

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 she or he knows,
but how much the student 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, dass 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 86 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 106 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
101.5 - 106	1,0
96.0 - 101	1,3
90.5 - 95.5	1,7
85.5 - 90.0	2,0
80.0 - 85.0	2,3
74.5 - 79.5	2,7
69.5 - 74.0	3,0
64.0 - 69.0	3,3
59.0 - 63.5	3,7
53.0 - 58.5	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	16	20	9	11	10

- Es sind ____ von 86 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.

Programmieren in R

1. Aufgabe

(2 Punkte)

Sie wollen Ihren Datensatz in **R** einlesen und stehen nun vor einem Problem. Sie stellen fest, dass die Hilfeseiten alle in englischer Sprache verfasst sind. Warum mag die Nutzung von Deutsch problematisch sein?

- A** ☐ Im Allgemeinen haben Programmiersprachen Probleme mit Umlauten und Sonderzeichen, die in der deutschen Sprache vorkommen. Eine Nutzung der englischen Sprache umgeht dieses Problem auf einfache Art.
- B** ☐ Die Spracherkennung von **R** ist nicht in der Lage Deutsch zu verstehen.
- C** ☐ **R** Pakete sind nur in englischer Sprache verfasst. Es macht keinen Sinn **R** daher in Deutsch zu bedienen.
- D** ☐ Programmiersprachen können nur englische Begriffe verarbeiten. Zusätzliche Pakete können zwar geladen werden, aber meist funktionieren diese Pakete nicht richtig. Deutsch ist International nicht bedeutend genug.
- E** ☐ Programmiersprachen haben Probleme mit Umlauten und Sonderzeichen der deutschen Sprache. Daher ist die Nutzung in Deutsch in den AGBs von **R** untersagt.

2. Aufgabe

(2 Punkte)

In **R** müssen die Daten eines Experiments in einem bestimmten Format vorliegen. Wir sprechen auch von dem log-Format der Daten. Welche Aussage zu dem Long-Format ist richtig?

- A** ☐ Das Long-Format hat die gleiche Anzahl an Spalten wie auch an Zeilen.
- B** ☐ Das Long-Format beschreibt in den Spalten die Beobachtungen sowie in den Zeilen die *unabhängigen* Beobachtungen.
- C** ☐ In den Spalten finden sich die einzelnen Beobachtungen und in den Zeilen die gemessenen Variablen (Y) sowie die experimentellen Faktoren X.
- D** ☐ Wichtig ist, dass sich in den Zeilen die Beobachtungen finden und in den Spalten die Variablen, wie die Messwerte Y sowie die experimentellen Faktoren X.
- E** ☐ In den Spalten sind die Beobachtungen in den Zeilen die Variablen, wie die Messwerte und experimentellen Faktoren.

3. Aufgabe

(2 Punkte)

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

- A** ☐ Die Funktionsreihenfolge ist wie folgt: `read_excel()` -> `mutate()` -> `ggplot()`. Dabei ist bei der Transformation der Daten darauf zu achten, dass keine Faktoren erstellt werden.
- B** ☐ Die Funktionsreihenfolge ist wie folgt: `read_excel()` -> `mutate()` -> `ggplot()`. Dabei ist bei der Transformation der Daten darauf zu achten, dass die Faktoren richtig erstellt werden.

- C** ☐ 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.
- D** ☐ Für eine explorativen Datenanalyse (EDA) in **R** 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 eigentliche EDA.
- E** ☐ 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.

Deskriptive Statistik & Explorative Datenanalyse

4. Aufgabe

(2 Punkte)

Gegeben ist y mit 8, 12, 15, 5 und 18. Berechnen Sie den Mittelwert und Standardabweichung.

- A** ☐ Es ergibt sich 12.6 +/- 2.61
- B** ☐ Sie erhalten 11.6 +/- 2.61
- C** ☐ Sie erhalten 11.6 +/- 5.22
- D** ☐ Es berechnet sich 11.6 +/- 27.3
- E** ☐ Sie erhalten 11.6 +/- 2.28

5. Aufgabe

(2 Punkte)

Wie lautet der Median, das 1st Quartile sowie das 3rd Quartile von y mit 30, 13, 15, 29, 32, 9, 2, 37, 26, 11 und 63.

- A** ☐ Es berechnet sich 24 [12; 33]
- B** ☐ Sie erhalten 26 [9; 30]
- C** ☐ Es ergibt sich 26 +/- 11
- D** ☐ Sie erhalten 26 [11; 32]
- E** ☐ Es ergibt sich 24 +/- 11

6. Aufgabe

(2 Punkte)

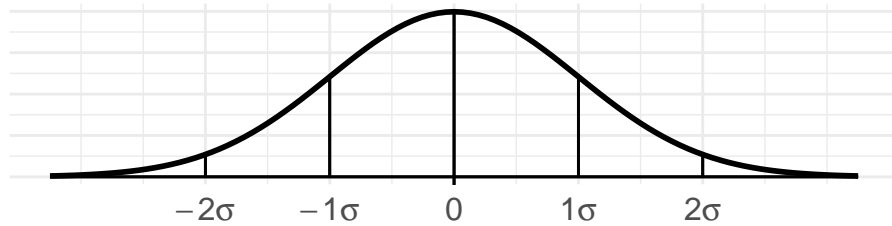
Um die Standardabweichung zu berechnen müssen wir folgende Rechenoperationen durchführen.

- A** ☐ Den Mittelwert berechnen und die Abstände quadrieren. Die Summe mit der Fallzahl multiplizieren.
- B** ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend subtrahieren wir die Fallzahl.
- C** ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend teilen wir durch die Fallzahl.
- D** ☐ 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.
- E** ☐ Wir berechnen erst den Mittelwert und dann die quadratischen Abstände zu dem Mittelwert. Diese quadratischen Abstände summieren wir auf und teilen am Ende durch die Fallzahl. Als letzten Schritt ziehen wir die quadratische Wurzel.

7. Aufgabe

(2 Punkte)


Die Normalverteilung wird durch den Mittelwert und die Standardabweichung definiert. Welche Aussage im Bezug zur Standardabweichung in einer Normalverteilung ist richtig?



- A** ☐ Die Fläche rechts von 2σ ist der p-Wert mit $Pr(D|H_0)$ in der obigen Abbildung.
- B** ☐ Die Fläche zwischen -1σ und 1σ ist 0.95 und 95% der Beobachtungen liegen somit zwischen $\bar{y} \pm \sigma$ in der obigen Verteilung.
- C** ☐ Die Fläche zwischen -2σ und 2σ ist 0.68 und 68% der Beobachtungen liegen somit zwischen $\bar{y} \pm \sigma$ in der obigen Verteilung.
- D** ☐ Die Fläche unter der Kurve ist 1, wenn die Nullhypothese falsch ist. Wenn die Nullhypothese gilt, dann ist die Fläche $1 - \alpha$. Somit ergibt sich auch eine Standardabweichung σ mit α gleich 0.05 in beiden Fällen.
- E** ☐ Die Fläche zwischen -2σ und 2σ ist 0.95 und 95% der Beobachtungen liegen somit zwischen $\bar{y} \pm \sigma$ in der obigen Verteilung.

8. Aufgabe

(2 Punkte)

Nachdem Sie eine ANOVA und die paarweisen t-Tests über das  Paket {emmeans} durchgeführt haben, müssen Sie Ihre Daten nochmal zur Überprüfung visualisieren. Sie entscheiden sich für den Boxplot. Welche statistischen Maßzahlen stellt der Boxplot dar?

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

9. Aufgabe

(2 Punkte)

Der Mittelwert \bar{y} und der Median \tilde{y} unterscheiden sich nicht in Ihren Feldexperiment zu Leistungssteigerung von Brokoli. 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 nicht unterscheiden, liegen vermutlich keine Outlier in den Daten vor.
- C** ☐ Wenn sich der Mittelwert und der Median nicht unterscheiden, liegen vermutlich Outlier in den Daten vor.
- D** ☐ Da sich der Mittelwert und der Median unterscheiden, ist der Datensatz nicht zu verwenden. Mittelwert und Median müssen gleich sein.
- E** ☐ Wenn sich der Mittelwert und der Median unterscheiden, liegen vermutlich keine Outlier in den Daten vor.

10. Aufgabe

(2 Punkte)

Mit einem Dotplot können Sie sehr gut die Verteilung von Daten visualisieren. Die empfohlene Mindestanzahl an Beobachtungen ist dabei?

- A** ☐ Wir sollten zwei bis fünf Beobachtungen mindestens pro Gruppe vorliegen haben.
- B** ☐ Wir sollten eine Beobachtung mindestens pro Gruppe vorliegen haben.
- C** ☐ Wir brauchen fünf oder mehr Beobachtungen.
- D** ☐ Die untere Grenze liegt bei einer Beobachtung.
- E** ☐ erhalten, sollten wir mindestens zwanzig Beobachtungen haben.

11. Aufgabe

(2 Punkte)

In der Statistik müssen wir häufig überprüfen, ob unser Outcome einer bestimmten Verteilung folgt. Meistens überprüfen wir, ob eine Normalverteilung vorliegt. Folgende drei Abbildungen eignen sich im Besonderen für die Überprüfung einer Verteilungsannahme an eine Variable.

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

12. Aufgabe

(2 Punkte)

Ihre Betreuung der Abschlussarbeit fragt überraschend in der letzten Besprechung, ob Ihre Messwerte einer Varianzhomogenität 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 Varianzhomogenität zur Anwendung?

- A** ☐ Einen Dotplot. Die Punkte müssen sich wie an einer Perlenschnur ausreihen. Eine Abweichung führt zur Ablehnung der Annahme einer Varianzhomogenität.
- B** ☐ 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.
- C** ☐ Nach dem Einlesen der Daten nutzen wir einen Barplot um zu schauen, ob alle Mittelwerte über alle Behandlungen in etwa gleich groß sind. Damit ist dann auch die Varianz in allen Behandlungen in etwa gleich.
- D** ☐ In einer explorativen Datenanalyse nutzen wir den Violinplot. Dabei sollte der Bauch am Rand liegen. Dann können wir von einer Varianzhomogenität ausgehen.
- E** ☐ 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 Varianzhomogenität ausgehen.

Statistische Testtheorie

13. Aufgabe

(2 Punkte)

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

- A** ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit nicht die Daten D zu beobachten sondern die Nullhypothese, wenn diese wahr ist.
- B** ☐ $Pr(D|H_0)$ stellt die Wahrscheinlichkeit die Daten D und somit die Teststatistik T_D zu beobachten dar, wenn die Nullhypothese wahr ist.

- C** ☐ $Pr(D|H_0)$ stellt die Wahrscheinlichkeit die Teststatistik T zu beobachten dar, wenn die Nullhypothese falsch ist.
- D** ☐ Die Wahrscheinlichkeit für die Nullhypothese, wenn die Daten wahr sind.
- E** ☐ Die Wahrscheinlichkeit der Daten unter der Nullhypothese in der Grundgesamtheit.

14. Aufgabe

(2 Punkte)

Das statistische Testen basiert auf dem Falsifikationsprinzip. Es besagt,

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

15. 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** ☐ In der Wissenschaft gibt es neben der Naturkonstante, die sich aus der Beobachtung der Welt ergibt, noch die Kulturkonstante, die von einer Gruppe Menschen selbstgewählt wird. Dabei ist $\alpha = 5\%$ eine Kulturkonstante und wurde somit eher zufällig gewählt.
- B** ☐ 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.
- C** ☐ 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.
- D** ☐ Auf einer Statistikkonferenz in Genf im Jahre 1942 wurde dieser Cut-Off nach langen Diskussionen festgelegt. Bis heute ist der Cut Off aber umstritten, da wegen dem 2. Weltkrieg viele Wissenschaftler nicht teilnehmen konnten.
- E** ☐ 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.

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

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

18. Aufgabe

(2 Punkte)

Sie sollen in Ihrer Abschlussarbeit die Relevanz und die Signifikanz in einer statistischen Maßzahl vereinen. Welche Aussage ist richtig?

- A** ☐ Das Δ . Durch die Effektstärke haben wir einen Wert für die Relevanz, die vom Anwender bewertet werden muss. Da Δ antiproportional zum p-Wert ist, bedeutet auch ein hohes Δ ein sehr kleinen p-Wert.
- B** ☐ Einem Konfidenzintervall. Das Konfidenzintervall bringt durch eine Visualisierung und zwei Intervallgrenzen die Möglichkeit mit, eine Relevanzschwelle neben der definierten Signifikanzschwelle zu definieren.
- C** ☐ 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.
- D** ☐ Der p-Wert. Durch den Vergleich mit α lässt sich über die Signifikanz entscheiden und der β -Fehler erlaubt über die Power eine Einschätzung der Relevanz.
- E** ☐ Das OR. Als Chancenverhältnis gibt es das Verhältnis von Relevanz und Signifikanz wieder.

19. Aufgabe

(2 Punkte)

Ein statistischer Test produziert für einen Gruppenvergleich einen p-Wert. Welche Aussage zusammen mit dem Signifikanzniveau α gleich 5% stimmt?

- A** ☐ Wir machen eine Aussage über die Flächen und der Kurve der Teststatistik, wenn die H_0 gilt. Dabei werden Wahrscheinlichkeiten verglichen, die durch die Flächen unter der Kurve repräsentiert werden.
- B** ☐ 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.
- C** ☐ Wir machen eine Aussage über die individuelle Wahrscheinlichkeit des Eintretens der Nullhypothese H_0 . Der p-Wert wird mit dem Signifikanzniveau verglichen und bewertet.
- D** ☐ Wir vergleichen mit dem p-Wert und dem Signifikanzniveau α Wahrscheinlichkeiten und damit die absoluten Werte auf einem Zahlenstrahl, wenn die H_0 gilt.
- E** ☐ 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.

20. Aufgabe

(2 Punkte)

Um die Ergebnisse eines statistischen Tests und die damit verbundene Theorie besser zu verstehen, kann eine Analogie zur Wettervorhersage genutzt werden. Welche Analogie zu der Testtheorie trifft am meisten zu?

- A** ☐ 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.
- B** ☐ In der Analogie des Niederschlags oder Regenmenge: ein statistischer Test gibt die Stärke eines Effektes wieder. Zum Beispiel, wie hoch ist der Mittelwertsunterschied.
- C** ☐ 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.
- D** ☐ In der Analogie der Sonnenscheindauer: Wie lange kann mit einem entsprechenden Effekt gerechnet werden? Die Wahrscheinlichkeit für den Effekt gibt der statistische Test wieder.
- E** ☐ In der Analogie der Regenwahrscheinlichkeit: ein statistischer Test gibt die Wahrscheinlichkeit für das Auftreten eines Ereignisses wieder. Die Stärke des Effektes wird nicht wiedergegeben.

21. Aufgabe

(2 Punkte)

Sie wollen eine Aussage über die untersuchte Population treffen. Dazu nutzen Sie einen statistischen Test. Können Sie eine valide Aussage aus einem statistischen Test erhalten?

- A** ☐ Ja, wir erhalten eine Aussage. Müssen aber das Individuum im Kontext der Population adjustieren.
- B** ☐ Nein, wir erhalten nur eine Aussage zu zwei Individuen. Ein statistischer Test liefert Informationen zu einem Individuum im Vergleich zu einem anderen Individuum.
- C** ☐ Nein, es ist nicht möglich die untersuchte Population mit einem t-Test auszuwerten. Wir erhalten dann leider keine Aussage zur Population.
- D** ☐ Nein, wir können die untersuchte Population nicht mit einer ANOVA auswerten. Wir erhalten keine Aussage zur Population. Wir können aber den Test adjustieren und so die Auswertung ermöglichen.
- E** ☐ Ja, die untersuchte Population können wir mit einem statistischen Test auswerten. Wir erhalten dann eine Aussage zur Population.

22. Aufgabe

(2 Punkte)

Sie haben die *Power* berechnet. Was sagt Ihnen dieser statistische Begriff aus?

- A** ☐ Die Power ist nicht in der aktuellen Testthorie mehr vertreten. Wir rechnen nur noch mit dem Fehler 1. Art.
- B** ☐ Es gilt $\alpha + \beta = 1$ und somit liegt β meist bei 95%.
- C** ☐ Alle statistischen Tests sind so konstruiert, dass die H_A mit 20% *bewiesen wird*. Die Power ist $1 - \beta$ mit β gleich 80% gesetzt.
- D** ☐ Die Power $1 - \beta$ wird auf 80% gesetzt. Damit liegt die Wahrscheinlichkeit für die H_0 bei 20%.
- E** ☐ 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*

23. 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 Ende wissen, ob das Ergebnis eines Experiments relevant für seine Forschung ist. Dafür kann der Effekt eines statistischen Tests genutzt werden. Damit beschreibt der Effekt den biologischen interpretierbaren Teil einer Ausgabe eines Tests. Zum Beispiel der Unterschied zwischen zwei Anteilen.
- B** ☐ 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.
- C** ☐ 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.
- D** ☐ Der Effekt eines statistischen Tests beschreibt die mathematisch 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.

24. Aufgabe

(2 Punkte)

Welche Aussage über die Entscheidung anhand der berechneten Teststatistik gegen die Nullhypothese ist richtig?

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

25. Aufgabe

(2 Punkte)

Ein statistischer Test benötigt für die richtige Durchführung Hypothesen H , sonst ist der Test nicht zu interpretieren. Welche Aussage ist richtig?

- A** ☐ Mit der Nullhypothese H_0 und der Alternativhypothese H_A oder H_1 gibt es zwei Hypothesen.
- B** ☐ Ein statistisches Hypothesenpaar gibt es. Zum einen die Nullhypothese und zum anderen die Alternativhypothese. Es ist aber nur notwendig die Alternative anzugeben, da die Nullhypothese nicht beim Testen benötigt wird.
- C** ☐ Es gibt ein statistisches Hypothesenpaar mit der Hypothese für und gegen die wissenschaftliche Fragestellung. Die Hypothesen werden H_{pro} und H_{contra} bezeichnet.
- 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** ☐ 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.

26. Aufgabe

(2 Punkte)

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

- A** ☐ Randomisierung ist die direkte Folge von Strukturgleichheit. Die Strukturgleichheit erlaubt es erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- B** ☐ Randomisierung erlaubt erst die Varianzen zu schätzen. Ohne eine Randomisierung ist die Berechnung von Mittelwerten und Varianzen nicht möglich. Dadurch lässt sich erst ein Experiment auswerten.
- C** ☐ Randomisierung sorgt für Strukturgleichheit und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- D** ☐ 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.
- E** ☐ Randomisierung war bis 1952 bedeutend, wurde dann aber in Folge besserer Rechnerleistung nicht mehr verwendet. Aktuelle Statistik nutzt keine Randomisierung mehr.

27. 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** ☐ Relevanz und Signifikanz haben nichts miteinander zu tun. Daher gibt es auch keinen Zusammenhang zwischen hoher Fallzahl ($n > 10000$) und einem signifikanten Test. Ein Effekt ist immer relevant und somit signifikant.
- B** ☐ 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.
- C** ☐ 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.
- D** ☐ Aktuell werden immer größere Datensätze erhoben. Dadurch wird auch die Varianz immer höher was automatisch zu mehr signifikanten Ergebnissen führt.
- E** ☐ Eine große Fallzahl führt zu mehr signifikanten Ergebnissen auch bei kleinen Effekten. Daher werden fast alle Vergleiche signifikant, wenn die Fallzahl nur groß genug wird.

28. Aufgabe

(2 Punkte)

In einem Zuchtexperiment messen wir die Ferkel verschiedener Sauen. Die Ferkel einer Muttersau sind daher im statistischen Sinne...

- A** ☐ Untereinander abhängig, wenn die Mütter ebenfalls miteinander verwandt sind. Erst die Abhängigkeit 2. Grades wird in der Statistik modelliert.
- B** ☐ Die Ferkel stammen von der gleichen Sau und sind somit untereinander unabhängig.
- C** ☐ Abhängig von der Stallanlage und des Experiments können die Ferkel abhängig oder unabhängig sein. Allgemein gilt, dass Ferkel von unterschiedlichen Sauen näher miteinander verwandt sind als Ferkel von gleichen Sauen. Das Fisher-Axiom.
- D** ☐ Die Ferkel stammen vom gleichen Muttertier und haben vermutlich eine ähnlichere Varianzstruktur als die Ferkel von anderen Sauen. Die Ferkel sind untereinander über die Mutter abhängig.
- E** ☐ Untereinander stark korreliert. Die Ferkel sind von einer Mutter und somit miteinander korreliert. Dies wird in der Statistik jedoch meist nicht modelliert.

Statistische Tests für Gruppenvergleiche

29. Aufgabe

(2 Punkte)

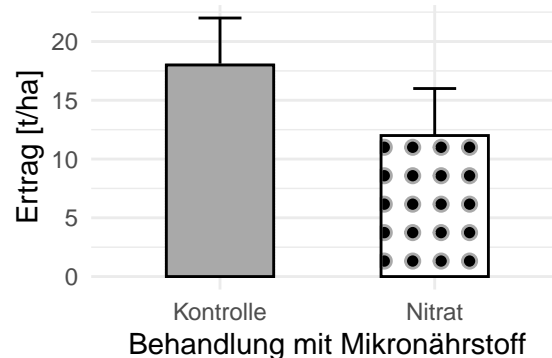
Nach einem Feldexperiment wollen Sie zwei Gruppen mit einem Welch t-Test vergleichen. Welche Aussage ist auch für den Student t-Test richtig?

- A** ☐ Der t-Test vergleicht zwei oder mehr Gruppen indem die Mittelwerte miteinander verglichen werden.
- B** ☐ Der t-Test vergleicht die Mittelwerte von zwei Gruppen.
- C** ☐ Der t-Test ist ein Vortest der ANOVA und basiert daher auf dem Vergleich von Streuungsparametern
- 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.

30. Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung des Mikronährstoff Nitrat auf den Ertrag in t/ha von Mango im Vergleich zu einer Kontrolle. Der Versuch wurde in 9 Parzellen pro Gruppe durchgeführt. Welche Aussage im Bezug auf eine statistische Auswertung ist richtig?



- A** ☐ Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt kann nicht bei einem t-Test aus Barplots bestimmt werden.
- B** ☐ 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.
- C** ☐ Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei -6.
- D** ☐ Die Barplots deuten auf keinen signifikanten Unterschied. Der Effekt liegt vermutlich bei -6 unter einer groben Abschätzung. Wir müssen aber eine ANOVA rechnen um den Effekt wirklich bestimmen zu können.
- E** ☐ Der Test deutet auf keinen signifikanten Unterschied hin. Der Effekt liegt vermutlich bei -6.

31. Aufgabe

(2 Punkte)

Sie rechnen einen gepaarten t-Test, da Ihre Beobachtungen verbunden sind. Welche der folgenden Aussagen 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** ☐ Abhängige Beobachtungen müssen gesondert in einem gepaarten t-Test modelliert werden. Wenn wiederholt an dem gleichen Tier oder Pflanze gemessen wird, dann bilden wir die Differenz zwischen den beiden Zeitpunkten. Auf den Differenzen rechnen wir den gepaarten t-Test.
- 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** ☐ Der gepaarte t-Test nutzt die Varianz der Beobachtungen jeweils paarweise und bildet dafür eine verbundene Stichprobe. Dieser Datensatz d dient dann zur Differenzbildung.

32. Aufgabe

(2 Punkte)

Nach einem Experiment mit fünf Weizensorten ergibt eine ANOVA ($p = 0.048$) einen signifikanten Unterschied für den Ertrag. Sie führen anschließend die paarweisen t-Tests für alle Vergleiche der verschiedenen Weizensorten 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 $p_{3-2} = 0.051$. Welche Aussage ist richtig?

- A** ☐ Der Fehler liegt in den t-Tests. Wenn eine ANOVA signifikant ist, dann muss zwangsweise auch ein t-Test signifikant sein.
- B** ☐ Die adjustierten p-Werte deuten in die richtige Richtung. Zusammen mit den nicht signifikanten rohen p-Werten ist von einem Fehler in der ANOVA auszugehen.
- C** ☐ Es gibt einen Fehler in der Varianzstruktur. Daher kann die ANOVA nicht richtig sein und paarweise t-Tests liefern das richtige Ergebnis.
- D** ☐ Das Beispiel kann so nicht auftreten, da die ANOVA und die t-Tests algorithmisch miteinander verknüpft sind.
- E** ☐ Das ist kein Wunder. Die ANOVA testet auf der gesamten Fallzahl und die paarweisen t-Tests verlieren immer eine oder mehr Gruppen als Fallzahl. Mit steigender Fallzahl sind mehr signifikante Unterschiede zu erwarten. Die p-Werte unterscheiden sich numerisch auch kaum.

ANOVA

33. 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 $\eta^2 = 0.31$. 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 ist damit mit dem R^2 aus der linearen Regression zu vergleichen und beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird.
- C** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen nicht erklärt wird. Somit der Rest an nicht erklärbarer Varianz.
- D** ☐ Die Berechnung von η^2 ist ein Wert für die Interaktion.
- 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.

34. Aufgabe

(2 Punkte)

Sie führen ein Feldexperiment durch um das Gewicht von Lauch zu steigern. Die Pflanzen wachsen unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen. Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.19$. Welche Aussage ist richtig?

- A** ☐ Es werden 81% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.

- B** ☐ Es werden 19% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- C** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 19% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 81%.
- D** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird. Daher werden 19% der Varianz durch die Behandlungsgruppen erklärt.
- E** ☐ 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.

35. Aufgabe

(2 Punkte)

Sie rechnen eine einfaktorielle ANOVA und erhalten eine Teststatistik. Nun müssen Sie diese Teststatistik interpretieren. Welche Aussage ist richtig?

- A** ☐ 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.
- B** ☐ 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.
- C** ☐ 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.
- D** ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 1 annähert kann die Nullhypothese nicht abgelehnt werden.
- E** ☐ 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.

36. Aufgabe

(2 Punkte)

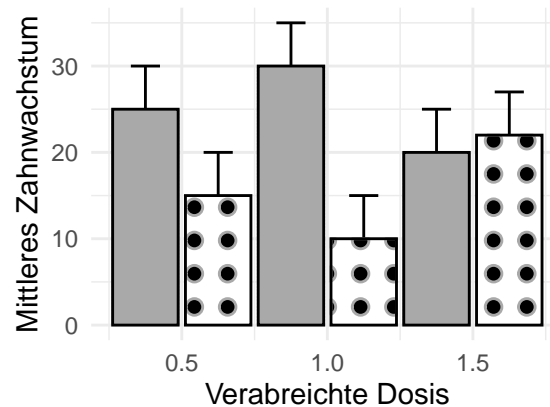
Viele statistische Verfahren nutzen eine Teststatistik um eine Aussage über den Zusammenhang zwischen der Grundgesamtheit und der Stichprobe abzubilden. Ein statistisches Testwerkzeug ist hierbei die ANOVA. Die ANOVA rechnet dabei...

- A** ☐ ... den Unterschied zwischen der Varianz aus verschiedenen Behandlungsgruppen und der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in den Gruppen exakt zu bestimmen.
- B** ☐ ... den Unterschied zwischen der Varianz in den verschiedenen Behandlungsgruppen und der Varianz in einer der Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in einer der Gruppen exakt zu bestimmen.
- C** ☐ ... den Unterschied zwischen der Varianz ausgelöst durch alle Behandlungsgruppen und der Varianz aus globalen Behandlungsgruppen der Kontrollen. Wenn die ANOVA nicht signifikant ist, muss ein Posthoc-Test ausgeschlossen werden.
- D** ☐ ... den Unterschied zwischen der Mittelwerte und der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist bekannt welcher Vergleich konkret unterschiedlich ist.
- E** ☐ ... den Unterschied zwischen der Varianz durch verschiedene Behandlungsgruppen unter der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, kann kein Effekt η^2 bestimmt werden.

37. Aufgabe

(2 Punkte)

Ein Versuch wurde an 70 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. Die folgende Abbildung enthält die Daten aus diesem Versuch zur Bewertung der Wirkung von Vitamin E auf das Zahnwachstum bei Hasen. Welche Aussage ist richtig, wenn Sie eine zweifaktorielle ANOVA rechnen?



- A ☐ Eine mittlere bis starke Interaktion liegt vor ($p \leq 0.05$)
- B ☐ Keine Interaktion liegt vor ($p \leq 0.05$).
- C ☐ Das Bestimmtheitsmaß R^2 ist klein.
- D ☐ Die Koeffizienten sind positiv ($\beta_0 > 0$; $\beta_1 > 0$).
- E ☐ Keine Korrelation liegt vor ($p \geq 0.05$).

Multiple Gruppenvergleiche

38. Aufgabe

(2 Punkte)

Sie haben folgende unadjustierten p-Werte gegeben: 0.03, 0.42, 0.34, 0.001 und 0.01. 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, 1, 1, 0.005 und 0.05. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- B ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.15, 2.1, 1.7, 0.005 und 0.05. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- C ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.15, 1, 1, 0.005 und 0.05. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.
- D ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.006, 0.084, 0.068, 2e-04 und 0.002. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- E ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.006, 0.084, 0.068, 2e-04 und 0.002. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.

39. Aufgabe

(2 Punkte)

Auf wissenschaftlichen Postern finden Sie unter Abbildungen häufig die Abkürzung *CLD*. Für welchen statistischen Fachbegriff steht die Abkürzung und wie interpretieren Sie ein *CLD*?

- A ☐ 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.

- B** ☐ Compound letter display. Gleichheit in dem Outcomes wird durch den gleichen Buchstaben oder Symbol dargestellt. Teilweise ist die Interpretation des Verbunds (eng. compound) herausfordernd, da wir ja nach dem Unterschied suchen.
- C** ☐ 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.
- D** ☐ Compact line display. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt. Früher wurden keine Buchstaben sondern eine durchgezogene Linie verwendet. Bei mehr als drei Gruppen funktioniert die Linie aber graphisch nicht mehr.
- E** ☐ Compact letter detection. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt.

40. 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 {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.
- B** ☐ 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.
- C** ☐ Da Sie für Ihre Bachelorarbeit einen Barplot mit CLD brauchen nutzen Sie das R Paket {emmeans} welches Ihnen schnell die notwendigen Informationen liefert um einen Barplot zu erstellen. Die Berechnung eines CLD ist hierbei auch einfach.
- D** ☐ 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.
- 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.

41. Aufgabe

(2 Punkte)

Bei einem multiplen Vergleich oder Posthoc Test kann es zu einer Besonderheit beim statistischen Testen kommen. Wie nennt man diese Besonderheit beim statistischen Testen und wie kann man mit ihr umgehen?

- A** ☐ Das globale Signifikanzniveau explodiert und erreicht Werte größer als Eins. Es kommt zu einer α -Inflation. Dagegen kann mit der Adjustierung der α -Werte nach Bonferroni vorgegangen werden.
- B** ☐ Beim multiplen Testen kann es zu einer α -Inflation 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 Welch das bekannteste Verfahren ist.
- C** ☐ Die Adjustierung der p-Werte nach Bonferroni erlaubt es gegen die β -Inflation vorzugehen, die häufig beim multiplen Testen auftritt. Das globale Powerniveau liegt nicht mehr bei 80% sondern sehr viel niedriger.
- D** ☐ 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.
- E** ☐ 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.

42. Aufgabe

(2 Punkte)

Sie rechnen mehrere t-Tests für einen multiplen Vergleich nachdem eine einfaktorielle ANOVA sich als signifikant herausgestellt hat. Welche Aussage im Bezug auf den Effekt ist richtig?

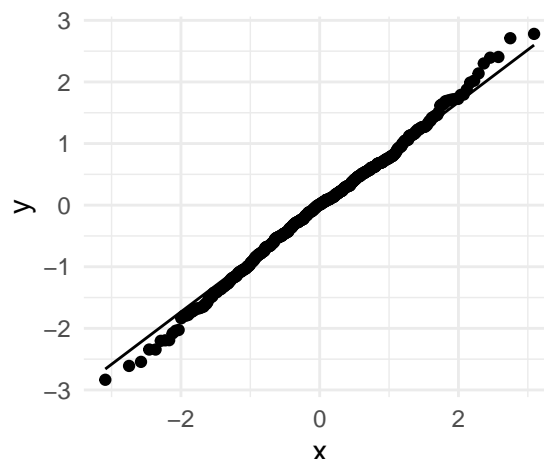
- A** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ adjustiert werden im Gegensatz zu den p-Werten.
- B** ☐ Beim multiplen Testen werden die Effekte der paarweisen Vergleiche ignoriert. Der Nachteil des multiplen Testens ist ja auch, dass wir am Ende keine Effekte mehr vorliegen haben. Eine ANOVA liefert hier bessere Informationen.
- C** ☐ 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.
- D** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ nicht adjustiert werden. Bei einem Effekt im multiplen Testen handelt es sich um eine Wahrscheinlichkeit für das Auftreten der Nullhypothese.
- E** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ nicht adjustiert werden im Gegensatz zu den p-Werten.

Lineare Regression & Korrelation

43. 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 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.
- B** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.
- C** ☐ 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.
- D** ☐ Wir betrachten die Punkte. Wenn die Punkte einigermaßen gleichmäßig verteilt liegen, dann gehen wir von normalen Residuen aus.
- E** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.

44. Aufgabe

(2 Punkte)

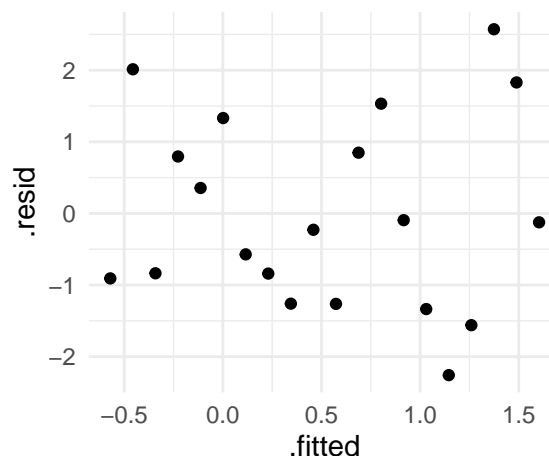
In Ihrer Abschlussarbeit wollen Sie ein kausales Modell rechnen. Jetzt stellt sich die Frage, was diese Entscheidung für Ihre Auswertung bedeutet. Welche Aussage ist richtig?

- A** ☐ Wir modellieren den Zusammenhang zwischen X und Y wenn ein kausales Modell rerechnet wird. Dabei kann der gesamte Datensatz genutzt werden. Eine Aufteilung wie in einem prädiktiven Modell ist nicht notwendig.
- B** ☐ Ein kausales Modell benötigt mindestens eine Fallzahl von über 100 Beobachtungen und darf keine fehlenden Werte beinhalten. Die Varianzkomponenten müssen homogen sein.
- C** ☐ Ein kausales Modell schliesst grundsätzlich lineare Modell aus. Es muss ein Graph gefunden werden, der alle Punkte beinhaltet. Erst dann kann das R^2 berechnet werden.
- D** ☐ Wenn ein kausales 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.
- E** ☐ Wir modellieren den Zusammenhang zwischen X und Y wenn ein kausales Modell rerechnet wird. Dabei kann nicht der gesamte Datensatz genutzt werden. Es wird ein Trainingsdatensatz zum Trainieren des Modells benötigt.

45. 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** ☐ Wenn wir die Nulllinie betrachten so müssen die Punkte gleichmäßig unter der Nulllinie liegen. Unser Modell erfüllt somit nicht die Annahme von normalverteilten Residuen mit einem Mittelwert von > 0 und einer Streuung von s .
- B** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Ein klares Muster ist zu erkennen und/oder einige Outlier sind zu beobachten.
- C** ☐ Die Punkte müssen gleichmäßig in dem positiven Bereich liegen. Dies ist hier klar nicht der Fall. Einzelne Ausreißer können beobachtet werden. Die Analyse ist gescheitert.
- D** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Diagonalen. Damit ist das Modell erfolgreich geschätzt worden.
- E** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Kein Muster ist zu erkennen und keine Outlier zu beobachten.

46. Aufgabe

(2 Punkte)

Welche Aussage über den Korrelationskoeffizienten ρ ist richtig?

- A** ☐ Korrelationskoeffizienten ρ liegt zwischen 0 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als Standardisierung verstanden werden.
- B** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen -1 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos.
- C** ☐ 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.
- D** ☐ 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.
- E** ☐ 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.

47. Aufgabe


(2 Punkte)


In einer linearen Regression kann es vorkommen, dass der Effekt repräsentiert durch den β Koeffizienten nicht so richtig von der Größenordnung zu dem p-Wert passen will. So liefert eine Untersuchung des Einflusses von der Fe_3O_4 -Konzentration in $[\mu\text{g}]$ im Wasser auf das Trockengewicht in $[\text{kg}]$ an Erbsen folgende Effekte und p-Werte: 0.00032 als p-Wert und einen $\beta_{\text{Fe}_3\text{O}_4}$ Koeffizienten von 7.4×10^{-6} . Welche Aussage ist richtig?

- A** ☐ Die Einheit der Fe_3O_4 -Konzentration ist zu klein gewählt. Die Erhöhung der Fe_3O_4 -Konzentration um 1 Einheit führt nur zu einem sehr winzigen Anstieg von $\beta_{\text{Fe}_3\text{O}_4}$ im Gewicht der Wasserlinsen. Die Einheit $[\mu\text{g}]$ muss besser gewählt werden.
- B** ☐ Wenn der Effekt $\beta_{\text{Fe}_3\text{O}_4}$ winzig ist, dann kann es an einer falsch gewählten Einheit liegen. Der Anstieg von einer Einheit in X führt ja zu einer Änderung von $\beta_{\text{Fe}_3\text{O}_4}$ in x . Wir müssen daher die Einheit von y entsprechend anpassen.
- C** ☐ Das Gewicht und die Fe_3O_4 -Konzentration korrelieren sehr stark, deshalb wird der $\beta_{\text{Fe}_3\text{O}_4}$ Koeffizient sehr klein. Mit einer ANOVA kann für die Korrelation korrigiert werden und der Effektschätzer passt dann zum p-Wert.
- 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** ☐ Die Fallzahl ist zu klein angesetzt. Je kleiner die Fallzahl ist, desto höher ist die Teststatistik und damit auch der p-Wert kleiner. Wir brauchen also mehr Fallzahl um den geringen Effekt noch signifikant zu kriegen.

48. Aufgabe

(2 Punkte)

Sie wollen nach der explorativen Datenanalyse (EDA) Ihre Daten in der Abschlussarbeit auswerten. Nach einiger Recherche finden Sie heraus, dass Sie zuerst die Daten mit der Funktion `lm()` in  modellieren müssen. Welche Anwendung folgt drauf?

- A** ☐ 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.
- B** ☐ Die Funktion `lm()` berechnet die Varianzstruktur für eine ANOVA. Dannach 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.
- C** ☐ Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich.

- D** ☐ Die Funktion `lm()` in **R** wird klassischerweise für die nicht-lineare Regression genutzt. Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt.
- E** ☐ 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.

49. Aufgabe

(2 Punkte)

Welche Aussage über das *generalisierte lineare Modell (GLM)* ist richtig?

- A** ☐ In **R** 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 **R** auch mit zusätzlich geladenen Paketen nicht möglich.
- B** ☐ Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsfamilien für das Y bzw. das Outcome in einer linearen Regression zu wählen.
- C** ☐ 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.
- D** ☐ Das GLM erlaubt auch nicht normalverteilte Residuen in der Schätzung der Regressionsgrade.
- E** ☐ Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsgruppen für das X bzw. die Einflussvariablen in einer linearen Regression zu wählen.

50. Aufgabe

(2 Punkte)

In einer Studie wollen Sie den Effektschätzer Risk ratio berechnen. Sie finden in Ihrem Experiment zur Behandlung von Klaueninfektionen bei Schafe in 3 Tieren Erkrankung der Klauen vor. 8 Tiere sind gesund. Welche Aussage ist richtig?

- A** ☐ Das Verhältnis der Chancen Risk ratio ergibt ein Chancenverhältnis von 0.27. Wir sind an der Chance krank zu sein interessiert.
- B** ☐ Das Verhältnis der Anteile Risk ratio ergibt ein Anteilsverhältnis von 0.27. Wir sind am Anteil der Kranken interessiert.
- C** ☐ Es ergibt sich ein Risk ratio von 0.38, da es sich um ein Anteil handelt.
- D** ☐ Es ergibt sich ein Risk ratio von 0.27, da es sich um eine Chancenverhältnis handelt.
- E** ☐ Das Verhältnis von Chancen Risk ratio ergibt ein Chancenverhältnis von 0.38.

Teil I.

Programmieren in R

51. Aufgabe

(9 Punkte)



Grundlegende Kenntnisse der Programmierung in R Nilufar muss ihrer Abschlussarbeit mit R arbeiten. Deshalb sitzt sie jetzt mit Ihnen zusammen und hat einige Fragen zu den Grundlagen in R an Sie! Na dann wollen Sie mal helfen. Immerhin will ihr Betreuer, dass R genutzt wird.

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

Sie antworten:

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

Sie antworten:

Nilufar: *Jetzt sehe ich wieder diese Tilde (~) in R. Wo brauchen wir diese denn nochmal?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

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

Sie antworten:

Nilufar: *In R hat sich doch mein Datensatz geändert. Warum ist er jetzt wieder so wie vorher, wenn ich den Datensatz wieder aufrufe? Was mache ich falsch?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

Nilufar: *Wenn ich was in R machen möchte, dann lade ich nochmal welche zwei R Pakete sehr häufig?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

52. Aufgabe

(9 Punkte)



Fortgeschrittene Kenntnisse der Programmierung in R 'Hm...am Ende ist dann R eigentlich gar nicht so schwer, wenn ich Hilfe habe.', meint Yuki stolz und lacht Sie an. Nur leider kennt er sich überhaupt nicht mit R aus! Das heißt, Sie müssen hier einmal Rede und Antwort stehen und helfen. Sonst wird es für Yuki dann in seiner Hausarbeit nichts mit der Auswertung und Abgabe. Das kann auch keine Lösung für Yuki und Sie sein. Immerhin haben Sie schon so viel gelernt.

Yuki fragt: *Wenn ich zwei Faktoren f_1 und f_2 habe, wie spezifiziere ich nochmal eine Interaktion in einem Modell?* **(1 Punkt)**

Sie antworten:

Yuki fragt: *Viele Abbildungen gehen ja auch in Excel, aber {ggplot} ist schon besser. Wie verbindet {ggplot} eigentlich nochmal die einzelnen Ebenen?* **(1 Punkt)**

Sie antworten:

Yuki 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:

Yuki fragt: *Oh, wie baue ich mir nochmal die Daten in R? Wie heißt das Dateiformat? Gerne einmal mit Beispiel!* **(1 Punkt)**

Sie antworten:

Yuki fragt: *Man kann doch die Funktion `emmeans()` von Varianzhomogenität auf Varianzheterogenität umstellen. Wie ging das noch gleich?* **(1 Punkt)**

Sie antworten:

Yuki fragt: *Was muss ich bei der Eingabe eines Datums in Excel beachten, wenn ich später die Exceldatei in R einlesen will? Wie lautet das Format?* **(1 Punkt)**

Sie antworten:

Yuki fragt: *Okay, und für was brauche ich nochmal die R Pakete {emmeans}, {ggplot} und {readxl}?* **(2 Punkte)**

Sie antworten:

Teil II.

Deskriptive Statistik & Explorative Datenanalyse

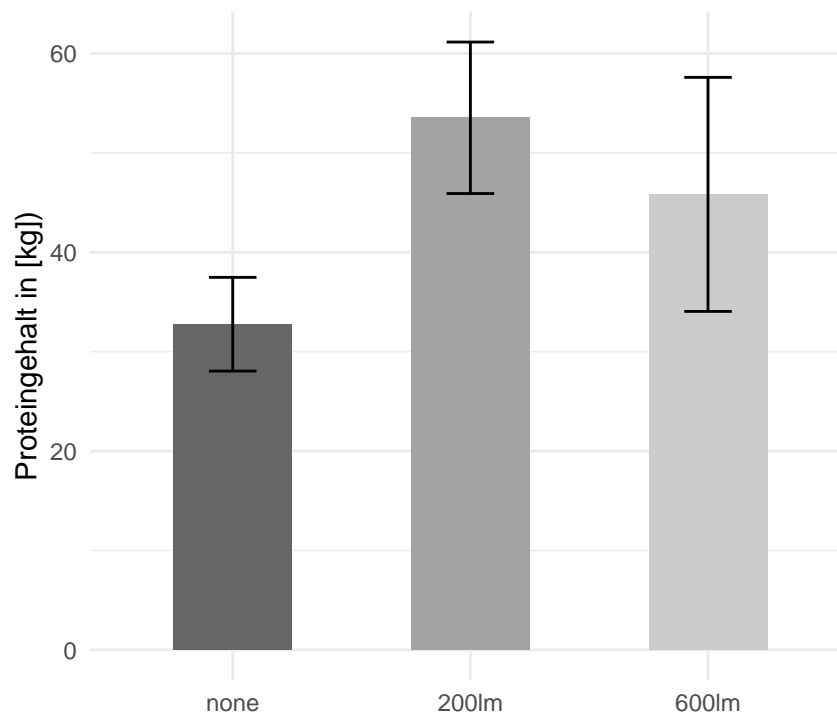
53. Aufgabe


(8 Punkte)


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



Zerforschen des Barplots Eine echte Herausforderung für ihn war schon immer die Erschöpfung gewesen. Ein leidiges Lied. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Jonas ist es eine Möglichkeit schneller ans Ziel zu gelangen. Jonas soll in seiner Hausarbeit Maiss untersuchen. Die Behandlung in seiner Hausarbeit werden verschiedene Lichtstufen (*none*, *200lm* und *600lm*) sein. Erheben wird Jonas als Endpunkt (Y) *Proteingehalt* benannt als *protein* in seiner Exceldatei. Von seiner Betreuer erhält er nun folgende Abbildung von Barplots, die er erstmal zur Übung nachbauen soll, bevor er mit dem eigentlichen Versuch beginnt.



Leider kennt sich Jonas mit der Erstellung von Barplots in  nicht aus. Deshalb braucht er 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 Jonas einen Unterschied zwischen den Behandlungen erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

54. Aufgabe

(8 Punkte)

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



Visualisierung des Barplots Alex steht vor einem ersten Problem, denn wenn es nach seiner Betreuerin geht, soll er in einem Freilandversuch Erbsen auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Alex liebt Starcraft. Darin kann er sich wirklich verlieren und immer wieder neu begeistern. Die Behandlung waren verschiedene Bewässerungstypen (*low*, *mid* und *high*). In seiner Excel-datei hat er den Outcome (Y) *Frischegewicht* als *freshmatter* aufgenommen. Nun soll Alex 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. Alex und die Gefälligkeit, eine unendliche Geschichte mit kniffligen Wendungen. Aber egal. Alex will später nochmal raus um zu Laufen. Druck ablassen, dass muss er auch.

treatment	freshmatter
high	43.8
mid	28.0
mid	39.1
mid	31.4
low	4.9
high	34.6
low	26.2
low	33.0
high	39.6
high	41.9
mid	48.3
low	19.9

Leider kennt sich Alex 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 Erbsen! Beschriften Sie die Achsen entsprechend! **(4 Punkte)**
3. Beschriften Sie *einen* Barplot mit den gängigen statistischen Maßzahlen! **(2 Punkte)**
4. Wenn Alex *keinen Effekt* zwischen den Behandlungen von Erbsen erwarten würde, wie sehen dann die Barplots aus? Antworten Sie mit einer Skizze der Barplots! **(1 Punkt)**

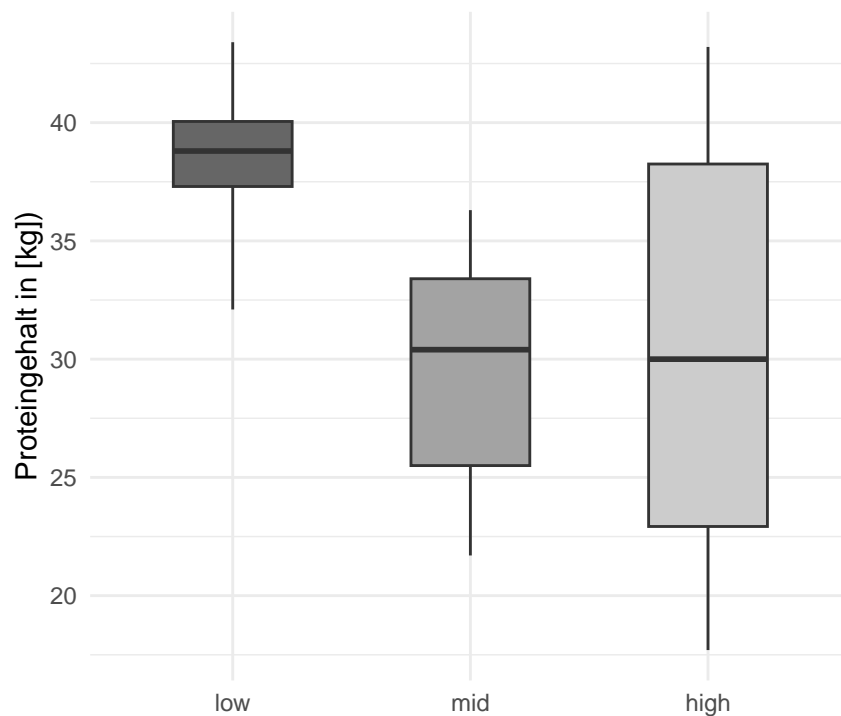
55. Aufgabe

(9 Punkte)

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



Zerforschen des Boxplots Tina steht vor einem ersten Problem, denn wenn es nach ihrem Betreuer geht, soll sie in einem Versuch in einer Klimakammer Maiss ausgewertet. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Tina liebt Astronomie. Darin kann sie sich wirklich verlieren und immer wieder neu begeistern. Das heißt erstmal überlegen für Tina. Aus den Boxen wummert Tocotronic und ihr Mund ist verklebt von Katjes. 'Herrlich', denkt Tina. Die Behandlung werden verschiedene Bewässerungstypen (*low*, *mid* und *high*) sein. In ihrer Exceldatei wird sie den Endpunkt (Y) *Proteingehalt* als *protein* aufnehmen. Vorab soll Tina aber einmal die folgenden Boxplots ihrem Betreuer nachbauen, damit sie den R Code schonmal für später vorliegen hat. Damit geht das Problem schon los. Eine echte Herausforderung für sie war schon immer die Wut gewesen. Ein leidiges Lied.



Leider kennt sich Tina mit der Erstellung von Boxplots in R 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 R üblichen Format! **(2 Punkte)**
4. Kann Tina einen Unterschied zwischen den Behandlungen erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

56. Aufgabe

(9 Punkte)

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



Visualisierung des Boxplots Boxplots sind bedeutend in der Darstellung von wissenschaftlichen Ergebnissen. Leider hat sich Paula nicht gemerkt, welche statistischen Maßzahlen für einen Boxplot erhoben werden müssen. Besser wäre was anderes gewesen. Am Ende dann doch besser Harry Potter. Wunderbar. Eine echte Ablenkung für Paula. Das ist in soweit doof, da nach ihrem Betreuer nun Boxplots aus ihren Daten gebaut werden sollen, bevor es mit dem statistischen Testen weitergeht. Anhand von Boxplots lässt sich eine Aussage über die Normalverteilung von Y treffen. Die Behandlung für Brokoli waren verschiedene Genotypen (AA und BB). Erfasst wurde von Paula als Endpunkt (Y) *Proteingehalt*. Paula hat dann *protein* in ihrer Exceldatei eintragen. Aber nur in passender Atmosphäre! Hm, lecker Smarties und dazu dann im Hintergrund Jagd auf roter Oktober laufen lassen.

treatment	drymatter
AA	49.3
BB	21.7
BB	29.8
AA	47.8
BB	21.5
BB	34.6
AA	40.6
AA	46.1
BB	19.6
AA	45.8
AA	44.5
AA	42.1
BB	26.0
BB	35.7
AA	49.1

Leider kennt sich Paula 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 Brokoli! 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 Brokoli erwarten würden, wie sehen dann die beiden Boxplots aus? Antworten Sie mit einer Skizze der Boxplots! **(1 Punkt)**

57. Aufgabe

(8 Punkte)

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



Visualisierung des Histogramm für kategoriale Daten In seiner Hausarbeit möchte Mark gerne die Daten aus einem Feldexperiment mit Lauch in einem Histogramm darstellen. Das Histogramm erlaubt ihm dabei Rückschlüsse auf die Verteilung über das Outcome (Y) zu treffen. Mark schmeißt noch eine Handvoll Marzipankugeln in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Andrea Berg. In seinem Experiment hat Mark die seltsamen Verdickungen gezählt. Es wäre einfacher, wenn da nicht noch was wäre. Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. Mark streichelt liebevoll der Hamster. Der Kopf ist in seinem Schloß vergraben um den Klang von Andrea Berg zu dämpfen.

Die seltsamen Verdickungen: 5, 3, 3, 5, 3, 6, 7, 6, 3, 6, 8, 4, 3, 5, 3, 6, 7, 3, 5, 5, 1, 5, 1, 8, 2, 2, 4, 5, 4, 4

Leider kennt sich Mark mit der Erstellung von Histogrammen überhaupt nicht aus. Deshalb braucht er 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* mehr als die Anzahl 4 zu beobachten! **(1 Punkt)**
5. Berechnen Sie aus den Daten die *Chance* mehr als die Anzahl 4 zu beobachten! **(1 Punkt)**

58. Aufgabe

(8 Punkte)

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



Visualisierung des Histogramm für kontinuierliche Daten In einem Gespräch mit ihrem Betreuer wird Paula gebeten seine Daten aus einem Leistungssteigerungsversuch mit Puten in einem Histogramm darzustellen. Aus den Boxen wummert White Lies und ihr Mund ist verklebt von Smarties. 'Herrlich', denkt Paula. In ihrem Experiment hat er die mittleren auffälligen Hautflecken erst fotografiert und dann ausgezählt. Laut ihrem Betreuer soll das Histogramm helfen, die Verteilung der die mittleren auffälligen Hautflecken zu bestimmen. Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für sie war schon immer der Perfektionismus gewesen. Ein leidiges Lied. Paula nickt im Takt von White Lies und bemerkt dabei gar nicht was die Ratte schon wieder anstellt.

Die mittleren auffälligen Hautflecken: 8.7, 10.4, 11.3, 13.5, 11.3, 8.6, 15, 10.1, 9.9, 8.1, 8.7, 12.4, 10.1, 10.8, 10.2, 10.1, 8.5, 12.6, 10, 10.9, 6.6, 6.9

Leider kennt sich Paula 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)**

59. Aufgabe

(10 Punkte)

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



Visualisierung des Scatterplots Wenn es nach Yuki ginge, wäre er schon längst fertig mit seiner Hausarbeit. Geht es aber nicht. Aus den Boxen wummert London Grammar und sein Mund ist verklebt von Reese's Peanut Butter Cups. 'Herrlich', denkt Yuki. In seiner Hausarbeit hatte er ein Feldexperiment im Oldenburger Land durchgeführt. Nach der Meinung seinem Betreuer sieht das jedoch etwas anders aus. Jetzt soll er doch noch eine explorative Datenanalyse für den Zusammenhang zwischen durchschnittlichen Anteil an Ton [%/l] und Trockengewicht [kg/ha] in Kartoffeln durchführen. Wie nervig! Yuki und die Faulheit, eine unendliche Geschichte mit kniffligen Wendungen. Da zwei kontinuierliche Variablen vorliegen, geht die explorative Datenanalyse leider nicht mit Boxplots oder Barplots. Dann was anderes. Irgendwie komisch, wenn er Matrix anmacht, dann ist das Minischwein eigentlich sofort vor dem Bildschirm und starrt hinein.

Durchschnittlichen Anteil an Ton [%/l]	Trockengewicht [kg/ha]
18.5	16.1
23.1	23.3
20.0	17.7
21.5	20.8
19.9	18.4
23.9	19.3
15.9	18.7
19.0	17.8
21.1	18.1
21.5	17.0

Leider kennt sich Yuki 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)**

60. Aufgabe

(10 Punkte)

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



Visualisierung des Mosaicplots In einen Versuch in einer Klimakammer im Oldenburger Land hatte Mark sich zum einen die Behandlung KI-gesteuert [ja/nein] und zum anderen die Messung Proteingehalt im Zielbereich [ja/nein] im Kontext von Kartoffeln angeschaut. Nun steht Mark vor dem Problem, dass er zwei kategoriale Variablen in seinem Projektbericht gemessen hat. Dazu kommt dann noch was anderes. Mark und die Unsicherheit, eine unendliche Geschichte mit kniffligen Wendungen. Da seine Betreuerin erstmal die langen Tabellen mit ja/nein in einer explorativen Datenanalyse zusammengefasst und präsentiert bekommen möchte bevor es überhaupt weitergeht, muss er jetzt eine Lösung finden. Was alles auch nicht einfacher macht. Am liebsten würde er ja was anderes machen. Mark liebt Geocaching. Darin kann er sich wirklich verlieren und immer wieder neu begeistern.

KI-gesteuert	Proteingehalt im Zielbereich	KI-gesteuert	Proteingehalt im Zielbereich
nein	nein	ja	ja
nein	nein	ja	ja
nein	ja	ja	nein
nein	nein	ja	nein
ja	ja	ja	nein
nein	nein	nein	ja
nein	nein	nein	nein
nein	ja	ja	ja
nein	ja	ja	nein
ja	nein	nein	ja
ja	ja	ja	nein
ja	ja	ja	ja
nein	ja	ja	ja
nein	nein	nein	nein
nein	ja	ja	ja
nein	ja	ja	nein

Leider kennt sich Mark 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)**

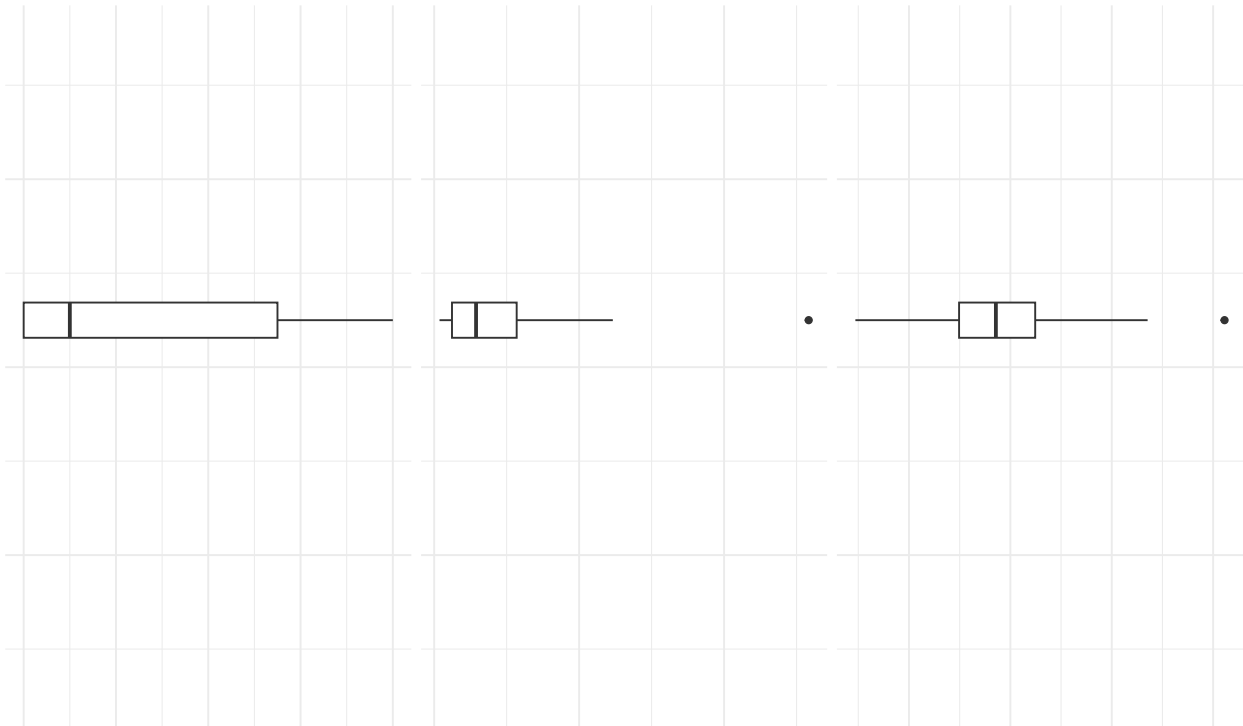
61. Aufgabe

(10 Punkte)

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



Visualisierung von Verteilungen 'Ich glaube, dass es sich hier wieder um so ein kryptisches Lernziel handelt, was nicht so gleich klar ist.', meint Paula und streichelt sanft die Ratte. Das Tier versucht dem strammen Griff zu entkommen, gibt aber auf. Tina sieht sich sehr genau die drei liegenden Boxplots an. 'Du weißt doch wie es heißt, *Frei ist, wer missfallen kann*.¹', merkt Paula nickend an. Das Ziel ist es zu verstehen, wie eine Verteilung anhand eines Boxplots bewertet werden kann. Paula und der Perfektionismus machen die Sache nicht einfacher.



Jetzt brauchen Paula und Tina 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

Teil III.

Statistisches Testen & statistische Testtheorie

62. Aufgabe

(9 Punkte)



Grundgesamtheit und experimentelle Stichprobe 'Grundlage des statistischen Testen ist das Verständnis von der Grundgesamtheit (eng. *population* oder *ground truth*) und der experimentellen Stichprobe (eng. *sample*).', liest Steffen laut aus dem Skript vor. Alex war kurz eingenickt und wird mit einem Stoß geweckt. 'Reiz dich zusammen und iss noch ein paar Oreos das hilft mir immer. Alleine komme ich hier nicht weiter.', tadelt Steffen Alex etwas zu forsch. 'War ne lange Nacht', mault Alex. Beide sollen in ihrer Abschlussarbeit einen statistischen Test interpretieren und versuchen die Grundlagen zu wiederholen. Alex war auf einem Konzert von Abba.

Leider kennen sich Steffen und Alex 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 ein aussagekräftiges Beispiel!* **(3 Punkte)**
3. Erweitern Sie das Schaubild um die Entstehung von $Pr(D|H_0)$! *Nutzen Sie hierfür als Veranschaulichung ein aussagekräftiges Beispiel!* **(3 Punkte)**

63. Aufgabe

(9 Punkte)



Das Nullritual - Die statistische Testtheorie 'Das Känguruh hat mir gerade zugezwinkert. Das macht mir Angst', bemerkt Jessica. Es reicht ja schon ein Problem. Wenn der Mangel nicht wäre, ja dann wäre wohl vieles möglich für Jessica! Aber so.. Da braucht es nicht noch Wahnvorstellungen. 'Ich glaube nicht, das Känguruhs zwinkern. Aber was Wichtigeres. Wo hast du eigentlich meine Schnapspralinen hingetan? Wir haben nur noch Schokobons von dir. Bäh!', antwortet Mark. Beide sind im Zoo und wollen sich von der statistische Testtheorie ablenken lassen. Eigentlich wollte ja Jessica stoppen wie lange Tiere pinkeln², scheiterte aber an einer Oma mit Stock, die die beiden beschimpfte.

Leider kennen sich Jessica und Mark 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)**

H_0 falsch (Unbekannte) Wahrheit Testentscheidung H_0 wahr

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 Monat 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!

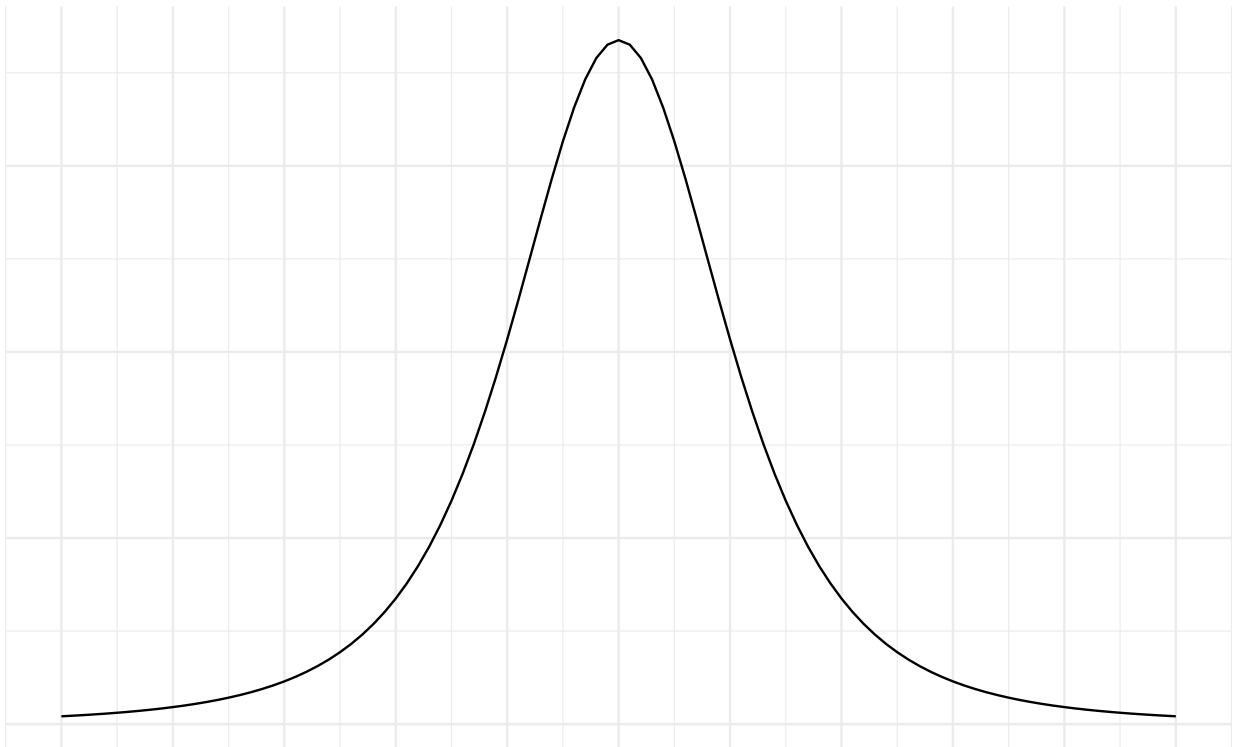


Visualisierung der Teststatistik T_D und dem p-Wert Jessica und Jonas wollten eigentlich einen Flug nach Mallorca buchen, sind jetzt aber dann doch dazu übergegangen nochmal die Aufgaben für die Statistiklausur durchzugehen. 'Kannst du mir nochmal an einer Visualisierung erklären, wie der Zusammenhang zwischen der Teststatistik aus den Daten T_D und dem p-Wert ist? Ich habe hier zig Fachbegriffe, kriege die aber nicht zusammen...', fragt Jessica. Jonas zuckt mit den Schultern. So genau hatte Jonas jetzt auch nicht aufgepasst. Da hilft dann eventuell das YouTube Video weiter. Jessica mapmft Schokobons und fragt sich, was das alles soll.

Leider kennen sich Jessica und Jonas 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 „ $A = 95\%$ “! **(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 signifikant p-Wert in die Abbildung! Begründen Sie Ihre Antwort! **(2 Punkte)**



65. Aufgabe

(10 Punkte)



Visualisierung des 95% Konfidenzintervalls 'Okay, für was war jetzt nochmal das 95% Konfidenzintervall gut?', fragt Jessica und schaut in das leere Gesicht von Mark. 'Keine Ahnung. Irgendwas mit Relevanz und Effekt oder Signifikanz. Da kannst du irgendwie was verbinden. Keine Ahnung warum', entgegnet Mark. 'Wir haben doch als Messwert *Energieverbrauch der Klimakammer* erhoben.', stellt Jessica fest. Jetzt haben beide das Problem, die möglichen 95% Konfidenzintervalle zu interpretieren.

Leider kennen sich Jessica und Mark 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 95% Konfidenzintervall mit niedriger Fallzahl n in der Stichprobe als der Rest 95% der Konfidenzintervalle
 - (b) Ein signifikantes, relevantes 99% 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 signifikantes, relevantes 95% Konfidenzintervall
 - (f) Ein nicht signifikantes, nicht relevantes 95% Konfidenzintervall



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 Dezembormorgen haben sich Nilufar und Jonas zum Lernen verabredet. Eine große Kanne Tee und Berge von Takis Blue Heat warten darauf gegessen zu werden. Nilufar 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 $KI_{1-\alpha}$?

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

Leider kennen sich Nilufar und Jonas 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 90%-igen Konfidenzintervall! Begründen Sie Ihre Visualisierung anhand der Formel des Konfidenzintervalls des t-Tests mathematisch! **(3 Punkte)**

Teil IV.

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 Mark ist im Wendland für einen Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) mit Spargel. Allein diese Tatsache ist für ihn eine Erzählung wert. Eine echte Herausforderung für ihn war schon immer die Unsicherheit gewesen. Ein leidiges Lied. Für seiner Hausarbeit musste er ein Feldexperiment mit Spargel durchführen und das sollte laut seiner Betreuerin an diesem Ort besonders gut gelingen, da man hier gut neue technische Anlagen und Behandlungen fernab der Bevölkerung testen könne. Zeugen gibt es hier jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn seine Behandlung Lichtstufen (*none* und *600lm*) und der Messwert Frischengewicht [kg/ha] sollen mit einem t-Test ausgewertet werden. Immerhin weiß er, dass sein Messwert einer Normalverteilung folgt. Hm..., was entspannendes wäre gut. Einfach mal raus um zu Reiten. Ohne Ziel und Uhr. Das ist was für Mark.

treatment	weight
ctrl	15.5
dose	9.8
dose	26.0
dose	16.5
ctrl	17.8
ctrl	19.0

Leider kennt sich Mark mit der Berechnung eines t-Tests überhaupt nicht aus. Deshalb braucht er 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 $T_{\alpha=5\%} = 2.68$ 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 Mark 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 Paula ist im Oldenburger Land für einen Versuch mit Lamas. Allein diese Tatsache ist für sie eine Erzählung wert. Wenn der Perfektionismus nicht wäre, ja dann wäre wohl vieles möglich für Paula! Aber so.. Für ihrer Hausarbeit musste sie ein Stallexperiment mit Lamas durchführen und das sollte laut ihrer Betreuerin an diesem Nichtort besonders gut gelingen. Ablenkung gibt es jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn ihre Behandlung Flüssignahrung (*ctrl* und *fIOW*) und der Messwert Gewichtszuwachs in der 1LW sollen mit einem t-Test ausgewertet werden. Immerhin weiß sie, dass ihr Messwert einer Normalverteilung folgt. Hm..., was entspannendes wäre gut. Um zu Fechten geht Paula dann später nochmal raus. Echte Entspannung.

Flüssignahrung	Gewichtszuwachs
fIOW	48.9
ctrl	20.2
ctrl	42.6
ctrl	64.1
fIOW	34.0
ctrl	51.5
fIOW	35.2
ctrl	46.5
ctrl	40.2
ctrl	32.6
fIOW	46.3
ctrl	26.2
fIOW	32.9
fIOW	54.8
ctrl	51.0
ctrl	35.4
fIOW	50.8
fIOW	42.4
fIOW	42.8
ctrl	39.8

Leider kennt sich Paula mit der Berechnung eines t-Tests überhaupt nicht aus. Deshalb braucht sie 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 $T_{\alpha=5\%} = 2.68$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *einen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann der Effekt? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Paula 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 testet einen normalverteilten Messwert (Y).', liest Mark laut. Das hilft jetzt auch nur bedingt weiter. Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. Laut seinem Betreuer ist zwar ihm Messwert Fettgehalt [%/kg] normalverteilt, aber wie rechnet er jetzt einen t-Test? Für seinen Projektbericht musste er ein Kreuzungsexperiment mit Milchvieh in der Uckermark durchführen. Als wäre das nicht schon anstrengend genug gewesen. Jetzt soll er auch noch testen, ob die Behandlung Genotypen (AA und BB) ein signifikantes Ergebnis liefert. Hm..., was spannendes wäre gut. Schon dutzende Male gesehen: Columbo. Aber immer noch großartig zusammen mit Marzipankugeln.

Genotypen	Fettgehalt
BB	12.4
AA	35.7
AA	29.3
AA	17.0
BB	22.3
AA	39.8
BB	29.1
BB	21.6
AA	25.0
AA	24.9
AA	41.3
AA	26.3
BB	26.9
BB	33.1
BB	51.7
AA	32.1
BB	35.9
BB	33.6
BB	40.9
AA	39.7
BB	21.1
AA	18.4

Leider kennt sich Mark 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 $T_{\alpha=5\%} = 2.04$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie das 95% 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 Mark 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 Alles voll mit Spargel. Das haben Steffen und Mark gemeinsam in einer Hausarbeit gemacht! Worum ging es aber konkret? Beide haben als ein normalverteiltes Outcome (Y) Trockengewicht [kg/ha] von Spargel bestimmt. Die Daten haben beide zusammen in einem Gewächshaus-experiment erhoben. In dem Experiment ging es um eine vorher/nachher Untersuchung an den gleichen Spargel. Als Behandlung wurde Beschattung (7d und 14d) eingesetzt. Nach der Meinung der Betreuerin muss hier ein gepaarter t-Test gerechnet werden. Steffen schaut nachdenklich zu Mark. Steffen und die Romantik, eine unendliche Geschichte mit kniffligen Wendungen.. Steffen denkt derweil angestrengt an Marzipankugeln. Im Hintergrund wummert Taylor Swift.

ID	treatment	freshmatter
7	7d	5.5
3	14d	31.6
1	14d	34.5
1	7d	3.7
8	7d	34.0
5	14d	27.0
6	14d	31.5
10	7d	13.4
7	14d	30.0
9	7d	5.0
6	7d	22.1
4	14d	33.9
2	7d	8.7
5	7d	39.6
11	7d	21.0
3	7d	12.9
2	14d	28.5
4	7d	39.3

Leider kennen sich Steffen und Mark 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 $T_{\alpha=5\%} = 2.68$ 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 Steffen 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 Almería. Spanien. Sonne und Strand. Paula und Mark haben ihren gemeinsamen Auslandsaufenthalt sichtlich genossen. Dann hatte sich auch noch angeboten ihre Abschlussarbeit gemeinsam in Almería durchzuführen. Es hätte sogar noch besser funktioniert, wenn Tina nicht die Wut ein paar Mal im Weg gestanden hätte und Mark nicht das Problem gehabt hätte der Perfektionismus zu händeln. Nun müssen jetzt alle Daten in R ausgewertet werden, da R international der Standard in der Datenauswertung ist und die Betreuer in Spanien nur R können. Während beide Tina Oliven mit Katjes füttern, hoffen Paula und Mark mehr Informationen von Tina über die seltsame R Ausgabe des t-Tests. Immerhin erinnern beide sich an die Behandlung Substrattypen (*torf* und *70p30n*) und das es um Lauch ging. Im Hintergrund wummert Tocotronic und Fotos zeigen Tina mit dem Hobby Astronomie.

```
##
## Two Sample t-test
##
## data: Chlorophyllgehalt by Substrattypen
## t = 1.4736, df = 19, p-value = 0.157
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -2.250984 12.960075
## sample estimates:
## mean in group torf mean in group 70p30n
## 42.05455 36.70000
```

Helfen Sie Tina bei der Interpretation des t-Tests! Sonst geht es auch für Paula und Mark 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 $T_{\alpha=5\%} = |2.09|$ 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 'Programmieren ist wie eine Sprache lernen. Man muss es nur machen, dann wird man mit der Zeit immer besser!', gibt Steffen zwinkernd zu Protokoll. Ein paar Mal hat sie schon die Romantik gehindert weiterzumachen. Das hilft jetzt Paula und Jessica nur bedingt, da beide jetzt die R Ausgabe interpretieren müssen und nicht vor drei Wochen, wo noch Zeit gewesen wäre. Beide mampfen konzentriert Smarties und Schokobons in sich hinein. Die beiden hatten in der Uckermark einen Versuch mit Puten in einem Kreuzungsexperiment durchgeführt. Das war schon anstrengend genug! 'Wir haben Protein/Fettrate [%/kg] gemessen, vielleicht hilft das ja...', meint Jessica leicht genervt. Alle starren auf die R Ausgabe des t-Tests. Im Hintergrund wummert Taylor Swift und man versteht kaum sein eigenes Wort. Jessica hofft, dass die Schlange von Steffen beruhigend wirkt.

```
##
## Two Sample t-test
##
## data: Protein/Fettrate by Lüftungssystem
## t = 0.8988, df = 17, p-value = 0.3813
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -4.358017 10.826906
## sample estimates:
## mean in group keins mean in group vorhanden
## 37.84444 34.61000
```

Helfen Sie Steffen bei der Interpretation des t-Tests! Sonst geht es auch für Paula und Jessica 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 Almería. Spanien. Sonne und Strand. Jonas und Tina haben ihren gemeinsamen Auslandsaufenthalt sichtlich genossen. Dann hatte sich auch noch angeboten ihre Abschlussarbeit gemeinsam in Almería durchzuführen. Es hätte sogar noch besser funktioniert, wenn Nilufar nicht die Erwartung ein paar Mal im Weg gestanden hätte und Tina nicht das Problem gehabt hätte die Erschöpfung zu händeln. Nun müssen jetzt alle Daten in R ausgewertet werden, da R international der Standard in der Datenauswertung ist und die Betreuer in Spanien nur R können. Während beide Nilufar Oliven mit Takis Blue Heat füttern, hoffen Jonas und Tina mehr Informationen von Nilufar über die seltsame R Ausgabe des t-Tests. Immerhin erinnern beide sich an die Behandlung Genotypen (AA und BB) und das es um Lauch ging. Im Hintergrund wummert Deichkind und Fotos zeigen Nilufar mit dem Hobby Hip Hop.

```
##
## Two Sample t-test
##
## data: Trockengewicht by Genotypen
## t = 2.532, df = 19, p-value = 0.02032
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
##  1.242567 13.091978
## sample estimates:
## mean in group AA mean in group BB
##      39.34000      32.17273
```

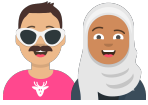
Helfen Sie Nilufar bei der Interpretation des t-Tests! Sonst geht es auch für Jonas und Tina nicht weiter.

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 Alles voll mit Kartoffeln. Aber das haben Alex und Nilufar eben gemeinsam in einer Abschlussarbeit gemacht! Worum ging es aber konkret? Beide haben als einen normalverteilten Endpunkt (Y) Frischegewicht [kg/ha] von Kartoffeln bestimmt. Die Daten haben beide zusammen in einem Versuch in einer Klimakammer erhoben. In dem Experiment ging es um eine vorher/nachher Untersuchung an den gleichen Kartoffeln. Als Behandlung wurde Ausgeizen (*ctrl* und *28d*) eingesetzt. Nach der Meinung der Betreuerin muss hier ein gepaarter t-Test gerechnet werden. Leider kennen sich beide nicht sehr gut in R aus. Alex hat einiges an Gummibärchen geholt, so dass beide die Zeit gut durchbringen werden. Dann geht Alex nochmal zum Sport. Einfach mal raus um zu Laufen. Ohne Ziel und Uhr. Das ist was für Alex.

```
##
## Paired t-test
##
## data: Frischegewicht by Ausgeizen
## t = -2.4381, df = 8, p-value = 0.04068
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -11.0694889 -0.3082888
## sample estimates:
## mean difference
## -5.688889
```

Jetzt brauchen Alex und Nilufar 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 V.

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?', Paula 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. Paula hatte sich in ein Stallexperiment verschiedene Zandern angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Ernährungszusatz (*ctrl*, *fedX* und *getIt*) und dem Messwert Schlachtgewicht [kg] gibt. Später wird noch Star Trek geguckt. Nilufar befürwortet das!

Ernährungszusatz	Schlachtgewicht
fedX	44
getIt	30
fedX	43
ctrl	45
ctrl	45
ctrl	43
ctrl	45
ctrl	46
fedX	44
getIt	28
ctrl	45
getIt	34
fedX	46
getIt	30
fedX	45
getIt	29
fedX	45
ctrl	46

Leider kennen sich Paula 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!
 - Den globalen Mittelwert β_0 **(1 Punkt)**
 - Die Mittelwerte der einzelnen Behandlungsstufen **(1 Punkt)**
 - Die Mittelwertsdifferenz der einzelnen Behandlungsstufen mit β_{ctrl} , β_{fedX} und β_{getIt} **(1 Punkt)**
 - Die Residuen oder Fehler mit 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.', Alex schaut Steffen fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Da hilft die Schlange von Steffen auch nur bedingt. Steffen tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren im Oldenburger Land um ein Gewächshausexperiment mit Erbsen durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Bewässerungstypen (*ctrl*, *low*, *mid* und *high*) und dem Messwert Proteingehalt [g/kg] gibt. Nachher wollen sich beide noch mit dem Hobby Klemmbau- steine von Steffen beschäftigen. Kennt Alex noch nicht, klingt aber interessant.

Leider kennen sich Alex und Steffen mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Des- halb brauchen beide bei der Erstellung Ihre Hilfe, die Schlange 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)
Bewässerungstypen	3	351.68			
error	24				
Total	27	1272.68			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.01$ 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 Nilufar und Paula schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten mit einer einfaktoriellen ANOVA auswerten damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse in den multiplen Gruppenvergleichen zu erwarten sind. Deshalb erstmal Smarties mampfen, die Paula mitgebracht hat. Nun möchte erstmal ihre Betreuung der Arbeit eine ANOVA Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Nilufar schaut Paula sehen erstmal gar nichts. Die beiden waren im Oldenburger Land um ein Feldexperiment mit Brokkoli durchzuführen. Dabei haben Nilufar und Paula den Messwert Trockengewicht [kg/ha] unter der Behandlung Lüftungssystemen und Folientunneln (*ctrl*, *storm*, *thunder* und *tornado*) ermittelt. Später wollen die beiden dann noch raus um zu Fechten.

Leider kennen sich Nilufar und Paula 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)
Lüftungssystemen	3	3906.54			
Error	26	471.46			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 2.98$ 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)

Lüftungssystemen	Fallzahl (n)	Mittelwert	Standardabweichung
ctrl	8	6.88	4.76
storm	8	8.88	1.73
thunder	6	8.83	2.23
tornado	8	33.88	6.17

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 Mark und hebt die Augenbraue. Mark hatte sich ein Gewächshausexperiment mit Erbsen angeschaut. Als wäre das nicht alles schon schwer genug. Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. Dabei ging es beim Experiment herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lichtstufen (*none*, *200lm*, *400lm* und *600lm*) und dem Messwert Frischegewicht [kg/ha] gibt. Nun möchte sein Betreuer seiner Hausarbeit erstmal eine ANOVA sehen und die Ergebnisse präsentiert bekommen. Und eigentlich will er ja was anderes... Schon dutzende Male gesehen: Columbo. Aber immer noch großartig zusammen mit Marzipankugeln.

```
## Analysis of Variance Table
##
## Response: Frischegewicht
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Lichtstufen  3 5495.1 1831.69   72.254 5.072e-13
## Residuals   27  684.5    25.35
```

Leider kennen sich Mark mit Berechnung einer einfaktoriellen 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. 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 Wie absolut ärgerlich. Jetzt stellt sich tatsächlich heraus, dass seinem Betreuer keine Ahnung von der zweifaktoriellen ANOVA hat. Woher soll Steffen jetzt das Wissen nehmen? Steffen mampft aus Frust noch eine Handvoll Oreos. Immerhin muss er ja noch mit seiner Abschlussarbeit dieses Jahr fertig werden. In ein Kreuzungsexperiment hatte er Schweinen mit der Behandlung Ernährungszusatz (*ctrl*, *fedX* und *getIt*) sowie der Behandlung Genotypen (*AA* und *BB*) im Oldenburger Land untersucht. Es wurde als Messwert Schlachtgewicht [kg] bestimmt. Jetzt muss er erstmal die zweifaktorielle ANOVA verstehen. Und eigentlich wollte Steffen doch noch zum Sport! Steffen will später nochmal raus um zu Ringen. Druck ablassen, dass muss er auch.

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)
Ernährungszusatz	3	315.8			
Genotypen	1	260.72			
Ernährungszusatz:Genotypen	3	590.25			
Error	18	246.42			

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

	$F_{\alpha=5\%}$
Ernährungszusatz	4.26
Genotypen	3.40
Ernährungszusatz:Genotypen	5.23

5. Was bedeutet ein signifikantes Ergebnis in einer zweifaktoriellen ANOVA? (2 Punkte)
6. Was sagt der Term *Ernährungszusatz:Genotypen* 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 Es ist schon kurz nach fünf und Jessica wird langsam nervös. Jessica wollte heute Abend noch ihre E-Sport Qualifikation schauen. Stattdessen versucht ihre Betreuerin die Ausgabe der zweifaktoriellen ANOVA zu visualisieren und zu überprüfen, ob es mit der Visualisierung der Daten als Boxplots zusammenpasst. Jessica hatte im Oldenburger Land einen Versuch in einer Klimakammer mit Erbsen durchgeführt. Es gab dabei zwei Behandlungen. Einmal Bewässerungstypen (*ctrl*, *low*, *mid* und *high*) sowie als zweite Behandlung Lichtstufen (*none*, und *600lm*). Gemessen wurde der Messwert (*Y*) Proteingehalt [g/kg]. So kompliziert kann das jetzt doch nicht sein! Eigentlich wollte Jessica nachher noch zum Sport. Jessica will später nochmal raus um Rad zu fahren. Druck ablassen, dass muss sie auch.

```
## Analysis of Variance Table
##
## Response: Proteingehalt
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Bewässerungstypen      2 324.59  162.297  23.2371 1.03e-05
## Lichtstufen            1   0.11    0.108   0.0154 0.902570
## Bewässerungstypen:Lichtstufen  2 216.47  108.233  15.4964 0.000122
## Residuals             18 125.72    6.984
```

Leider kennt sich Jessica 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 Die Spinne dreht durch und verwüstet Tinas Palme zu kleinen Schnetzeln. Aber dafür hat sie jetzt keine Zeit. Tina muss verstehen wie die Formeln der ANOVA und des t-Tests miteinander zusammen hängen und was das verbindene Konzept ist. Tina dreht Toco-tronic lauter, damit die Spinne sie nicht mehr stört. Die Palme leidet still. Was hat Tina eigentlich gemacht? In ein Stallexperiment wurden Lamas mit der Behandlung Flüssignahrung (*ctrl*, *superIn* und *flOw*) sowie der Behandlung Lüftungssystem (*keins* und *thunder*) untersucht. Das hilft der Palme auch nicht mehr. Aber das ist nicht das einzige Problem von Tina. Tina und die Wut, eine unendliche Geschichte mit kniffligen Wendungen.

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 Tina mit dem Zusammenhang zwischen der ANOVA und dem t-Test nicht aus. Deshalb braucht sie 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 In einen Versuch in einer Klimakammer wurden Erdbeeren mit der Behandlung Düngestufen (*ctrl*, *low*, *mid* und *high*) sowie der Behandlung Substrattypen (*torf*, *70p30n*) untersucht. Es wurde als Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] bestimmt. Jetzt startet Jessica mit auf die  Ausgabe einer zweifaktoriellen ANOVA. Leider startet ihr Betreuer in der gleichen Art Jessica zurück an. Es liegt anscheinend eine signifikante Interaktion vor? '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 ihre Hausarbeit. Jessica hätte doch nichts mit Erdbeeren machen sollen. Erdbeeren – was soll das auch bedeutendes sein? Eigentlich wollte Jessica nachher noch einen Film schauen. Das Verrückte ist, dass die Hündin Herr der Ringe wirklich liebt. Das ist Jessica sehr recht, denn sie braucht Entspannung.

Leider kennen sich Jessica und ihr Betreuer mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion. Deshalb braucht sie bei der Erstellung Ihre Hilfe, sonst wird es heute Abend mit seinem Hobby Warhammer 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 'Mit der einfaktoriellen ANOVA lassen sich flott die Gruppen in einer Behandlungen vergleichen, wenn wir normalverteilte Daten und Varianzhomogenität vorliegen haben!', ihr Betreuer scheint die einfaktoriellen ANOVA zu verstehen. Warum jetzt sie jetzt nochmal alles wiederkauen muss, wird Yuki echt nicht so klar. Wenn es doch so klar ist? 'Wir haben jetzt bei der ANOVA einen p-Wert mit 0.071 raus sowie eine F-Statistik F_D mit 1.61 berechnet. Nach den Boxplots müsste sich eigentlich ein Unterschied zwischen *mid* und *ctrl* ergeben. Der Unterschied ist in {emmeans} auch signifikant mit einem p-Wert von 0.045. Wie kann das sein?', fragt Yuki etwas provokant und dreht London Grammar leiser. Yuki war in der Uckermark und hatte dort einen Versuch in einer Klimakammer mit Maiss durchgeführt. Die Komune wo sie untergekommen war, war cool gewesen. Dort gab es selbstgemachte Reese's Peanut Butter Cups aus Vollkorn! Nur jetzt muss eben das Experiment fertig ausgewertet werden. Yuki hatte eine Behandlungen Bewässerungstypen (*ctrl*, *low*, *mid* und *high*) auf Maiss angewendet. Gemessen wurde der Messwert (Y) Trockengewicht [kg/ha]. Dabei wurden die Daten D erhoben. Jetzt muss das hier zu einem Ende kommen! Yuki hat schon genug Probleme. Wenn die Faulheit nicht wäre, dann wäre es einfacher.

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 ihr Betreuer 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 VI.

Multiple Gruppenvergleiche

84. Aufgabe

(12 Punkte)

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



Adjustierung multipler Vergleiche In einen Versuch in einer Klimakammer mit Spargel wurde die Behandlung Genotypen (00, AA, AB und BB) gegen die Ergebnisse einer früheren Studie von Almar et al. (2012) verglichen. Im Rahmen des Experiments haben Jessica und Alex verschiedene Student t-Tests für den Mittelwertsvergleich für den Messwert Frischegewicht [kg/ha] gerechnet. Es ergab sich dann die folgende Tabelle der rohen p-Werte für die Vergleiche zu Almar et al. (2012). Jetzt sollen die beiden einmal schauen, was in den Daten so drin ist.

Rohen p-Werte	Adjustierte p-Werte	Nullhypothese ablehnen?
0.340		
0.230		
0.030		
0.001		

Leider kennen sich Jessica und Alex 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) Paula betrachtet in sich gekehrt die Poster vor dem Büro von ihrem Betreuer. Viele der explorativen Abbildungen sagen ihr etwas. Die Barplots und die Boxplots könnte sie dann schon nachbauen. Das macht sie dann zuversichtlich die Abschlussarbeit auch hinzukriegen. Etwas komischer sind die seltsamen Buchstaben über den Barplots. Paula betrachtet ein Poster das sich mit Puten beschäftigt. Genotypen (00, AA, AB und BB) und Schlachtgewicht [kg] wurden dort bestimmt. So richtig schlau, wird sie daraus nicht.

Behandlung	Compact letter display
00	a
AA	b
AB	a
BB	b

Leider kennen sich Paula mit dem *Compact letter display (CLD)* überhaupt nicht aus. Deshalb braucht sie 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 Jessica hatte in die Projektbericht ein Kreuzungsexperiment durchgeführt. Soweit so gut. Dabei hat sie sich mit Lamas beschäftigt. Angeblich der neueste heiße Kram... aber das ist wiederum was anderes. So richtig mitgenommen hat Jessica das Thema dann doch nicht. Hat sie sich doch mit Bestandsdichte (*effizient*, *standard*, *eng* und *kontakt*) und Protein/Fettrate [%/kg] schon eine Menge an Daten angeschaut. Nach ihr Betreuer soll sie nun ein CLD bestimmen. Weder weiß sie was ein CLD ist, noch war ihr erster Gedanke mit Köln und die LGBTQ Community richtig... Als erstes solle sie die Gruppen nach absteigender Effektstärke sortieren. Was immer das jetzt bringen soll.

Bestandsdichte	Fallzahl (n)	Mittelwert	Standardabweichung
effizient	9	5.17	2.40
standard	8	11.65	2.64
eng	8	5.15	2.16
kontakt	9	14.27	4.12

Leider kennen sich Jessica mit dem *Compact letter display (CLD)* überhaupt nicht aus. Deshalb braucht sie 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 Jessica und Alex! **(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 Mark aus und rollt entnervt mit seinen Augen. Mark 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. Mark findet einfach keine Möglichkeit ein CLD in SPSS zu erhalten. Aber sein Betreuer möchte unbedingt ein CLD. Sonst wird es mit der Abgabe nichts. Dabei hatte er schon wirklich eine Menge gemacht! Mark hatte sich zwei Variablen mit Genotypen (00, AA, AB und BB) und Schlachtgewicht [kg] in einen Leistungssteigerungsversuch mit Hühnern angeschaut. Wo kriegt er jetzt ein CLD her? Dann eben per Hand aus der Matrix der p -Wert. Mark stöhnt...

	00	AA	AB	BB
00	1.0000000	0.0030905	0.4299008	0.0069354
AA	0.0030905	1.0000000	0.0186585	0.8142810
AB	0.4299008	0.0186585	1.0000000	0.0371324
BB	0.0069354	0.8142810	0.0371324	1.0000000

Leider kennen sich Mark 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 Mark und Alex! **(2 Punkte)**

Teil VII.

Der Chi-Quadrat-Test & Der diagnostische Test

88. Aufgabe

(10 Punkte)

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



Den Chi-Quadrat-Test berechnen Am Ende war es für Jessica in ihrer Abschlussarbeit dann doch kein normalverteiltes Outcome. Das was jetzt etwas doof, da er sich auf eine ANOVA gefreut hatte. Dann noch schnell David Bowie auf das Ohr und los gehts. Prinzipiell ginge das auch irgendwie, aber nun möchte ihr Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Jessica hatte sich in ein Stallexperiment $n = 142$ Beobachtungen von Zandern angeschaut. Dabei hat sie als Behandlung *Automatische Fütterung* [ja/nein] bestimmt und zum anderen die Variable *Schlachtgewicht im Zielbereich* [ja/nein] ermittelt. Jetzt muss Jessica mal schauen, wie sie das jetzt rechnet. Eigentlich wollte Jessica nachher noch zum Sport. Jessica will später nochmal raus um Rad zu fahren. Druck ablassen, dass muss sie auch.

	56	11	
	23	52	

Leider kennt sich Jessica 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 $\chi^2_{\alpha=5\%} = 4.56$! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Skizzieren Sie in einer Abbildung die χ^2 -Verteilung, wenn die H_0 wahr ist! Ergänzen Sie $\chi^2_{\alpha=5\%}$ und χ^2_D in der Abbildung! Beachten Sie folgenden Informationen zur χ^2 -Verteilung. Die χ^2 -Verteilung hat ein Maxima bei $\chi^2 = 6$ sowie ein Minima bei $\chi^2 = 11$. **(2 Punkte)**

89. Aufgabe

(10 Punkte)

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



Den Chi-Quadrat-Test in einem Fragebogen Mark hatte sich gleich von Beginn an in seinem Projektbericht für eine Umfrage im Marketing interessiert. Jetzt geht es um den Haupt- und Nebenerwerb von Erlebnishöfen in Norddeutschland. Viele Höfe haben angefangen auch Großkatzen zu halten, damit mehr Kunden auf die Höfe kommen. Für den Verband der Großkatzenbesitzer e.V. möchte er nun einen Fragebogen zur Zukunftsfähigkeit Schritt für Schritt auswerten. Dabei teilt er zuerst die Antwortenden in die beiden Gruppen 'Höfe mit Großkatzen [ja]' und 'Höfe mit Großkatzen [nein]' ein. Daraufhin möchte er für folgende Frage *f5verband* einmal auswerten, ob es einen Unterschied zwischen den beiden Höfen mit oder ohne Großkatzen gibt.

Sehen Sie die Notwendigkeit von genetisch veränderten Großkatzen zur Steigerung der Attraktivität des Erlebnishofes?

Mark kratzt sich an seinem Kopf. Wie soll man eine Tabelle mit so vielen Zahlen sinnvoll auswerten? Schnell noch ein paar Marzipankugeln einwerfen und los gehts!

<i>f5verband</i>	trifft gar nicht zu	trifft nicht zu	weder noch	trifft zu	trifft voll zu	
ja	3	8	12	21	11	
nein	1	11	25	9	4	

Leider kennt sich Mark mit der Berechnung eines χ^2 -Test auf einer Frage in einem Fragebogen überhaupt nicht aus. Deshalb braucht er 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. Berechnen Sie die Teststatistik χ_D^2 eines Chi-Quadrat-Test! Ignorieren Sie Zellbelegungen kleiner gleich fünf in der Berechnung von χ_D^2 ! **(2 Punkte)**
4. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem $\chi_{\alpha=5\%}^2 = 23.83$! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Visualisieren Sie die 2x5 Kreuztabelle ohne die Berücksichtigung der Antwortkategorie 'weder noch'! **(2 Punkte)**
6. Berechnen Sie den Effektschätzer *Cramers V* auf der 2x5 Kreuztabelle! **(1 Punkt)**
7. Interpretieren Sie den berechneten Effektschätzer *Cramers V* unter Berücksichtigung der Fragestellung! **(1 Punkt)**

90. Aufgabe

(10 Punkte)

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



Den Chi-Quadrat-Test mit Effektmaß berechnen Jonas hat sich ein Herz gefasst und war für seiner Hausarbeit in die Niederlande gegangen. Das war eine super Zeit in der er viel gelernt hat. Klar gab es auch die ein oder andere Besonderheit, aber das gehört hier eher nicht hin. Dann noch schnell Iron Maiden auf das Ohr und los gehts. Jonas ist schon eine ganze Zeit im Büro, da seine Betreuerin möchte, dass er jetzt auf seinen Daten mit $n = 188$ Beobachtungen von Lamas einen χ^2 -Test rechnet. Das ginge, da er als Behandlung *Ökologisch* [ja/nein] bestimmt und zum anderen die Variable *Protein/Fettrate im Zielbereich* [ja/nein] ermittelt hat. Wie genau, das ist jetzt eine andere Frage. Eigentlich wollte Jonas nachher noch einen Film schauen. Irgendwie komisch, wenn er Mission Impossible anmacht, dann ist das Meerschweinchen eigentlich sofort vor dem Bildschirm und starrt hinein.

	56	51	
	43	38	

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. Ergänzen Sie die Tabelle um die fehlenden Informationen! **(1 Punkt)**
3. Berechnen Sie die Teststatistik eines Chi-Quadrat-Test! **(2 Punkte)**
4. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem $\chi^2_{\alpha=5\%} = 5.61$! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effektschätzer *Cramers V* auf der 2x2 Kreuztabelle! **(1 Punkt)**
6. Welchen Wertebereich kann der Effektschätzer *Cramers V* annehmen? Wann liegt kein Effekt und wann ein starker Effekt vor? **(2 Punkte)**
7. Interpretieren Sie den berechneten Effektschätzer *Cramers V* unter Berücksichtigung der Fragestellung! **(1 Punkt)**

91. Aufgabe

(10 Punkte)

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



Den Chi-Quadrat-Test konzeptionell verstehen Am Ende hätte Alex dann doch einen normalverteilten Endpunkt in seiner Abschlussarbeit nehmen sollen. Dann noch schnell Gummibärchen zur Stärkung und los gehts. Vor ihm liegen jetzt die Daten von zwei Variablen als Kategorien oder wie es in R so schön heißt, als Faktoren. Aber immerhin, hofft er das was bei den Daten rausgekommen ist. Gezählt hat Alex einiges mit $n = 122$ Beobachtungen von Milchvieh. Zum einen hat er als Behandlung *Automatische Fütterung* [ja/nein] bestimmt und zum anderen die Variable *Schlachtgewicht im Zielbereich* [ja/nein] ermittelt. Nun möchte seine Betreuerin gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Am Ende des Tages möchte er dann noch sein Hobby Starcraft genießen. Das muss auch mal sein!

			64
			58
	61	61	122

Leider kennt sich Alex 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. 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)**

92. Aufgabe

(10 Punkte)

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



Der Chi-Quadrat-Test in R Mark hat sich ein Herz gefasst und war für seinem Projektbericht in die Niederlande gegangen. Das war eine super Zeit in der er viel gelernt hat. Klar gab es auch die ein oder andere Besonderheit, aber das gehört hier eher nicht hin. Dann noch schnell Columbo starten und los gehts mit der Kraft von Marzipankugeln. Mark ist schon eine ganze Zeit im Büro, da seine Betreuerin möchte, dass er jetzt auf seinen Daten mit $n = 122$ Beobachtungen von Lauch einen χ^2 -Test rechnet. Das ginge, da er als Behandlung *KI-gesteuert* [*ja/nein*] bestimmt und zum anderen die Variable *Trockengewicht über Zielwert* [*ja/nein*] ermittelt hat. Wie genau, das ist jetzt eine andere Frage. Nach seinem Experiment erhielt er folgende 2x2 Kreuztabelle aus seinen erhobenen Daten.

```
##               KI-gesteuert
## Trockengewicht über Zielwert ja nein
##               ja      6    12
##               nein   10     7
```

Dann rechnete Mark 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: Trockengewicht 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 Mark 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)**

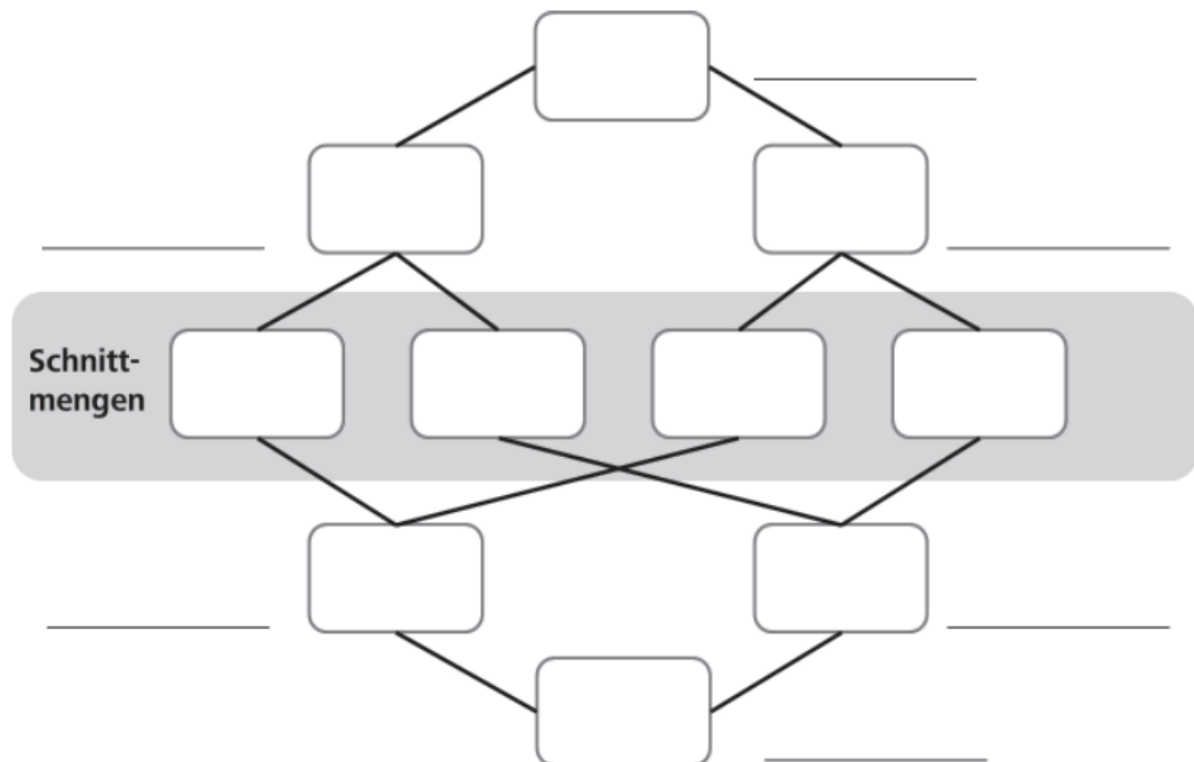
93. Aufgabe

(11 Punkte)

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




Den diagnostische Test am Doppelbaum berechnen Jonas liest laut vor. 'Die Prävalenz von Klauenseuche bei Brokkoli wird mit 2% angenommen. In 80% 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 Schorf an 2000 Brokkoli mit einem diagnostischen Test durch.' Jessica klappt die Kinnlade runter. In der Stille dudelt Iron Maiden. Jonas schaut kompetent und schmeißt sich mit offenen Mund Schokobons an den Kopf vorbei.



Leider kennen sich Jonas und Jessica 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)**

(11 Punkte)



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 VIII.

Lineare Regression & Korrelation

95. 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 Yuki. 'Ich sehe nur eine Zahlen und keine Punkte. Wie soll ich da denn jetzt eine Gerade durchzeichnen?', fragt Jonas. Yuki atmet schwer ein. Die beiden hatten ein Feldexperiment in der Uckermark mit Lauch durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlichen Niederschlag [ml/w] und Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Jetzt will die Betreuung von den beiden einmal die Visualisierung der Daten und auch gleich noch die lineare Regression gerechnet bekommen.

Durchschnittlichen Niederschlag [ml/w]	Chlorophyllgehalt (SPAD-502Plus) [SPAD]
22.9	21.3
30.1	36.8
26.8	32.5
28.8	28.7
30.5	33.5
26.3	29.6
24.7	27.7
24.9	32.1
22.4	25.4

Leider kennen sich Yuki 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)**

96. 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 Mark. Alex schaut fragend zurück. 'Keine Ahnung... das ist jetzt jedenfalls keine Abbildung von irgendwas sondern Zahlen in einer Tabelle...', antwortet Alex leicht angespannt. Die beiden hatten ein Kreuzungsexperiment im Oldenburger Land mit Milchvieh durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: mittlere Anzahl an weißen Blutkörperchen [LEU/ml] und Schlachtgewicht [kg]. Jetzt haben die beiden eigentlich alles zusammen. *Eigentlich...*, denn mit der R Ausgabe haben beide jetzt ein Problem.

term	estimate	std.error	t statistic	p-value
(Intercept)	5.84	1.66		
Mittlere Anzahl	-0.30	0.16		

Leider kennen sich Mark und Alex 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. 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)**

97. Aufgabe

(11 Punkte)

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



Interpretation der Ergebnisse einer linearen Regression in R 'Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen!', ruft Steffen. 'Ich sehe nur Kauderwelsch und keine Punkte. Wie soll ich da denn jetzt eine Gerade durchzeichnen? Und warum überhaupt? War das unsere Fragestellung?', fragt Jonas. Steffen atmet schwer ein und starrt auf die R Ausgabe der Funktion `lm()`. Die beiden hatten ein Feldexperiment im Emsland mit Erdbeeren durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Anteil an Ton [%/l] und Frischegewicht [kg/ha]. Jetzt will die Betreuung von den beiden die Interpretierung der Daten in Form einer linearen Regression gerechnet bekommen. Das haben beide in R gemacht, aber wie soll das jetzt gehen? Das mit der Interpretation?

```
##
## Call:
## Frischegewicht ~ Durchschnittlicher_Anteil
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.97052 -0.67736  0.07321  0.79656  2.18951
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.2370      1.4442   1.549  0.12968
## Durchschnittlicher_Anteil  0.4666      0.1405   3.320  0.00199
##
## Residual standard error: 1.174 on 38 degrees of freedom
## Multiple R-squared:  0.2249, Adjusted R-squared:  0.2045
## F-statistic: 11.02 on 1 and 38 DF,  p-value: 0.001993
```

Leider kennen sich Steffen und Jonas 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)**

98. Aufgabe

(10 Punkte)

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



Interpretation der Ergebnisse einer Korrelationsanalyse in R 'Ich glaube ich bringe da was durcheinander. Ich möchte eine Gerade durch die Punkte zeichnen oder doch eine Korrelation berechnen?', merkt Tina laut an. 'Ich sehe keine Punkte... das ist doch eine Ausgabe in R. Überhaupt, darum geht es doch gar nicht in meinem Versuch. Ich wollte doch keine Gerade zeichnen?.', antwortet Tina sich sichtlich übernächtigt selber. Das Verrückte ist, dass die Spinne Indiana Jones wirklich liebt. Das ist Tina sehr recht, denn sie braucht Entspannung. Die Nacht war zu lang und überhaupt. Wenn die Wut nicht wäre, ja dann wäre wohl vieles möglich für Tina! Aber so.. Tina hatte ein Feldexperiment in der Uckermark mit Erbsen durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche Regenwurmdichte [Anzahl/l] und Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Jetzt will sie erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der R Ausgabe möglich sein.

```
##  
## Pearson's correlation  
##  
## data: Durchschnittliche Regenwurmdichte and Chlorophyllgehalt  
## t = -8.8093, df = 8, p-value = 2.168e-05  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.9889147 -0.8052156  
## sample estimates:  
## cor  
## -0.9521276
```

Leider kennt sich Tina 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)**

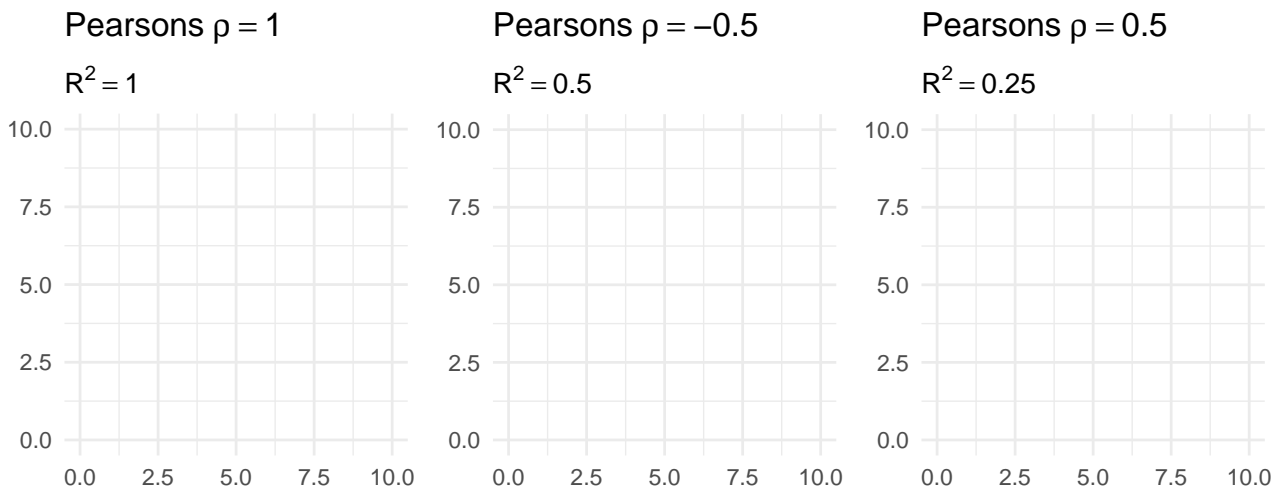
99. Aufgabe

(12 Punkte)

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



Visualisierung der Korrelation und des Bestimmtheitsmaßes 'Hm..., drei leere Abbildungen. Was soll ich da jetzt machen?', fragt sich Tina. Tina kennt sich nur begrenzt bis gar nicht mit der linearen Regression und Korrelation aus. Dafür mit etwas anderem. Am Ende dann doch besser Astronomie. Wunderbar. Eine echte Ablenkung für Tina. Aber das hilft hier auch nur so halb, die Aufgabe zu lösen und mehr über den Korrelationskoeffizienten zu erfahren.



Visualisierung der Korrelation und des Bestimmtheitsmaßes Leider kennt sich Tina 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)**

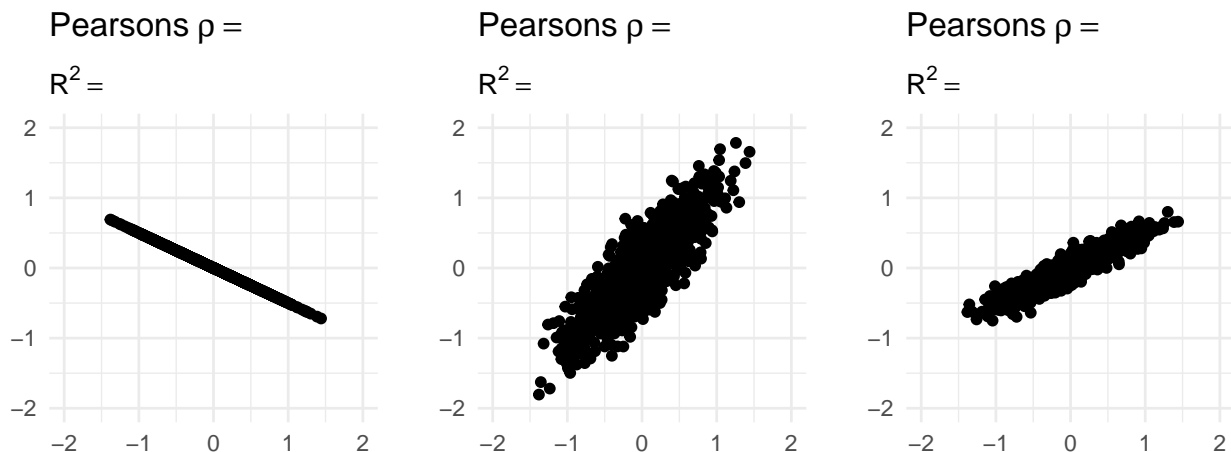
100. Aufgabe

(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 Nilufar. Es ist schon spät. Und das hat einen Grund. Hm, lecker Takis Blue Heat und dazu dann im Hintergrund Star Trek laufen lassen. . Nilufar überlegt, aber ihre Gedanken sind etwas zäh. 'Was soll das hier alles bedeuten?', fragt sich Nilufar. Irgendwie ist ihr nicht klar wie sie ρ -Werte oder R^2 -Werte abschätzen soll. Alles nicht so einfach. Wenn die Erwartung nicht wäre, ja dann wäre wohl vieles möglich für Nilufar! Aber so..



Leider kennt sich Nilufar mit der Korrelationsanalyse und der linearen Regression überhaupt nicht aus. Deshalb braucht sie 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)**

101. 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. Schon dutzende Male gesehen: Alien. Aber immer noch großartig zusammen mit Gummibärchen. Aber sowas hilft ihm natürlich hier nicht. Da schmeißt sich Alex noch ein paar Gummibärchen in den Mund und kaut los.

Durchschnittlicher Bewegungsscore	Gewichtszuwachs	\hat{y}	ϵ
23.8	10.4	24.4	
24.1	11.0	25.3	
30.0	13.1	28.6	
29.0	15.4	32.0	
29.3	14.4	30.6	
34.8	14.1	30.1	
28.9	15.0	31.4	
27.9	12.2	27.1	
16.7	5.3	16.7	
21.6	9.5	23.0	
31.4	12.9	28.3	

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)**

102. Aufgabe


(12 Punkte)

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



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

Leider kennt sich Steffen 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 IX.

Experimentelles Design




103. Aufgabe

(16 Punkte)



Einfache experimentelle Designs Die Schlange macht mal wieder Randalen in Steffens Zimmer und rennt davon! Paula und Jessica sind bei Steffen im Wendland wo der neue, bessere Versuch stattfinden soll. Dabei soll in einem Gewächshausexperiment im Wendland mit Spargel durchgeführt werden. Auf dem Tisch stapeln sich Oreos aus Vollkorndinkelmehl. Eine Spezialität der Komune hier. Paula hasst Vollkorn wie Schokobons geliebt werden. In dem neuen Versuch geht es um den Zusammenhang zwischen der Behandlung Düngestufen (*ctrl*, *low*, *mid* und *high*) und dem Messwert Trockengewicht [kg/ha]. Immerhin ist der Messwert normalverteilt, was einiges einfacher macht. Was es nicht so einfacher macht ist, dass Jessica als zusätzliche Herausforderung noch der Mangel mitgebracht hat. Daher entscheiden sich alle drei für ein *Complete randomized design (CRD)*. 'Naja, so viel einfacher ist es dann doch nicht...', merkt Jessica an und sucht die Schlange.

Leider kennen sich Steffen, Paula und Jessica 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 notwendigen 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)**

104. Aufgabe

(20 Punkte)



Fortgeschrittene experimentelle Designs Steffen und Yuki sind bei Mark um sich Hilfe für eine Versuchsplanung in **R** zu holen. Im Hintergrund läuft viel zu laut Taylor Swift. Dabei geht es um den Zusammenhang zwischen der Behandlung Düngestufen (*ctrl*, *low*, *mid* und *high*) sowie Genotypen (*AA*, und *BB*) sowie vier Blöcken und dem Messwert Frischegewicht [kg/ha] in Spargel. Der Versuch soll in einem Freilandversuch im Oldenburger Land durchgeführt werden. Nach der Dozentin ist der Messwert Frischegewicht [kg/ha] normalverteilt. Die beiden entschieden sich für ein faktorielles Versuchsdesign. Im ersten Schritt überlegt Mark ein komplexeres experimentelles Design zu probieren. Daher entscheiden sich alle drei für ein *Split plot design* oder auch *Spaltanlage* mit Berücksichtigung einer Interaktion. Das sollte für den anfang erstmal reichen. 'Und jetzt, was machen wir jetzt?', Yuki schaut die anderen beiden mit großen Augen an. Die zucken mit der Schulter. Alle mampfen Oreos.

Leider kennen sich Mark, Steffen und Yuki mit dem *Split plot design* oder auch *Spaltanlage* überhaupt 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 für das faktorielle Versuchsdesign in **R**! **(2 Punkte)**
5. Erstellen Sie das statistische Modell in der in **R** üblichen Schreibweise für 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 mögliche Abbildung im Kontext der wissenschaftlichen Fragestellung! Beschriften Sie die Abbildung! **(3 Punkte)**
8. Ergänzen Sie zu der Abbildung ein mögliches Ergebnis des multiplen Gruppenvergleichs! Welche Annahme hinsichtlich der Modellierung haben Sie getroffen? Begründen Sie Ihre Antwort! **(3 Punkte)**

Teil X.

Forschendes Lernen

Zerforschen einer wissenschaftlichen Veröffentlichung

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

- Sánchez, M., et al. (2022). Hoverfly pollination enhances yield and fruit quality in mango under protected cultivation. *Scientia Horticulturae*, 304, 111320. [\[Link\]](#)
- 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. *Plants*, 11(8), 1010. [\[Link\]](#)
- Selle, P. H., et al. (2010). Implications of sorghum in broiler chicken nutrition. *Animal Feed Science and Technology*, 156(3-4), 57-74. [\[Link\]](#)
- Wu, G., et al. (2004). Arginine nutrition in neonatal pigs. *The Journal of Nutrition*, 134(10), 2783S-2790S. [\[Link\]](#)
- Graham, J., E., et al. (2024) Stock assessment models overstate sustainability of the world's fisheries. *Science* 385, 860-865. [\[Link\]](#)

In der Prüfung erhalten Sie keinen Auszug der wissenschaftlichen Veröffentlichung! Die Veröffentlichungen werden als bekannt in der Prüfung vorgelegt. Sie haben sich vorab Notizen und Anmerkungen auf Ihrem Spickzettel gemacht.

Zerforschen eines wissenschaftlichen Datensatzes

Das forschende Lernen basiert zum anderen auf den folgenden wissenschaftlichen Datensätzen und deren vertiefende Analyse werden als bekannt vorausgesetzt. Die Datensätze werden über ILIAS bereitgestellt.

- salad_fert_weight.xlsx
- weight_gain_pig.xlsx
- flowercolor_data.xlsx
- chickentype_class.xlsx

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 auf Ihrem Spickzettel gemacht.

105. Aufgabe



(20 Punkte)

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



Zerforschen einer wissenschaftlichen Veröffentlichung 'Uff', denkt Tina. Das ist jetzt doch etwas umfangreicher. Tina soll die wissenschaftliche Veröffentlichung *Sánchez, M., et al. (2022). Hoverfly pollination enhances yield and fruit quality in mango under protected cultivation* einmal zusammenfassen. Die Arbeit soll als eine Vorlage für ihre eigene Arbeit dienen. Daher möchte ihr Betreuer, dass sie einmal die Veröffentlichung in einer PowerPoint Präsentation zusammenfasst. 'Das ist jetzt aber doch umfangreicher als gedacht.', mault Tina in sich hinein. Schnell nochmal ein paar Katjes zur Stärkung gegessen. Das wird dann vermutlich heute Abend nichts mehr mit ihrem Hobby Astronomie. Die Spinne schaut mitleidig.

Leider kennt sich Tina mit dem Lesen einer wissenschaftlichen Veröffentlichung mit Fokus auf die Statistik überhaupt nicht aus. Deshalb braucht sie 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 'Uff', denkt Yuki Das ist jetzt doch etwas umfangreicher. Yuki soll die wissenschaftlichen Daten *data3* einmal sinnvoll analysieren und zusammenfassen. Die Daten sollen als eine Vorlage für ihre eigene Arbeit und Experiment dienen. Daher möchte ihre Betreuerin, dass sie einmal die Daten in einer PowerPoint Präsentation zusammenfasst. 'Das ist jetzt aber doch umfangreicher als gedacht.', mault Yuki in sich hinein und mampft noch ein paar Reese's Peanut Butter Cups. Dann starrt sie eine Weile in ihren Laptop. Das wird dann vermutlich heute Abend nichts mehr mit Matrix

Leider kennt sich Yuki mit der Analyse eines wissenschaftlichen Datensatzes überhaupt nicht aus. Deshalb braucht sie 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 Mark, Yuki, Alex und Jessica waren auf Exkursion in Brandenburg 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 1mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 12m in Bruthöhe hatte.

1. Wie groß war der Durchmesser in m der Eiche im Jahr 1840 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 195cm, 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 *sehr bequem mit Abstand* 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 10cm 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 *H* bestehend aus Jonas, Steffen, Tina und Nilufar hat sich zusammengefunden um den ersten Versuch zu planen. In einem Experiment wollen sie die Wuchshöhe von 180 Sonnenblumen 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 Sonnenblumen einpflanzen und müssen dafür Substrat bestellen. Zum anderen muss die Projektgruppe die Sonnenblumen 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 *H* bei der Planung des Versuches!

1. Skizzieren Sie den Versuchsplan auf zwei Tischen im Gewächshaus! **(2 Punkte)**
2. Berechnen Sie die benötigte Anzahl an Pflanztöpfen, wenn Sie Randpflanzen mit berücksichtigen wollen! **(1 Punkt)**
3. Berechnen Sie die benötigte (a) *Pflanztopf*fläche in m^2 sowie die (b) *Tisch*fläche in m^2 gegeben der Anzahl an Pflanztöpfen inklusive Randpflanzen am Anfang der Keimungsphase! **(4 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! **(2 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 Tina bringt ein neues, tolles Projekt mit in die Lerngruppe *Sinnlos im Studium* bestehend aus ihr, Nilufar, Yuki sowie Mark. Um die Energiekosten ihres Betriebes zu senken, will sie eine Solaranlage auf den Schweinestall montieren lassen. Dafür hat sie ihren Stall ausgemessen und findet folgende Maße wieder. Die vordere Seite des Schweinestall hat eine Höhe h_v von $6.5m$. Die hintere Seite des Schweinestall hat eine Höhe h_b von $11m$. Der Schweinestall hat eine Tiefe t von $12m$ und eine Breite b von $40m$. 'Sag mal Tina, ist das eine Matheaufgabe oder rechnen wir hier gerade für dich kostenlos als menschliche Computer Sachen für deinen Betrieb?', fragt Yuki mit erhobenen Augenbrauen. Mark und Nilufar nicken zustimmend.

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

1. Skizzieren Sie den Schweinestall 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 Schweinestall! **(3 Punkte)**

Ebenfalls plant Tina 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. Das Fundament des zylindrischen Methantanks kann nur ein Gewicht von maximal $5t$ aushalten bevor der Tank wegbricht. Tina rechnen eine Sicherheitstoleranz von 15% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $240kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $270kg/m^3$. Tina betreibt ihre Anlage bei $-85^\circ C$.

3. Extrapolieren Sie die effektive Dichte des Methans in dem Methantank! Welche Annahme haben Sie getroffen? **(1 Punkt)**
4. Berechnen Sie wie viel Kubikmeter m^3 in den Methantank gefüllt werden 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 puderrote 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 Edeka 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 4 Blaubeerschalen 7.96 Euro kosten, wie viel kosten 5 Schalen? **(2 Punkte)**
2. Wenn Sie die 5 Blaubeerschalen gekauft haben, wie viele Aligatorbirnen zu je 2.59 EUR können Sie sich dann noch für 50 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 Edeka über Gemüse erfährt.

- Ein Kilo Strauchtomaten benötigt 180l Wasser. Eine Strauchtomate wiegt 90 - 130g.
 - Ein Kilo Salat benötigt 140l Wasser. Ein Salatkopf wiegt 280 - 510g.
 - Ein Kilo Avocado benötigt 980l Wasser. Eine Avocado wiegt 120 - 380g.
 - Ein Kilo Blaubeeren benötigt 830l Wasser. Eine Blaubeere wiegt 3.2 - 3.6g.
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 2021 blieben die Erträge von Blaubeeren mit 7.5×10^4 t in dem prognostizierten Rahmen. Die Menge steigerte sich um 6.1%. Die Exporte für Avocados fielen in dem gleichen Zeitraum um 22.1% auf 1.9×10^5 t.

4. Wie viele Liter Wasser hat Chile in dem Exportjahr 2020 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 48 Liter Wasser pro Kopf für den täglichen Bedarf. In *Deutschland* liegt der Verbrauch bei 35 - 115 Liter pro Waschgang einer Waschmaschine und 9 - 14 Liter pro Spülgang.

5. Mit der rationierten Wassermenge aus Chiles Anbaugeländen 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

(10 Punkte)

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



Stichworte: Kardaschow-Skala • Dyson-Sphäre • Hohlerte • 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 Paula 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 67 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 Hohlerte 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 9.81 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.235 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.44 g/cm^3 . Das Gewicht von einem heute lebenden asiatischen Elefanten liegt bei 3t bis 5t und das Gewicht von einem Brachiosaurus bei bis zu 30t.

1. Welchen Durchmesser müsste die Erde vor 67 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 67 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 67 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 87% Wasserstoffkernen mit einer molaren Masse von 1.08 g/mol , 10% Heliumkernen mit 4.11 g/mol sowie 3% weiteren Atomkernen mit 69.18 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 7 cm^{-3} pro Sekunde.

Lösen Sie den folgenden Aufgabenteil mit einer aussagekräftigen Skizze!

3. Berechnen Sie die Anzahl an massebehafteten Teilchen des Sonnenwindes, die die gesamte Erde pro Sekunde treffen! **(2 Punkte)**
4. 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 Gänse in Mastställen sagen. Das ist natürlich etwas ungünstig, den 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 Alex, Yuki, Steffen und Paula aber als vorsorgliche Gänse-Halter:innen nicht⁷. Gemeinsam sind die Vier in einer Projektgruppe gelandet. Betrachten wir also gemeinsam einmal das Platzangebot (eng. *space allowance*, abk. *SA*) der Gänse 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 Gans plus dem benötigten Radius für das Verhalten i .
- R_i dem notwendigen Abstand zu den Nachbarn für das Verhalten i .
- i dem Verhalten: (1) standing, (2) preening, (3) drinking/eating und (4) walking.

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)
standing	11cm; 27cm; 3.5%	22cm; 19cm; 12.1%	44cm; 13cm; 10.1%
preening	50cm; 39cm; 4.3%	1cm; 16cm; 2.1%	25cm; 19cm; 4.1%
drinking/eating	23cm; 17cm; 3.6%	37cm; 5cm; 6.3%	32cm; 18cm; 6.3%
walking	45cm; 6cm; 13.2%	25cm; 20cm; 14.2%	14cm; 41cm; 12.3%

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

1. Skizzieren Sie die Werte r_i , R_i und A_i für zwei nebeneinander agierende Gänse für ein Verhalten i . Nutzen Sie hierfür vereinfachte geometrische Formen! **(2 Punkte)**
2. 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)**
3. Ergänzen Sie eine Spalte mit dem benötigten Platz A für das jeweilige Verhalten, welches sich aus den mittleren Werten ergibt! **(1 Punkt)**
4. Berechnen Sie das benötigte Platzangebot SA für alle betrachteten Verhalten! **(1 Punkt)**
5. Sie entnehmen der Literatur folgende Aussage zur Verteilung der Gänse 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 Jessica und Yuki 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 Jessica ziemlich plötzlich, so dass Yuki die Reese's Peanut Butter Cups aus dem Mund fallen. 'Nein, und ehrlich gesagt bin ich auch ziemlich müde...'. Das ist jetzt aber Jessica egal, denn sie möchte folgende Sachlage diskutieren. Und Jessica 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 42 Tagen die ersten Symptome ein; die ersten Toten sind nach 72 Tagen zu beklagen; nach 115 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 $7000\mu\text{g}/50\text{g}$ Vitamin C. Der Bedarf liegt bei 110mg 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 22 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 Alex 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. Alex mampft ungerührt Gummibärchen. Folgenden Gedankengang nutzt Paula um dem Film zu entkommen. Die Sonne hat eine aktuelle, angenommene Masse von $2 \times 10^{30} \text{ kg}$. Wenn die Sonne nun am Ende ihrer Lebenszeit zu einem schwarzen Loch mit dem Radius von 3500m 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 $6.165 \cdot 10^{-11} \text{ m}^3 (\text{kg} \cdot \text{s}^2)^{-1}$

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! **(2 Punkte)**
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 $2.7 \times 10^8 \text{ m/s}$ 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! Erstellen Sie dafür eine Datentabelle mit fünf Werten für den Radius r ! **(3 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 Jessica und Nilufar wandern durch den Teuteburgerwald um mal vom Studium runterzukommen. 'Kennst du eigentlich Enrico Fermi?', fragt Jessica 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?'. Nilufar schaut sie irritiert und interessiert an. Die beiden hat das Problem gepackt. Deshalb wollen Jessica und Nilufar 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 zwei Voyager 1 ähnliche Sonden mit der Geschwindigkeit von $6.2587 \times 10^4 \text{ km/h}$ los um sich auf den erreichten Planeten selbst zu replizieren. Nach 250 Jahren ist die Replikation abgeschlossen und wiederum zwei Sonden werden ausgesendet. Gehen Sie von 3.57 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 2×10^{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 vermutet Alter der Erde von 4.1×10^9 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 9×10^7 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 an! 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 51 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 44 Königsellen. Eine Königselle misst 52.4cm.

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 4cm 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 5 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 Rückenschmerzen 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 110 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 8° ändert! **(2 Punkte)**

117. Aufgabe

(10 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. Tina und Steffen schwingen sich auf ihre Cachermobile um mit 16km/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. Tina und Steffen 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 Tina und Steffen für die Planung der Route zu Verfügung¹⁰.

Ort	Cache	Wertung (S T G)
A	GCYCARA	2.5 5.0 Normal
B	GCJLPMP	4.0 2.0 Normal
C	GCEHQ0R	2.0 4.5 Mikro
D	GCVPFGE	1.5 2.5 Klein
E	GC0J7NV	5.0 3.5 Mikro

Im Weiteren sind den beiden folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \vec{AC} ist 4km. Im Weiteren ist Ihnen der Entfernungsvektor \vec{CB} mit 4.5km bekannt. Der Entfernungsvektor \vec{BE} ist das 1.2-fache des Entfernungsvektor \vec{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 25° nördlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 50° 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 C 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! **(5 Punkte)**
2. Welche Strecke in km legen Sie bei der Bewältigung der abwärts Terrainchallenge zurück? **(2 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

$$\text{Suchzeit} = 0.15 + 0.15 \cdot \text{Schwierigkeit}$$

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

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

118. Aufgabe

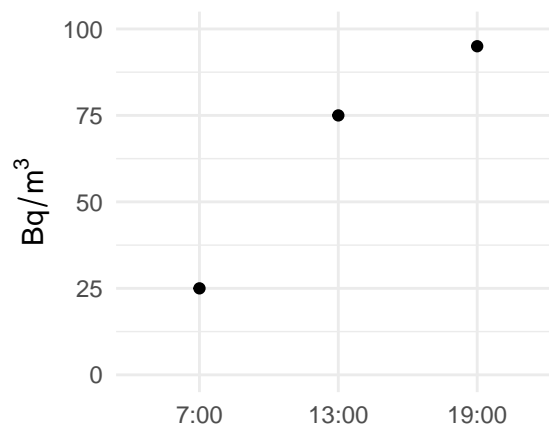
(10 Punkte)

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 Steffen das Ausmaß der Radonbelastung in seinem 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 19:00 bestimmt er dreimal automatisch die Radonbelastung in seinem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung¹². Leider helfen die Messwerte Steffen überhaupt nicht weiter. Sie müssen also helfen!



- Wie lange dauert es in Stunden bis Sie eine kritische Belastung von $380 Bq/m^3$ in Ihrem ungelüfteten Kellerraum erreicht haben? **(2 Punkte)**

Radon zerfällt mit einer Halbwertszeit von 1.8d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 143d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

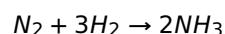
- Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $380 Bq/m^3$ auf unter $90 Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	77.1	28.4	
Sauerstoff	19.5	15.8	
Kohlenstoffdioxid	0.045	12.5	

- 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 Steffen sein etwas pappiges Toastbrot mampfen kommt Steffen die Dokumentation über Brot aus Luft in den Sinn. Steffen 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.

- Welche Masse an Ammoniak in Kilogramm kg können Sie aus einem Kubikmeter m^3 Luft unter normalen Umweltbedingungen gewinnen? **(2 Punkte)**
- Wie viel 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 Tina war nach Ladenschluss bei IKEA die Regale einzuräumen. Dabei ist Tina in der Auslage der Sonderangebote das Necronomicon¹⁴ in die Hände gefallen. Nun ist sie ein Magier der Zeichen geworden! Also eigentlich kann Tina nur Mathe und das dämliche Necronomicon hat sie in die Vergangenheit geschleudert... aber gut, was tut man nicht alles im Jahr 194 n. Chr. für den neuen Lehnsherren Henry dem Roten. Tina baut natürlich einen Schrottkugelturm um sich den Horden der Finsternis mit genug Schrott erwehren zu können! Tina 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 $40mg$ 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 $4.1mm$. Die Löcher im Sieb erlauben ein Tropfendurchmesser von $1.7mm$. Welchen Durchmesser in mm haben Ihre produzierten Bleikugeln? **(3 Punkte)**

Sie haben jetzt die 1.2×10^6 Bleikugeln zusammen. Blei hat eine Dichte von $10.32g/cm^3$.

4. Wie schwer in Kilogramm kg sind die 1.2×10^6 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. 900 Bleikugeln produzieren wollen und die Bleikugel im Fall $1.4cm$ 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 Tina mit dem Krokodilreservat in Down Under nicht geklappt. War vielleicht auch nicht so die beste Idee... aber dafür hat Tina eine neue Eingebung! Oder wie es Mike Tyson zugeschrieben wird: »Ich wurde nie niedergeschlagen, ich war immer am Aufstehen!«. Daher macht Tina jetzt einen Großhandel mit Kaninchenfleisch und damit dem teuersten Fleisch in Australien auf. Moment, hopsen hier nicht, seit Thomas Austin im Jahr 1875 ungefähr 26 Kaninchen entlassen hat, Millionen von Kaninchen rum? Wieso ist das Kaninchenfleisch dann so exklusiv? Tina wird stutzig und frag Sie, dem mal mathematisch nachzugehen!¹⁵

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

$$f(t) = 1.1 \times 10^{10} - 10^9 \cdot 1.9^{-0.3 \cdot t + 2.7}$$

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 9 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 20 Kaninchen ausgebracht. Da die Kaninchen keine Maßnahmen gegen die Ausbreitung vornehmen können, verläuft die Ausbreitung mit einem wöchentlichen Wachstumsfaktor von 1.4 nach folgender Formel.

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

3. Wie viele Wochen benötigen die Viren um theoretisch die gesamte Kaninchenpopulation nach 7 Jahren Wachstum zu durchseuchen? **(1 Punkt)**

Das Myxoma Virus und das RHDV töten 99.7% der Kaninchenpopulation innerhalb weniger Wochen.

4. Wie lange in Jahren dauert es bis eine Kaninchenpopulation nach einer Viruspandemie wieder auf 50% der gesättigten Kaninchenpopulation angewachsen ist? **(2 Punkte)**

Thomas Austin entließ die Kaninchen im äußersten Süden von Australien. Australien hat eine West-Ost-Ausdehnung von 4300km und eine Nord-Süd-Ausdehnung von knapp 3500km. 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. 45\$ pro Tier.

6. In Ihrem Stall leben 1000 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!



Lüneburger Heide. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer des Esels Fridolin und Jessica. Grünes Gras unter Jessicas Füßen und ein strammer Wind im Gesicht, egal wohin sie schaut. Ein schmatzendes Geräusch ertönt unter Jessica. Jessica sinniert, sollte sie ihre weiten Graslandschaften jetzt schon düngen? Dafür benötigt Jessica die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Jessica 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 $0.6\times$, Februar mit $0.75\times$ 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	1.2
01. Mrz 2023	2.7
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 Jessicas Pink Lady Plantage werden Sie beide auf dem Trecker von einer Gruppe elektrifizierter Rentner abgedrängt. Der Trecker muss wieder aus dem Graben! Fridolin und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Fridolin mit 230N . Die elektrifizierten Rentner bringen eine Kraft von 190N 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 Fridolin 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 Fridolin und die Rentner mit einem 40° Winkel mit ihrem Seil an dem Trecker. Welche Kraft wird aufgebracht? **(2 Punkte)**
7. Mit welcher Beschleunigung ziehen Sie den 1.2t 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 Yuki, Steffen, Jessica und Tina sitzen im Bus. Wenn man sich zu spät anmeldet, dann ist die Exkursion nicht so toll. Yuki hatte den Anderen in der Lerngruppe zu spät Bescheid gesagt. 'Was denn, bin ich eure Nanny oder was?!', entfährt es Yuki 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, 5-mal-gefaltete, 0.5mm, 40-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 40cm 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 110m 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 110m 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

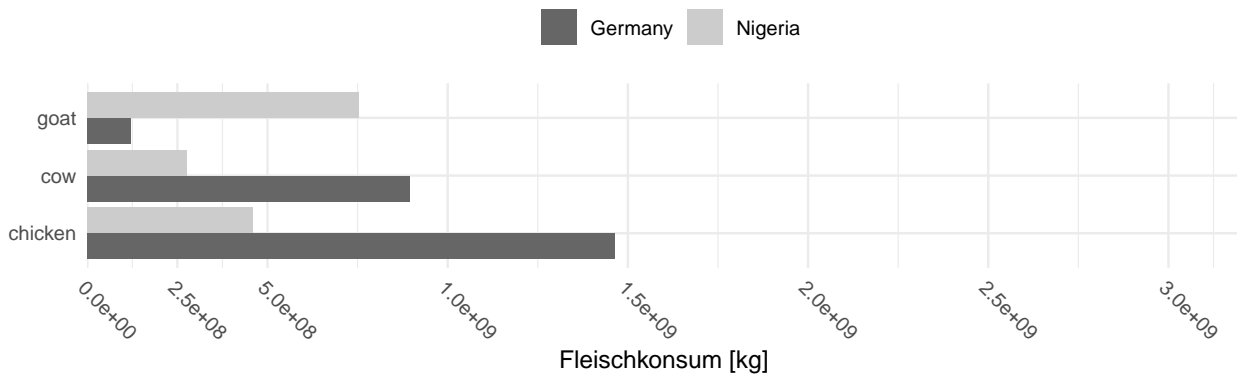
(10 Punkte)

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



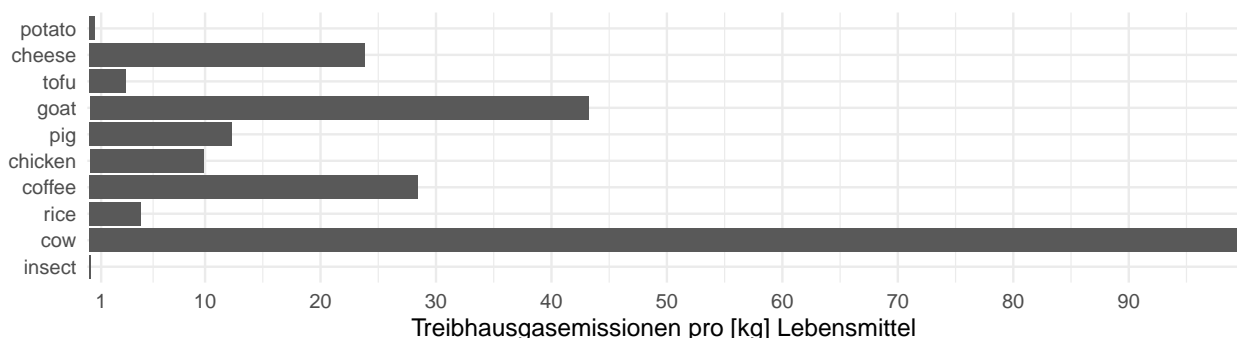
Ein Pfund Insekten, bitte! 'Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹⁶.', merkt Paula an. Die Lerngruppe um Jonas, Nilufar und Mark sind bei Paula um mal was außergewöhnliches zu essen. Um den Sinn der Nahrungsumstellung zu verdeutlichen, vergleicht Paula 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. 8×10^7 Menschen in Deutschland und ca. 1.84×10^8 Menschen in Nigeria. Mit den Informationen wollen Sie und Paula 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)**

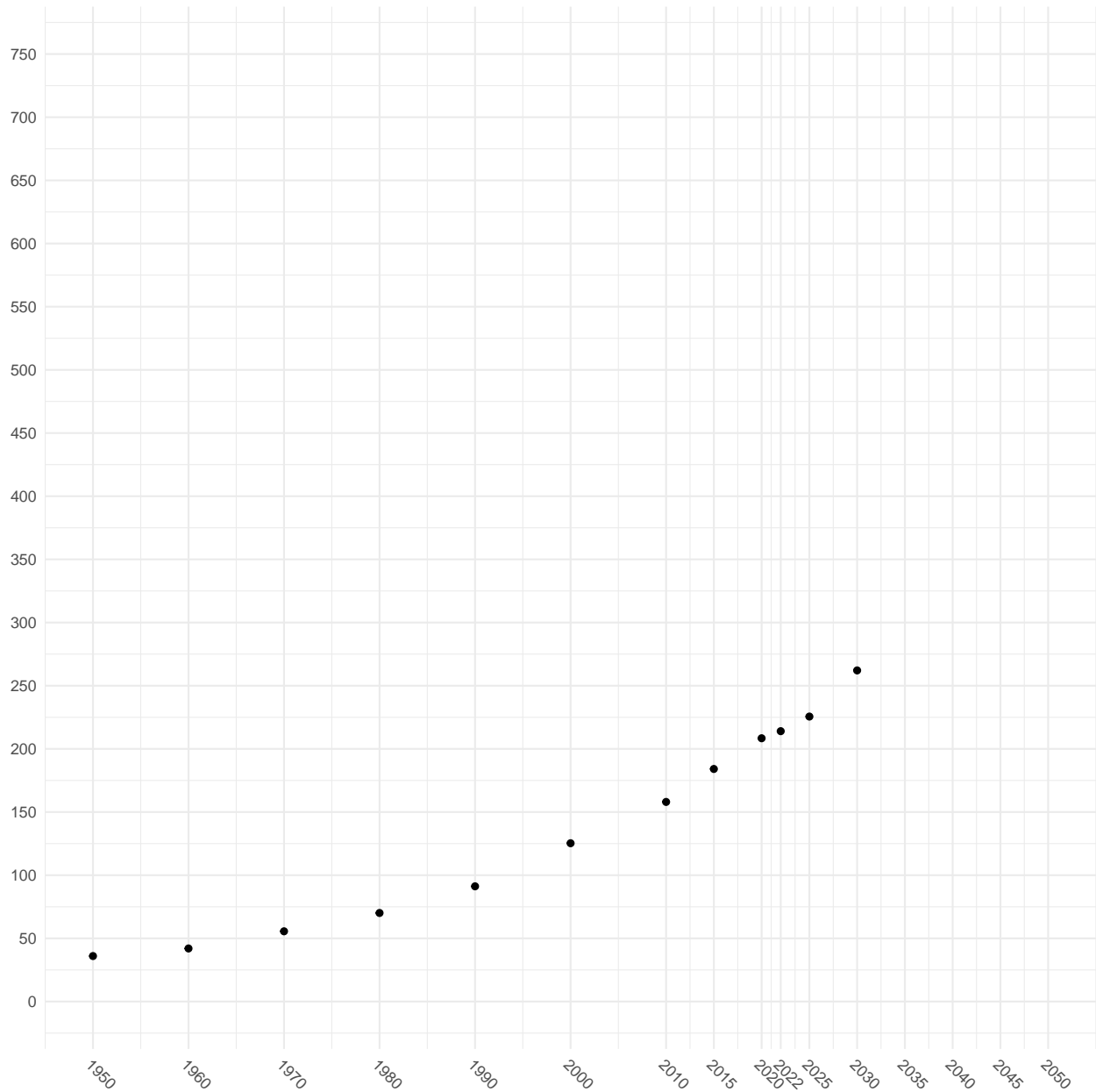
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



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!



Stichworte: Immunsystem – Muskel vs. Interpol • Inzidenz • Prävalenz

Tödliche Seuche AIDS – Die rätselhafte Krankheit Irritiert legt Nilufar die historische Ausgabe des Spiegels aus den 80zigern beiseite. Nilufar und Mark sind bei ihrem HNO-Arzt und wollen einen AIDS-Test machen lassen. Woanders leider keinen Termin gekriegt... Immerhin denken die beiden über Nachwuchs nach und da geht es eben nur durch ungeschützten Sex. Was wissen Nilufar und Mark nun aber über AIDS und dem diagnostischen AIDS-Test, den die beiden nun machen werden? Leider zu wenig. Da brauchen dann Nilufar und Mark mal wieder Ihre Hilfe bei der Interpretation eines diagnostischen Tests!

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.8% 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 $n = 3 \times 10^4$ 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 Nilufar und Mark, dass beim diagnostischen Testen *True Positives* (TP), *True Negatives* (TN), *False Positives* (FP) und *False Negatives* (FN) auftreten. Das verstehen beiden so noch nicht und deshalb stellen Sie für Nilufar und Mark 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! Nilufar, Alex und Paula sitzen bei Jonas und hören sich etwas über Network-Marketing an. Jonas ist jetzt im Network-Marketing tätig. 'Jetzt reicht es. Wir sind eine Lerngruppe und du versuchst uns hier abzuziehen!', poltert Nilufar und fährt fort, 'Ich erklär dir mal, wie falsch du liegst!'. 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. 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? Nilufar und Sie wollen hier einmal in die Untiefen des »passiven Einkommens« abtauchen und die Lerngruppe vor Schlimmeren bewahren¹⁸!

Das Jahr 2022 war das erfolgreichste Jahr in der Geschichte von Healthy Herbs Manufacture International (HeHeMan). Das Unternehmen steigerte den Umsatz um rund 15 Prozent von 280 Millionen Euro im Jahr 2021. Doch wie viel kommt bei den Partnern an? Laut HeHeMan habe das Unternehmen 3.6×10^5 aktive Partner.

1. Berechnen Sie zuerst den Umsatz der Firma HeHeMan 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 25%? **(1 Punkt)**

Das von Jonas zu vermarkende Produkt, hinter dem Jonas voll steht, kostet 150EUR pro Einheit im Direktverkauf. Die direkte Provision für die erste Stufe beträgt 20%. Für die zweite, dritte und vierte Stufe betragen die indirekten Provisionen jeweils 3%, 2% und 1.5%. Jeder von Jonas angeworbener »Partner« wirbt wiederum vier Partner für sich selbst an. Pro Monat werden im Schnitt vier Einheiten vom Produkt verkauft. Jonas will nun 3200EUR im Monat *passiv* – also durch indirekte Provisionen – erwirtschaften. Kann das klappen? Sie sind zusammen mit Nilufar skeptisch.

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)**

Jonas musste zum Einstieg bei HeHeMan Einheiten des Produkts für 6000EUR kaufen. Diese Einheiten kann Jonas nur direkt verkaufen. Das ganze Wohnzimmer ist voll davon. Leider musste Jonas den Kauf über einen Kredit über 6.2% p.a. über 48 Monate finanzieren. Sie schütteln den Kopf und klären Jonas über Zinsen auf.

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 Jessica, Jonas und Steffen sitzen bei Paula nachdem sich alle begeistert in der Serie *Stranger Thinks* verloren haben. Alle drei wollen jetzt einmal bei Paula *Höhlen & Drachen* ausprobieren. Um Geld zu sparen, das Zeug kostet echt, wurde etwas an den Regeln gebastelt. Schnell stellen die Drei fest, dass hier ganz schön viele unterschiedliche Würfel durch die Gegend fliegen. Daher müssen Sie sich jetzt einiges an Fragen klären damit Paula nicht so alleine ist.

In dem Spiel hat Jessica nun auf einmal 7 achtseitige Würfel ($7d8$) zum würfeln in der Hand. Wenn Jessica eine 8 würfelt, hat Jessica 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)**

Jonas betrachtet nun aufmerksam die ausufernden Ausrüstungstabellen. Jonas wird aber geholfen und muss 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 vierseitigen Würfel plus 5 ($1d4+5$) als Schaden bevorzugen? Begründen Sie Ihre Antwort mathematisch! **(1 Punkt)**

Jetzt wird es immer wilder, da Jonas und Steffen sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass der Rettungswurf gegen den zaubernden Hexer funktioniert. Jonas und Steffen haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.7$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.7$. Sie haben aber mitgezählt und festgestellt, dass in 50 von 100 Fällen der 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 Das war zu viel für Tina gestern. Die Lerngruppe mit Yuki und Steffen ging viel zu lang. Während sie wegdämmert, kommen in ihr seltsame Bilder hoch. '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 Tina fragt sich, ob sie nicht doch lieber bezahlter Gast bei Barbara Salesch hätten sein sollten. Aber Tina braucht das Geld und jetzt heißt es Spielschows farmen! Erstmal eine Kaffeemaschine von Mitropa gewinnen. Ein Kandidat gewinnt die Kaffeemaschine von Mitropa, wenn nicht alle Kandidaten überbieten (eng. *outbid*). Mit Ihnen bilden Yuki und Steffen das Team der drei Kandidaten. Tina braucht dringend Ihre Hilfe in ihren Wahnträumen. Sie gehen wie in *Interception* rein!¹⁹

Name	$P(\text{win})$	$P(\text{outbid})$
Yuki	0.2	0.11
Steffen	0.3	0.12

1. Mit welcher Wahrscheinlichkeit gewinnen Sie die Kaffeemaschine von Mitropa, wenn keiner der Kandidaten überbietet? **(1 Punkt)**
2. Wenn Ihre überbietungswahrscheinlichkeit $P(\text{outbid})$ bei 0.076 liegt, mit welcher Wahrscheinlichkeit gewinnt *keiner* die Kaffeemaschine von Mitropa? **(1 Punkt)**

Glücksrad für Arme auf der Kirmes! Leider hat es für Maren Gilzer nicht gereicht. Deshalb sind Sie jetzt mit Tina auf der Kirmes und spielen mit einem Typen in einem Tentakelkostüm um das große Geld. Das Glücksrad hat 24 Felder. Sie beide drehen das Glücksrad zweimal. Auf 10 Feldern gewinnen Tina und Sie 4000EUR 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 6000EUR? **(1 Punkt)**

Im Fiebertraum von Tina reisen sie beide im Zug nach Köln um bei »Geh aufs Ganze!« mitzuspielen. Tina und Sie schaffen es tatsächlich ins Finale und können als Hauptgewinn ein Auto hinter einer der drei Türen gewinnen. Und was braucht man mehr als ein Auto in einem Fiebertraum?

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«! Berechne Sie dazu die Wahrscheinlichkeiten der einzelnen Pfade. Lohnt sich ein Wechsel der anfangs gewählte Tür? Begründen Sie Ihre Antwort mathematisch! **(2 Punkte)**

¹⁹South Park Inception Spoof – Wunderbare South Park Folge

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} : I-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.