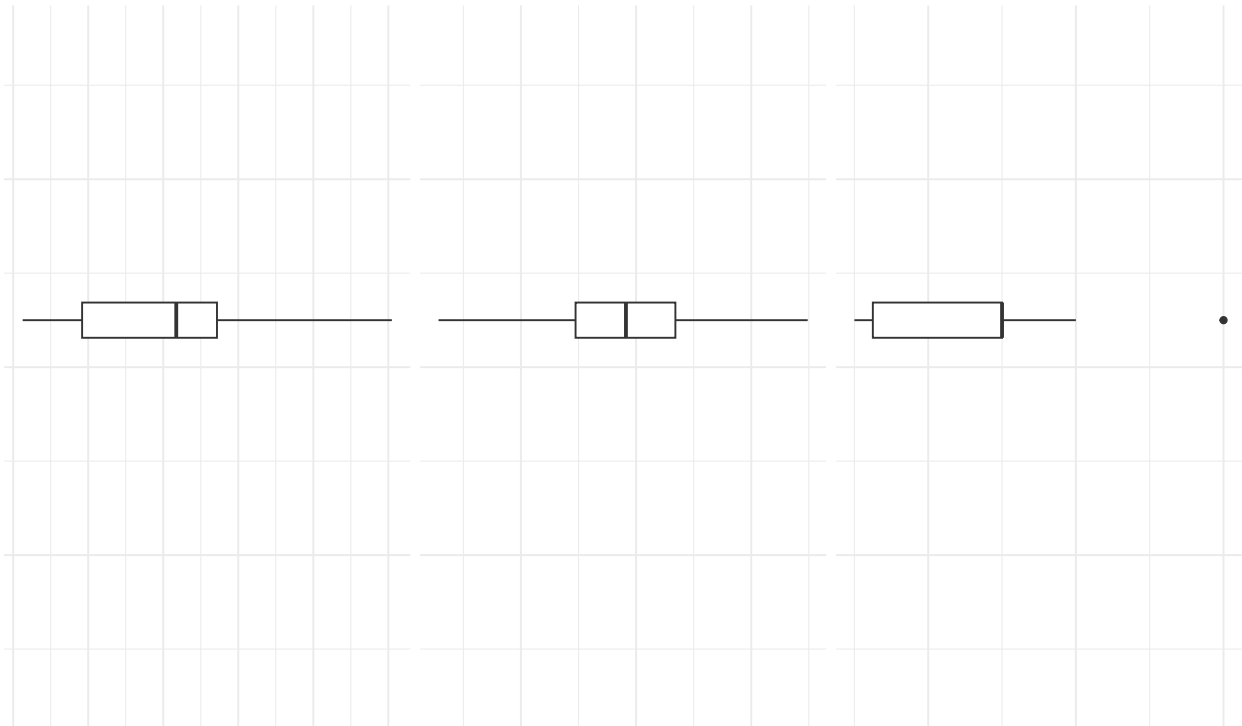
59. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung von Verteilungen 'Was hast du dir denn da hingeklebt? *Frei ist, wer missfallen kann.*¹', liest Mark vom Kühlschrank vor. Mark und Alex sitzen zusammen in der Küche und versuchen zu verhindern, dass der Hamster den Biomüll mampft. 'Können wir uns auf die etwas kryptische Aufgabe konzentrieren?', nöhlt Alex. Die beiden schauen angestrengt auf die drei Boxplots. Das Ziel ist es zu verstehen, wie eine Verteilung anhand eines Boxplots bewertet werden kann. Alex und die Gefälligkeit machen die Sache nicht einfacher.



Jetzt brauchen Mark und Alex Ihre Hilfe bei der Abschätzung einer Verteilung anhand von Boxplots um ihre Arbeit dann in diesem Semester noch abschließen zu können.

1. Zeichnen Sie über die Boxplots die entsprechende zugehörige Verteilung! **(3 Punkte)**
2. Zeichnen Sie unter die Boxplots die entsprechende zugehörige Beobachtungen als Stiche! **(3 Punkte)**
3. Wie viel Prozent der Beobachtungen fallen in das IQR? Ergänzen Sie die Abbildung entsprechend um den Bereich! **(2 Punkte)**
4. Wie viel Prozent der Beobachtungen fallen in $\bar{y} \pm 1s$ und $\bar{y} \pm 2s$ unter der Annahme einer Normalverteilung? **(2 Punkte)**

¹Oschmann, A. (2024) Mädchen stärken: Stärken fördern, Selbstwert erhöhen und liebevoll durch Krisen begleiten. Goldegg Verlag

60. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung der Normalverteilung Alex und die Gefälligkeit machen die Sache mit dem Studium nicht einfacher. Immerhin ist noch Tina zur Hilfe mit dabei. Tina hat Katjes mitgebracht und Tocotronic aufgedreht. Das ist immerhin eine Ablenkung. Nicht so gut wie Starcraft, aber immerhin etwas. Jetzt sollen die beiden diese komische Aufgabe lösen. Es geht um verschiedene Normalverteilungen. Anscheinend hängen Normalverteilungen vom Mittelwert \bar{y} und der Standardabweichung s ab. 'Wozu brauchen wir nochmal Normalverteilungen?', entfährt es Alex. Durch das Mampfen von Tina versteht er kein Wort der Antwort. Tina lächelt.

Jetzt brauchen Alex und Tina Ihre Hilfe bei der Abschätzung einer Verteilung um ihre Arbeit dann in diesem Semester noch abschließen zu können.

1. Skizzieren Sie zwei Normalverteilungen mit $\bar{y}_1 \neq \bar{y}_2$ und $s_1 \neq s_2$! **(3 Punkte)**
2. Beschriften Sie die Normalverteilungen mit den statistischen Maßzahlen! **(2 Punkte)**
3. Liegt Varianzhomogenität oder Varianzheterogenität vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. In welchen Bereich fallen 68% bzw. 95% der Beobachtungen in einer Normalverteilung? Ergänzen Sie die Bereiche in einer Normalverteilung! **(2 Punkte)**
5. Ergänzen Sie unter einer der Normalverteilungen den entsprechenden Boxplot! **(1 Punkt)**

61. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung der Normalverteilung und der Poissonverteilung 'Was zum Geier?!', entfährt es Alex und schaut dabei Jonas an. Dabei nimmt sogar die Katze Reißaus und versteckt sich unter dem Bett. 'Wir sollen eine Normalverteilung mit einem Mittelwert von $\bar{y}_1 = 3$ und einer Standardabweichung von $s_1 = 1$ zeichnen. Sowie eine weitere Normalverteilung mit einem Mittelwert von $\bar{y}_2 = 1$ und einer Standardabweichung von $s_2 = 1$. Keine Ahnung wie das geht. Darunter sollen dann noch eine Poissonverteilung mit einem Mittelwert von $\lambda_1 = 1$ sowie einer weiteren Poissonverteilung mit einem Mittelwert von $\lambda_2 = 15$ gezeichnet werden.', meint Jonas sichtlich genervt und mampft noch ein paar Snickers. Jonas und die Erschöpfung machen die Suche nach der Lösung nicht einfacher. Im Hintergrund spielt viel zu leise Iron Maiden, die diesmal Jonas ausgewählt hat.

Zwei Normalverteilungen

Zwei Poissonverteilungen

Jetzt brauchen Alex und Jonas Ihre Hilfe bei der Abschätzung einer Verteilung um ihre Arbeit dann in diesem Semester noch abschließen zu können.

1. Skizzieren Sie die zwei Normalverteilungen und zwei Poissonverteilungen! **(4 Punkte)**
2. Achten Sie auf die entsprechende Skalierung in den jeweiligen Abbildungen! **(2 Punkte)**
3. Ergänzen Sie unter einer Normalverteilung den entsprechenden Boxplot! **(1 Punkt)**
4. Ergänzen Sie unter einer Poissonverteilung den entsprechenden Boxplot! **(1 Punkt)**
5. Geben Sie ein Beispiel für ein Outcome y , welches einer Normalverteilung folgt! **(1 Punkt)**
6. Geben Sie ein Beispiel für ein Outcome y , welches einer Poissonverteilung folgt! **(1 Punkt)**

Teil II.

Statistisches Testen & statistische Testtheorie

62. Aufgabe

(9 Punkte)



Grundgesamtheit und experimentelle Stichprobe An einem schwülen Sommernachmittag sitzen Alex und Paula in einem Eiskaffee und wollen sich auf die Klausur vorbereiten. In fast allen Fragen geht es ja um die Interpretation eines statistischen Tests. Daher wollen die beiden jetzt nochmal nacharbeiten, was die Grundlagen der Stichprobe (eng. *sample*) und der Grundgesamtheit (eng. *population* oder *ground truth*) sind. Alex hat sich Gummibärchen Eisbecher bestellt und Paula bleibt lieber bei einem Smarties Eis. 'Irre, was die Lebensmittelindustrie alles auf die Beine kriegt', merkt Paula an und Alex schüttelt anerkennend den Kopf.

Leider kennen sich Alex und Paula mit der Grundgesamtheit und der Stichprobe überhaupt nicht aus. Daher sind Sie gefragt!

1. Nennen Sie das statistische Verfahren und zwei konkrete Beispiele zur Durchführung um von einer Grundgesamtheit auf eine Stichprobe zu gelangen! **(3 Punkte)**
2. Erklären Sie den Zusammenhang zwischen Stichprobe und Grundgesamtheit an einem Schaubild! Beschriften Sie das Schaubild entsprechend! *Nutzen Sie hierfür als Veranschaulichung die Körpergröße von Männern oder Frauen aus den Gummibärchendaten!* **(3 Punkte)**
3. Erweitern Sie das Schaubild um die Entstehung von $Pr(D|H_0)$! *Nutzen Sie hierfür als Veranschaulichung zusätzlich die Gruppierungsvariable „Modul“ aus den Gummibärchendaten!* **(3 Punkte)**

63. Aufgabe

(9 Punkte)



Das Nullritual - Die statistische Testtheorie 'Kicken ist der beste Sport, den es gibt.', meint Nilufar. Jonas entgegnet, 'Ich empfehle ja immer allen Schwimmen.' Die beiden sind im Zoo und diskutieren, ob Pinguine Knie haben. Eigentlich wollten beide nochmal die statistische Testtheorie durchgehen, sind dann aber auf abenteuerlichen Wege im Zoo gelandet. Nilufar starrt wie hypnotisiert auf einen strullenden Elefanten und stoppt die Zeit.² 'Du bist so peinlich.', entfährt es Jonas.

Leider kennen sich Nilufar und Jonas mit statistischen Testtheorie, auch Null-Ritual genannt, überhaupt nicht aus. Geschweige denn mit der Visualisierung als Kreuztabelle.

1. Tragen Sie folgende statistische Fachbegriffe zur statistischen Testtheorie korrekt eine selbst erstellte Kreuztabelle ein! **(3 Punkte)**

Richtige Entscheidung Testentscheidung H_0 abgelehnt (Unbekannte) Wahrheit

2. Ergänzen Sie Ihre erstellte Kreuztabelle um vier weitere, passende Fachbegriffe zur statistischen Testtheorie! **(2 Punkte)**

Die Entscheidungsfindung durch einen statistischen Test kann auch durch die Analogie zu einem Feuermelder abgebildet werden. Dabei symbolisiert der Feuermelder den statistischen Test und es soll getestet werden, ob ein Feuer ausgebrochen ist.

3. In der Analogie des Feuermelders, wie lautet der α -Fehler? **(1 Punkt)**
4. In der Analogie des Feuermelders, wie lautet der β -Fehler? **(1 Punkt)**
5. Wenn der Feuermelder einmal pro Tag messen würde, wie oft würde der Feuermelder mit einem α von 5% in einem Jahr Alarm schlagen? Begründen Sie Ihre Antwort! **(2 Punkte)**

²Yang, P. J., et al. (2014). Duration of urination does not change with body size. Proceedings of the National Academy of Sciences, 111(33), 11932-11937.

64. Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

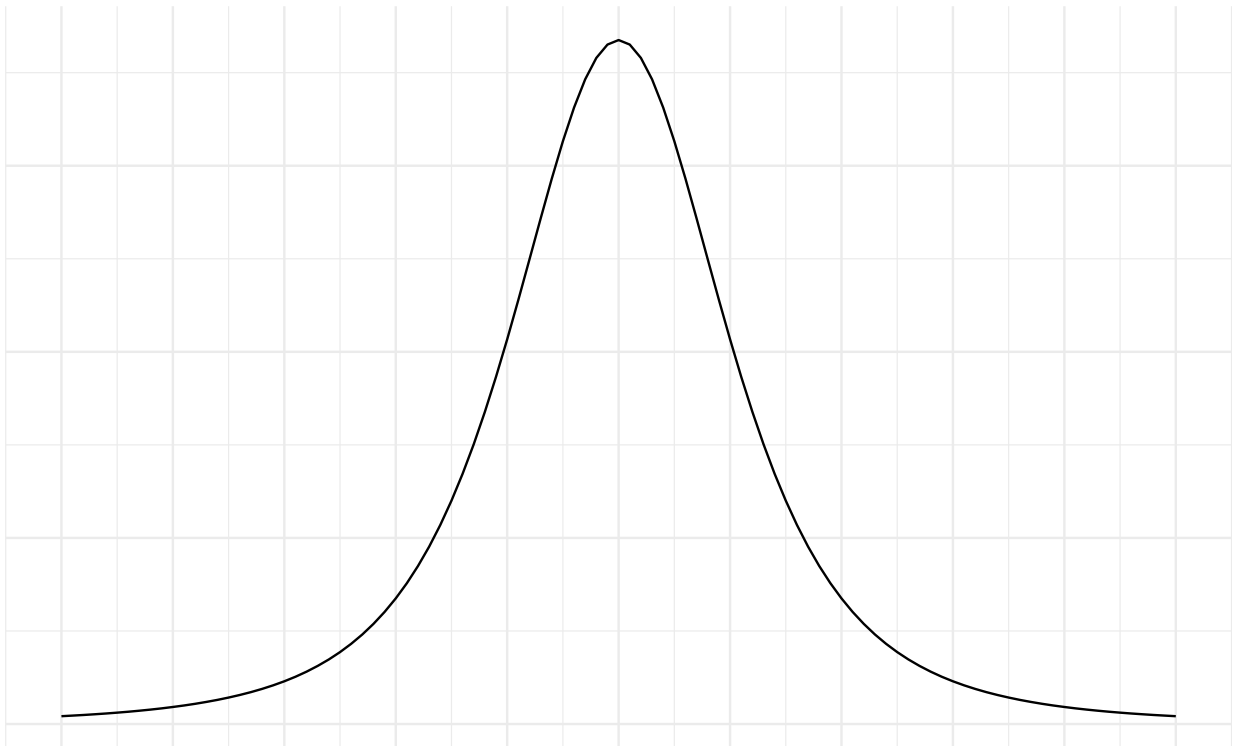


Visualisierung der Teststatistik T_D und dem p-Wert 'Wir sollen die Teststatistik T_D und dem p-Wert visualisieren, da mit einer Visualisierung vieles verständlicher wird!', ruft Mark um Andrea Berg zu übertönen. 'Ich weiß nicht, was das jetzt helfen soll. Können wir nicht einfach schauen, ob der p-Wert kleiner als das Signifikanzniveau α gleich 5% ist? Und gut ist?', merkt Nilufar an, was aber im Refrain von Andrea Berg untergeht. Mark nickt im Beat. 'Wir haben hier eine t-Verteilung unter der Annahme der Nullhypothese!', singt er.

Leider kennen sich Mark und Nilufar mit der Visualisierung der Teststatistik T_D und dem p-Wert überhaupt nicht aus und brauchen daher Ihre Hilfe!

Beachten Sie, dass im Folgenden keine numerisch korrekte Darstellung verlangt wird! Es gilt Erkennbarkeit vor Genauigkeit!

1. Ergänzen Sie eine beschriftete x-Achse! **(1 Punkt)**
2. Ergänzen Sie „ $\bar{y}_1 = \bar{y}_2$ “! **(1 Punkt)**
3. Ergänzen Sie „95%“! **(1 Punkt)**
4. Zeichnen Sie $T_{\alpha=5\%}$ in die Abbildung! **(1 Punkt)**
5. Zeichnen Sie das Signifikanzniveau α in die Abbildung! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Zeichnen Sie $+T_D$ in die Abbildung! **(1 Punkt)**
7. Zeichnen Sie einen nicht signifikanten p-Wert in die Abbildung! Begründen Sie Ihre Antwort! **(2 Punkte)**



65. Aufgabe

(10 Punkte)



Visualisierung des 95% Konfidenzintervalls 'So, was haben wir gemacht? Wir haben einen t-test für den Vergleich der Mittelwerte gerechnet.', meint Paula. Alex schaut fragend. 'Hatten wir nicht alles zu einer Kontrolle verglichen? Das war doch so!', ruft Alex laut aus. 'Wir haben doch Messwert *Pilzbefall nach Fungizid* erhoben.', stellt Paula fest. Jetzt haben beide das Problem, die möglichen 95% Konfidenzintervalle zu interpretieren.

Leider kennen sich Paula und Alex mit der Visualisierung des 95% Konfidenzintervall überhaupt nicht aus.

1. Beschriften Sie die untenstehende Abbildung mit der Signifikanzschwelle! Begründen Sie Ihre Antwort! **(2 Punkte)**
2. Ergänzen Sie eine *in den Kontext passende* Relevanzschwelle! Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Skizzieren Sie in die untenstehende Abbildung sechs einzelne Konfidenzintervalle (a-f) mit den jeweiligen Eigenschaften! **(6 Punkte)**
 - (a) Ein signifikantes, relevantes 95% Konfidenzintervall
 - (b) Ein signifikantes, relevantes 90% Konfidenzintervall.
 - (c) Ein signifikantes, nicht relevantes 95% Konfidenzintervall
 - (d) Ein 95% Konfidenzintervall mit höherer Fallzahl n in der Stichprobe als der Rest der 95% Konfidenzintervalle
 - (e) Ein nicht signifikantes, nicht relevantes 95% Konfidenzintervall
 - (f) Ein 95% Konfidenzintervall mit niedriger Fallzahl n in der Stichprobe als der Rest 95% der Konfidenzintervalle



66. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Zusammenhang zwischen dem Effekt, der Streuung sowie der Fallzahl An einem kalten Dezembermorgen haben sich Paula und Jessica zum Lernen verabredet. Eine große Kanne Tee und Berge von Smarties warten darauf gegessen zu werden. Paula liest laut vor:

Beim statistischen Testen gibt es einen Zusammenhang zwischen dem Effekt, der Streuung sowie der Fallzahl. Gegeben sei die Formel für den Student t-Test auf den die folgenden Überlegungen basieren sollen. Welche Auswirkung hat die Änderungen der jeweiligen statistischen Maßzahl des Effekts Δ , der Streuung s und der Fallzahl n auf die Teststatistik T_D , den p-Wert $Pr(D|H_0)$ sowie dem Konfidenzintervall $KI_{1-\alpha}$?

Jessica hebt die Augenbraue. 'Irgendwie sagt mir die Aufgabe jetzt mal so gar nichts. Was soll da gemacht werden?', merkt Jessica an und ergänzt: 'Schauen wir doch erstmal zur Entspannung Herr der Ringe, den Film habe ich extra nochmal mitgebracht.' Paula ist der Idee nicht abgeneigt und auch die Ratte kommt unter dem Sofa hervor um mitzuschauen.

Leider kennen sich Paula und Jessica mit dem Zusammenhang zwischen dem Effekt, der Streuung sowie der Fallzahl überhaupt nicht aus.

1. Visualisieren Sie den Zusammenhang zwischen der Teststatistik T_D und dem p-Wert $Pr(D|H_0)$ für sich verändernde T_D -Werte! Geben Sie dafür ein numerisches Beispiel in dem Sie drei T_D -Werte und deren Einfluss auf den p-Wert vergleichen! **(3 Punkte)**
2. Füllen Sie die untenstehende Tabelle aus in dem Sie die Änderung der statistischen Maßzahlen auf die Teststatistik, den p-Wert sowie das Konfidenzintervall in einem Wort oder Symbol beschreiben! **(4 Punkte)**

	T_D	$Pr(D H_0)$	$KI_{1-\alpha}$		T_D	$Pr(D H_0)$	$KI_{1-\alpha}$
$\Delta \uparrow$				$\Delta \downarrow$			
$s \uparrow$				$s \downarrow$			
$n \uparrow$				$n \downarrow$			

3. Visualisieren Sie ein 95%-iges Konfidenzintervall im Vergleich zu einem 99%-igen Konfidenzintervall! Begründen Sie Ihre Visualisierung anhand der Formel des Konfidenzintervalls des t-Tests mathematisch! **(3 Punkte)**

Teil III.

Der Student t-Test, Welch t-Test & gepaarter t-Test

67. Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Berechnung des Student t-Test oder Welch t-Test Der t-Test. Nilufar erschauert. Wenn die Erwartung nicht wäre, ja dann wäre wohl vieles möglich für Nilufar! Aber so.. Ein mächtiges Werkzeug ist der t-Test in den Händen desjenigen, der einen normalverteilten Messwert (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Nilufar überhaupt aus? Nilufar schmeißt noch eine Handvoll Takis Blue Heat in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Deichkind. Nilufar hat einen Leistungssteigerungsversuch mit Milchvieh durchgeführt um eine neue technische Versuchsanlage zu testen. Bei dem Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) wurde die Behandlung Bestandsdichte (*Verordnung* und *Erhht*) an den Milchvieh getestet und dabei wurde geschaut, ob der Versuch überhaupt technisch klappen könnte. Gemessen hat Nilufar dann als Messwert Schlachtgewicht [kg]. Warum der Versuch in der Uckermark für ihre Abschlussarbeit stattfinden musste, ist ihr bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Schlachtgewicht [kg]?

treatment	weight
dose	10.3
ctrl	14.6
ctrl	17.4
dose	16.6
ctrl	14.0
dose	18.9

Leider kennt sich Nilufar mit der Berechnung eines t-Tests überhaupt nicht aus. Deshalb braucht sie bei der Berechnung Ihre Hilfe!

1. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
2. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 2.04$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des Welch t-Tests! **(1 Punkt)**
5. Formulieren Sie eine Antwort an Nilufar über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

68. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Berechnung des Student t-Test Der t-Test. Alex erschauert. Wenn die Gefälligkeit nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so.. Ein mächtiges Werkzeug ist der t-Test in den Händen desjenigen, der einen normalverteilten Endpunkt (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Alex überhaupt aus? Alex schmeißt noch eine Handvoll Gummibärchen in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Abba. Alex hat ein Gewächshausexperiment mit Kartoffeln durchgeführt. Dabei wurde die Behandlung Bewässerungstypen (*low* und *high*) an den Kartoffeln getestet. Gemessen hat Alex dann als Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Warum der Versuch im Wendland für seiner Hausarbeit stattfinden musste, ist ihm bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Chlorophyllgehalt (SPAD-502Plus) [SPAD]?

Bewässerungstypen	Chlorophyllgehalt
high	23.6
high	27.6
low	30.7
high	24.2
low	36.6
low	31.4
low	22.1
low	31.0
low	36.3
low	36.5
high	24.2
low	36.7
high	26.7
low	33.3
high	23.2
low	29.3
high	26.1
high	22.4

Leider kennt sich Alex mit der Berechnung eines t-Tests überhaupt nicht aus. Deshalb braucht er bei der Berechnung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *keinen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann der Effekt? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Alex über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

69. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Berechnung des Welch t-Test Der t-Test. Alex erschauert. Alex und die Gefälligkeit, eine unendliche Geschichte mit kniffligen Wendungen. Ein mächtiges Werkzeug ist der t-Test in den Händen desjenigen, der ein normalverteiltes Outcome (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Alex überhaupt aus? Alex schmeißt noch eine Handvoll Gummibärchen in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Abba. Alex hat ein Kreuzungsexperiment mit Hühnern durchgeführt. Dabei wurde die Behandlung Bestandsdichte (*Verordnung* und *Erhöht*) an den Hühnern getestet. Gemessen hat Alex dann als Messwert Schlachtgewicht [kg]. Warum der Versuch im Oldenburger Land für seinen Projektbericht stattfinden musste, ist ihm bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Schlachtgewicht [kg]?

Bestandsdichte	Schlachtgewicht
Verordnung	50.7
Erhöht	24.6
Verordnung	33.8
Erhöht	38.8
Erhöht	25.8
Verordnung	30.0
Erhöht	20.4
Erhöht	23.2
Verordnung	49.9
Verordnung	64.0
Verordnung	52.4
Verordnung	58.9
Erhöht	29.0
Erhöht	41.6
Verordnung	52.1
Verordnung	43.5
Erhöht	16.6
Verordnung	41.6
Verordnung	46.2
Erhöht	23.8

Leider kennt sich Alex mit der Berechnung eines t-Tests überhaupt nicht aus. Deshalb braucht er bei der Berechnung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie das 99% Konfidenzintervall. Welche Annahmen haben Sie getroffen? **(2 Punkte)**
6. Nennen Sie den statistischen Grund, warum Sie sich zwischen einem Student t-Test und einem Welch t-Test entscheiden müssen! **(1 Punkt)**
7. Formulieren Sie eine Antwort an Alex über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

70. Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Berechnung des gepaarten t-Test Es gibt ja immer die Möglichkeit sich Hilfe zu holen. Das geht natürlich auch immer in einer Hausarbeit. Deshalb arbeiten Tina und Nilufar gemeinsam an einer Hausarbeit. Das macht dann auch die Analyse ihres Hauptversuches einfacher. Zwar hat jeder von ihnen noch ein Subthema, aber auch da kann man sich ja helfen. In dem Hauptversuch wurde Folgendes von den beiden gemacht. Tina und Nilufar haben sich Schweinen angeschaut. Dabei geht um Zusammenhang zwischen Flüssignahrung (1l/d und 5l/d) und Protein/Fettrate [%/kg]. Jetzt sollen beide einen gepaarten t-Test rechnen. Es würde auch besser funktionieren, wenn Tina nicht die Wut im Weg stehen würde und Nilufar nicht das Problem hätte die Erwartung zu händeln. Gott sei Dank haben beide genug Katjes und Takis Blue Heat auf dem Tisch aufgetürmt.

ID	treatment	freshmatter
6	5l/d	47.1
3	1l/d	36.2
1	5l/d	50.7
8	1l/d	38.6
7	1l/d	46.8
3	5l/d	38.1
5	5l/d	48.7
2	5l/d	47.3
4	1l/d	36.8
9	1l/d	39.1
6	1l/d	30.0
1	1l/d	27.1
2	1l/d	29.4
7	5l/d	49.9
10	1l/d	33.9
5	1l/d	39.8
4	5l/d	46.8

Leider kennen sich Tina und Nilufar mit der Berechnung eines gepaarten t-Tests überhaupt nicht aus. Deshalb brauchen sie beide bei der Berechnung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines gepaarten t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Schätzen Sie den p -Wert des gepaarten t-Tests ab! Begründen Sie Ihre Antwort mit einer Skizze! **(2 Punkte)**
6. Formulieren Sie eine Antwort an Tina über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

71. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interpretation des t-Tests in R - die Teststatistik und der p-Wert 'Mit dem R Paket {emmeans} können wir gleich die Gruppenvergleiche rechnen und uns das *compact letter displac*' ausgeben lassen!', verkündet Alex sichtlich stolz. Ein paar Mal hat sie schon die Gefälligkeit gehindert weiterzumachen. 'Nach Meinung der Betreuerin soll es aber nur erstmal ein t-Test sein. Und die Ausgabe ist schon wirr genug.', merkt Yuki an. Yuki und Paula sind bei Alex um sich in R helfen zu lassen. Im Hintergrund wummert Abba. Paula streichelt zur Beruhigung die Katze von Alex. Die beiden waren 1 Monate im Oldenburger Land um einen Versuch mit Hühnern in einem Leistungssteigerungsversuch durchzuführen. Ziel war es das Outcome (Y) Fettgehalt [%/kg] zu bestimmen. Alex überlegt, ob er die beiden nicht noch auf den Film *Alien* einlädt oder dann doch lieber raus geht um zu Laufen? Vielleicht will ja Paula mit. Besser als der Film.

```
##
## Two Sample t-test
##
## data: Fettgehalt by Ernährungszusatz
## t = 2.6662, df = 14, p-value = 0.01844
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
##  2.225812 20.536093
## sample estimates:
## mean in group ctrl mean in group fedX
##           36.36667           24.98571
```

Helfen Sie Alex bei der Interpretation des t-Tests! Sonst geht es auch für Yuki und Paula nicht weiter.

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.14|$ einzeichnen! **(4 Punkte)**
5. Beschriften Sie die Abbildung! **(1 Punkt)**
6. Berechnen Sie den Effekt des t-Tests! **(1 Punkt)**

72. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interpretation des t-Tests in R - das 95% Konfidenzintervall Alex und Steffen sind bei Jessica um sich Hilfe in R zu holen. Im Hintergrund wummert David Bowie. Die beiden hatten zwar schon erste Kontakte mit R sind sich aber unsicher bei der Interpretierung der Ausgabe eines t-Tests für ihren gemeinsamen Versuch. Es würde auch besser funktionieren, wenn Jessica nicht der Mangel im Weg stehen würde und Steffen nicht das Problem hätte die Gefälligkeit zu händeln. In einer Hausarbeit haben beide zusammen Spargel untersucht. Dabei ging es um den Zusammenhang zwischen der Behandlung Genotypen (AA und BB) und dem Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Der Versuch wurde in einem Feldexperiment im Emsland durchgeführt. Nach der Betreuerin ist der Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] normalverteilt und ein t-Test passt daher. Das wird jetzt nicht mehr angezweifelt...Jessica überlegt, ob sie die beiden nicht noch auf den Film *Herr der Ringe* einlädt.

```
##
## Two Sample t-test
##
## data: Chlorophyllgehalt by Genotypen
## t = -3.3951, df = 14, p-value = 0.004355
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -19.438072 -4.386928
## sample estimates:
## mean in group AA mean in group BB
##          35.1750          47.0875
```

Helfen Sie Jessica bei der Interpretation des t-Tests! Sonst geht es auch für Alex und Steffen nicht weiter.

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie das sich ergebende 95% Konfidenzintervall! **(2 Punkte)**
5. Beschriften Sie die Abbildung und das 95% Konfidenzintervall entsprechend! **(2 Punkte)**
6. Interpretieren Sie den Effekt des 95% Konfidenzintervalls! **(2 Punkte)**

73. Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interpretation des t-Tests in R - die Visualisierung 'Programmieren ist wie eine Sprache lernen. Man muss es nur machen, dann wird man mit der Zeit immer besser!', gibt Tina zwinkernd zu Protokoll. Ein paar Mal hat sie schon die Wut gehindert weiterzumachen. Das hilft jetzt Nilufar und Paula nur bedingt, da beide jetzt die R Ausgabe interpretieren müssen und nicht vor drei Wochen, wo noch Zeit gewesen wäre. Beide mampfen konzentriert Takis Blue Heat und Smarties in sich hinein. Die beiden hatten in der Uckermark einen Versuch mit Zandern in einem Leistungssteigerungsversuch durchgeführt. Das war schon anstrengend genug! 'Wir haben Fettgehalt [%/kg] gemessen, vielleicht hilft das ja...', meint Nilufar leicht genervt. Alle starren auf die R Ausgabe des t-Tests. Im Hintergrund wummert Tocotronic und man versteht kaum sein eigenes Wort. Paula hofft, dass die Spinne von Tina beruhigend wirkt.

```
##
## Two Sample t-test
##
## data: Fettgehalt by Elterlinie
## t = 1.804, df = 14, p-value = 0.09279
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -1.034358 11.984358
## sample estimates:
## mean in group Standard      mean in group Xray
##                41.7875                36.3125
```

Helfen Sie Tina bei der Interpretation des t-Tests! Sonst geht es auch für Nilufar und Paula nicht weiter.

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie die sich ergebenden Boxplot! Welche Annahmen an die Daten haben Sie getroffen? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Skizzieren Sie die sich ergebenden Barplots! **(2 Punkte)**
6. Berechnen Sie den Effekt des t-Tests! **(1 Punkt)**

74. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interpretation des gepaarten t-Tests in R Yuki und Jessica haben sich dazu entschieden zusammenzuarbeiten. Das sollte alles etwas einfacher machen. Jeder hat zwar ein getrenntes Themenfeld aber den Hauptversuch machen beide gemeinsam. Das hat sich schonmal als gut Idee soweit herausgestellt. Wenn die Faulheit nicht wäre, ja dann wäre wohl vieles möglich für Yuki! Aber so.. In einer Abschlussarbeit sollen beide herausfinden, ob es einen Zusammenhang zwischen Ausgeizen (*ctrl* und *28d*) und Trockengewicht [kg/ha] gibt. Die Besonderheit ist hierbei, dass die Messungen an der gleichen Beobachtung stattfinden. Beide messen also zweimal an den gleichen Brokkoli. Hier muss dann wohl auf einen normalverteilten Messwert (Y) ein gepaarter t-Test gerechnet werden. Leider kennen sich beide nicht sehr gut in R aus. Aber wenigstens haben beide eine Menge an Reese's Peanut Butter Cups.

```
##
## Paired t-test
##
## data: Trockengewicht by Ausgeizen
## t = -4.7009, df = 8, p-value = 0.00154
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -24.461459 -8.360763
## sample estimates:
## mean difference
## -16.41111
```

Jetzt brauchen Yuki und Jessica Ihre Hilfe bei der Berechnung eines gepaarten t-Tests in R um ihre Arbeit dann in diesem Semester noch abschließen zu können.

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie das sich ergebende 95% Konfidenzintervall! **(2 Punkte)**
5. Interpretieren Sie den Effekt des gepaarten t-Tests! **(2 Punkte)**
6. Skizzieren Sie den sich ergebenden Boxplot der Differenzen! Welche Annahmen an die Daten haben Sie getroffen? Begründen Sie Ihre Antwort! **(2 Punkte)**

Teil IV.

Die einfaktorielle & zweifaktorielle ANOVA

75. Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung der einfaktoriellen ANOVA 'Uff... die einfaktorielle ANOVA. Und wir können jetzt anhand der Visualisierung sehen, ob da schon was signifikant ist?', Mark hebt die Augenbraue. 'Ja, können wir. Dafür müssen wir aber erstmal in {ggplot} uns die Daten anschauen. Oder wir zeichnen es flott mit der Hand. Geht auch.', meint Nilufar dazu. Die beiden hatten sich auf einem Konzert von Deichkind kennengelernt. Mark hatte sich in ein Gewächshausexperiment verschiedene Erdbeeren angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lichtstufen (*none*, 200lm und 600lm) und dem Messwert Trockengewicht [kg/ha] gibt. Später wird noch Star Trek geguckt. Nilufar befürwortet das!

Lichtstufen	Trockengewicht
600lm	45
200lm	47
600lm	44
none	30
none	27
200lm	44
600lm	46
none	31
600lm	46
200lm	45
200lm	47
600lm	45
none	30
none	29
200lm	44

Leider kennen sich Mark und Nilufar mit Darstellung einer einfaktoriellen ANOVA überhaupt nicht aus.

1. Erstellen Sie eine Visualisierung der Datentabelle! Beschriften Sie die Abbildung! **(2 Punkte)**
2. Benennen Sie die Visualisierung mit dem korrekten, statistischen Fachbegriff! **(1 Punkt)**
3. Zeichnen Sie folgende statistischen Maßzahlen passend ein!
 - Globale Mittelwert: β_0 **(1 Punkt)**
 - Mittelwerte der einzelnen Behandlungsstufen: $\bar{y}_{0.5}$, $\bar{y}_{1.5}$ und $\bar{y}_{2.5}$ **(1 Punkt)**
 - Mittelwertsdifferenz der einzelnen Behandlungsstufen: $\beta_{0.5}$, $\beta_{1.5}$ und $\beta_{2.5}$ **(1 Punkt)**
 - Residuen oder Fehler: ϵ **(1 Punkt)**
4. Liegt ein *vermutlicher* signifikanter Unterschied vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Schätzen Sie die Effekte der Behandlungsstufen! **(2 Punkte)**

76. Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Ergebnistabelle der einfaktoriellen ANOVA 'Als erstes bauen wir uns aus unsere Daten die ANOVA Tabelle dann sehen wir schon, ob unser Gruppenvergleich in der ANOVA signifikant ist.', Mark schaut Jessica fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Da hilft die Hündin von Jessica auch nur bedingt. Jessica tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren in der Uckermark um einen Versuch in einer Klimakammer mit Kartoffeln durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lüftungssystemen und Folientunneln (*ctrl*, *storm*, *thunder* und *tornado*) und dem Messwert Trockengewicht [kg/ha] gibt. Nachher wollen sich beide noch mit dem Hobby Warhammer von Jessica beschäftigen. Kennt Mark noch nicht, klingt aber interessant.

Leider kennen sich Mark und Jessica mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe, die Hündin reicht als Hilfe nicht aus!

1. Formulieren Sie die wissenschaftliche Fragestellung! (1 Punkt)
2. Formulieren Sie das statistische Hypothesenpaar! (1 Punkt)
3. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Lüftungssystemen	3	6293.05			
error	25	1351.09			
Total	28				

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 2.99$ ab. Begründen Sie Ihre Antwort! (2 Punkte)
5. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? (2 Punkte)

77. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Die einfaktoriellen ANOVA und der Student t-Test 'Als erstes bauen wir uns aus unsere Daten die ANOVA Tabelle dann sehen wir schon, ob unser Gruppenvergleich in der ANOVA signifikant ist.', Yuki schaut Jessica fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Jessica schmeißt sich noch ein paar Schokobons in den Rachen. Beide tun sich sehr schwer mit der einfaktoriellen ANOVA. Nun möchte erstmal ihre Betreuung der Arbeit eine ANOVA Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Beide waren in der Uckermark um einen Leistungssteigerungsversuch mit Hühnern durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Elterlinie (*ctrl*, *Standard*, *Yray* und *Xray*) und dem Messwert Schlachtgewicht [kg] gibt. Später wollen die beiden dann noch raus um Rad zu fahren.

Leider kennen sich Yuki und Jessica mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! (1 Punkt)
2. Formulieren Sie das statistische Hypothesenpaar! (1 Punkt)
3. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Elterlinie	3	719.96			
Error	27	356.04			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 2.96$ ab. Begründen Sie Ihre Antwort! (2 Punkte)
5. Was bedeutet ein signifikantes Ergebnis in einer einfaktoriellen ANOVA? (1 Punkt)
6. Berechnen Sie *einen* Student t-Test für den *vermutlich* signifikantesten Gruppenvergleich anhand der untenstehenden Tabelle mit $T_{\alpha=5\%} = 2.03$. Begründen Sie Ihre Auswahl! (3 Punkte)

Elterlinie	Fallzahl (n)	Mittelwert	Standardabweichung
ctrl	6	10.50	3.56
Standard	7	15.71	4.99
Yray	9	3.89	4.04
Xray	9	4.44	1.24

7. Gegebenen der ANOVA Tabelle war das Ergebnis des Student t-Tests zu erwarten? Begründen Sie Ihre Antwort! (2 Punkte)

78. Aufgabe

(9 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Die einfaktorielle ANOVA in R 'Uff... die einfaktorielle ANOVA und R. Nicht so einfach... Was sagt mir jetzt die Ausgabe der ANOVA und wo sehe ich, ob da was signifikant ist?', denkt Nilufar und hebt die Augenbraue. Nilufar hatte sich ein Feldexperiment mit Spargel angeschaut. Als wäre das nicht alles schon schwer genug. Eine echte Herausforderung für sie war schon immer die Erwartung gewesen. Ein leidiges Lied. Dabei ging es beim Experiment herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lüftungssysteme (*ctrl*, *storm*, *thunder* und *tornado*) und dem Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] gibt. Nun möchte ihr Betreuer ihrer Abschlussarbeit erstmal eine ANOVA sehen und die Ergebnisse präsentiert bekommen. Und eigentlich will sie ja was anderes... Auf seinem Second Screen läuft Star Trek und Nilufar schaufelt Takis Blue Heat. Nicht effizient, aber gut.

```
## Analysis of Variance Table
##
## Response: Chlorophyllgehalt
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Lüftungssysteme  3 3940.1 1313.36  41.937 6.626e-10
## Residuals       25  782.9   31.32
```

Leider kennen sich Nilufar mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Interpretieren Sie das Ergebnis der einfaktoriellen ANOVA! **(2 Punkte)**
4. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? **(2 Punkte)**
5. Skizzieren Sie eine Abbildung, der dem obigen Ergebnis der einfaktoriellen ANOVA näherungsweise entspricht! **(3 Punkte)**

79. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Ergebnistabelle der zweifaktoriellen ANOVA Steffen steht in der Uckermark. Und das ist noch langweiliger als es sich anhört. Wäre es nur so spannend wie bei seinen Kommilitonen, die in Almería waren. Ödnis wohin man nur blickt. Oder eben Brokkoli. Die Schlange durchbohrt ihn mit leeren Blick. 'Woher zum Teufel!', entfährt es ihm. Aber da ist es schon weg. Ja, darum geht es in seinem Projektbericht. Und wäre das nicht noch alles genug, ist sein Experiment auch noch als einen Versuch in einer Klimakammer komplex geraten. Es wurde der Messwert Proteingehalt [g/kg] mit dem Behandlung Lüftungssysteme (*ctrl*, *storm*, *thunder* und *tornado*) sowie der Behandlung Lichtstufen (*none*, und *600lm*) untersucht. 'Hmpf', denkt Steffen und ruft 'Und jetzt!?' in die Leere. Und eigentlich wollte Steffen doch noch seinem Hobby nachgehen! Klemmbausteine. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Steffen denkt gerne über Klemmbausteine nach.

Leider kennen sich Steffen mit Berechnung einer zweifaktoriellen ANOVA überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! (1 Punkt)
2. Formulieren Sie das statistische Hypothesenpaar! (1 Punkt)
3. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Lüftungssysteme	3	212.85			
Lichtstufen	1	46.68			
Lüftungssysteme:Lichtstufen	3	354.03			
Error	18	269.32			

4. Schätzen Sie den p-Wert der Tabelle ab. Begründen Sie Ihre Antwort! (3 Punkte)

	$F_{\alpha=5\%}$
Lüftungssysteme	4.26
Lichtstufen	3.40
Lüftungssysteme:Lichtstufen	5.23

5. Was bedeutet ein signifikantes Ergebnis in einer zweifaktoriellen ANOVA? (2 Punkte)
6. Was sagt der Term *Lüftungssysteme:Lichtstufen* aus? Interpretieren Sie das Ergebnis! (2 Punkte)

80. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Die zweifaktorielle ANOVA in R In ein Freilandversuch wurden Kartoffeln mit der Behandlung Düngestufen (*ctrl*, *low*, *mid* und *high*) sowie der Behandlung Lichtstufen (*none*, und *600lm*) untersucht. Es wurde als Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] bestimmt. Jetzt starrt Tina mit auf die R Ausgabe einer zweifaktoriellen ANOVA. Leider starrt ihre Betreuerin in der gleichen Art Tina zurück an. Das wird ein langer Nachmittage, denkt sie sich und kreuselt ihren Mund. 'Und was machen wir jetzt?' entfährt es ihr überrascht entnervt. Immerhin war geht es ja um ihre Hausarbeit. Tina hätte doch nichts mit Kartoffeln machen sollen. Kartoffeln – was soll das auch bedeutendes sein? Eigentlich wollte Tina nachher noch einen Film schauen. Irgendwie komisch, wenn sie Indiana Jones anmacht, dann ist die Spinne eigentlich sofort vor dem Bildschirm und starrt hinein.

```
## Analysis of Variance Table
##
## Response: Chlorophyllgehalt
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Düngestufen      2 366.86  183.428  12.2034 0.0004472
## Lichtstufen       1  162.73   162.729   10.8263 0.0040671
## Düngestufen:Lichtstufen  2  140.53    70.264    4.6746 0.0231710
## Residuals       18  270.56    15.031
```

Leider kennt sich Tina mit Berechnung einer zweifaktoriellen ANOVA überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Interpretieren Sie das Ergebnis der einfaktoriellen ANOVA! **(3 Punkte)**
4. Zeichnen Sie eine Abbildung, der dem obigen Ergebnis der zweifaktoriellen ANOVA näherungsweise entspricht! **(5 Punkte)**

81. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Zusammenhang zwischen der ANOVA und dem t-Test Steffen schaut konzentriert auf die Formeln der ANOVA und des t-Tests. In seinem Experiment wurde als Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] bestimmt. Wenn die Romantik nicht wäre, ja dann wäre wohl vieles möglich für Steffen! Aber so.. Dann wäre es nicht noch komplizierter. In ein Freilandversuch wurden Kartoffeln mit der Behandlung Genotypen (AA, AB und BB) sowie der Behandlung Lichtstufen (*none*, und 600lm) untersucht. Beide Verfahren müssen etwas miteinander zu tun haben und sein Betreuer möchte das jetzt auch noch verstehen. Im Hintergrund läuft leise Harry Potter auf seinem Second Screen. Immerhin hat er die beiden Formeln vorliegen.

Gegebene Formeln

$$F_D = \frac{MS_{treatment}}{MS_{error}} \quad T_D = \frac{\bar{y}_1 - \bar{y}_2}{s_p \cdot \sqrt{2/n_g}}$$

Leider kennen sich Steffen mit dem Zusammenhang zwischen der ANOVA und dem t-Test nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Welche statistische Maßzahl testet der t-Test, welche die ANOVA? Begründen Sie Ihre Antwort! **(2 Punkte)**
2. Erklären Sie den Zusammenhang zwischen der F_D Statistik und T_D Statistik! **(2 Punkte)**
3. Visualisieren Sie in einer 2x2 Tafel den Zusammenhang von $MS_{treatment}$ und MS_{error} ! **(2 Punkte)**
4. Beschriften Sie die erstellte 2x2 Tafel mit signifikant und nicht signifikant! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Nennen Sie das numerische Minimum der F-Statistik F_D ! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Wenn die F-Statistik F_D minimal ist, welche Aussage erhalten Sie über die Nullhypothese? Begründen Sie Ihre Antwort! **(2 Punkte)**

82. Aufgabe


(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interaktion in der zweifaktoriellen ANOVA 'Mit der zweifaktoriellen ANOVA lässt sich die Interaktion zwischen den beiden Behandlungen nachweisen!', sein Betreuer scheint die zweifaktoriellen ANOVA zu verstehen. Warum jetzt er jetzt nochmal alles wiederkauen muss, wird Jonas echt nicht so klar. Wenn es doch so klar ist? Jonas war in der Uckermark und hatte dort ein Feldexperiment mit Brokkoli durchgeführt. Die Komune wo er untergekommen war, war cool gewesen. Nur jetzt muss eben das Experiment fertig ausgewertet werden. Es liegt anscheinend eine signifikante Interaktion vor? Jonas hatte zwei Behandlungen auf Brokkoli angewendet. Einmal Düngestufen (*ctrl*, *low*, *mid* und *high*) sowie als zweite Behandlung Genotypen (*AA*, und *BB*). Gemessen wurde der Messwert (*Y*) Trockengewicht [kg/ha]. Jetzt muss das hier zu einem Ende kommen! Eigentlich wollte Jonas nachher noch einen Film schauen. Das Verrückte ist, dass das Meerschweinchen Mission Impossible wirklich liebt. Das ist Jonas sehr recht, denn er braucht Entspannung.

Leider kennen sich Jonas und sein Betreuer mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion. Deshalb braucht er bei der Erstellung Ihre Hilfe, sonst wird es heute Abend mit seinem Hobby Stricken nichts mehr!


1. Visualisieren Sie folgende mögliche Interaktionen zwischen den Behandlungen! Beschriften Sie die Abbildung! **(4 Punkte)**
 - a) Keine Interaktion liegt vor.
 - b) Eine schwache Interaktion liegt vor.
 - c) Eine starke Interaktion liegt vor.
2. Erklären Sie den Unterschied zwischen den verschiedenen Interaktionen! **(2 Punkte)**
3. Welche statistische Maßzahl betrachten Sie für die Bewertung der Interaktion? **(1 Punkt)**
4. Skizzieren Sie die notwendigen Funktionen in  für eine Post-hoc Analyse! **(2 Punkte)**
5. Wenn eine signifikante Interaktion in den Daten vorliegt, wie ist dann das weitere Vorgehen? Berücksichtigen Sie auch die Funktion `emmeans()`! **(2 Punkte)**

83. Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Zusammenhang zwischen der ANOVA und dem Post-hoc-Test In einen Leistungssteigerungsversuch wurden die Daten D von Puten mit der Behandlung Bestandsdichte (*standard*, *eng*, *weit* und *kontakt*) untersucht. Es wurde als Messwert Gewichtszuwachs in der 1LW bestimmt. Jetzt starrt Yuki mit auf die  Ausgabe einer einfaktoriellen ANOVA. Anscheinend gibt es ein Problem mit der Annahme der Normalverteilung und der Varianzhomogenität der ANOVA in den Daten. 'Wir haben jetzt bei der ANOVA einen p-Wert mit 0.061 raus sowie eine F-Statistik F_D mit 1.51 berechnet. Nach den Boxplots müsste sich eigentlich ein Unterschied zwischen *weit* und *kontakt* ergeben. Der Unterschied ist in {emmeans} auch signifikant mit einem p-Wert von 0.036. Wie kann das sein?', meint Yuki ungläubig. Leider starrt ihre Betreuerin in der gleichen Art Yuki zurück an. Yuki hat schon genug Probleme. Wenn die Faulheit nicht wäre, dann wäre es einfacher. Das wird ein langer Nachmittag, denkt sie sich und kreuselt ihren Mund. 'Und was machen wir jetzt?' entfährt es ihr überrascht entnervt. Immerhin war geht es ja um die Projektbericht. Yuki hätte doch nichts mit Puten machen sollen. Puten – was soll das auch bedeutendes sein? Hoffentlich kommt sie noch zum Konzert von London Grammar. Nervös krault sie das Minischwein.

Gegebene Formeln

$$MS_{treatment} = \frac{SS_{treatment}}{df_{treatment}} \quad MS_{error} = \frac{SS_{error}}{df_{error}} \quad F_D = \frac{MS_{treatment}}{MS_{error}}$$

Leider kennen sich Yuki und ihre Betreuerin mit der Interpretation einer ANOVA überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe und die Zeit wird knapp.

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Was bedeutet eine signifikante ANOVA für die beobachteten Daten D ? **(1 Punkt)**
4. Visualisieren Sie den Unterschied zwischen Varianzhomogenität und Varianzheterogenität anhand der Daten D ! Beschriften Sie die Abbildung! **(2 Punkte)**
5. Visualisieren Sie für die Daten D die Verletzung der Annahme der Varianzhomogenität der ANOVA unter zu Hilfenahme von Boxplots! Beschriften Sie die Abbildung! **(2 Punkte)**
6. Welche Auswirkung hat die Verletzung der Annahme der Varianzhomogenität für die Teststatistik F_D der ANOVA? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Erklären Sie abschließend die Diskrepanz zwischen den Ergebnis der ANOVA und dem paarweisen Gruppenvergleich in {emmeans}! **(2 Punkte)**

Teil V.

Multiple Gruppenvergleiche

84. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Adjustierung multipler Vergleiche In ein Stallexperiment mit Zandern wurde die Behandlung Flüssignahrung (*ctrl*, *superIn*, *compostIn* und *fIOW*) gegen die Ergebnisse einer früheren Studie von Totsdorf et al. (2018) verglichen. Im Rahmen des Experiments haben Alex und Yuki verschiedene Student t-Tests für den Mittelwertsvergleich für den Messwert Gewichtszuwachs in der 1LW gerechnet. Es ergab sich dann die folgende Tabelle der rohen p-Werte für die Vergleiche zu Totsdorf et al. (2018). Jetzt sollen die beiden einmal schauen, was in den Daten so drin ist.

Rohen p-Werte	Adjustierte p-Werte	Nullhypothese ablehnen?
0.012		
0.760		
0.030		
0.020		

Leider kennen sich Alex und Yuki mit der Adjustierung von p -Werten und dem Signifikanzniveau α überhaupt nicht aus. Deshalb brauchen die beiden bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Füllen Sie die Spalte *Adjustierte p-Werte* nach der Bonferoni-Methode aus! **(2 Punkte)**
4. Entscheiden Sie, ob nach der Adjustierung die Nullhypothese abgelehnt werden kann! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Wie ist Ihr Vorgehen, wenn Sie anstatt der p -Werte das Signifikanzniveau α adjustieren? **(2 Punkte)**
6. Erklären Sie warum die p -Werte oder das Signifikanzniveau α bei multiplen Vergleichen adjustiert werden müssen! **(2 Punkte)**
7. Würden Sie die Adjustierung der p -Werte oder die Adjustierung des Signifikanzniveaus α vorziehen? Begründen Sie Ihre Antwort! **(2 Punkte)**

85. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Compact Letter Displays (CLD) Steffen sitzt schon etwas länger bei seiner Betreuerin. So langsam macht Steffen sich Gedanken, ob er nicht doch mal anmerken sollte, dass er von CLD noch nie was gehört hat. Aber noch kann gelauscht werden, ein Ende ist erstmal nicht in Sicht! Steffen hatte in seiner Hausarbeit einen Versuch in einer Klimakammer durchgeführt. Deshalb sitzt er hier. Also eigentlich nein, deshalb nicht. Steffen will fertig werden. Hat er sich doch mit Lüftungssysteme (*ctrl*, *storm*, *thunder* und *tornado*) und Trockengewicht [kg/ha] schon eine Menge angeschaut. Steffen beugt sich leicht nach vorne. Nein, doch keine Pause. Weiter warten auf eine Lücke im Fluss...

Behandlung	Compact letter display
ctrl	a
storm	a
thunder	a
tornado	b

Leider kennen sich Steffen mit dem *Compact letter display (CLD)* überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich anhand des *Compact letter display (CLD)* ergebenden Barplots! **(2 Punkte)**
4. Ergänzen Sie das *Compact letter display (CLD)* zu den Barplots! **(1 Punkt)**
5. Erklären Sie *einen* Vorteil und *einen* Nachteil des *Compact letter display (CLD)*! **(2 Punkte)**
6. Erstellen Sie eine Matrix mit den paarweisen *p*-Werten eines Student t-Tests, die sich näherungsweise aus dem *Compact letter display (CLD)* ergeben würde! Begründen Sie Ihre Antwort! **(3 Punkte)**

86. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Berechnung des Compact Letter Displays (CLD) anhand von t-Tests Steffen hatte in seine Hausarbeit ein Feldexperiment durchgeführt. Soweit so gut. Dabei hat er sich mit Kartoffeln beschäftigt. Angeblich der neueste heiße Kram... aber das ist wiederum was anderes. So richtig mitgenommen hat Steffen das Thema dann doch nicht. Hat er sich doch mit Substrattypen (*kompost*, *torf*, *40p60n* und *70p30n*) und Proteingehalt [g/kg] schon eine Menge an Daten angeschaut. Nach seine Betreuerin soll er nun ein CLD bestimmen. Weder weiß er was ein CLD ist, noch war sein erster Gedanke mit Köln und die LGBTQ Community richtig... Als erstes solle er die Gruppen nach absteigender Effektstärke sortieren. Was immer das jetzt bringen soll.

Substrattypen	Fallzahl (n)	Mittelwert	Standardabweichung
kompost	8	10.46	2.25
torf	8	15.68	1.43
40p60n	8	15.05	4.01
70p30n	7	15.03	2.61

Leider kennen sich Steffen mit dem *Compact letter display (CLD)* überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich ergebenden Barplots! **(1 Punkt)**
4. Berechnen Sie die Matrix der *p*-Werte anhand von Student t-Tests! **(4 Punkte)**
5. Ergänzen Sie das *Compact letter display (CLD)* zu den gezeichneten Barplots! Begründen Sie Ihre Antwort! **(4 Punkte)**
6. Interpretieren Sie das *Compact letter display (CLD)* für Steffen und Yuki! **(1 Punkt)**

87. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Berechnung des Compact Letter Displays (CLD) anhand der Matrix der p-Werte 'Oh, nee!', ruft Yuki aus und rollt entnervt mit seinen Augen. Yuki hatte seine gesamte Analyse in SPSS gerechnet. Das war ja auch alles in Ordnung. Abbildungen haben geklappt und auch die statistischen Tests gingen dann irgendwie doch. Aber das CLD nicht. Yuki findet einfach keine Möglichkeit ein CLD in SPSS zu erhalten. Aber seine Betreuerin möchte unbedingt ein CLD. Sonst wird es mit der Abgabe nichts. Dabei hatte er schon wirklich eine Menge gemacht! Yuki hatte sich zwei Variablen mit Genotypen (00, AA, AB und BB) und Frischengewicht [kg/ha] in einen Versuch in einer Klimakammer mit Brokkoli angeschaut. Wo kriegt er jetzt ein CLD her? Dann eben per Hand aus der Matrix der p -Wert. Yuki stöhnt...

	00	AA	AB	BB
00	1.0000000	0.0000270	0.0222101	0.0000055
AA	0.0000270	1.0000000	0.0071719	0.2847865
AB	0.0222101	0.0071719	1.0000000	0.0009646
BB	0.0000055	0.2847865	0.0009646	1.0000000

Leider kennen sich Yuki mit dem *Compact letter display* (CLD) überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich anhand der Matrix der p -Werte ergebenden Barplots! **(2 Punkte)**
4. Ergänzen Sie das *Compact letter display* (CLD)! Begründen Sie Ihre Antwort! **(4 Punkte)**
5. Interpretieren Sie das *Compact letter display* (CLD) für Yuki und Yuki! **(2 Punkte)**

Teil VI.

Der Chi-Quadrat-Test & Der diagnostische Test

88. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Den Chi-Quadrat-Test berechnen Paula hat sich ein Herz gefasst und war für ihrer Hausarbeit in die Niederlande gegangen. Das war eine super Zeit in der sie viel gelernt hat. Klar gab es auch die ein oder andere Besonderheit, aber das gehört hier eher nicht hin. Dann noch schnell White Lies auf das Ohr und los gehts. Paula ist schon eine ganze Zeit im Büro, da ihre Betreuerin möchte, dass sie jetzt auf ihren Daten mit $n = 126$ Beobachtungen von Kartoffeln einen χ^2 -Test rechnet. Das ginge, da sie als Behandlung *Pestizideinsatz* [ja/nein] bestimmt und zum anderen die Variable *Frischegewicht über Zielwert* [ja/nein] ermittelt hat. Wie genau, das ist jetzt eine andere Frage. Eigentlich wollte Paula nachher noch einen Film schauen. Wenn Jagd auf roter Oktober läuft, dann ist die Ratte nicht mehr da. Aber jetzt braucht sie mal Entspannung!

	56	19	
	13	38	

Leider kennt sich Paula mit der Berechnung eines χ^2 -Test für kategoriale Daten überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Ergänzen Sie die Tabelle um die fehlenden Informationen! **(1 Punkt)**
3. Visualisieren Sie den Zusammenhang zwischen den beiden kategorialen Variablen! **(2 Punkte)**
4. Berechnen Sie die Teststatistik eines Chi-Quadrat-Test! **(2 Punkte)**
5. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem $\chi^2_{\alpha=5\%} = 3.841$! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Skizzieren Sie die χ^2 -Verteilung, wenn die H_0 wahr ist! Ergänzen Sie $\chi^2_{\alpha=5\%}$ und χ^2_D in der Abbildung! **(2 Punkte)**
7. Berechnen Sie den Effektschätzer *Cramers V*! Interpretieren Sie den Effektschätzer! **(2 Punkte)**

89. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Der Chi-Quadrat-Test konzeptionell verstehen 'Der χ^2 -Test auf einer 2x2-Kreuztabelle berechnet.', liest Paula in ihrer Mitschrift. So richtig helfen tut ihr das jetzt eherlichweise dann doch nicht. Dann noch schnell Smarties zur Stärkung und los gehts. Paula hatte sich in ein Kreuzungsexperiment $n = 113$ Beobachtungen von Fleischrindern angeschaut. Dabei hat er als Behandlung *Automatische Fütterung [ja/nein]* bestimmt und zum anderen die Variable *Schlachtgewicht im Zielbereich [ja/nein]* ermittelt. Am Ende möchte dann ihre Betreuerin gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Am Ende des Tages möchte sie dann noch ihr Hobby Harry Potter genießen. Das muss auch mal sein!

			48
			65
	58	55	113

Leider kennt sich Paula mit der Berechnung eines χ^2 -Test für kategoriale Daten überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Ergänzen Sie die Tabelle um die fehlenden Informationen! **(1 Punkt)**
3. Ergänzen Sie die Felder innerhalb der 2x2 Kreuztabelle, so dass *kein* signifikanter Effekt zu erwarten wäre! **(2 Punkte)**
4. Begründen Sie Ihr Vorgehen an der Formel des Chi-Quadrat-Tests. Erklären Sie Ihr Vorgehen an einem Beispiel! **(2 Punkte)**
5. Visualisieren Sie den Zusammenhang zwischen den beiden kategorialen Variablen! **(2 Punkte)**
6. Was ist die Mindestanzahl an Beobachtungen je Zelle? Wenn in einer der Zellen weniger Beobachtungen auftreten, welchen Test können Sie anstatt des Standard Chi-Quadrat-Tests anwenden? **(2 Punkte)**

90. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Der Chi-Quadrat-Test in R Am Ende war es für Jonas in seinem Projektbericht dann doch kein normalverteiltes Outcome. Das was jetzt etwas doof, da er sich auf eine ANOVA gefreut hatte. Dann noch schnell Mission Impossible starten und los gehts mit der Kraft von Snickers. Prinzipiell ginge das auch irgendwie, aber nun möchte sein Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Jonas hatte sich in ein Gewächshausexperiment $n = 113$ Beobachtungen von Erdbeeren angeschaut. Dabei hat er als Behandlung *Pestizideinsatz* [ja/nein] bestimmt und zum anderen die Variable *Frischegewicht über Zielwert* [ja/nein] ermittelt. Jetzt muss Jonas mal schauen, wie er das jetzt rechnet. Nach seinem Experiment erhielt er folgende 2x2 Kreuztabelle aus seinen erhobenen Daten.

```
##                               Pestizideinsatz
## Frischegewicht über Zielwert ja  nein
##                               ja    6   12
##                               nein 10    7
```

Dann rechnete Jonas den Fisher-Exakt-Test auf der 2x2-Kreuztabelle in R und erhielt folgende R Ausgabe der Funktion `fisher.test()`.

```
##
## Fisher's Exact Test for Count Data
##
## data: Frischegewicht über Zielwert
## p-value = 0.1811
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.07075797 1.67575062
## sample estimates:
## odds ratio
##  0.3611682
```

Leider kennt sich Jonas mit der Berechnung eines χ^2 -Test für kategoriale Daten überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Visualisieren Sie den Zusammenhang zwischen den beiden kategorialen Variablen! **(2 Punkte)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie das sich ergebende 95% Konfidenzintervall! **(2 Punkte)**
5. Beschriften Sie die Abbildung des 95% Konfidenzintervalls! **(1 Punkt)**
6. Interpretieren Sie das *Odds ratio* im Kontext der wissenschaftlichen Fragestellung! **(2 Punkte)**

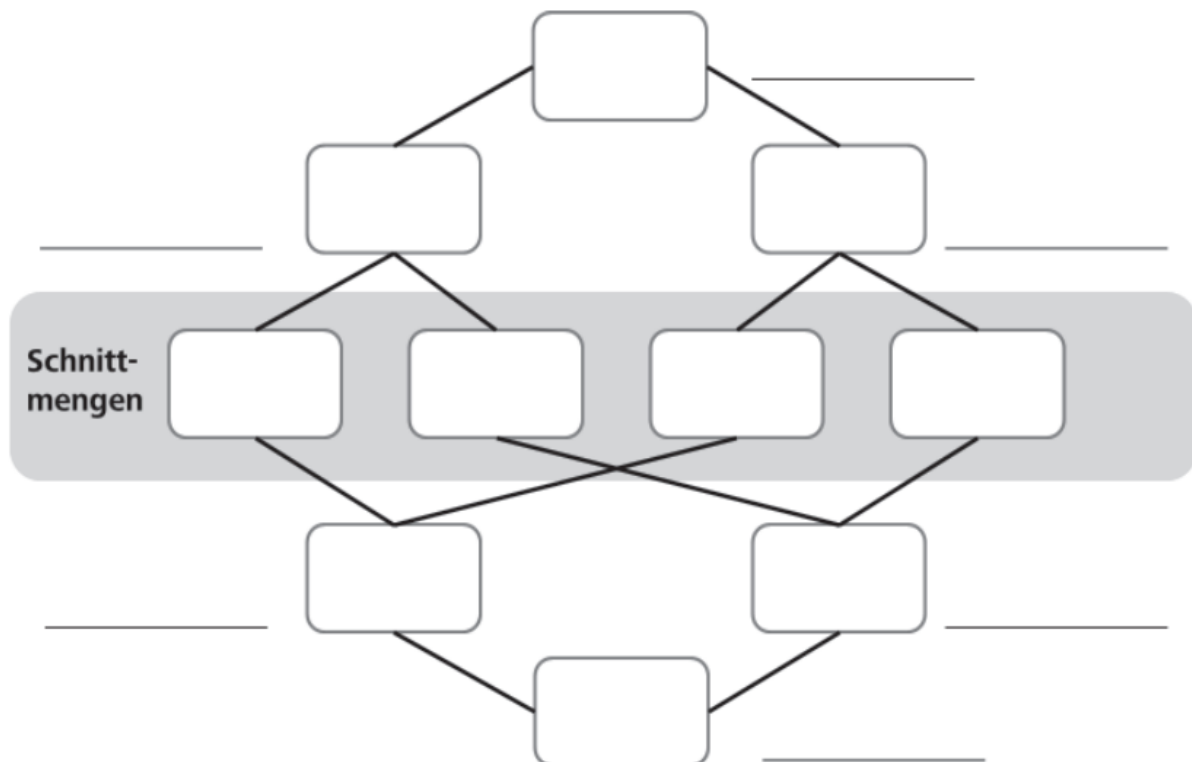
91. Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Den diagnostische Test am Doppelbaum berechnen Yuki liest laut vor. 'Die Prävalenz von Klauenseuche bei Brokkoli wird mit 4% angenommen. In 85% der Fälle ist ein Test positiv, wenn das Pflanze erkrankt ist. In 8.0% der Fälle ist ein Test positiv, wenn das Pflanze nicht erkrankt ist und somit gesund ist. Wir führen einen Test auf Kräuselkrankheit an 1000 Brokkoli mit einem diagnostischen Test durch.' Paula klappt die Kinnlade runter. In der Stille dudelt London Grammar. Yuki schaut kompetent und schmeißt sich mit offenen Mund Smarties an den Kopf vorbei.



Leider kennen sich Yuki und Paula mit dem diagnostischen Testen überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

1. Beschriften Sie die Äste des Doppelbaumes, mit denen Ihnen bekannten Informationen! **(2 Punkte)**
2. Beschriften Sie den Doppelbaum! **(2 Punkte)**
3. Füllen Sie freien Felder des Doppelbaums aus! **(4 Punkte)**
4. Berechnen Sie die Wahrscheinlichkeit $Pr(K^+|T^+)$! **(2 Punkte)**
5. Was sagt Ihnen die Wahrscheinlichkeit $Pr(K^+|T^+)$ aus? **(1 Punkt)**

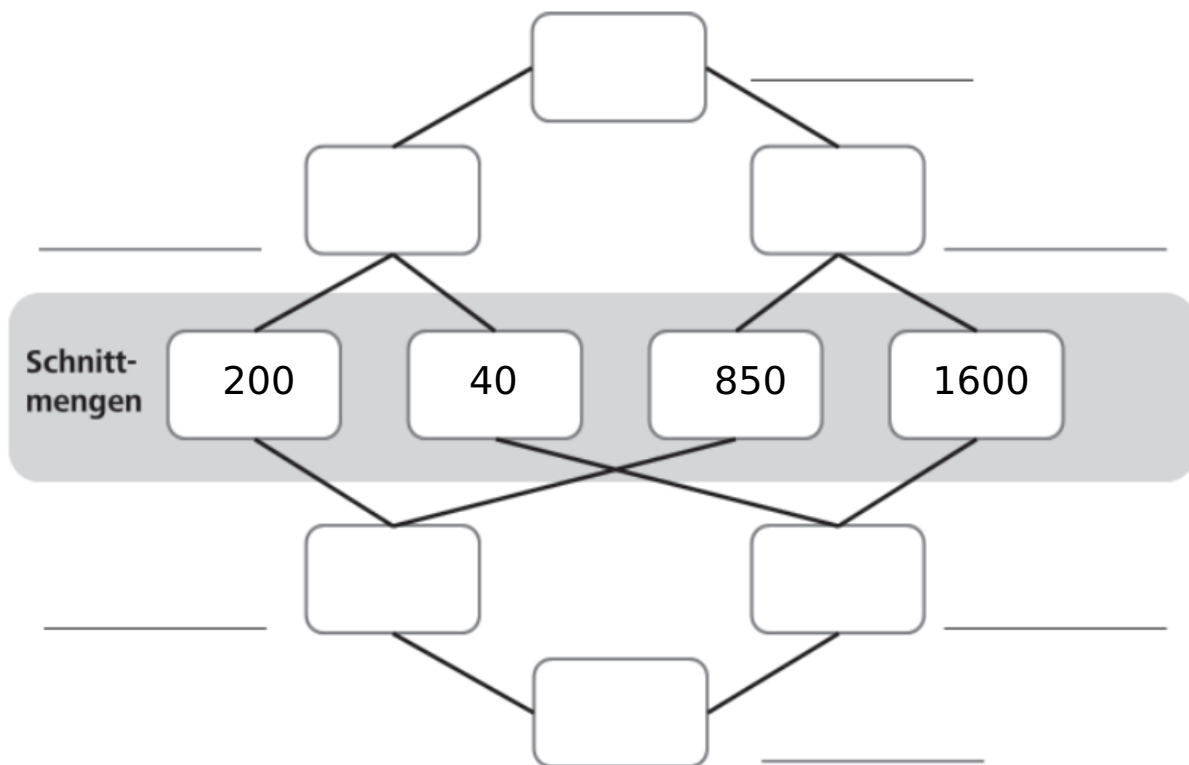
92. Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Der diagnostische Test und statistische Maßzahlen 'Was ist denn das?', entfährt es Jessica. 'Hm... ich glaube es handelt sich um einen Doppelbaum, den wir beim diagnostischen Testen brauchen.', meint Nilufar und dreht Herr der Ringe auf dem Second Screen etwas leiser. Was jetzt beide von einem diagnostischen Test haben, ist ihnen auch nicht klar. Es ist ja schon alles komplex genug und der Mangel von Jessica macht es heute auch nicht mehr einfacher. 'Es geht um Escherichia coli (E. coli) an Lamas.', stellt Nilufar fest. Eigentlich wollte Nilufar eher los um zu Kicken. Das wird aber wohl nichts mehr.



Leider kennen sich Jessica und Nilufar mit dem diagnostischen Testen überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

1. Beschriften Sie den Doppelbaum! **(2 Punkte)**
2. Füllen Sie freien Felder des Doppelbaums aus! **(4 Punkte)**
3. Berechnen Sie die Wahrscheinlichkeit $Pr(K^+|T^+)$! **(2 Punkte)**
4. Berechnen Sie die Prävalenz für Klauenseuche! **(1 Punkt)**
5. Berechnen Sie die Sensitivität und Spezifität des diagnostischen Tests! Erstellen Sie dafür zunächst eine 2x2 Kreuztabelle! **(2 Punkte)**

Teil VII.

Lineare Regression & Korrelation

93. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung der linearen Regression 'Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen!', ruft Paula. 'Ich sehe nur eine Zahlen und keine Punkte. Wie soll ich da denn jetzt eine Gerade durchzeichnen?', fragt Jonas. Paula atmet schwer ein. Die beiden hatten ein Stallexperiment im Oldenburger Land mit Hühnern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Tagestemperatur [C/d] und Gewichtszuwachs in der 1LW. Jetzt will die Betreuung von den beiden einmal die Visualisierung der Daten und auch gleich noch die lineare Regression gerechnet bekommen.

Durchschnittlicher Tagestemperatur [C/d]	Gewichtszuwachs in der 1LW
20.4	22.5
31.6	34.6
30.5	33.8
28.1	31.0
21.3	23.4
27.0	31.6
30.7	33.4
21.4	25.3
19.9	24.3
23.6	28.3
19.9	23.1
17.5	18.6

Leider kennen sich Paula und Jonas mit der linearen Regression für kontinuierliche Daten überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!


1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Erstellen Sie eine Visualisierung für die Datentabelle. Beschriften Sie die Achsen! **(2 Punkte)**
3. Schätzen Sie die Regressionsgleichung aus der obigen Abbildung ab! **(2 Punkte)**
4. Beschriften Sie die Grade mit den statistischen Maßzahlen der linearen Regressionsgleichung! **(2 Punkte)**
5. Liegt ein Zusammenhang zwischen x und y vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Wenn kein Zusammenhang zu beobachten wäre, wie würde die Grade aussehen? *Antworten Sie mit einer Skizze der Geraden!* **(1 Punkt)**

94. Aufgabe


(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interpretation der Ergebnisse einer linearen Regression 'Hä? Was ist denn das? Hatten wir das als Aufgabe eine lineare Regression zu rechnen? Wir bauen aus kontinuierlichen Daten eine Abbildung und interpretieren diese dann?', fragt Steffen. Paula schaut fragend zurück. 'Keine Ahnung... das ist jetzt jedenfalls keine Abbildung von irgendwas sondern Zahlen in einer Tabelle...', antwortet Paula leicht angespannt. Die beiden hatten ein Freilandversuch im Teutoburgerwald mit Kartoffeln durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Niederschlag [ml/w] und Proteingehalt [g/kg]. Jetzt haben die beiden eigentlich alles zusammen. *Eigentlich...*, denn mit der  Ausgabe haben beide jetzt ein Problem.

term	estimate	std.error	t statistic	p-value
(Intercept)	1.61	2.32		
Durchschnittlicher Niederschlag	2.06	0.23		

Leider kennen sich Steffen und Paula mit der linearen Regression für kontinuierliche Daten in  überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Erstellen Sie eine Visualisierung der `lm()`-Ausgabe. Beschriften Sie die Achsen! **(2 Punkte)**
3. Beschriften Sie die Visualisierung mit den statistischen Maßzahlen der `lm()`-Ausgabe! **(2 Punkte)**
4. Formulieren Sie die Regressionsgleichung! **(1 Punkt)**
5. Ergänzen Sie die t Statistik in der `lm()`-Ausgabe! **(2 Punkte)**
6. Ergänzen Sie den p-Wert in der `lm()`-Ausgabe mit $T_{\alpha=5\%} = 1.96$! **(2 Punkte)**
7. Interpretieren Sie den p-Wert im Kontext der wissenschaftlichen Fragestellung! **(1 Punkt)**
8. Wie groß ist der Effekt im Kontext der wissenschaftlichen Fragestellung? **(1 Punkt)**

95. Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interpretation der Ergebnisse einer linearen Regression in R 'Hä? Was ist denn das? Das wird ja immer wilder! Hatten wir das als Aufgabe eine lineare Regression zu rechnen? Wir bauen aus kontinuierlichen Daten eine Abbildung und interpretieren diese dann?', fragt Jonas. Yuki schaut fragend zurück. 'Keine Ahnung... das ist jetzt jedenfalls keine Abbildung von irgendwas sondern eine R Ausgabe mit ganz wilden Bezeichnungen...', antwortet Yuki leicht angespannt. Die beiden hatten ein Freilandversuch im Emsland mit Lauch durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Niederschlag [ml/w] und Proteingehalt [g/kg]. Jetzt haben die beiden eigentlich alles zusammen. *Eigentlich...*, denn mit der R Ausgabe haben beide jetzt ein Problem.

```
##
## Call:
## Proteingehalt ~ Durchschnittlicher_Niederschlag
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0686 -0.7677  0.1105  1.0391  3.2898
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.5441     1.6685  -0.326   0.746
## Durchschnittlicher_Niederschlag  2.2387     0.1600  13.987 6.76e-16
##
## Residual standard error: 1.468 on 35 degrees of freedom
## Multiple R-squared:  0.8483, Adjusted R-squared:  0.8439
## F-statistic: 195.6 on 1 and 35 DF, p-value: 6.761e-16
```

Leider kennen sich Jonas und Yuki mit der linearen Regression für kontinuierliche Daten in R überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Wie groß ist der Effekt im Kontext der wissenschaftlichen Fragestellung? **(2 Punkte)**
3. Interpretieren Sie die p -Werte im Kontext der wissenschaftlichen Fragestellung! **(2 Punkte)**
4. Visualisieren Sie die Verteilung der Residuen! **(2 Punkte)**
5. Ist die Annahme der Normalverteilung erfüllt? Begründen Sie die Antwort! **(2 Punkte)**
6. Erklären Sie *kurz* den Begriff R-squared! Was sagt Ihnen der Wert aus? **(2 Punkte)**

96. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Interpretation der Ergebnisse einer Korrelationsanalyse in R 'Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen...', denkt Jessica. 'Ich sehe nur Kauderwelsch und keine Punkte. Ich glaube das war jetzt doch eine Korrelation, die ich rechnen sollte. Und warum überhaupt? War das unsere Fragestellung?', denkt sich Jessica. Jessica atmet schwer ein und starrt auf die R Ausgabe der Funktion `cor.test()`. Das hilft alles nur begrenzt. Irgendwie komisch, wenn sie Herr der Ringe anmacht, dann ist die Hündin eigentlich sofort vor dem Bildschirm und starrt hinein. Jessica hatte ein Freilandversuch im Emsland mit Lauch durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche UV-Einstrahlung [UV/d] und Proteingehalt [g/kg]. Jetzt will die Betreuung von ihr die Interpretierung der Daten in Form einer Korrelation berechnet bekommen. Das hat Jessica in R gemacht, aber wie soll das jetzt gehen? Das mit der Interpretation? Eine echte Herausforderung für sie war schon immer der Mangel gewesen. Ein leidiges Lied.

```
##  
## Pearson's correlation  
##  
## data: Durchschnittliche UV-Einstrahlung and Proteingehalt  
## t = 4.7369, df = 8, p-value = 0.00147  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
##  0.4983744 0.9660024  
## sample estimates:  
##      cor  
## 0.8585855
```

Leider kennt sich Jessica mit der Korrelationsanalyse in R überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Erstellen Sie eine Visualisierung für den Korrelationskoeffizienten! Beschriften Sie die Abbildung! **(2 Punkte)**
4. Nennen Sie die zwei Eigenschaften des Korrelationskoeffizienten! **(2 Punkte)**
5. Interpretieren Sie den Korrelationskoeffizienten hinsichtlich des Effekts und der Signifikanz! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Visualisieren Sie das 95% Konfidenzintervall! Beschriften Sie die Abbildung! **(2 Punkte)**

97. Aufgabe

(12 Punkte)

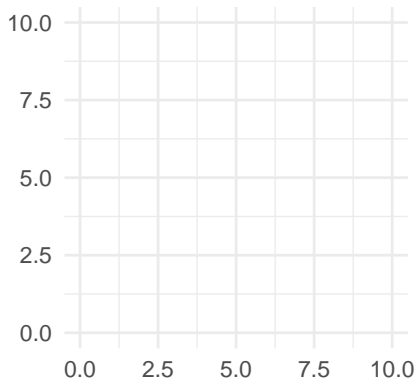
Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



'Im Folgenden sind die drei leeren Abbildungen zu füllen.', liest Nilufar und denkt nach. Nilufar kennt sich nur begrenzt bis gar nicht mit der linearen Regression und Korrelation aus. Dafür mit etwas anderem. Am Ende dann doch besser Hip Hop. Wunderbar. Eine echte Ablenkung für Nilufar. Aber das hilft hier auch nur so halb. Daher mampft sie noch ein paar Takis Blue Heat

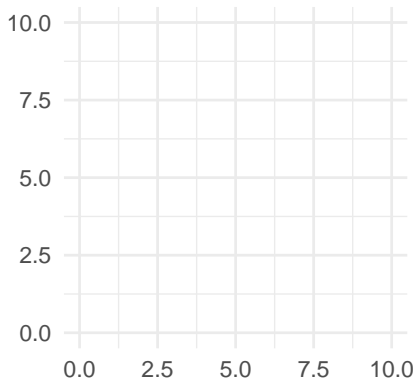
Pearsons $\rho = 0$

$R^2 = 0.5$



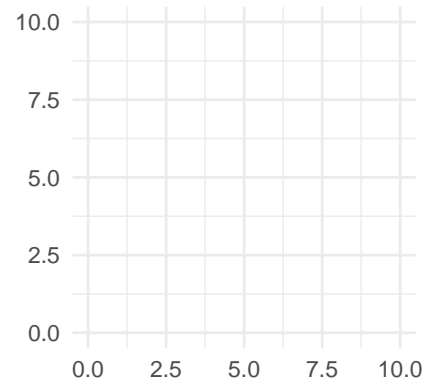
Pearsons $\rho = -0.75$

$R^2 = 0.75$



Pearsons $\rho = -0.5$

$R^2 = 1$



Visualisierung der Korrelation und des Bestimmtheitsmaßes Leider kennt sich Nilufar mit der Korrelationsanalyse und der linearen Regression überhaupt nicht aus. Deshalb braucht sie bei der Auswertung Ihre Hilfe!

1. Zeichnen Sie für die ρ -Werte eine Gerade in die entsprechende Abbildung! **(3 Punkte)**
2. Zeichnen Sie für die R^2 -Werte die entsprechende Punktwolke um die Gerade! **(3 Punkte)**
3. Nennen Sie die zwei Eigenschaften des Korrelationskoeffizienten! **(2 Punkte)**
4. Interpretieren Sie die R^2 -Werte für die jeweilige Gerade! **(2 Punkte)**
5. Warum müssen Sie ein R^2 -Wert berechnen, wenn Sie die einfachere Möglichkeit der visuellen Überprüfung haben? Begründen Sie Ihre Antwort! **(2 Punkte)**

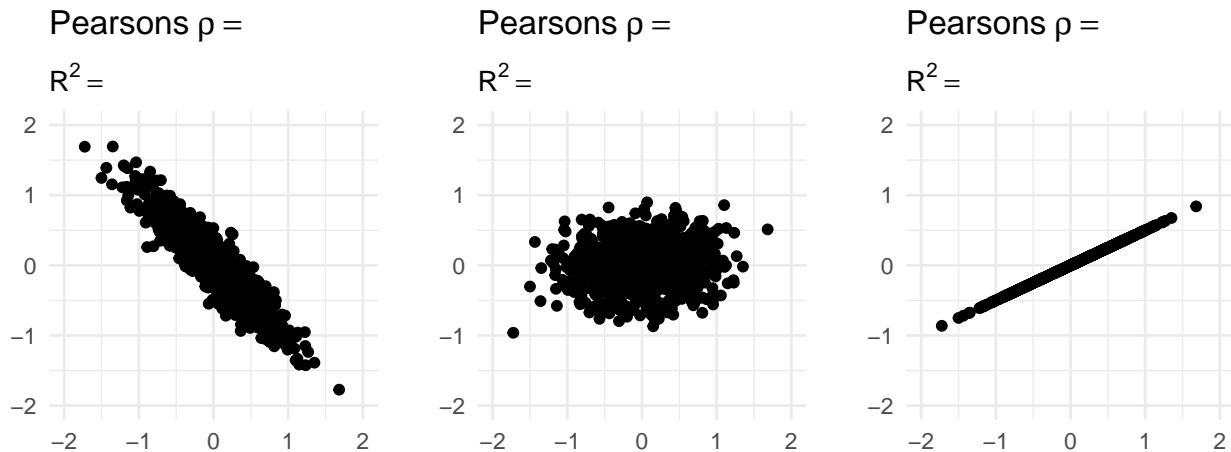
98. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Schätzen der Korrelation und des Bestimmtheitsmaßes Der Bildschirm strahlt blau in das Gesicht von Mark. Es ist schon spät. Und das hat einen Grund. Hm, lecker Marzipankugeln und dazu dann im Hintergrund Columbo laufen lassen. . Mark überlegt, aber seine Gedanken sind etwas zäh. 'Was soll das hier alles bedeuten?', fragt sich Mark. Irgendwie ist ihm nicht klar wie er ρ -Werte oder R^2 -Werte abschätzen soll. Alles nicht so einfach. Mark und die Unsicherheit, eine unendliche Geschichte mit kniffligen Wendungen.



Leider kennt sich Mark mit der Korrelationsanalyse und der linearen Regression überhaupt nicht aus. Deshalb braucht er bei der Auswertung Ihre Hilfe!

1. Schätzen Sie die ρ -Werte in den Abbildungen! **(2 Punkte)**
2. Schätzen Sie die R^2 -Werte in den Abbildungen! **(2 Punkte)**
3. Interpretieren Sie die R^2 -Werte für die jeweilige Gerade! **(2 Punkte)**
4. Was ist der optimale R^2 -Wert im Kontext einer wissenschaftlichen Fragestellung? Begründen Sie Ihre Antwort an einem Beispiel! **(2 Punkte)**
5. Was ist der optimale ρ -Wert im Kontext einer wissenschaftlichen Fragestellung? Begründen Sie Ihre Antwort an einem Beispiel! **(2 Punkte)**
6. Erklären Sie die Aussage "*Correlation does not imply causation!*" an einem Beispiel! **(2 Punkte)**

99. Aufgabe

(11 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Modellgüte der linearen Regression 'Oh! Residuen. Die waren wichtig um zu wissen, ob eine Modellierung funktioniert hat! Da schauen wir uns dann mit der Funktion `augment()` die Werte der einzelnen Residuen an. Oder gleich den Residuenplot...da sehen wir dann... ja was eigentlich?', verkündet Alex stolz. Leider hat Alex vergessen wie der `R` Code für den Residuenplot geht. Alex hatte anderes im Kopf. Hm, lecker Gummibärchen und dazu dann im Hintergrund Alien laufen lassen. Aber sowas hilft ihm natürlich hier nicht. Da schmeißt sich Alex noch ein paar Gummibärchen in den Mund und kaut los.

Durchschnittliche Regenwurmdichte	Trockengewicht	\hat{y}	ϵ
27.4	10.1	25.3	
32.3	14.2	31.8	
25.6	8.7	22.9	
23.9	10.0	25.0	
20.8	9.0	23.5	
28.5	12.1	28.4	
28.4	12.9	29.8	
24.5	10.0	25.1	
34.9	15.7	34.4	

Leider kennt sich Alex mit der linearen Regression überhaupt nicht aus. Deshalb braucht er bei der Auswertung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Ergänzen Sie die Werte der Residuen ϵ in der obigen Tabelle! **(2 Punkte)**
3. Zeichnen Sie den Boxplot der Residuen ϵ . Beschriften Sie die Abbildung! **(2 Punkte)**
4. Zeichnen Sie den Residualplot. Beschriften Sie die Abbildung! **(2 Punkte)**
5. Gibt es auffällige Werte anhand des Residualplots? Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Erklären Sie die Eigenschaft eines statistischen Modells, welche mit dem Residualplot überprüft wird! Begründen Sie Ihre Antwort anhand einer Visualisierung! **(2 Punkte)**

100. Aufgabe


(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Visualisierung des Regressionskreuzes Alex hat einen Leistungssteigerungsversuch mit Schweinen durchgeführt. Soweit so gut. Dann war er bei seiner Betreuerin. Leider war der Schritt nicht so hilfreich. Eine echte Herausforderung für ihn war schon immer die Gefälligkeit gewesen. Ein leidiges Lied. Aber es muss ja weitergehen. Alex hatte dann in seiner Abschlussarbeit einfach zu viele Endpunkte gemessen und ist jetzt vollkommen durcheinander, welche Analyse er nun wie rechnen soll. Naja, dann heißt es jetzt eben Abba aufdrehen und darüber nachdenken, was hier eigentlich gemacht wurde. Alex fängt einfach an und nimmt den ersten Endpunkt Schlachtgewicht [kg]. Dann kann er sich voran arbeiten. Später dann noch raus um zu Laufen um mal zu entspannen und vielleicht ist Steffen auch da. Wäre toll.

Leider kennt sich Alex mit dem Kontext der linearen Regression überhaupt nicht aus. Deshalb braucht er bei der Auswertung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Zeichnen Sie die Zeile des Regressionskreuzes für den Endpunkt mit drei Feldern! Beschriften Sie die Abbildung! **(4 Punkte)**
3. Ergänzen Sie die entsprechenden statistische Methoden zur Analyse in jedem Feld! **(2 Punkte)**
4. Formulieren Sie die Nullhypothese für die statistische Methode in jedem Feld! **(2 Punkte)**
5. Ergänzen Sie die entsprechenden Funktionen in  zur Analyse in jedem Feld! **(2 Punkte)**
6. Welchen Effekt erhalten Sie in jedem Feld? Geben Sie ein Beispiel! **(2 Punkte)**


Teil VIII.

Experimentelles Design




101. Aufgabe

(16 Punkte)



Einfache experimentelle Designs Jessica und Nilufar sind bei Mark um sich Hilfe für eine Versuchsplanung in  zu holen. Im Hintergrund läuft viel zu laut Andrea Berg. Dabei geht es um den Zusammenhang zwischen der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) und dem Messwert Fettgehalt [%/kg] in Hühnern. Der Versuch soll in einem Leistungssteigerungsversuch im Oldenburger Land durchgeführt werden. Nach dem Dozenten ist der Messwert Fettgehalt [%/kg] normalverteilt. Die beiden entschieden sich für ein faktorielles Versuchsdesign. Im ersten Schritt überlegt Mark ein einfaches experimentelles Design zu probieren. Daher entscheiden sich alle drei für ein *Complete randomized design (CRD)*. Das sollte für den Anfang erstmal reichen. 'Und jetzt, was machen wir jetzt?', Nilufar schaut die anderen beiden mit großen Augen an. Die zucken mit der Schulter. Alle mampfen Marzipankugeln.

Leider kennen sich Mark, Jessica und Nilufar mit dem *Complete randomized design (CRD)* überhaupt nicht aus. Deshalb brauchen die Drei bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Skizzieren Sie das faktorielle Versuchsdesign! **(3 Punkte)**
4. Skizzieren Sie eine Datentabelle für das faktorielle Versuchsdesign in ! **(2 Punkte)**
5. Erstellen Sie das statistische Modell in der in  üblichen Schreibweise für eine ANOVA! Skizzieren Sie die notwendige Funktionen in ! **(3 Punkte)**
6. Skizzieren Sie die weitere Datenanalyse hinsichtlich eines multiplen Gruppenvergleiches! **(2 Punkte)**
7. Skizzieren Sie eine mögliche Abbildung im Kontext der wissenschaftlichen Fragestellung! Beschriften Sie die Abbildung! **(2 Punkte)**
8. Ergänzen Sie zu der Abbildung ein mögliches Ergebnis des multiplen Gruppenvergleichs! Begründen Sie Ihre Antwort! **(2 Punkte)**

102. Aufgabe

(20 Punkte)



Fortgeschrittene experimentelle Designs Neuer Versuch neues Glück! Tina und Yuki sind bei Nilufar um sich Hilfe für eine Versuchsplanung in **R** zu holen. Im Hintergrund läuft viel zu laut Columbo. Daher hat der Hamster schon lange reiaus genommen. In dem neuen Versuch geht es um den Zusammenhang zwischen der Behandlung Elterlinie (*ctrl*, *Standard*, *Yray* und *Xray*) sowie Bestandsdichte (*standard* und *kontakt*) sowie drei Blcken und dem Messwert Protein/Fettrate [%/kg] in Fleischrindern. Der Versuch soll in einem Stallversuch im Emsland durchgefhrt werden. Immerhin ist der Messwert normalverteilt, was einiges einfacher macht. Was es nicht so einfacher macht ist, dass Tina noch als zustzliche Herausforderung etwas anderes umtreibt: der Mangel. Im ersten Schritt berlegt Nilufar ein komplexeres experimentelles Design zu probieren. Daher entscheiden sich alle drei fr ein *Strip plot design* oder auch *Streifenanlage*. Ob es das jetzt einfacher macht?

Leider kennen sich Nilufar, Tina und Yuki mit dem *Strip plot design* oder auch *Streifenanlage* berhaupt nicht aus. Deshalb brauchen die Drei bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistische Hypothesenpaare! **(2 Punkte)**
3. Skizzieren Sie das faktorielle Versuchsdesign! **(3 Punkte)**
4. Skizzieren Sie eine Datentabelle fr das faktorielle Versuchsdesign in **R**! **(2 Punkte)**
5. Erstellen Sie das statistische Modell in der in **R** blichen Schreibweise fr eine ANOVA! Skizzieren Sie die notwendige Funktionen in **R**! **(4 Punkte)**
6. Skizzieren Sie die weitere Datenanalyse hinsichtlich eines multiplen Gruppenvergleiches! **(2 Punkte)**
7. Skizzieren Sie eine mgliche Abbildung im Kontext der wissenschaftlichen Fragestellung! Beschriften Sie die Abbildung! **(3 Punkte)**
8. Ergnzen Sie zu der Abbildung ein mgliches Ergebnis des multiplen Gruppenvergleichs! Welche Annahme hinsichtlich der Modellierung haben Sie getroffen? Begrnden Sie Ihre Antwort! **(3 Punkte)**

Teil IX.

Programmieren in R

103. Aufgabe

(9 Punkte)



Grundlegende Kenntnisse der Programmierung in R 'Hm. R ist eigentlich gar nicht so schwer, wenn man die Grundlagen kann.', meint Jessica ganz zuversichtlich. Nur leider kennt sie sich überhaupt nicht mit R aus! Das heißt, Sie müssen hier einmal Rede und Antwort stehen und helfen.

Jessica: *Was ist eigentlich ein Faktor in R?* (1 Punkt)

Sie antworten:

Jessica: *In R gibt es Objekte, Wörter und Funktionen. Wie unterscheiden sich diese voneinander?* (1 Punkt)

Sie antworten:

Jessica: *Ich verstehe den Unterschied zwischen library() und Packages nicht. Warum gibt es die?* (1 Punkt)

Sie antworten:

Jessica: *Wie sieht der Pipe-Operator aus und was ist seine Funktion? Gerne mit Beispiel!* (1 Punkt)

Sie antworten:

Jessica: *Ich habe den Namen der Funktion, die intern Daten speichert, vergessen. Was waren da nochmal die Vorteile?* (1 Punkt)

Sie antworten:

Jessica: *Warum gibt es eigentlich das RStudio und R? Wie unterscheiden sich beide voneinander?* (1 Punkt)

Sie antworten:

Jessica: *Ich sehe überall dieses c(). Was ist denn deren Nutzen?* (1 Punkt)

Sie antworten:

Jessica: *Wo nutzen wir nochmal die Tilde (~) in R. Das war irgendwie voll wichtig. Wo wird diese genutzt?* (1 Punkt)

Sie antworten:

Jessica: *Ich habe gehört, dass es Vorteile gibt R zu nutzen. Nenne mir mal einen Vorteil!* (1 Punkt)

Sie antworten:

104. Aufgabe

(9 Punkte)



Fortgeschrittene Kenntnisse der Programmierung in R 'Unter den Blinden ist der Einäuge König!', ruft Ihnen Jonas entgegen. Das können Sie schon nicht mehr hören. Nur weil Sie einmal gesagt haben, dass Sie sich schonmal mit R beschäftigt haben, stehen hier alle Schlange. Aber gut, das hat Sie dann doch vorangebracht. Leider kennt sich Jonas auch wieder überhaupt nicht mit R aus aber sein Betreuer möchte gerne, dass die Auswertung in R gemacht wird. Da müssen Sie dann wohl mal nochmal ran und helfen.

Jonas fragt: *Das R Paket {ggplot} erlaubt es ja schöne Abbildungen zu bauen. Wie verbindet das Paket nochmal die einzelnen Ebenen einer Abbildung? (1 Punkt)*

Sie antworten:

Jonas fragt: *Ich habe den Faktor f_1 und f_2 und möchte den Faktor f_1 getrennt in jedem Level des Faktors f_2 auszuwerten. Wie geht das in der Funktion `emmeans()`? (1 Punkt)*

Sie antworten:

Jonas fragt: *Ich möchte ein CLD erstellen. Welche Funktionen muss ich in welcher Reihenfolge nutzen? (2 Punkte)*

Sie antworten:

Jonas fragt: *Wie spezifizieren wir nochmal eine Interaktion in einem Modell mit zwei Faktoren f_1 und f_2 ? (1 Punkt)*

Sie antworten:

Jonas fragt: *Datumsangaben sind schwierig, da es nur ein gültiges Format gibt, was zwischen Programmen funktioniert. Wie lautet das Format? (1 Punkt)*

Sie antworten:

Jonas fragt: *ANOVA in R ist ja nicht so kompliziert. Welche beiden Funktionen brauche ich nochmal in welcher Reihenfolge? (1 Punkt)*

Sie antworten:

Jonas fragt: *Wenn ich Daten in R mit Gruppen eingelesen habe, welche Funktion nutze ich dann meistens als erstes und warum muss ich das machen? Was muss ich da machen? (2 Punkte)*

Sie antworten:

Teil X.

Forschendes Lernen

Das forschende Lernen basiert zum einen auf den folgenden wissenschaftlichen Veröffentlichungen. Für die Prüfung wird die vertiefende Kenntnis der folgenden Veröffentlichungen vorausgesetzt.

In der Prüfung erhalten Sie einen Auszug der wissenschaftlichen Veröffentlichung. Für die Einarbeitung in die Veröffentlichung ist in der Prüfung ausdrücklich keine Zeit vorgesehen.

- Sánchez, M., Velásquez, Y., González, M., & Cuevas, J. (2022). Hoverfly pollination enhances yield and fruit quality in mango under protected cultivation. *Scientia Horticulturae*, 304, 111320. [[Link](#)]
- Petersen, F., Demann, J., Restemeyer, D., Ols, H. W., Westendarp, H., Appenroth, K. J., & Ulbrich, A. (2022). Influence of light intensity and spectrum on duckweed growth and proteins in a small-scale, re-circulating indoor vertical farm. *Plants*, 11(8), 1010. [[Link](#)]
- Selle, P. H., Cadogan, D. J., Li, X., & Bryden, W. L. (2010). Implications of sorghum in broiler chicken nutrition. *Animal Feed Science and Technology*, 156(3-4), 57-74. [[Link](#)]
- Wu, G., Knabe, D. A., & Kim, S. W. (2004). Arginine nutrition in neonatal pigs. *The Journal of Nutrition*, 134(10), 2783S-2790S. [[Link](#)]

Das forschende Lernen basiert zum anderen auf den folgenden wissenschaftlichen Datensätzen und deren vertiefende Analyse werden als bekannt vorausgesetzt. Die Teilaufgaben der Aufgaben stellen nur eine zufällige Auswahl an möglichen Fragen dar. Die Datensätze werden über ILIAS bereitgestellt.

In der Prüfung erhalten Sie keinen Auszug aus den wissenschaftlichen Daten. Die Datensätze werden als bekannt in der Prüfung vorgelegt. Sie haben sich vorab Notizen und Anmerkungen gemacht.

- bar

105. Aufgabe



(20 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Zerforschen einer wissenschaftlichen Veröffentlichung Yuki hält die wissenschaftliche Veröffentlichung Petersen, F., et al. (2022). *Influence of light intensity and spectrum on duckweed growth and proteins in a small-scale, re-circulating indoor vertical farm* unter einem Schnaufen in die Luft. 'Worum geht es denn eigentlich in dieser Arbeit?', fragt er stirnrunzelnd und wirft die Arme in die Luft, da hilft dann auch nicht mehr die beruhigende Wirkung von London Grammar. Yuki soll die Veröffentlichung nutzen um das eigene Experiment zu planen. Als eine Vorlage sozusagen. Daher möchte sein Betreuer, dass er einmal die Veröffentlichung sinnvoll zusammenfasst. Das sollte dann doch etwas aufwendiger werden. Das wird dann vermutlich heute Abend nichts mehr mit seinem Hobby Orchideen. Das Minischwein schaut mitleidig.

Leider kennt sich Yuki mit dem Lesen einer wissenschaftlichen Veröffentlichung mit Fokus auf die Statistik überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe! Glücklicherweise kennen Sie die wissenschaftliche Veröffentlichung schon im Detail und können sofort helfen.

1. Erläutern Sie die wissenschaftliche Fragestellung der wissenschaftlichen Veröffentlichung anhand des OCAR Prinzips nach Schimel (2012)³ **(4 Punkte)**
2. Nennen Sie die untersuchten Endpunkte in der wissenschaftlichen Veröffentlichung! Wie lautet der primäre Endpunkt? **(2 Punkte)**
3. Erstellen Sie das statistische Modell in der in  üblichen Schreibweise! **(2 Punkte)**
4. Nennen Sie eine Auswahl an bedeutenden statistischen Maßzahlen in der wissenschaftlichen Veröffentlichung! **(1 Punkt)**
5. Interpretieren Sie die Hauptaussage der wissenschaftlichen Veröffentlichung hinsichtlich der Signifikanz für den primären Endpunkt! **(2 Punkte)**
6. Interpretieren Sie die Hauptaussage der wissenschaftlichen Veröffentlichung hinsichtlich der Effektstärke für den primären Endpunkt! **(2 Punkte)**
7. Diskutieren Sie die ökonomische Relevanz der Hauptaussage der wissenschaftlichen Veröffentlichung im Bezug auf Signifikanz und Effektstärke für den primären Endpunkt! **(1 Punkt)**
8. Skizzieren Sie für den primären Endpunkt den sich ergebenden Datensatz in  für eine ausgewählte Abbildung! **(2 Punkte)**
9. Skizzieren Sie einen möglichen Versuchsplan für den primären Endpunkt! **(2 Punkte)**
10. Schätzen Sie die benötigte Fallzahl für ein zukünftiges Experiment anhand der Ergebnisse in der wissenschaftlichen Veröffentlichung für den primären Endpunkt! **(2 Punkte)**

³Schimel, J. (2012). Writing science: how to write papers that get cited and proposals that get funded. OUP USA.

106. Aufgabe


(20 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Zerforschen eines wissenschaftlichen Datensatzes Vor dem Start der eigenen Arbeit möchte sein Betreuer, dass Steffen einmal die wissenschaftlichen Daten *data3* sinnvoll zusammenfasst. Dann würde die eigene Arbeit auch leichter von der Hand gehen und Steffen hätte dann schon eine Vorlage um die eigenen erhobenen Daten in eine Tabelle eintragen zu können. 'Das ist jetzt aber umfangreicher als gedacht!', schnauft er und runzelt die Stirn als er in seinen Laptop starrt. Dabei isst er noch ein paar Oreos. Das wird dann vermutlich heute Abend nichts mehr mit Harry Potter

Leider kennt sich Steffen mit der Analyse eines wissenschaftlichen Datensatzes überhaupt nicht aus. Deshalb braucht er bei der Auswertung Ihre Hilfe! Glücklicherweise kennen Sie den wissenschaftlichen Datensatz aus Ihren eigenen Analysen schon im Detail und können sofort helfen.

1. Formulieren Sie die wissenschaftliche Fragestellung des Datensatzes in Form einer PowerPoint Folie! **(2 Punkte)**
2. Nennen Sie zwei Besonderheiten des Datensatzes! Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Nennen Sie die untersuchten Endpunkte in dem Datensatz! Wie lautet der primäre Endpunkt für die Auswertung? **(2 Punkte)**
4. Skizzieren Sie die großen Analysebereiche der Statistik! Beschriften Sie die Abbildungen! **(2 Punkte)**
5. In welchen der großen Analysebereiche der Statistik fällt die Auswertung des primären Endpunktes? Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Skizzieren Sie eine ikonische Abbildung für den primären Endpunkt im Kontext der wissenschaftlichen Fragestellung! **(2 Punkte)**
7. Erstellen Sie das statistische Modell in der in  üblichen Schreibweise! **(2 Punkte)**
8. Skizzieren Sie die Datenanalyse hinsichtlich der Signifikanz für den primären Endpunkt! **(2 Punkte)**
9. Skizzieren Sie die Berechnung der Effektstärke für den primären Endpunkt! **(2 Punkte)**
10. Skizzieren Sie einen möglichen Versuchsplan für den primären Endpunkt! **(2 Punkte)**

Teil XI.

Mathematik

107. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Herodot – der Schimmel aus Ivenack Die Lerngruppe *Die Pantoffeltieren* bestehend aus Alex, Stefan, Yuki und Paula waren auf Exkursion in Mecklenburg-Vorpommern und haben dort Folgendes erarbeitet. Während der Besetzung Mecklenburgs durch die Franzosen kamen Napoleon die Geschichten des berühmten Apfelschimmels Herodot aus Ivenack zu Gehör. Herodot lief zwar niemals Rennen, war aber eines der berühmtesten Pferde dieser Zeit. Napoleon selbst gab den Auftrag, diesen Schimmel durch die Armee nach Frankreich zu bringen. Der Legende nach sollen Arbeiter den Schimmel im hohlen Stamm einer 1000-jährigen Eiche aus Ivenack vor den Franzosen versteckt haben. Doch Herodot verriet sein Versteck durch lautes Wiehern, woraufhin die französische Armee den Schimmel beschlagnahmte und nach Frankreich führte⁴. Jetzt wollen die vier herausfinden: "Konnten die Ivenacker den Apfelschimmel Herodot vor dem Zugriff von Napoleon in der 1000-jährigen Eiche verstecken?"

Helfen Sie der Lerngruppe *Die Pantoffeltieren* bei der Beantwortung der Forschungsfrage! Gehen Sie von einem radialen Wachstum der 1000-jährigen Eiche von 0.8mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 11m in Brusthöhe hatte.

1. Wie groß war der Durchmesser in m der Eiche im Jahr 1810 als Herodot in der Eiche versteckt werden sollte? **(2 Punkte)**
2. Skizzieren Sie in einer Abbildung einen linearen Zusammenhang und einen exponentiellen Zusammenhang für das Wachstum der 1000-jährigen Eiche. Erklären Sie die Auswirkungen der Entscheidung für linear oder exponentiell auf Ihre Berechnungen! **(2 Punkte)**

Herodot hatte eine Schulterhöhe von 190cm , eine Breite von 90cm sowie eine Länge von 250cm .

3. Berechnen Sie das effektive Volumen von Herodot in m^3 , welches Herodot in der 1000-jährigen Eiche einnehmen würde! **(2 Punkte)**

Es wurde berichtet, dass sich Herodot in der 1000-jährigen Eiche *bequem* um die eigene Achse drehen konnte.

4. Berechnen Sie die Dicke der Eichenwand in cm ! Verdeutlichen Sie Ihre Berechnungen an einer aussagekräftigen Skizze für Pferd und Eiche! **(2 Punkte)**
5. Unter einer Dicke der Eichenwand von 15cm bricht die Eiche zusammen. Beantworten Sie die Forschungsfrage! Begründen Sie Ihre Antwort! **(2 Punkte)**

⁴Die Quelle der Inspiration für die Aufgabe war eine Fahrt an die Ostsee und folgender Artikel: [Entdecke das erste Nationale Naturmonument Deutschlands - Ivenacker Eichen und Hutewald](#)

108. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Von Töpfen auf Tischen Die Projektgruppe F bestehend aus Jonas, Jessica, Steffen und Alex hat sich zusammengefunden um den ersten Versuch zu planen. In einem Experiment wollen sie die Wuchshöhe von 120 Stockrosen bestimmen. Bevor die Vier überhaupt mit dem Experiment beginnen können, gibt es aber ein paar Abschätzungen über die Kosten und den Aufwand zu treffen. Zum einen müssen sie die Stockrosen einpflanzen und müssen dafür Substrat bestellen. Zum anderen muss die Projektgruppe die Stockrosen auch bewegen und in ein Gewächshaus platzieren. Die Töpfe für die Keimung haben einen Durchmesser von 8.5cm und eine Höhe von 7cm. Der Kubikmeterpreis für Torf liegt bei 280 EUR.

Helfen Sie der Projektgruppe F bei der Planung des Versuches!

1. Skizzieren Sie den Versuchsplan auf *vier* Tischen im Gewächshaus! **(2 Punkte)**
2. Berechnen Sie die benötigte Anzahl an Pflanztöpfen, wenn Sie Randpflanzen mit berücksichtigen wollen! **(1 Punkt)**
3. Welche *Pflanztopf*fläche in m^2 gegeben der Anzahl an Pflanztöpfen inklusive Randpflanzen benötigen Sie im Gewächshaus am Anfang der Keimungsphase? **(3 Punkte)**
4. Berechnen Sie die benötigte Menge an Torf in Liter l , die Sie für das Befüllen der Pflanztöpfe benötigen! Gehen Sie von *einem Zylinder* für die Pflanztöpfe aus! **(3 Punkte)**
5. Berechnen Sie die Kosten in EUR für Ihre Torfbestellung! **(1 Punkt)**

109. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Solar- & Biogasanlagen Jessica bringt ein neues, tolles Projekt mit in die Lerngruppe *Die Kühe auf dem Deich* bestehend aus ihr, Tina, Jonas sowie Alex. Um die Energiekosten ihres Betriebes zu senken, will sie eine Solaranlage auf den Hühnerstall montieren lassen. Dafür hat sie ihren Stall ausgemessen und findet folgende Maße wieder. Die vordere Seite des Hühnerstall hat eine Höhe h_v von $6.5m$. Die hintere Seite des Hühnerstall hat eine Höhe h_b von $9.5m$. Der Hühnerstall hat eine Tiefe t von $15m$ und eine Breite b von $30m$. 'Sag mal Jessica, ist das eine Matheaufgabe oder rechnen wir hier gerade für dich kostenlos als menschliche Computer Sachen für deinen Betrieb?', fragt Jonas mit erhobenen Augenbrauen. Alex und Tina nicken zustimmend.

Wenn die Lerngruppe nicht will, dann müssen Sie bei der Planung helfen!

1. Skizzieren Sie den Hühnerstall auf dem die Solaranlage montiert werden soll! Ergänzen Sie die Angaben für die Höhen h_v , h_b , die Tiefe t und die Breite b des Stalls! **(2 Punkte)**
2. Berechnen Sie die Fläche der schrägen, neuen Solaranlage auf dem Hühnerstall! **(3 Punkte)**

Ebenfalls plant Jessica eine neue Biogasanlage für ihren Betrieb. Der neue Methantank hat einen Radius r von $1.8m$. Leider gibt es ein paar bauliche Beschränkungen auf dem Grundstück. Ihr Fundament des zylindrischen Methantanks kann nur ein Gewicht von maximal $12t$ aushalten bevor der Tank wegbricht. Jessica rechnen eine Sicherheitstoleranz von 25% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $220kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $270kg/m^3$. Jessica betreibt ihre Anlage bei $-92^\circ C$.

3. Extrapolieren Sie die effektive Dichte des Methans in Ihrem Methantank! Welche Annahme haben Sie getroffen? **(1 Punkt)**
4. Berechnen Sie wie viel Kubikmeter m^3 Sie in den Methantank füllen können, bevor das Fundament nachgibt! **(2 Punkte)**
5. Berechnen Sie die maximale Höhe h_{max} in m für den gefüllten Methantank mit dem Radius r , bevor das Fundament wegbricht! **(2 Punkte)**

110. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Stichworte: Riesenfaultier • Evolution der Avocado • Bluetooth • Blauzahn • Colonia Dignidad • ODESSA • Rattenlinie • Adolf Eichmann

Aligatorenbirnen und Blaubeeren “Sind Sie ein Riesenfaultier oder warum kaufen Sie so viele Aligatorenbirnen?”, spricht es hinter Ihnen. Irritiert drehen Sie sich um und blicken in das puderröte Gesicht von Alex. “Wieso?“, entfährt es Ihnen und Sie bereuen sogleich die Frage. Sofort werden Sie zu einem Whiteboard voller roter Schnüre geschliffen und müssen folgenden mathematischen untermauerten Argumenten im Lidl über sich ergehen lassen. Da kommen Sie nicht mehr raus, also können Sie auch gleich mitmachen. Das Problem liegt in Chile⁵. Tja, die Deutschen und Südamerika.

Zuerst werden Ihre Fähigkeiten getestet, der Mathematik folgen zu können. Oder berechnen Sie gerade den Einkauf von Alex?

1. Wenn 5 Blaubeerschalen 7.95 Euro kosten, wie viel kosten 8 Schalen? **(2 Punkte)**
2. Wenn Sie die 8 Blaubeerschalen gekauft haben, wie viele Aligatorbirnen zu je 2.89 EUR können Sie sich dann noch für 200 EUR leisten? **(1 Punkt)**

Das Whiteboard beinhaltet folgende Liste mit Informationen zum Wasserverbrauch bei der Produktion von Gemüse aus Chile. Seltsam, was man so alles in einem Lidl über Gemüse erfährt.

- Ein Kilo Strauchtomaten benötigt 190l Wasser. Eine Strauchtomate wiegt 100 - 120g.
- Ein Kilo Salat benötigt 130l Wasser. Ein Salatkopf wiegt 310 - 490g.
- Ein Kilo Avocado benötigt 1050l Wasser. Eine Avocado wiegt 120 - 400g.
- Ein Kilo Blaubeeren benötigt 880l Wasser. Eine Blaubeere wiegt 3.2 - 3.7g.

3. Berechnen Sie den Wasserverbrauch für die Produktion für jeweils eine Strauchtomate, einem Salat, einer Avocado und einer Blaubeeren. Stellen Sie das Ergebnis als Tabelle dar! **(3 Punkte)**

Chile exportiert im großem Ausmaß Blaubeeren und Avocados. In dem Exportjahr 2023 blieben die Erträge von Blaubeeren mit 7.7×10^4 t in dem prognostizierten Rahmen. Die Menge steigerte sich um 6.1%. Die Exporte für Avocados stiegen in dem gleichen Zeitraum um 17.3% auf 2.1×10^5 t.

4. Wie viele Kubikmeter Wasser hat Chile in dem Exportjahr 2022 exportiert? **(2 Punkte)**

Chile ist eines der wenigen Länder der Welt, die ihr Wasser komplett privatisiert haben. Derzeit sind nur zwei Prozent des Wassers des Landes für den häuslichen Verbrauch vorgesehen. In den Dörfern der Anbauregionen versorgen Tankwagen die Bevölkerung jede Woche mit Wasser, es gibt etwa 52 Liter Wasser pro Kopf für den täglichen Bedarf. In *Deutschland* liegt der Verbrauch bei 35 - 115 Liter pro Waschgang einer Waschmaschine und 10 - 15 Liter pro Minute Duschen.

5. Mit der rationierten Wassermenge aus Chiles Anbaugebieten können Sie in *Deutschland* wie oft Ihren Bedarf stillen? **(1 Punkt)**

Das alles hätten Sie nicht von Alex erwartet. Ganz schön viele Informationen wurden da zusammengetragen.

6. Nennen Sie eine *Datenquelle* im Internet, wo Sie mehr Informationen zu landwirtschaftlichen Daten oder klimatischen, wirtschaftlichen und gesellschaftlichen Daten erhalten! **(1 Punkt)**

⁵Die Quelle der Inspiration für die Aufgabe waren folgende Reportagen: “Bis zum letzten Tropfen” in AMNESTY – Magazin der Menschenrechte vom August 2021 und “Wasserknappheit in Chile: Eine Folge der Privatisierung?” in Die Welternährung dem Fachjournal der Welthungerhilfe vom April 2022.

111. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Stichworte: Kardaschow-Skala • Dyson-Sphäre • Hohlerde • Entropie • Proton $r_p = 1.7 \times 10^{-15}$ • Wasserstoff $r_H = 5.3 \times 10^{-11}$

Die Dampfnudelerde “Was für einen Unsinn!“, rufen Sie. Jetzt haben Sie auf Empfehlung von von Yuki kostbaren Schlaf prokrastiniert um einem Ernährungswissenschaftler auf YouTube über die Erde als Dampfnudel zu lauschen. Irgendwie passt es dann doch mit der Analogie. Übermüdet müssen Sie darüber nachdenken, warum vor 66 Millionen Jahren die Dinosaurier - so groß sie auch waren - nicht von der Schwerkraft zu Boden gerissen wurden. Hat der Dampfplauderer etwa recht und war die Schwerkraft vor Millionen von Jahren eine andere? Sind deshalb alle Lebewesen auf der Erde *heutzutage* so viel kleiner, weil die Schwerkraft größer ist als damals? War die Erde kleiner und hatte weniger Masse? Oder ist es nur ein Rechenfehler wie bei der Theorie der Hohlerde von Edmond Halley aus dem 17.–18. Jahrhundert? Müde reiben Sie sich die Augen. So wird es nichts mehr mit dem Schlafen, dann können Sie auch mal etwas rechnen⁶.

Betrachten wir die Schwerkraft oder Gewichtskraft, die auf Lebewesen damals und heute gewirkt haben soll. Nehmen Sie für die Fallbeschleunigung g der Erde *heutzutage* einen Wert von 9.65 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.289 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.21 g/cm^3 . Das Gewicht von einem heute lebenden asiatischen Elefanten liegt bei 3t bis 5t und das Gewicht von einem Tyrannosaurus rex (T. rex) bei 4.5t bis 8t.

1. Welchen Durchmesser müsste die Erde vor 66 Millionen Jahren gehabt haben, wenn Dinosaurier und Elefanten die gleiche Gewichtskraft \vec{F}_G damals und heute erfahren hätten? *Beantworten Sie die Frage anhand der folgenden Teilaufgaben!*
 - a) Berechnen Sie die Fallbeschleunigung von vor 66 Millionen Jahren unter der obigen Annahme gleich wirkender Gewichtskraft \vec{F}_G auf Elefant und Dinosaurier! **(1 Punkt)**
 - b) Berechnen Sie Masse der heutigen Erde! **(2 Punkte)**
 - c) Schließen Sie über die Masse auf den Durchmesser der Erde vor 66 Millionen Jahren! **(2 Punkte)**
2. Beantworten Sie die Eingangsfrage mit 1-2 Antwortsätzen! **(1 Punkt)**

Die Distanz zwischen Sonne und Erde entspricht 1.05 astronomische Einheiten (AE). Die Einheit 1 AE wird mit $1.5 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 79% Wasserstoffkernen mit einer molaren Masse von 1.08 g/mol , 12% Heliumkernen mit 4.11 g/mol sowie 9% weiteren Atomkernen mit 89.32 g/mol . Die Teilchendichte bei Eintritt in die Erdatmosphäre liegt zwischen 0.4 bis 100 Teilchen cm^{-3} pro Sekunde mit einer mittleren Teilchendichte von 5 cm^{-3} pro Sekunde.

Lösen Sie den folgenden Aufgabenteil mit einer aussagekräftigen Skizze!

4. Berechnen Sie die Anzahl an massebehafteten Teilchen des Sonnenwindes, die die gesamte Erde pro Sekunde treffen! **(2 Punkte)**
5. Berechnen Sie die Anzahl an massebehafteten Teilchen des Sonnenwindes, die die Sonne pro Sekunde in alle Richtungen aussendet! **(2 Punkte)**
6. Berechnen Sie die Masse, die die Erde pro Jahr durch die *massebehafteten* Teilchen des Sonnenwind zunimmt! **(2 Punkte)**

⁶Die Quelle der Inspiration für die Aufgabe war folgender Artikel: "Skeptische Anmerkungen — Die Erde als Dampfnudel" in Der Humanistische Pressedienst

112. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



'Entschuldigung, ist das Ihre Feder in meinem Auge?' So hört man häufiger höfliche Puten in Mastställen sagen. Das ist natürlich etwas ungünstig, denn dann kommt es zu Picken und Kannibalismus. Denn wenn der Nachbar nervt, dann muss zu Maßnahmen gegriffen werden. Kennt jeder aus einer mittelmäßigen Wohngemeinschaft. Das wollen Paula, Alex, Jonas und Steffen aber als vorsorgliche Puten-Halter:innen nicht⁷. Gemeinsam sind die Vier in einer Projektgruppe gelandet. Betrachten wir also gemeinsam einmal das Platzangebot (eng. *space allowance*, abk. *SA*) der Puten für vier Tätigkeiten und versuchen die notwendige Fläche zu optimieren. Wie immer gibt es dafür eine mathematische Formel:

$$SA = \sum_{i=1}^n (A_i \times PB_i) \quad A_i = \pi \times (r_i + R_i)^2$$

mit

- SA dem benötigten Platzangebot aller aufsummierten Verhalten i .
- A_i dem benötigten Platz für ein Verhalten i .
- PB_i dem Anteil des Auftretens eines Verhaltens i .
- r_i dem Radius Pute plus dem benötigten Radius für das Verhalten i .
- R_i dem notwendigen Abstand zu den Nachbarn für das Verhalten i .
- i dem Verhalten: (1) wingflapping, (2) foraging incl. scratching, (3) preening und (4) wing/leg stretching.

In der folgenden Tabelle 1 sind die Werte für r_i , R_i und PB_i für ein spezifisches Verhalten i aus drei wissenschaftlichen Veröffentlichungen dargestellt.

	Aldridge et al. (2021)	Baxter et al. (2022)	Jacobs et al. (2019)
wingflapping	27cm; 26cm; 0.2%	30cm; 31cm; 0.2%	32cm; 19cm; 0.9%
foraging incl. scratching	37cm; 26cm; 8.1%	30cm; 21cm; 3.6%	31cm; 21cm; 5.1%
preening	37cm; 13cm; 6.2%	42cm; 23cm; 3.2%	27cm; 28cm; 1.2%
wing/leg stretching	32cm; 23cm; 1.2%	35cm; 29cm; 1%	27cm; 30cm; 0.1%

Leider kennen sich die Vier nicht so gut mit der Berechnung aus! Daher brauchen die Vier Ihre Hilfe!

1. Erstellen Sie eine zusammenfassende Tabelle mit den mittleren Werten für r , R und PB aus der obigen Tabelle 1 für die jeweiligen Verhalten! **(3 Punkte)**
2. Ergänzen Sie eine Spalte mit dem benötigten Platz A für das jeweilige Verhalten, welches sich aus den mittleren Werten ergibt! **(1 Punkt)**
3. Berechnen Sie das benötigte Platzangebot SA für alle betrachteten Verhalten! **(1 Punkt)**
4. Skizzieren Sie die Werte r_i , R_i und A_i für zwei nebeneinander agierende Puten für ein Verhalten i . Nutzen Sie hierfür vereinfachte geometrische Formen! **(2 Punkte)**
5. Sie entnehmen der Literatur folgende Aussage zur Verteilung der Puten in der Fläche A : „Assuming, that the animals will optimally and equally distribute in an area A , we observe a small part, which is not covered. This area is called ω and is calculated with $\omega = \frac{A}{0.9069}$.“ Veranschaulichen Sie die Fläche ω in einer aussagekräftigen Abbildung! **(1 Punkt)**
6. Ein Tier braucht Platz für sich selbst. Berechnen Sie nun die Körperfläche a , die ein Tier einnimmt. Welche Annahmen haben Sie für die Berechnung der Körperfläche getroffen? **(2 Punkte)**

⁷Die Quelle der Inspiration für die Aufgabe war der folgende wissenschaftliche Artikel: [EFSA Panel on Animal Health and Welfare, et al. \(2023\) Welfare of broilers on farm. EFSA Journal 21.2.](#)

113. Aufgabe

(8 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Nelken von den Molukken Mark und Paula waren gemeinsam in Berlin und sitzen nun im IC nach Amsterdam um zurück nach Osnabrück zu fahren. 'Weißt du was ich mich frage?', entfährt es Mark ziemlich plötzlich, so dass Paula die Smarties aus dem Mund fallen. 'Nein, und ehrlich gesagt bin ich auch ziemlich müde...'. Das ist jetzt aber Mark egal, denn er möchte folgende Sachlage diskutieren. Und Mark hat jetzt 3 Stunden Zeit. Plus Verspätung. In der Ausstellung *Europa und das Meer* im Deutschen Historischen Museum in Berlin gab es folgendes Zitat über die Probleme der frühen Hochseeschifffahrt.

»Ohne ausreichende Zufuhr von Vitamin C stellen sich nach 45 Tagen die ersten Symptome ein; die ersten Toten sind nach 60 Tagen zu beklagen; nach 105 Tagen rafft die Skorbut eine ganze Schiffsbesatzung dahin«

Ferdinand Magellan stach im Jahre 1519 in See um eine Passage durch den südamerikanischen Kontinent zu finden. Zu seiner Flotte gehörten fünf Schiffe - das Flaggschiff *Trinidad*, die *San Antonio*, die *Victoria*, die *Concepción* und die *Santiago* - mit einer Besatzung von insgesamt 218 Mann.

1. Stellen Sie den Verlauf der Anzahl an Matrosen auf einem Schiff der Flotte in der Form einer Überlebenszeitkurve dar! Beschriften Sie die Achsen entsprechend! **(2 Punkte)**
2. Was ist die Besonderheit der Überlebenszeitkurve? Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Schätzen Sie die Überlebenswahrscheinlichkeit nach 90 Tagen aus Ihrer Abbildung ab! **(1 Punkt)**

Der Chronist an Bord der *Trinidad*, Antonio Pigafetta, schrieb in seinem Bericht '[...] Um nicht Hungers zu sterben, aßen wir das Leder, mit dem die große Rahe zum Schutz der Taue umwunden war.' Insbesondere die Mannschaft der *Concepción* erlitt große Verluste durch die Skorbut bei der Überquerung des Pazifiks, da durch Erkundungsfahrten weniger Zeit blieb, um wilden Sellerie aufzunehmen. Wilder Sellerie enthält $6000\mu\text{g}/50\text{mg}$ Vitamin C. Der Bedarf liegt bei 115mg pro Tag für Männer.

3. Berechnen Sie die notwendige Menge in t an aufzunehmenden wilden Sellerie auf die *Concepción* für die ununterbrochene Fahrt von drei Monate und 18 Tage über den Pazifik! **(3 Punkte)**
4. Skizzieren Sie die Überlebenszeitkurve für die *Concepción* im Vergleich zu der Überlebenszeitkurve der *Trinidad*! Beschriften Sie die Achsen! **(2 Punkte)**

114. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Event Horizon – Am Rande des Universums Paula ist bei Yuki um gemeinsam *Event Horizon – Am Rande des Universums* zu streamen. Das war jetzt nicht die beste Idee. Denn Paula kann Horror überhaupt nicht ab. Deshalb flüchtet sie sich in Logik um ihre Emotionen zu bändigen. Yuki mampft ungerührt Reese's Peanut Butter Cups. Folgenden Gedankengang nutzt Paula um dem Film zu entkommen. Die Sonne hat eine aktuelle, angenommene Masse von $2 \times 10^{27} \text{ kg}$. Wenn die Sonne nun am Ende ihrer Lebenszeit zu einem schwarzen Loch mit dem Radius von 5000m kollabiert, wird die Sonne 45% der aktuellen Masse verloren haben. Ein Lichtteilchen mit der Masse m_f und der Fluchtgeschwindigkeit v_f will dem schwarzen Loch entkommen. An folgende Formeln erinnert sich Paula für die kinetische Energie des Lichtteilchens E_{kin} und der Gravitationsenergie des schwarzen Lochs E_{grav} ⁸.

$$E_{kin} = \frac{1}{2} m_f v_f^2 \quad E_{grav} = \frac{G m_s m_f}{r_s}$$

mit

- m_f , gleich der Masse [kg] des fliehenden Objektes
- m_s , gleich der Masse [kg] des stationären Objekts
- r_s , gleich dem Radius [m] des stationären Objekts
- G , gleich der Gravitationskonstante mit $6.274 \cdot 10^{-11} \text{ m}^3 (\text{kg} \cdot \text{s}^2)^{-1}$

Im Folgenden wollen wir Paula bei der Ablenkung helfen und uns mit der Frage beschäftigen, ob das Lichtteilchen der Gravitation des schwarzen Lochs entkommen kann.

1. Geben Sie die Formel für die Fluchtgeschwindigkeit v_f an! **(1 Punkt)**
2. Überprüfen Sie Ihre umgestellte Formel nach v_f anhand der Einheiten! **(1 Punkt)**
3. Berechnen Sie die notwendige Fluchtgeschwindigkeit v_f des Lichtteilchens mit den angegebenen Informationen! **(2 Punkte)**
4. Gehen Sie von einer Lichtgeschwindigkeit von $2.8 \times 10^8 \text{ m/s}$ aus. Kann das Lichtteilchen der Gravitation des schwarzen Lochs entkommen? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Stellen Sie den Zusammenhang zwischen dem sich verringernden Radius r des schwarzen Lochs bei gleichbleibender Masse m_s und der notwendigen Fluchtgeschwindigkeit v_f in einer Abbildung dar! **(2 Punkte)**
6. Eine Kirchenglocke und ein Lolli stürzen aus großer und gleicher Höhe in ein schwarzes Loch. Welches der beiden Objekte überschreitet zuerst den Ereignishorizont des schwarzen Loches? Begründen Sie Ihre Antwort mathematisch! **(2 Punkte)**

⁸Die Quelle der Inspiration für die Aufgabe war ein Montagnachtfilm: [Event Horizon – Am Rande des Universums](#)

115. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Stichworte: Great filter • SETI • WOW-Signal • 5-Sigma • Voyager 1 • Voyager 2

Das Fermi Paradoxon Tina und Jonas wandern durch den Teutoburgerwald um mal vom Studium runterzukommen. 'Kennst du eigentlich Enrico Fermi?', fragt Tina und fährt ohne die Antwort abzuwarten fort, 'Er war ein berühmter Kernphysiker! Enrico Fermi diskutierte 1950 auf dem Weg zum Mittagessen im Los Alamos National Laboratory mit seinen Kollegen angebliche UFO-Sichtungen und fragte schließlich: »Where is everybody?«. Warum seien weder Raumschiffe anderer Weltraumbewohner noch andere Spuren extraterrestrischer Technik zu beobachten?'. Jonas schaut sie irritiert und interessiert an. Die beiden hat das Problem gepackt. Deshalb wollen Tina und Jonas das Paradoxon mal mathematisch untersuchen! Wie lange würde eine außerirdische Zivilisation benötigen um die gesamte Milchstraße zu besuchen, wenn das maximale Reisetempo die Geschwindigkeit der Voyager 1 Sonde wäre?⁹

Die beiden treffen folgende Annahmen. Eine außerirdische Zivilisation schickt *drei* Voyager 1 ähnliche Sonden mit der Geschwindigkeit von $6.3587 \times 10^4 \text{ km/h}$ los um sich auf den erreichten Planeten selbst zu replizieren. Nach 250 Jahren ist die Replikation abgeschlossen und wiederum *drei* Sonden werden ausgesendet. Gehen Sie von 6.23 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 1.5×10^{11} Sterne in der Milchstraße. Nehmen Sie eine Lichtgeschwindigkeit von $2.8 \times 10^8 \text{ m/s}$ an.

1. Skizzieren Sie in einer Abbildung die ersten vier Schritte der Vervielfältigung der Sonden in der Galaxie! Beschriften Sie die Abbildung mit der Dauer und der Anzahl an Sonden für jeden Schritt der Vervielfältigung! **(4 Punkte)**
2. Berechnen Sie die theoretische Anzahl an Vervielfältigungsschritten die benötigt werden um mit *einem einzigen Vervielfältigungsschritt* die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
3. Berechnen Sie die Dauer, die eine außerirdische Zivilisation annährungsweise benötigt um die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
4. Bei einem vermuteten Alter der Erde von 4.3×10^9 Jahren, wie oft war dann eine Sonde einer außerirdischen Zivilisation schon zu Besuch? Korrigieren Sie Ihre Antwort mit dem Wissen, dass sich die Kontinentalplatten einmal alle 1.2×10^8 Jahre vollständig im Erdinneren umgewandelt haben! **(2 Punkte)**

⁹Die Quelle der Inspiration für die Aufgabe war folgender Wikipediaeintrag: [Fermi-Paradoxon](#)

116. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Pyramiden bauen Es stehen die niedersächsischen Pyramidentage! Sie und Alex sind auf abenteuerlichen Wegen für den Bau der Pyramiden zuständig. Zu allem Überfluss handelt es sich auch noch eine *Reenactment* Veranstaltung. Thema der diesjährigen Pyramidentage sind die Pyramiden von Meroe, die den Königen und Königinnen des historischen Reiches von Kusch in Nubien, dem heutigen Sudan, als Grabstätten dienten. Die Pyramiden in Meroe fallen durch ihren steilen Winkel von 70 Grad im Vergleich zu den ägyptischen Pyramiden mit 53 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 44 Königsellen. Eine Königselle misst 52.2cm.

Lösen Sie diese Aufgabe mit Hilfe einer Skizze der Pyramide. Bezeichnen Sie Seiten und die Winkel der Pyramide entsprechend!

1. Bei der Königspyramide von Meroe soll eine Seitenlänge der Grundfläche 44 Königsellen lang sein. Welche Höhe der Königspyramide in m ergibt sich? **(1 Punkt)**
2. Die Außenflächen der Pyramide soll begrünt werden. Für die Bepflanzung muss eine 3cm dicke Torfschicht auf die Pyramide aufgebracht werden. Berechnen Sie die ungefähre Menge an benötigten Torf in m^3 ! **(2 Punkte)**

Wie in jedem guten *Reenactment* gibt es viel Oberschicht, aber nur 2 Sklaven, die Ihnen und Alex bei dem Befüllen der Pyramide mit Schutt zu Seite stehen. Leider haben Ihre Sklaven zu allem Überfluss auch noch chronische Knieschmerzen entwickelt, als die Sklaven von der anstehenden Aufgabe erfahren haben. Gehen Sie daher von einer Effizienz der Sklaven von 85% aus. In eine Schubkarre passen 100 Liter.

3. Wie oft müssen Ihre maladen Sklaven die Rampe mit der Schubkarre zur Spitze der Pyramide hochfahren um die Pyramide mit Schutt zu füllen? **(1 Punkt)**
4. Berechnen Sie die Länge der Rampe zur Spitze der Pyramide mit einem Anstellwinkel von 10° ! **(2 Punkte)**
5. Wie weit reicht Ihre Rampe vom Fuß der Pyramide in die niedersächsische Landschaft? **(2 Punkte)**

Bei der Besichtigung der Pyramide teilt Ihnen der leicht übergewichtige Pharao (Nebenberuf *Mittelständler*) mit, das die Pyramide zu steil sei und somit nicht in die niedersächsische Landschaft passen würde. Sie müssen nochmal ran.

6. Die Grundfläche der Pyramide ändert sich nicht. Berechnen Sie die Änderung der Höhe in Königsellen, wenn sich der Anstellwinkel der Pyramide um 6° ändert! **(2 Punkte)**

117. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Geocaching – Von Satelliten und Plastikdosen Es ist Wochenende und das Wetter ist sweet. Alex und Nilufar schwingen sich auf ihre Cachermobile um mit 19km/h, geleitet von modernster Satellitentechnologie und einem Supercomputer aus dem Jahr 2000 in den Händen, Plastikdosen in der Natur und an sehenswerten Orten zu finden. Alex und Nilufar wollen diesmal endlich die abwärts Terrainchallenge durchführen. Die Reihenfolge der Caches nach Terrainwertung gibt daher die von den beiden abzufahrenden Orte vor. Die Terrain- und Schwierigkeitswertungen laufen von 1 (leichteste Wertung) bis 5 (schwierigste Wertung) in 0.5 Schritten. Folgende Informationen zu den Orten und den entsprechenden Caches stehen Alex und Nilufar für die Planung der Route zu Verfügung¹⁰.

Ort	Cache	Wertung (S T G)
A	GCJRHPF	1.5 4.5 Klein
B	GCDCCDC	2.0 3.0 Klein
C	GCPRWMG	3.5 3.5 Normal
D	GCTXNUND	4.5 5.0 Normal
E	GCSW2KT	4.0 2.0 Mikro

Im Weiteren sind den beiden folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \overrightarrow{AB} ist 4km. Im Weiteren ist Ihnen der Entfernungsvektor \overrightarrow{CB} mit 6.5km bekannt. Der Entfernungsvektor \overrightarrow{BE} ist das 1.2-fache des Entfernungsvektor \overrightarrow{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 40° nördlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 60° östlich. Vom Ort B betrachtet, bilden die Orte C und D einen rechten Winkel am Ort B. Der Ort B liegt auf gerader Linie zwischen den Orten C und E. Somit liegt der Ort E nördlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort D Ihre Cachertour.

Leider sind die beiden sehr schlecht im Navigieren und Entfernungen ausrechnen. Die beiden brauchen Ihre Hilfe!

1. Lösen Sie diese Aufgabe mit Hilfe einer aussagekräftigen Skizze der Orte und Caches. Bezeichnen Sie die Strecken und die Winkel Ihrer Skizze entsprechend! **(2 Punkte)**
2. Welche Strecke in km legen Sie bei der Bewältigung der abwärts Terrainchallenge zurück? **(5 Punkte)**
3. Gehen Sie von einer zusätzlichen Suchzeit in Stunden für die Caches an den jeweiligen Orten zur reinen Reisezeit mit Ihrem Cachermobil aus. Die Suchzeit in Stunden für *jeden einzelnen* Cache wird durch die Funktion

$$\text{Suchzeit} = 0.2 + 0.18 \cdot \text{Schwierigkeit}$$

beschreiben. Wie lange in Stunden benötigen Sie um die abwärts Terrainchallenge zu erfüllen? **(3 Punkte)**

4. An der höchsten Schwierigkeit müssen Sie angeln. Ihre Angel ist ausgefahren 5m lang. Erreichen Sie einen Cache in der Höhe von 6.8m? Berechnen Sie dazu Ihre maximale mögliche Angelhöhe! Welche Annahmen mussten Sie treffen um die Aufgabe zu lösen? **(2 Punkte)**

¹⁰Die Quelle der Inspiration für die Aufgabe war folgende Tätigkeit: [Geocaching – Mach mit bei der weltweit größten Schatzsuche](#).

118. Aufgabe

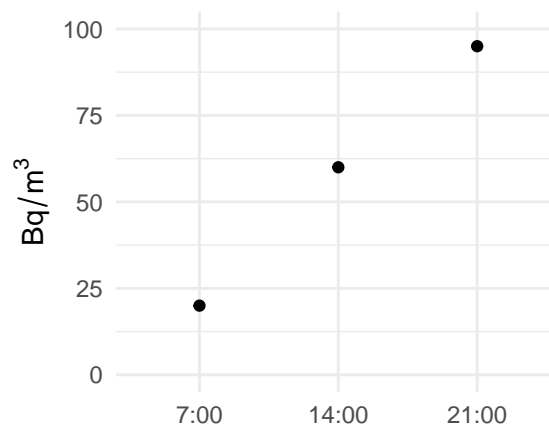
(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Stichworte: Brot aus Luft • Walöl • Haber-Bosch-Verfahren • 1. Weltkrieg • 40% N im menschlichen Körper • Positivist

Die atmende Wand und Brot aus Luft Als Kellerkind¹¹ vom Dorf will Paula das Ausmaß der Radonbelastung in ihrem Kellerzimmer bestimmen und lüften daher nicht. Passt schon. Spart dann auch Energie und lüften wird sowieso überschätzt. Während einer Messperiode von 7:00 Uhr bis 21:00 bestimmt sie dreimal automatisch die Radonbelastung in ihrem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung¹². Leider helfen die Messwerte Paula überhaupt nicht weiter. Sie müssen also helfen!



1. Wie lange dauert es in Stunden bis Sie eine kritische Belastung von $300 Bq/m^3$ in Ihrem ungelüfteten Kellerraum erreicht haben? **(2 Punkte)**

Radon zerfällt mit einer Halbwertszeit von 3.5d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 160d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

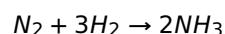
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $300 Bq/m^3$ auf unter $90 Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	77.1	27.9	
Sauerstoff	20.45	15.8	
Kohlenstoffdioxid	0.045	12.5	

3. Rechnen Sie die Volumenprozent (Vol-%) der Umgebungsluft in die entsprechenden ppm-Werte um und ergänzen Sie die berechneten ppm-Werte in die Tabelle! **(1 Punkt)**

Während Paula ihr etwas pappiges Toastbrot mampfen kommt Paula die Dokumentation über Brot aus Luft in den Sinn. Paula denkt darüber ein wenig nach. Für die Umwandlung von Stickstoff N_2 mit Wasserstoff H_2 zu Ammoniak NH_3 gilt folgende Reaktionsgleichung¹³:



Ein Mol eines beliebigen Gases hat bei normalen Umweltbedingungen ein Volumen von 22.4 Liter.

4. Welche Masse an Ammoniak in Kilogramm kg können Sie aus einem Kubikmeter m^3 Luft unter normalen Umweltbedingungen gewinnen? **(2 Punkte)**
5. Wieviel Ammoniak in mol erhalten Sie aus einem Kubikmeter Luft? **(1 Punkt)**

¹¹Tocotronic - Electric Guitar als passende Untermauerung für diese Aufgabe.

¹²Die Quelle der Inspiration für die Aufgabe war folgender Artikel: [Atmende Wand](#)

¹³Die Quelle der Inspiration für die Aufgabe war folgender Artikel: [Haber-Bosch-Verfahren – Brot aus Luft](#)

119. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Armee der Finsternis Der Studentenjob von Paula war nach Ladenschluss bei IKEA die Regale einzuräumen. Dabei ist Paula in der Auslage der Sonderangebote das Necronomicon¹⁴ in die Hände gefallen. Nun ist sie eine Magierin der Zeichen geworden! Also eigentlich kann Paula nur Mathe und das dämliche Necronomicon hat sie in die Vergangenheit geschleudert... aber gut, was tut man nicht alles im Jahr 301 n. Chr. für den neuen Lehnsherren Henry dem Roten. Paula baut natürlich einen Schrottkugelturm um sich den Horden der Finsternis mit genug Schrott erwehren zu können! Paula stehen zwei mächtige magische Formeln zur Unterstützung zu Verfügung. Leider wird das nicht reichen, deshalb müssen Sie hier auch noch durch Zeit und Raum helfen!

$$E_{kin} = \frac{1}{2} \cdot m \cdot v^2 \quad E_{pot} = m \cdot g \cdot h$$

mit

- m , gleich der Masse [kg] des Objekts
- h , gleich der Höhe [m] des ruhenden Objekts
- v , gleich der Geschwindigkeit [m/s] des Objekts
- g , gleich der Erdbeschleunigung mit $9.81 \frac{m}{s^2}$

Als erstes müssen Sie die Höhe des zu bauenden Schrottkugelturmes bestimmen. Hierfür ist wichtig zu wissen, dass sich die Bleitropfen mit einem Gewicht von $20mg$ zu gleichförmigen Bleikugeln bei einer Geschwindigkeit von $14m/s$ bilden.

1. Wie hoch müssen Sie den Schrottkugelturm bauen lassen, damit sich runde Bleikugeln durch die Fallgeschwindigkeit von $14m/s$ bilden? **(3 Punkte)**

Ihre erstellten Schrottkugeln sind leider zu groß und somit sind zu wenige Schrottkugeln in einer Ladung. Damit können Sie die Armee der Finsternis nicht aufhalten. Die Sachlage müssen Sie einmal mathematisch untersuchen.

2. Nennen Sie die beiden geometrischen Formen aus denen sich näherungsweise ein Tropfen zusammensetzt! Erstellen Sie eine beschriftete Skizze des Tropfens! **(2 Punkte)**
3. Sie messen eine Länge des Tropfens von $3.2mm$. Die Löcher im Sieb erlauben ein Tropfendurchmesser von $1.7mm$. Welchen Durchmesser in mm haben Ihre produzierten Bleikugeln? **(3 Punkte)**

Sie haben jetzt die 6.1×10^4 Bleikugeln zusammen. Blei hat eine Dichte von $12.63g/cm^3$.

4. Wie schwer in Kilogramm kg sind die 6.1×10^4 produzierten Bleikugeln, die Sie jetzt auf die Burgmauer transportieren müssen? **(1 Punkt)**

Am Ende müssen Sie noch die Produktion von dem Bleischrott im Turm optimieren.

5. Wie groß in cm^2 ist Ihr quadratisches Sieb am oberen Ende des Turms, wenn Sie pro Fall ca. 700 Bleikugeln produzieren wollen und die Bleikugel im Fall $1.2cm$ Abstand haben müssen? **(1 Punkt)**

¹⁴Ein wirklich gefährliches Buch ist: *Du bist genug: Vom Mut, glücklich zu sein* von Fumitake Koga und Ichiro Kishimi

120. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Armee der Kaninchen Leider hat es bei Alex mit dem Krokodilreservat in Down Under nicht geklappt. War vielleicht auch nicht so die beste Idee... aber dafür hat Alex eine neue Eingebung! Oder wie es Mike Tyson zugeschrieben wird: »Ich wurde nie niedergeschlagen, ich war immer am Aufstehen!«. Daher macht Alex jetzt einen Großhandel mit Kaninchenfleisch und damit dem teuersten Fleisch in Australien auf. Moment, hopsen hier nicht, seit Thomas Austin im Jahr 1859 ungefähr 28 Kaninchen entlassen hat, Millionen von Kaninchen rum? Wieso ist das Kaninchenfleisch dann so exklusiv? Alex wird stutzig und frag Sie, dem mal mathematisch nachzugehen!¹⁵

Forscherinnen fand folgende Sättigungsfunktion für das jährliche Wachstum der gesamten Kaninchenpopulation im westlichen Australien.

$$f(t) = 8 \times 10^9 - 1.2 \times 10^9 \cdot 2.1^{-0.2 \cdot t + 2.4}$$

1. Skizzieren Sie die Sättigungsfunktion *annäherungsweise* in einer Abbildung! **(1 Punkt)**
2. Wie viele Kaninchen können nach der Sättigungsfunktion maximal im westlichen Australien leben? Ergänzen Sie den Wert in Ihrer Abbildung! **(2 Punkte)**
3. Wie viele Millionen Kaninchen leben nach der Sättigungsfunktion nach 18 Jahren auf dem australischen Kontinent? **(1 Punkt)**

Um den Kaninchen Einhalt zu gebieten wurde das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) in 18 Kaninchen ausgebracht. Da die Kaninchen keine Maßnahmen gegen die Ausbreitung vornehmen können, verläuft die Ausbreitung mit einem wöchentlichen Wachstumsfaktor von 2 nach folgender Formel.

$$N(t) = N(0) \cdot a^t$$

3. Wie viele Wochen benötigen die Viren um theoretisch die gesamte Kaninchenpopulation nach 16 Jahren Wachstum zu durchseuchen? **(1 Punkt)**

Das Myxoma Virus und das RHDV töteten 99.9% der Kaninchenpopulation innerhalb weniger Wochen.

4. Wie lange in Jahren dauert es bis eine Kaninchenpopulation nach einer Viruspandemie wieder auf 30% der gesättigten Kaninchenpopulation angewachsen ist? **(2 Punkte)**

Thomas Austin entließ die Kaninchen im äußersten Westen von Australien. Australien hat eine West-Ost-Ausdehnung von 4300km und eine Nord-Süd-Ausdehnung von knapp 3700km. Die Kaninchen breiten sich radial mit einer Geschwindigkeit von 7.8km pro Jahr aus.

5. Wie lange dauert es in Jahren bis die Kaninchen jeden Ort in Australien erreicht haben? *Lösen Sie die Aufgabe unter der Verwendung einer schematischen Skizze!* **(2 Punkte)**

Eine jährliche Impfung gegen das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) kosten 12\$ pro Tier und der durchführende Arzt verlangt ca. 40\$ pro Tier.

6. In Ihrem Stall leben 1200 Mastkaninchen. Mit welchen jährlichen Zusatzkosten für die Impfungen der Kaninchen müssen Sie daher kalkulieren? **(1 Punkt)**

¹⁵Die Quelle der Inspiration für die Aufgabe war der folgendes YouTube Video: [Incredible Stories – Why don't they eat wild rabbits in Australia? They have millions of them! The reason is surprising...](#)

121. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Ostfriesland. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer des Esels Frida und Mark. Grünes Gras unter Marks Füßen und ein strammer Wind im Gesicht, egal wohin er schaut. Ein schmatzendes Geräusch ertönt unter Mark. Mark sinniert, sollte er seine weiten Graslandschaften jetzt schon düngen? Dafür benötigt Mark die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Mark sieht nicht ein, Geld für einen Agrarmetrologen zu bezahlen, wenn auch Sie mitrechnen können. Also rechnen Sie beide mit folgenden Informationen zu Monatsmultiplikatoren des GLT-Wertes: Januar mit $0.5\times$, Februar mit $0.7\times$ und März mit $1.1\times$. Sie haben noch im letzten Jahr folgende Temperaturen gemessen.

Datum	C°
01. Jan 2023	0.1
01. Feb 2023	1.5
01. Mrz 2023	3.5
01. Apr 2023	5.8

1. Erstellen Sie eine Skizze aus den Informationen aus der Temperaturtabelle! **(1 Punkt)**
2. Stellen Sie die linearen Funktionen $f_1(t)$, $f_2(t)$ und $f_3(t)$ aus der obigen Temperaturtabelle auf! **(1 Punkt)**
3. Bestimmen Sie die Stammfunktionen $F_1(t)$, $F_2(t)$ und $F_3(t)$ für Ihre linearen Funktionen aus der obigen Temperaturtabelle! **(1 Punkt)**
4. Osterglocken beginnen ab einer GLT von 210°C zu blühen. An welchem Tag im 1. Quartal des Jahres 2023 war dies der Fall? *Ignorieren Sie ein eventuelles Schaltjahr in Ihrer Berechnung.* **(4 Punkte)**

Auf dem Weg zu Marks Jonagoldplantage werden Sie beide auf dem Trecker von einer Gruppe elektrifizierter Renter abgedrängt. Der Trecker muss wieder aus dem Graben! Frida und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Frida mit 120N . Die elektrifizierten Renter bringen eine Kraft von 210N auf.

Lösen Sie diese Aufgabe mit Hilfe einer aussagekräftigen Skizze der Kraftvektoren. Bezeichnen Sie die Kraftvektoren und die Winkel Ihrer Skizze entsprechend!

5. Im ersten Versuch legen Sie das Seil für Frida lotrecht über einen Ast oberhalb des Treckers. Die Rentner ziehen in einer geraden Linie über die Böschung hinweg am anderen Seil. Welche Kraft wird aufgebracht? **(2 Punkte)**
6. Im zweiten Versuch ziehen Frida und die Rentner mit einem 50° Winkel mit ihrem Seil an dem Trecker. Welche Kraft wird aufgebracht? **(2 Punkte)**
7. Mit welcher Beschleunigung ziehen Sie den 1.5t schweren Trecker *jeweils* aus dem Graben, wenn $F = m \cdot a$ gilt? **(1 Punkt)**

122. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



In der Kartonagenfabrik Mark, Jessica, Jonas und Alex sitzen im Bus. Wenn man sich zu spät anmeldet, dann ist die Exkursion nicht so toll. Mark hatte den Anderen in der Lerngruppe zu spät Bescheid gesagt. 'Was denn, bin ich eure Nanny oder was?!', entfährt es Mark nachdem die vorwurfsvollen Blicke schon eine Weile auf ihm lasten. Also geht es eben mit Rektor Skinner und Mrs. Krabappel in die Kartonagenfabrik. Wie schon im vorherigen Semester... In der Kartonagenfabrik angekommen erfahren die Vier, dass die Kartons zum Versand von Nägeln nicht hier zusammengebaut werden sondern dass sich die Endfertigung in Flint, Michigan befindet. Unter anderem wird dort der berühmte *Doppelt gewellte, 8-mal-gefaltete, 0.5mm, 30-cm-Karton* durch Falzung hergestellt. Beim letzten Mal war Rektor Skinner die Stimmung zu schlecht und deshalb geht es erst nach Hause, wenn ein paar Aufgaben gelöst sind. Martin gefällt das. An dem Vorrat an Zigaretten von Mrs. Krabappel meinen alle wenig Zuversicht zu erkennen.

Jetzt heißt es Kartons optimieren, wenn Sie auch nochmal nach Hause wollen. Warum jetzt *Sie* mit dabei sind, lassen wir mal weg. Der nun zu optimierende, flache Karton hat eine Länge von 30cm und eine Breite von 22cm. Die Kartonagenmaschine in Flint soll dann einen quadratischen Eckenausschnitt der Länge x falzen.

1. Erstellen Sie eine Skizze des Kartonblattrohlings! Beschriften Sie die Skizze mit den entsprechenden Längenangaben **(1 Punkt)**
2. Berechnen Sie die Falztiefe x für ein maximales Volumen des flachen Kartons! **(3 Punkte)**
3. Welches Volumen in Liter ergibt sich mit der von Ihnen berechneten Falztiefe x ? **(1 Punkt)**
4. Sie wollen noch einen bündig mit dem Boden abschließenden Deckel für den Karton stanzen lassen. Wie groß ist die Fläche des Kartondeckelblattrohlings in cm^2 ? **(2 Punkte)**

Rektor Skinner möchte sich gerne wieder in seinem Vorgarten aufhalten und nicht die ganze Zeit von Bart mit Erdnüssen beworfen werden. Deshalb möchte er einen geräumigen Teil seines Vorgartens einzäunen. Ein Teil der Umzäunung bildet seine Vorderhauswand. Wegen Lieferschwierigkeiten stehen Rektor Skinner nur 120m Zaun zu Verfügung. Auch hier sollen Sie mal helfen, sonst fährt der Bus Sie nicht nach Hause. Sie wollen nun die maximale Fläche des abgeschirmten Vorgartens in Abhängigkeit der Seitenlängen bei der Verwendung von 120m Zaun bestimmen!

5. Welche Seitenlängen für den Zaun ergeben sich für die maximale Fläche des abgeschirmten Vorgartens? **(2 Punkte)**
6. Berechnen Sie die Fläche des abgeschirmten Vorgartens! **(1 Punkt)**

123. Aufgabe

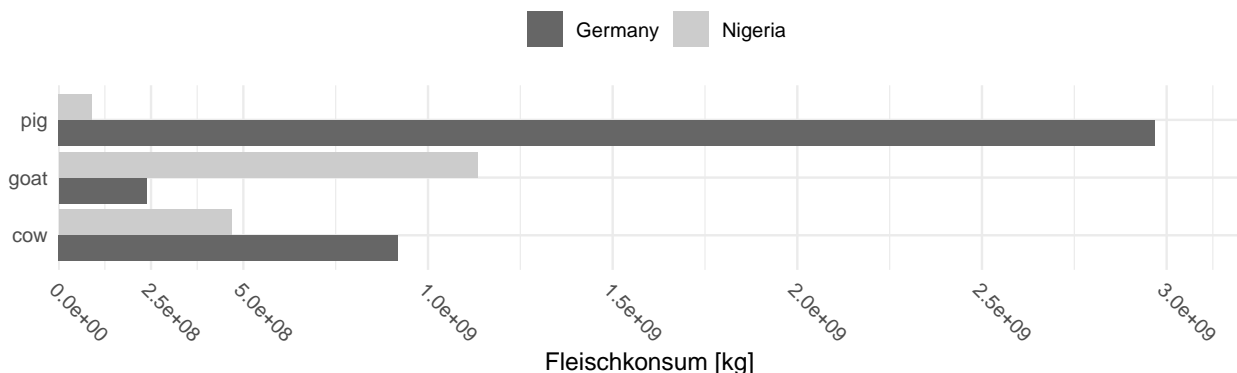
(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



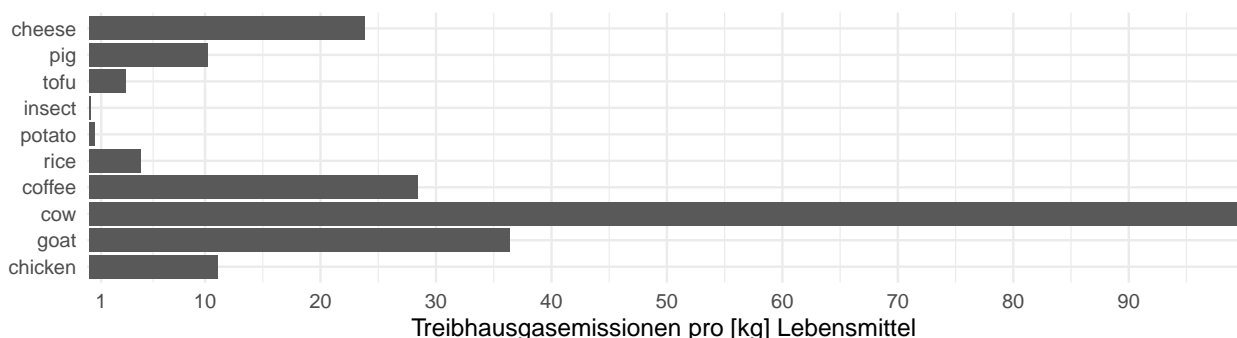
'Ein Pfund Insekten, bitte!' 'Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹⁶.', merkt Nilufar an. Die Lerngruppe mit Steffen, Paula und Jessica sind bei Nilufar um mal was außergewöhnliches zu essen. Um den Sinn der Nahrungsumstellung zu verdeutlichen, vergleicht Nilufar einmal Deutschland mit Nigeria. Nigeria hat eine der am schnellsten wachsenden Bevölkerungen der Welt und wird vermutlich im Jahr 2100 zu den Top 5 der bevölkerungsreichsten Länder zählen. Im Jahr 2020 leben ca. 8.2×10^7 Menschen in Deutschland und ca. 1.8×10^8 Menschen in Nigeria. Mit den Informationen wollen Sie und Nilufar mit der Überzeugungsarbeit anfangen und dann eine Prognose für den Fleischkonsum im Jahr 2050 zu treffen.

Im Folgenden ist Abbildung des Fleischkonsums im Jahr 2020 in Deutschland und Nigeria in [kg] einmal dargestellt.



1. Stellen Sie den Fleischkonsum in Deutschland und Nigeria im Jahr 2020 *pro Kopf* in einer aussagekräftigen Tabelle dar! **(2 Punkte)**
2. Ergänzen Sie in der Tabelle eine Spalte in der Sie für den Fleischkonsum in Nigeria auf Deutschland normieren, daher ins Verhältnis Nigeria/Deutschland, setzen! **(1 Punkt)**

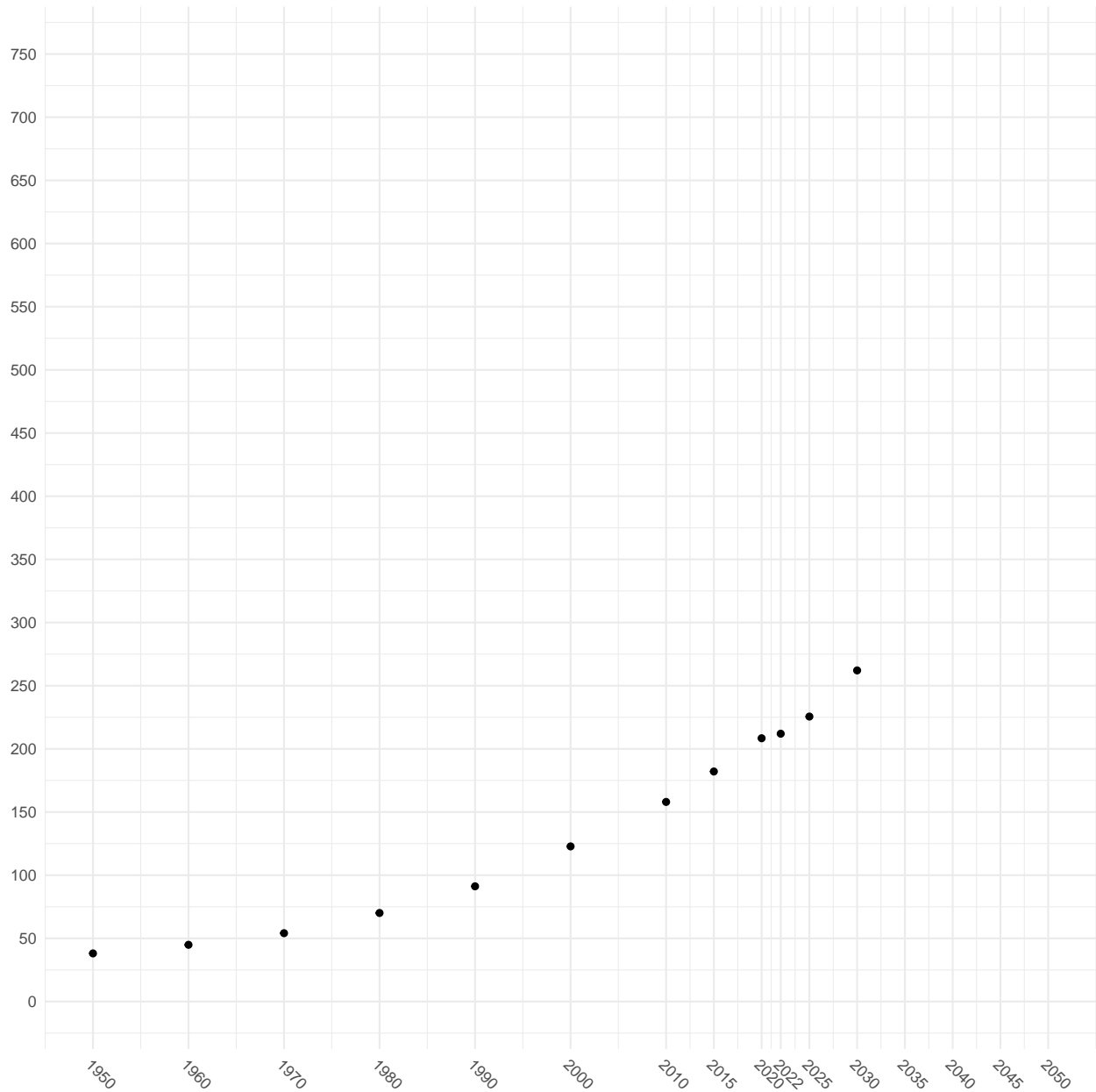
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



3. Stellen Sie in einer Tabelle die Treibhausgasemissionen an CO₂ pro Kopf, die durch den Fleischkonsum in Deutschland und Nigeria im Jahr 2020 entstehen, dar! Ergänzen Sie auch hier das Verhältnis Nigeria zu Deutschland! **(2 Punkte)**

¹⁶Die Quelle der Inspiration für die Aufgabe war der folgende Artikel aus dem Spiegel: [Acht Milliarden - sind wir bald zu viele Menschen auf der Erde?](#)

In der folgenden Abbildung sehen Sie die Bevölkerungsentwicklung [Millionen] in Nigeria von 1950 bis ins Jahr 2030 fortgeführt.



4. Schätzen Sie graphisch die zu erwartende Bevölkerung [Millionen] in Nigeria im Jahr 2050, die sich anhand der Informationen aus der Abbildung ergibt!
 - a) Ohne Berücksichtigung der Covid-19-Pandemie! **(1 Punkt)**
 - b) Unter Berücksichtigung der Covid-19-Pandemie! **(1 Punkt)**
5. Berechnen Sie den geschätzten Fleischkonsum von Nigeria im Jahr 2050 unter der Annahme 80%-iger Angleichung der Lebensbedingungen zu Deutschland im Jahr 2020! **(1 Punkt)**
6. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria im Jahr 2050 im Vergleich zum Jahr 2020, der sich durch den angeglichenen Fleischkonsum ergibt! **(1 Punkt)**
7. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria, wenn die gesamte Proteinaufnahme durch Insekten ersetzt würde! **(1 Punkt)**

124. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Tödliche Seuche AIDS – Die rätselhafte Krankheit Irritiert legen Sie die historische Ausgabe des Spiegels aus den 80zigern beiseite. Sie sind bei Ihrem Orthopäden und wollen einen AIDS-Test machen lassen. Woanders leider keinen Termin gekriegt... Immerhin denken Sie und Ihr Partner über Nachwuchs nach und da geht es eben nur durch ungeschützten Sex. Was wissen Sie nun aber über AIDS und dem diagnostischen AIDS-Test, den Sie nun machen werden?

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.01% angenommen. In 96% der Fälle ist ein HIV-Test positiv, wenn der Patient erkrankt ist. In 1% der Fälle ist ein HIV-Test positiv, wenn der Patient *nicht* erkrankt ist und somit gesund ist. Sie stutzen. Wie wahrscheinlich ist es denn eigentlich an AIDS erkrankt zu sein (K^+), wenn Sie einen positiven AIDS-Test vorliegen haben (T^+)? Gehen Sie für die folgenden Berechnungen von $n = 4 \times 10^4$ Patienten mit einem diagnostischen Test für AIDS aus. Sie nehmen sich also einen Kuli und fangen an auf der historischen Ausgabe des Spiegels zu rechnen¹⁷.

1. Welche Wahrscheinlichkeit Pr wollen Sie berechnen? **(1 Punkt)**
2. Zeichnen Sie einen Häufigkeitsdoppelbaum zur Bestimmung der gesuchten Wahrscheinlichkeit Pr ! **(2 Punkte)**
3. Beschriften Sie den Häufigkeitsdoppelbaum, mit denen Ihnen bekannten Informationen zu der AIDS Erkrankung und dem AIDS-Test! **(1 Punkt)**
4. Füllen Sie den Häufigkeitsdoppelbaum mit den sich ergebenden, absoluten Patientenzahlen n aus! **(2 Punkte)**
5. Berechnen Sie die gesuchte Wahrscheinlichkeit Pr ! **(1 Punkt)**

Bei dem folgenden Arztgespräch erfahren Sie, dass beim diagnostischen Testen *True Positives (TP)*, *True Negatives (TN)*, *False Positives (FP)* und *False Negatives (FN)* auftreten. Das verstehen Sie so noch nicht und deshalb stellen Sie für sich den Zusammenhang in einer 2x2 Kreuztabelle dar.

6. Tragen Sie TP , TN , FP und FN in eine 2x2 Kreuztabelle ein. Beschriften Sie die Tabelle entsprechend! **(1 Punkt)**
7. Berechnen Sie die Sensitivität und Spezifität des diagnostischen Tests für AIDS! Füllen Sie dafür die 2x2 Kreuztabelle mit den Informationen aus dem Häufigkeitsdoppelbaum aus! **(2 Punkte)**
8. Was beschreibt die Sensitivität und die Spezifität im Bezug auf die Gesunden und Kranken? Stellen Sie beide diagnostische Maßzahlen als Wahrscheinlichkeiten Pr dar! **(2 Punkte)**

¹⁷Die Quelle der Inspiration für die Aufgabe war der folgende wissenschaftlicher Artikel: Binder et al. (2022) Von Baumdiagrammen über Doppelbäume zu Häufigkeitsnetzen – kognitive Überlastung oder didaktische Unterstützung? Journal für Mathematik-Didaktik, 1-33

125. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Network-Marketing oder Schneeballschlacht! Eine Möglichkeit, leicht Geld zu verdienen, ist es anderen Menschen für Geld zu versprechen, wie man leicht reich werden kann. Am besten natürlich ohne viel Aufwand und ortsunabhängig. Dann wollen wir mal loslegen. Schnell ein YouTube-Werbevideo gedreht und auf geht es mit unserem Network-Marketing. Aber Moment, wie funktioniert Network-Marketing eigentlich und was hat das alles mit einer Schneeballschlacht zu tun? Wir wollen hier einmal in die Untiefen des „passiven Einkommens“ abtauchen¹⁸.

Das Jahr 2022 war das erfolgreichste Jahr in der Geschichte von Direct Finanzanlagen Left/Right (D-FL/R). Das Unternehmen steigerte den Umsatz um rund 24 Prozent von 275 Millionen Euro im Jahr 2021. Doch wie viel kommt bei den Partnern an? Laut D-FL/R habe das Unternehmen 3.8×10^5 aktive Partner.

1. Berechnen Sie zuerst den Umsatz der Firma D-FL/R im Jahr 2022! **(1 Punkt)**
2. Wie viel von dem Umsatz im Jahr 2022 wird im Durchschnitt von jedem aktiven Partner erwirtschaftet? **(1 Punkt)**
3. Welche *monatlicher* Umsatz ergibt sich dadurch im Durchschnitt für jeden aktiven Partner bei einer direkten Provision von 35%? **(1 Punkt)**

Ihr zu vermarkendes Produkt, hinter dem Sie voll stehen, kostet 75EUR pro Einheit im Direktverkauf. Die direkte Provision für die erste Stufe beträgt 35%. Für die zweite, dritte und vierte Stufe betragen die indirekten Provisionen jeweils 2%, 1% und 0.5%. Jeder Ihrer angeworbenen „Partner“ wirbt wiederum fünf Partner für sich selbst an. Pro Monat werden im Schnitt fünf Einheiten vom Produkt verkauft. Sie wollen nun 4100EUR im Monat *passiv* – also durch indirekte Provisionen – erwirtschaften.

4. Ergänzen Sie die folgende Tabelle mit den obigen Informationen! **(2 Punkte)**

Stufe	Anzahl Partner	Umsatz/Stufe	Provision
1	Sie selber		
2			
3			
4			

5. Wie viele Partner müssen Sie auf der 2 Stufe anwerben um Ihr passives Einkommen durch indirekte Provision zu erreichen? Wie viele Menschen arbeiten am Ende indirekt für Sie? Stellen Sie den Zusammenhang graphisch dar! **(3 Punkte)**

Sie mussten zum Einstieg bei D-FL/R Einheiten des Produkts für 2250EUR kaufen. Diese Einheiten können Sie nur direkt verkaufen. Leider mussten Sie den Kauf über einen Kredit über 6% p.a. über 36 Monate finanzieren.

6. Berechnen Sie die Gesamtsumme, die Sie als Kredit abbezahlen müssen! **(2 Punkte)**
7. Wie viele Einheiten müssen Sie pro Monat verkaufen um die anfallenden Zinsen durch die direkte Provision zu erwirtschaften? **(1 Punkt)**
8. Wie lange in Monaten benötigen Sie um den Kredit durch die direkte Provision abzubezahlen? **(1 Punkt)**

¹⁸Die Quellen der Inspiration für die Aufgabe waren folgendes YouTube Video: [Simplicissimus – Die meistgesuchte Betrügerin der Welt](#) und der Artikel: [Deutschlandfunk Kultur – Die Illusion, schnell reich zu werden](#)

126. Aufgabe

(10 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Höhlen & Drachen Nachdem Sie sich begeistert in der Serie *Stranger Thinks* verloren haben, wollen Sie bei einem Ihrer Freunde einmal *Höhlen & Drachen* ausprobieren. Um Geld zu sparen, das Zeug kostet echt, wurde etwas an den Regeln gebastelt. Schnell stellen Sie fest, dass hier ganz schön viele unterschiedliche Würfel durch die Gegend fliegen. Daher müssen Sie sich jetzt einiges an Fragen stellen.

In dem Spiel haben Sie nun auf einmal 6 vierseitige Würfel ($6d4$) zum würfeln in der Hand. Wenn Sie eine 4 würfeln, haben Sie einen Erfolg.

1. Berechnen Sie die Wahrscheinlichkeit *genau* 5 Erfolge zu erzielen! **(2 Punkte)**
2. Berechnen Sie die Wahrscheinlichkeit keinen Erfolg zu erzielen! **(1 Punkt)**

Sie betrachten nun aufmerksam die ausufernden Ausrüstungstabellen. Ihnen wird aber geholfen und Sie müssen sich jetzt nur zwischen der Axt oder dem Schwert entscheiden.

3. Würden Sie die Axt mit zwei vierseitigen Würfeln ($2d4$) als Schaden oder das Schwert mit einem sechseitigen Würfel plus 4 ($1d6+4$) als Schaden bevorzugen? Begründen Sie Ihre Antwort mathematisch! **(1 Punkt)**

Jetzt wird es immer wilder, da Sie sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass Ihr Rettungswurf gegen den zaubernden Hexer funktioniert. Sie haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.6$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.9$. Sie haben mitgezählt und festgestellt, dass in 45 von 100 Fällen Ihr Rettungswurf bei einem erfolgreichen Zauber funktioniert hat.

4. Erstellen Sie eine 2×2 Kreuztabelle mit den Ereignissen A und B sowie den Gegenereignissen \bar{A} und \bar{B} mit einem $\Omega = 100$. Beachten Sie hierbei die entsprechenden Wahrscheinlichkeiten für die Ereignisse A und B ! **(2 Punkte)**
5. Bestimmen Sie $Pr(A \cap B)$! **(1 Punkt)**
6. Erstellen Sie ein Baumdiagramm mit den passenden Informationen aus der 2×2 Kreuztabelle! **(2 Punkte)**
7. Bestimmen Sie Wahrscheinlichkeit $Pr(A|B)$, dass Ihr Rettungswurf gelingt, wenn der Hexer erfolgreich gezaubert hat! **(1 Punkt)**

127. Aufgabe

(12 Punkte)

Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Retrocheck im TV „Und hier ist sie wieder, die Show der fantastischen Preise. Seien Sie mit dabei, wenn es wieder heißt: Der Preis ist heiß!“, ertönt es und Sie fragen sich, ob Sie nicht doch lieber bezahlter Gast bei Barbara Salesch hätten sein sollten. Aber Sie brauchen das Geld und jetzt heißt es Spielschows farmen! Erstmal eine Kaffemaschine von Mitropa gewinnen. Ein Kandidat gewinnt die Kaffemaschine von Mitropa, wenn nicht alle Kandidaten überbieten (eng. *outbid*). Mit Ihnen bilden Annegret und Thorsten das Team der drei Kandidaten.

Name	$P(\text{win})$	$P(\text{outbid})$
Annegret	0.3	0.076
Thorsten	0.2	0.12

1. Mit welcher Wahrscheinlichkeit gewinnen Sie die Kaffemaschine von Mitropa, wenn keiner der Kandidaten überbietet? **(1 Punkt)**
2. Wenn Ihre Überbietungswahrscheinlichkeit $P(\text{outbid})$ bei 0.043 liegt, mit welcher Wahrscheinlichkeit gewinnt *keiner* die Kaffemaschine von Mitropa? **(1 Punkt)**

Glücksrad für Arme auf der Kirmes! Leider hat es für Maren Gilzer nicht gereicht. Deshalb sind Sie jetzt auf der Kirmes und spielen mit einem einäugigen Piraten um das große Geld. Das Glücksrad hat 22 Felder. Sie drehen das Glücksrad zweimal. Auf 10 Feldern gewinnen Sie 3000EUR sonst 2000EUR. Ganz schön viel Geld und ganz schön zwielichtig hier...

3. Skizzieren Sie das Glücksrad und ergänzen Sie die Wahrscheinlichkeiten! **(1 Punkt)**
4. Zeichnen Sie das zugehörige Baumdiagramm für das zweimalige Drehen! Ergänzen Sie die Wahrscheinlichkeiten und die entsprechenden Ereignisse **(2 Punkte)**
5. Mit welcher Wahrscheinlichkeit gewinnen Sie 5000EUR? **(1 Punkt)**

Nach Ihrem Fiebertraum reisen Sie im Zug nach Köln um bei „Geh aufs Ganze!“ mitzuspielen. Sie schaffen es tatsächlich ins Finale und können als Hauptgewinn ein Auto hinter einer der drei Türen gewinnen.

6. Bevor die Show beginnt, wird das Auto hinter eine zufällig bestimmte Tür gestellt. Mit welcher Wahrscheinlichkeit wird jeweils eine der drei Türen ausgewählt? Zeichnen Sie ein Baumdiagramm! **(1 Punkt)**
7. Mit welcher Wahrscheinlichkeit wählen Sie sofort die Tür mit dem Auto? Erweitern Sie das Baumdiagramm entsprechend! **(1 Punkt)**
8. Der Moderator öffnet nun eine der nicht gewählten Türen, aber natürlich nicht die mit dem Auto. Mit welcher Wahrscheinlichkeit steht das Auto hinter der anderen Tür? Erweitern Sie das Baumdiagramm entsprechend! **(2 Punkte)**
9. Lösen Sie nun das „Ziegenproblem“! Berechnen Sie dazu die Wahrscheinlichkeiten der einzelnen Pfade. Lohnt sich ein Wechsel der anfangs gewählten Tür? Begründen Sie Ihre Antwort mathematisch! **(2 Punkte)**

Teil XII.

Angewandte Nutztier- und Pflanzenwissenschaften (M.Sc.)

128. Aufgabe

(6 Punkte)

Vergleichen Sie die Standardabweichung mit dem Standardfehler und grenzen Sie die beiden Kennzahlen voneinander ab.

129. Aufgabe

(8 Punkte)

Ihnen liegt folgendes Varianzanalysemodell mit der üblichen Beschreibung zur Auswertung des Merkmals fett- und eiweißkorrigierte Milchleistung pro Kuh und Jahr in kg vor:

$$Y_{ijkl} = \mu + Var_i + EKA_j + VarEKA_{ij} + V_k + b(L_{ij} - L) + e_{ijkl}$$

mit

- Y_{ijkl} : l-te Beobachtung
- μ : Populationsmittel
- Var_i : fixer Effekt der i-ten Variante (i: Kontrolle, Versuchsgruppe 1, Versuchsgruppe 2)
- EKA_j : fixer Effekt der j-ten Erstkalbealtergruppe (j: $EKA \leq 25$ Monate, $EKA > 25$ Monate)
- $VarEKA_{ij}$: fixer Effekt der Interaktion Variante x Erstkalbealtergruppe
- V_k : zufälliger Effekt des Vaters
- $b(L_{ij} - L)$: lineare Kovariable Laktationsnummer
- e_{ijkl} : zufälliger Restfehler

Erläutern Sie anhand dieses Beispiels die Begriffe fixer Effekt, Interaktion, zufälliger Effekt und Kovariable und grenzen Sie diese Begriffe voneinander ab.

130. Aufgabe

(6 Punkte)

Wie bestimmen Sie die richtige Stichprobengröße? Welche Kennzahlen / statistische Maßzahlen benötigen Sie dabei und nennen Sie die Voraussetzungen.