

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

Fakultät Agrarwissenschaften und Landschaftsarchitektur (AuL)

Klausurfragen Bio Data Science

für Pflichtmodule

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

(Prüfungsleistung der Wahlpflichtmodule ist eine Portfolioprüfung)

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

Wintersemester 2024/25



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



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

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



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

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



Nach der Ausbildung wollte Nilufar eigentlich gleich anfangen zu arbeiten, aber nach einem Praktikum und der Probezeit stellte sie fest, dass es ohne einen Hochschulabschluss schwer wird Führungsverantwortung zu übernehmen. Mit Menschen kann sie schon immer und dann auch eigene Projekte mit anderen verwirklichen, 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 92 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 112 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
106.5 - 112	1,0
101.5 - 106	1,3
95.5 - 101	1,7
90.0 - 95.0	2,0
84.5 - 89.5	2,3
79.0 - 84.0	2,7
73.5 - 78.5	3,0
67.5 - 73.0	3,3
62.0 - 67.0	3,7
56.0 - 61.5	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	9	9	12	20	20	12	10

- Es sind ____ von 92 Punkten erreicht worden.

Multiple Choice Aufgaben

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

ANOVA

1. Aufgabe

(2 Punkte)

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

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

2. Aufgabe

(2 Punkte)

Sie führen ein Feldexperiment durch um das Gewicht von Erdbeeren zu steigern. Die Pflanzen wachsen unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen. Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.17$. Welche Aussage ist richtig?

- A** ☐ Es werden 83% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.
- B** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird. Daher werden 17% der Varianz erklärt.
- C** ☐ Es werden 17% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- D** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 17% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 83%.
- 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.

3. Aufgabe

(2 Punkte)


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

- A** ☐ Die ANOVA berechnet die F-Statistik aus den SS Behandlung geteilt durch die SS Fehler.
- B** ☐ Wenn die F-Statistik kleiner als der kritische Wert ist kann die Nullhypothese nicht abgelehnt werden. Die F-Statistik ist der Quotient der MS der Behandlung durch die MS des Fehlers.
- C** ☐ 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.

- D** ☐ Wenn die F-Statistik höher ist als der kritische Wert kann die Nullhypothese nicht abgelehnt werden. Die F-Statistik ist die Differenz der MS der Behandlung durch die MS des Fehlers.
- E** ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 1 annähert kann die Nullhypothese nicht abgelehnt werden.

4. Aufgabe

(2 Punkte)

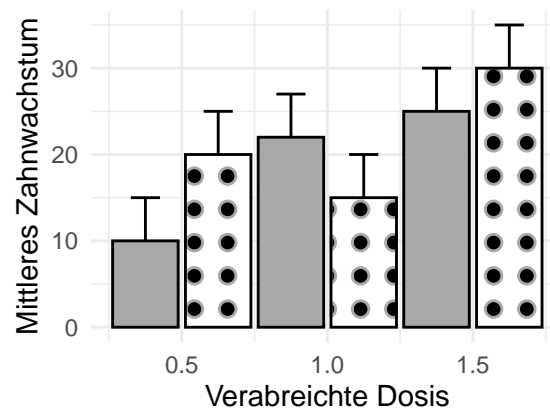
Wenn Sie mehr als zwei Gruppen als Behandlungen vorliegen haben, dann kann ein einfacher t-Test nicht für den globalen Vergleich genutzt werden. Sie entscheiden sich für eine ANOVA in . Die ANOVA analysiert dabei...

- A** ☐ ... den Unterschied zwischen der Varianz durch verschiedene Behandlungsguppen unter der Varianz über alle Behandlungsguppen. Wenn die ANOVA signifikant ist, kann kein Effekt η^2 bestimmt werden.
- B** ☐ ... den Unterschied zwischen der F-Statistik anhand der Varianz der Gruppen. Wenn die F-Statistik exakt 0 ist, kann die Nullhypothese abgelehnt werden.
- C** ☐ ... den Unterschied zwischen zwei paarweisen Mittelwerten aus verschiedenen Behandlungsguppen. Wenn die signifikant ist, ist daher bekannt welcher Vergleich konkret unterschiedlich ist.
- D** ☐ ... den Unterschied zwischen der Mittelwerte und der Varianz aus verschiedenen Behandlungsguppen. Wenn die ANOVA signifikant ist, ist bekannt welcher Vergleich konkret unterschiedlich ist.
- E** ☐ ... den Unterschied zwischen der globalen Varianz und der Varianz aus verschiedenen Behandlungsguppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.

5. Aufgabe

(2 Punkte)

In einer Studie zur Bewertung der Wirkung von Vitamin D auf das Zahnwachstum bei Kanarienvögel entstand folgende Abbildung. Der Versuch wurde an 53 Tieren durchgeführt, wobei jedes Tier eine von drei Vitamin-C-Dosen (0.5, 1 und 1.5 mg/Tag) über eine von zwei Verabreichungsmethoden erhielt. Welche Aussage ist im Bezug auf eine zweifaktorielle ANOVA richtig?



- A** ☐ Keine Korrelation liegt vor ($p \geq 0.05$).
- B** ☐ Das Bestimmtheitsmaß R^2 ist klein.
- C** ☐ Eine positive Interaktion liegt vor ($p \leq -0.5$)
- D** ☐ Eine mittlere bis starke Interaktion liegt vor ($p \leq 0.05$)
- E** ☐ Keine Interaktion liegt vor ($p \leq 0.05$).

Deskriptive Statistik & Explorative Datenanalyse

6. Aufgabe

(2 Punkte)

Gegeben ist y mit 14, 11, 4, 9 und 18. Berechnen Sie den Mittelwert und Standardabweichung.

- A ☐ Es berechnet sich 12.2 ± 27.7
- B ☐ Es berechnet sich 11.2 ± 5.26
- C ☐ Sie erhalten 11.2 ± 2.29
- D ☐ Es ergibt sich 10.2 ± 13.85
- E ☐ Es ergibt sich 12.2 ± 2.63

7. Aufgabe

(2 Punkte)

Wie lautet der Median, das 1st Quartile sowie das 3rd Quartile von y mit 13, 24, 14, 20, 15, 18, 20, 12, 17, 18 und 42.

- A ☐ Sie erhalten 18 [14; 20]
- B ☐ Sie erhalten 18 [12; 18]
- C ☐ Sie erhalten 18 ± 20
- D ☐ Es ergibt sich 19 ± 14
- E ☐ Es berechnet sich 19 [15; 21]

8. Aufgabe

(2 Punkte)

Sie überlegen Ihre Daten mit einem Boxplot zu visualisieren. Was ist die minimale Anzahl an Beobachtungen pro Gruppe ?

- A ☐ 2-5 Beobachtungen.
- B ☐ Die untere Grenze liegt bei einer Beobachtung.
- C ☐ Mindestens 20 Beobachtungen.
- D ☐ Die Mindestanzahl liegt bei fünf Beobachtungen.
- E ☐ 1 Beobachtung.

9. Aufgabe

(2 Punkte)

Um die Standardabweichung zu berechnen müssen wir folgende Rechenoperationen durchführen.

- A ☐ 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.
- B ☐ Den Mittelwert berechnen und die Abstände quadrieren. Die Summe mit der Fallzahl multiplizieren.
- C ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend subtrahieren wir die Fallzahl.
- D ☐ 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.
- E ☐ 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.

10. Aufgabe

(2 Punkte)

Der Barplot stellt folgende statistische Maßzahlen in einer Abbildung dar. Damit gehört der Barplot zu einem der am meisten genutzten statistischen Verfahren zur Visualisierung von Daten.

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

11. Aufgabe

(2 Punkte)

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

- A ☐ Der Mittelwert und der Median sollten gleich sein, wenn keine Outlier in den Daten vorliegen.
- B ☐ Da sich der Mittelwert und der Median unterscheiden, liegen vermutlich keine Outlier in den Daten vor. Wir verwenden den Datensatz so wie er ist.
- C ☐ Da sich der Mittelwert und der Median nicht unterscheiden, liegen vermutlich Outlier in den Daten vor. Wir untersuchen den Datensatz nach auffälligen Beobachtungen.
- D ☐ Wenn sich der Mittelwert und der Median unterscheiden, liegen vermutlich keine Outlier in den Daten vor.
- E ☐ Der Mittelwert und der Median sollten sich unterscheiden sein, wenn Outlier in den Daten vorliegen.

12. Aufgabe

(2 Punkte)

Sie wollen eine ANOVA im Anschluss an Ihr Feldexperiment rechnen. Dafür muss Ihr gemessener Endpunkt die Annahme einer Normalverteilung genügen. Zur Überprüfung können Sie folgende Visualisierung nutzen. Welche entsprechende Regel zur Abschätzung der Annahme einer Normalverteilung kommt zur Anwendung?

- A ☐ Einen Barplot. Die Mittelwerte müssen alle auf einer Höhe liegen. Die Fehlerbalken haben hier keine Informationen.
- B ☐ Einen Violinplot. Der Bauch der Violine muss hierbei einen höheren Wert annehmen als der Steg der Violine. Dann kann die Annahme einer Normalverteilung angenommen werden.
- C ☐ Wir erstellen uns für jede Behandlung einen Boxplot und schauen, ob die Box und damit das IQR für jede Behandlung gleich groß ist.
- D ☐ Nach der Erstellung eines Boxplots schauen wir, ob der Median in der Mitte der Box liegt. Dabei ist der Median als dicke Linie dargestellt und die Box ist das IQR.
- E ☐ Einen Dotplot. Die Punkte müssen sich wie an einer Perlenschnur audreihen. Eine Abweichung führt zur Ablehnung der Annahme einer Normalverteilung.

13. Aufgabe

(2 Punkte)

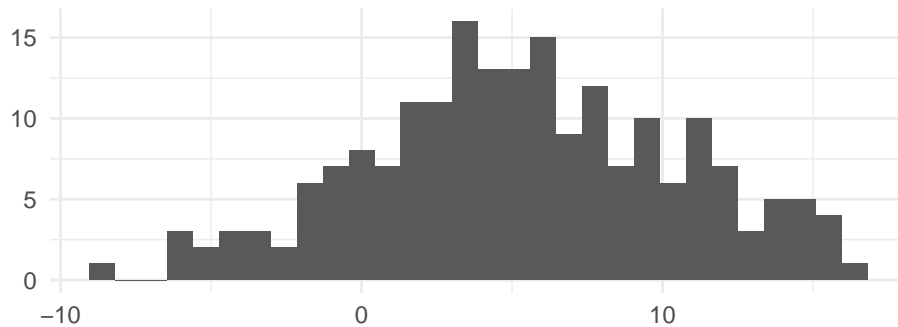
In der Statistik müssen wir häufig überprüfen, ob unser Outcome einer bestimmten Verteilung folgt. Meistens überprüfen wir, ob eine Normalverteilung vorliegt. Folgende drei Abbildungen eignen sich im Besonderen für die Überprüfung einer Verteilungsannahme an eine Variable.

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

14. Aufgabe

(2 Punkte)

Bevor Sie in Ihrer Abschlussarbeit einen statistischen Test rechnen, wollen Sie einmal betrachten, welcher Verteilung Ihre $n = 200$ geernteten Pflanzen folgen. Welche Verteilung ist abgebildet?



- A** ☐ In dem Histogramm ist eine Ordinalverteilung dargestellt.
- B** ☐ Wir haben eine Poisson-Verteilung vorliegen.
- C** ☐ Es handelt sich um eine Normalverteilung.
- D** ☐ Eine Standardnormalverteilung.
- E** ☐ Wir haben eine Gammaverteilung vorliegen.

Lineare Regression & Korrelation

15. Aufgabe

(2 Punkte)

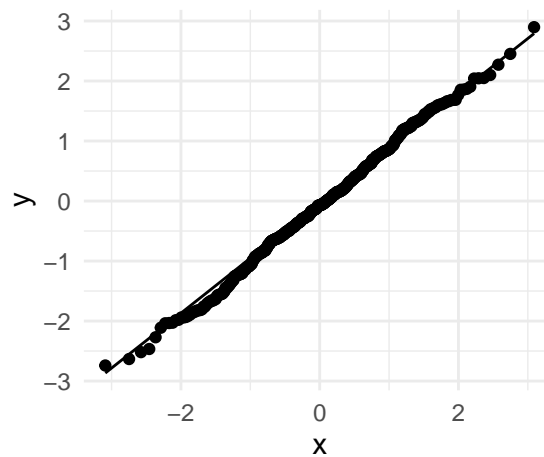
In Ihrer Abschlussarbeit wollen Sie ein kausales Modell rechnen. Jetzt stellt sich die Frage, was diese Entscheidung für Ihre Auswertung bedeutet. Welche Aussage ist richtig?

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

16. Aufgabe

(2 Punkte)

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

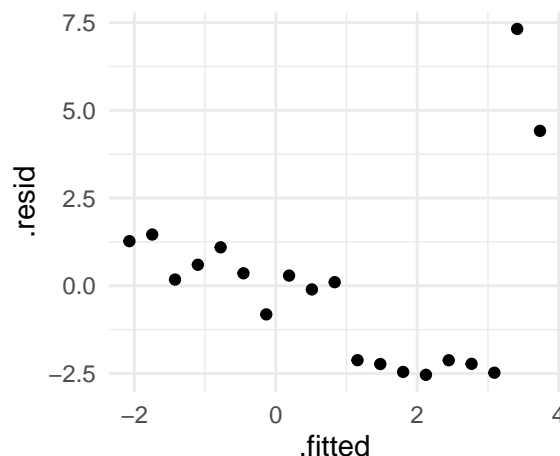


- A** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.
- B** ☐ Wir betrachten die Punkte. Wenn die Punkte einigermaßen gleichmäßig verteilt liegen, dann gehen wir von normalen Residuen aus.
- C** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Geraden.
- D** ☐ Wir betrachten die Gerade und dabei insbesondere die beiden Enden der Gerade in dem IQR, also dem ersten und dritten Quartile. Hier sollten die Punkte auf der Geraden liegen, dann ist die Annahme an die Normalverteilung der Residuen erfüllt.
- E** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.

17. Aufgabe

(2 Punkte)

Sie rechnen eine linearen Regression und erhalten folgende Abbildung der Residuen (.resid). Welche Aussage ist richtig?



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

- D** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Vereinzelte Punkte liegen oberhalb bzw. unterhalb der Geraden um die 0 Linie weiter entfernt. Ein klares Muster ist zu erkennen.
- E** ☐ 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 .

18. Aufgabe

(2 Punkte)

Sie berechnen in Ihrer Abschlussarbeit den Korrelationskoeffizienten ρ . Welche Aussage über den Korrelationskoeffizienten ρ ist richtig?

- A** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen -1 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos.
- B** ☐ 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.
- C** ☐ Der Korrelationskoeffizienten ρ zeigt keinen Zusammenhang zwischen zwei Variablen x und y bei einem Wert von 0. Einen negativen Zusammenhang Richtung -1 und somit auch einen positiven Zusammenhang Richtung 1. Je größer die Zahl allgemein, desto stärker der Effekt.
- 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 standardisierte, statistische Maßzahl, die zwischen 0 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos. Eine Signifikanz kann nicht nachgewiesen werden.

19. Aufgabe


(2 Punkte)


Sie haben ein Feldexperiment mit Wasserlinsen durchgeführt und wollen nun in einer simplen linearen Regression den Einfluss der CO_2 -Konzentration in $[\mu g]$ im Wasser auf den absoluten Proteingehalt in $[kg]$ untersuchen. Sie erhalten einen β_{CO_2} Koeffizienten von 6.9×10^{-7} und einen p -Wert mit 0.00051. Welche Aussage zu der Signifikanz und dem Effekt ist richtig?



- A** ☐ Das Gewicht und die CO_2 -Konzentration korrelieren sehr stark, deshalb wird der β_{CO_2} Koeffizient sehr klein. Mit einer ANOVA kann für die Korrelation korrigiert werden und der Effektschätzer passt dann zum p -Wert.
- B** ☐ Wenn der Effekt β_{CO_2} winzig ist, dann kann es an einer falsch gewählten Einheit liegen. Der Anstieg von einer Einheit in X führt ja zu einer Änderung von β_{CO_2} in x . Wir müssen daher die Einheit von y entsprechend anpassen.
- C** ☐ 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.
- D** ☐ Die Einheit der CO_2 -Konzentration ist zu klein gewählt. Dadurch sehen wir den sehr kleinen p -Wert. Der p -Wert und die Einheit von der CO_2 -Konzentration hängen antiproportional zusammen.
- E** ☐ Manchmal ist die Einheit der Einflussvariable X zu klein gewählt, so dass der Anstieg von 1 Einheit in X zu einer zu kleinen Änderung in y führt. Daher kann der Effekt β_{CO_2} sehr klein wirken, aber auf einer anderen Einheit sehr viel größer sein. Der p -Wert wird auf einer einheitslosen Teststatistik bestimmt.

20. Aufgabe

(2 Punkte)

Nachdem Sie Ihr Experiment abgeschlossen haben, stehen Sie vor der Frage wie Sie Ihre Daten modellieren sollen. In der Beispielauswertung von Ihrem Betreuenden finden Sie die Funktion `lm()` in . Welche Aussage ist richtig?


- A** ☐ Die Funktion `lm()` in  ist der erste Schritt für einen Gruppenvergleich. Danach kann eine ANOVA oder aber ein multipler Vergleich in `{emmeans}` gerechnet werden. In der Funktion `lm()` werden die Gruppenmittelwerte bestimmt.

- B** ☐ Die Funktion `lm()` berechnet die Varianzstruktur für eine ANOVA. Danach kann dann über eine explorative Datenanalyse nochmal eine Signifikanz berechnet werden. Sollte vor der Verwendung der Funktion `lm()` schon eine EDA gerechnet worden sein, so ist die Analyse wertlos.
- C** ☐ Ist die Einflussvariable X ein Faktor so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich. Die Funktion `lm()` kann dabei eigentlich weggelassen werden, wird aber traditionell gerechnet.
- D** ☐ Die Funktion `lm()` in  wird klassischerweise für die nicht-lineare Regression genutzt. Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt.
- E** ☐ Die Funktion `lm()` in  ist der letzte Schritt für einen Gruppenvergleich. Vorher kann eine ANOVA oder aber ein multipler Vergleich in `{emmeans}` gerechnet werden. In der Funktion `lm()` werden die Gruppenvarianzen bestimmt.

21. Aufgabe

(2 Punkte)

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

- A** ☐ In  ist mit dem *generalisierten linearen Modell (GLM)* eine Modellierung implementiert, die neben der klassischen Normalverteilung auch die Poissonverteilung für Zähldaten oder die Binomialverteilung für 0/1-Daten modellieren kann.
- B** ☐ Das GLM ist eine allgemeine Erweiterung der linearen Regression auf die Normalverteilung.
- C** ☐ Das GLM ist eine Vereinfachung des LM in R. Mit dem GLM lassen sich polygonale Regressionen rechnen. Somit stehen neben der Normalverteilung noch weitere Verteilungen zu Verfügung.
- D** ☐ Das GLM ist ein faktisch maschineller Lernalgorithmus, der selbstständig die Verteilungsfamilie für Y wählt.
- E** ☐ Das GLM erlaubt auch nicht normalverteilte Residuen in der Schätzung der Regressionsgrade.

Vermischte Themen

22. Aufgabe


(2 Punkte)

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






- A** ☐ Strukturgleichheit ist durch Randomisierung gegeben. Somit kann von der Stichprobe auf die Grundgesamtheit geschlossen werden
- B** ☐ Randomisierung ist die direkte Folge von Strukturgleichheit. Die Strukturgleichheit erlaubt es erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- C** ☐ Randomisierung erlaubt erst die Mittelwerte zu schätzen. Ohne Randomisierung keine Mittelwerte. Ohne Mittelwerte keine Varianz und somit auch kein statistischer Test.
- D** ☐ Randomisierung bringt starke Unstrukturiertheit in das Experiment und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- E** ☐ Durch eine Randomisierung können wir nicht von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.

23. Aufgabe

(2 Punkte)



Sie wollen Ihren Datensatz in  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** ☐ 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.

- B** ☐ Die  Pakete sind nur in englischer Sprache verfasst. Das ist aber nicht der Hauptgrund, denn  hat wie alle Programmiersprachen Probleme mit Umlauten und Sonderzeichen.
- C** ☐ Programmiersprachen haben Probleme mit Umlauten und Sonderzeichen der deutschen Sprache. Daher ist die Nutzung in Deutsch in den AGBs von  untersagt.
- D** ☐ Es gibt keinen Grund nicht auch deutsche Wörter zu verwenden. Es ist ein Stilmittel.
- E** ☐  Pakete sind nur in englischer Sprache verfasst. Es macht keinen Sinn  daher in Deutsch zu bedienen.

24. Aufgabe

(2 Punkte)

Nachdem Sie Ihr Feldexperiment als Vorversuch für Ihre Abschlussarbeit abgeschlossen haben, wollen Sie in einer explorativen Datenanalyse (EDA) in  einmal schauen, ob Sie überhaupt Effekte der Behandlung vorliegen haben. Welche Reihenfolge von Schritten müssen Sie in  durchführen, damit Sie eine EDA rechnen können?

- A** ☐ Für eine explorativen Datenanalyse (EDA) in  müssen wir als erstes die Daten über `read_excel()` einlesen. Danach müssen wir schauen, dass wir die Zeilen richtig über `mutate()` transformiert haben. Insbesondere müssen Variablen mit kontinuierlichen Werten in einen Faktor umgewandelt werden. Am Ende nutzen wir die Funktion `ggplot()` für die eigentliche EDA.
- B** ☐ Wir lesen die Daten über eine generische Funktion `read()` ein und müssen dann die Funktion `ggplot()` nur noch installieren. Dann haben wir die Abbildungen als *.png vorliegen.
- C** ☐ 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.
- D** ☐ Wir lesen als erstes die Daten über `read_excel()` ein, transformieren die Spalten über `mutate()` in die richtige Form und können dann über `ggplot()` uns die Abbildungen erstellen lassen. Wichtig ist, dass wir keine Faktoren sondern nur numerische Variablen vorliegen haben.
- E** ☐ Wir transformieren die Spalten über `mutate()` in ein `tibble` und können dann über `ggplot()` uns die Abbildungen erstellen lassen. Dabei beachten wir das wir keine Faktoren in den Daten haben.

25. Aufgabe

(2 Punkte)

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

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

26. Aufgabe

(2 Punkte)

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

- A** ☐ 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.
- B** ☐ Untereinander abhängig, wenn die Mütter ebenfalls miteinander verwandt sind. Erst die Abhängigkeit 2. Grades wird in der Statistik modelliert.
- C** ☐ Untereinander abhängig. Die Ferkel stammen von einem Muttertier und haben vermutlich eine ähnliche Varianzstruktur.
- D** ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.
- E** ☐ Untereinander stark korreliert. Die Ferkel sind von einer Mutter und somit miteinander korreliert. Dies wird in der Statistik jedoch meist nicht modelliert.

27. Aufgabe

(2 Punkte)

In einer Studie wollen Sie den Effektschätzer Odds ratio berechnen. Sie finden in Ihrem Experiment zur Behandlung von Klaueninfektionen bei Schweinen in 5 Tieren Erkrankung der Klauen vor. 8 Tiere sind gesund. Welche Aussage ist richtig?

- A** ☐ Es ergibt sich ein Odds ratio von 0.38, da es sich um eine Chancenverhältnis handelt.
- B** ☐ Der Anteil der Gesunden wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Odds ratio von 0.38.
- C** ☐ Das Verhältnis der Anteile Odds ratio ergibt ein Anteilsverhältnis von 0.38. Wir sind am Anteil der Kranken interessiert.
- D** ☐ Es ergibt sich ein Odds ratio von 1.6, da es sich um ein Anteil handelt.
- E** ☐ Es ergibt sich ein Odds ratio von 0.62, da es sich um eine Chancenverhältnis handelt

28. Aufgabe

(2 Punkte)

Historisch gesehen ergibt sich ein Problem, wenn Sie mit sehr großen Datensätzen, wie in der Bio Data Science üblich, rechnen. Warum ist es ein Problem, wenn Ihre Datensätze sehr groß werden hinsichtlich der Bewertung anhand der Signifikanz?

- A** ☐ Aktuell werden immer größere Datensätze erhoben. Dadurch wird auch die Varianz immer höher was automatisch zu mehr signifikanten Ergebnissen führt.
- 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** ☐ Riesige Datensatz haben mehr Fallzahl was zur α -Inflation führt. Durch eine Adjustierung kann dem Problem entgegengewirkt werden.
- D** ☐ Eine erhöhte Fallzahl führt automatisch zu mehr signifikanten Ergebnissen auch wenn der Effekt klein ist und damit nicht relevant. Dadurch sind die Informationen zur Signifikanz in riesigen Datensätzen schwer zu verwerten, da fast alle Vergleiche signifikant sind.
- E** ☐ 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.

Multiple Gruppenvergleiche

29. Aufgabe

(2 Punkte)

Sie haben folgende unadjustierten p-Werte gegeben: 0.89, 0.01, 0.03, 0.001 und 0.34. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 1, 0.05, 0.15, 0.005 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.178, 0.002, 0.006, 2e-04 und 0.068. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 1, 0.05, 0.15, 0.005 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 4.45, 0.05, 0.15, 0.005 und 1.7. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.178, 0.002, 0.006, 2e-04 und 0.068. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.

30. Aufgabe

(2 Punkte)

Die Abkürzung *CLD* steht für welches statistische Verfahren? Welche folgende Beschreibung der Interpretation ist korrekt?

- 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** ☐ 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.
- C** ☐ Compact letter detection. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt.
- D** ☐ Compact letter display. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt. Teilweise ist die Interpretation des CLD herausfordernd, da wir ja nach dem Unterschied suchen.
- E** ☐ Contrast letter display. Unterschiede in den Behandlungen werden durch den gleichen Buchstaben oder Symbol dargestellt. Die Interpretation des CLD führt häufig in die Irre.

31. Aufgabe

(2 Punkte)

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

- A** ☐ Das R Paket `{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.
- B** ☐ 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.
- C** ☐ 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.
- 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 `{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.

32. Aufgabe

(2 Punkte)


Bei einem Posthoc-Test kann es zu einer überraschenden Besonderheit beim statistischen Testen kommen. Wie lautet der Fachbegriff und wie kann mit der überraschenden Besonderheit umgegangen werden?

- 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 β -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 20%. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist.
- C** ☐ Beim multiplen Testen kann es zu Varianzheterogenität kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5%. Daher müssen die p-Werte entsprechend adjustiert werden. Das Verfahren nach Welch, bekannt aus dem t-Test, ist hier häufig anzuwenden.

- D** ☐ 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.
- E** ☐ Die Adjustierung der p-Werte nach Bonferroni erlaubt es gegen die α -Inflation vorzugehen, die häufig beim multiplen Testen auftritt. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern sehr viel höher. Das ist der Grund warum die p-Werte entsprechend adjustiert werden müssen.

33. Aufgabe

(2 Punkte)

In Ihrer Bachelorarbeit werten Sie einen einfaktoriellen Versuch aus. Dafür rechnen Sie in  zunächst eine ANOVA und schließen dann einen multiplen Vergleich mit t-Tests an. Welche Aussage über die Effekte in Ihrem Versuch ist richtig?

- A** ☐ 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.
- B** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ adjustiert werden im Gegensatz zu den p-Werten.
- C** ☐ Beim multiplen Testen kann es zu einer Δ -Deflation kommen. Das globale Relevanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die Δ -Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der Δ -Werte nach Bonferroni das bekannteste Verfahren ist. Die Δ -Werte werden durch die Anzahl an Vergleichen geteilt.
- D** ☐ 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.
- E** ☐ Beim multiplen Testen muss der Effekt, hier der Mittelwertsunterschied Δ aus den paarweisen t-Tests, nicht adjustiert werden.

Statistische Testtheorie

34. Aufgabe

(2 Punkte)

Welche Aussage zum mathematischen Ausdruck $Pr(D|H_0)$ ist richtig?

- A** ☐ Die Wahrscheinlichkeit der Daten unter der Nullhypothese in der Grundgesamtheit.
- B** ☐ Die Inverse der Wahrscheinlichkeit unter der die Nullhypothese nicht mehr die Alternativhypothese überdeckt.
- C** ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit der Alternativhypothese und somit $1 - Pr(H_A)$
- D** ☐ Die Wahrscheinlichkeit für die Nullhypothese, wenn die Daten wahr sind.
- E** ☐ $Pr(D|H_0)$ stellt die Wahrscheinlichkeit die Daten D und somit die Teststatistik T_D zu beobachten dar, wenn die Nullhypothese wahr ist.

35. Aufgabe

(2 Punkte)

Das statistische Testen basiert auf dem Falsifikationsprinzip. Es besagt,

- A** ☐ ... dass Modelle meist falsch sind und selten richtig.
- B** ☐ ... dass ein schlechtes Modell durch das Falsifikationsprinzip durch ein weniger schlechtes Modell ersetzt wird.
- C** ☐ ... dass Fehlerterme in statistischen Modellen nicht verifiziert werden können.
- D** ☐ ... dass ein minderwertiges Modell durch ein minderwertiges Modell ersetzt wird. Es gilt das Verifikationsprinzip nach Karl Popper.
- E** ☐ ... dass in der Wissenschaft immer etwas falsch sein muss. Sonst gebe es keinen Fortschritt.

36. Aufgabe

(2 Punkte)

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

- A** ☐ 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.
- B** ☐ Der Begründer der modernen Statistik, R. Fischer, hat die Grenze simuliert und berechnet. Dadurch ergibt sich dieser optimale Cut-Off.
- C** ☐ 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.
- D** ☐ 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.
- E** ☐ 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.

37. Aufgabe

(2 Punkte)

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

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

38. Aufgabe

(2 Punkte)

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

H_0 ablehnen obwohl die H_0 gilt

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

39. Aufgabe

(2 Punkte)

Sie lesen eine wissenschaftliche Arbeit, die damit wirbt, dass Effekte und Signifikanz nicht separat dargestellt sind, sondern in einer statistischen Maßzahl zusammen. Welche Aussage ist richtig?

- A** ☐ Einem Konfidenzintervall. Das Konfidenzintervall bringt durch eine Visualisierung und drei Intervallgrenzen die Möglichkeit mit, eine Relevanzschwelle neben der Signifikanzschwelle und der α -Schwelle zu definieren.
- B** ☐ 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.
- C** ☐ 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.
- D** ☐ Über das Konfidenzintervall. Das Konfidenzintervall inkludiert eine Entscheidung über die Relevanz und zusätzlich kann über die Visualisierung des Konfidenzintervalls eine Signifikanzschwelle vom Forschenden definiert werden.
- E** ☐ Das Konfidenzintervall. Durch die Visualisierung des Konfidenzintervalls kann eine Relevanzschwelle vom Anwender definiert werden. Zusätzlich erlaubt das Konfidenzintervall auch eine Entscheidung über die Signifikanz.

40. Aufgabe

(2 Punkte)

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

- A** ☐ Wir vergleichen mit dem p -Wert und dem Signifikanzniveau α Wahrscheinlichkeiten und damit die Flächen unter der Kurve der Teststatistik, wenn die H_0 gilt.
- B** ☐ Wir machen ein Aussage über die Flächen und zwischen den Kurve der Teststatistiken der Hypothesen H_0 und H_A , wenn die H_0 gilt. Dabei werden Wahrscheinlichkeiten verglichen, die durch die Flächen unter der Kurve repräsentiert werden.
- C** ☐ 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.
- D** ☐ Wir vergleichen die Effekte des p -Wertes mit den Effekten der Signifikanzschwelle unter der Annahme der Nullhypothese. Dabei gilt, dass wir die Nullhypothese nur ablehnen können anhand des Falsifikationsprinzips.
- E** ☐ Wir schauen, ob der p -Wert 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.

41. Aufgabe

(2 Punkte)

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

- A** ☐ In der Analogie der Regenwahrscheinlichkeit in einem bestimmten Gebiet: ein statistischer Test gibt die Wahrscheinlichkeit für ein Ereignis in einem Experiment mit den Daten D wieder und lässt sich kaum verallgemeinern.
- B** ☐ 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.
- C** ☐ In der Analogie der Maximaltemperatur: Was ist der maximale Unterschied zwischen zwei Gruppen. Wir erhalten hier eine Aussage über die Spannweite und den maximalen Effekt.
- D** ☐ Die Analogie der Regenwahrscheinlichkeit: der statistische Test erlaubt es die Wahrscheinlichkeit für Regen abzuschätzen jedoch nicht die Menge und somit den Effekt.
- E** ☐ In der Analogie des Niederschlags oder Regenmenge: ein statistischer Test gibt die Stärke eines Effektes wieder. Zum Beispiel, wie hoch ist der Mittelwertsunterschied.

42. Aufgabe

(2 Punkte)

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

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

43. Aufgabe

(2 Punkte)

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

- A** ☐ 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
- B** ☐ Alle statistischen Tests sind so konstruiert, dass die H_A mit 20% bewiesen wird. Die Power ist $1 - \beta$ mit β gleich 80% gesetzt.
- C** ☐ Die Power wird berechnet und ist keine Eigenschaft des Tests. Die Power wird auf 80% gesetzt und beschreibt mit welcher Wahrscheinlichkeit H_0 bewiesen wird
- D** ☐ Es gilt $\alpha + \beta = 1$ und somit liegt β meist bei 95%.
- E** ☐ Die Power $1 - \beta$ wird auf 80% gesetzt. Damit liegt die Wahrscheinlichkeit für die H_0 bei 20%.

44. Aufgabe

(2 Punkte)

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

- A** ☐ Der Forschende muss am Anfang wissen, ob das Ergebnis eines Experiments relevant für seine Forschung ist. Dafür kann der Effekt eines statistischen Tests genutzt werden oder auch der Prähoc-Test. Damit beschreibt der Effekt den biologischen interpretierbaren Teil eines Experiments vor der Durchführung. Zum Beispiel der Unterschied zwischen zwei Mittelwerten.
- B** ☐ Durch den Effekt erfahren wir die statistische interpretierbare Ausgabe eines statistischen Tests. Zum Beispiel das η^2 aus einer ANOVA. Damit können wir die Signifikanz direkt mit dem Effekt verbinden. Am Ende muss der Forschende aber entscheiden, ob der Effekt entsprechend seinen Erwartungen als bedeutet zu bewerten ist.
- C** ☐ 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.
- D** ☐ 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.
- E** ☐ Durch den Effekt erfahren wir die biologisch interpretierbare Ausgabe eines statistischen Tests. Zum Beispiel das η^2 aus einer ANOVA. Damit können wir die Relevanz direkt mit dem Effekt verbinden. Am Ende muss der Forschende aber entscheiden, ob der Effekt entsprechend seinen Erwartungen als bedeutet zu bewerten ist.

45. Aufgabe

(2 Punkte)

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

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

46. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit müssen Sie für die statistischen Tests im Anhang Ihrer Arbeit die Hypothesen H formulieren. Welche Aussage über Hypothesen H ist richtig

- A** ☐ 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.
- B** ☐ Mit der Nullhypothese H_A und der Alternativhypothese H_0 gibt es zwei Hypothesen, die aber selten genutzt werden.
- C** ☐ Mit der Nullhypothese H_0 und der Alternativhypothese H_A oder H_1 gibt es zwei Hypothesen.
- D** ☐ Es gibt - bedingt durch das das Falsifikationsprinzip - ein Set von k Nullhypothesen, die iterative gegen $k - 1$ Alternativhypothesen getestet werden.
- 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.

Statistische Tests für Gruppenvergleiche

47. 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 Gruppen indem die Mittelwerte miteinander verglichen werden.
- B** ☐ Der t-Test vergleicht zwei oder mehr Gruppen indem die Mittelwerte miteinander verglichen werden.
- C** ☐ Der t-Test vergleicht die Varianzen von mindestens zwei oder mehr Gruppen
- D** ☐ 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.
- E** ☐ Der t-Test testet generell zu einem erhöhten α -Niveau von 20%.

48. Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung des Mikronährstoff Eisen auf den Ertrag in t/ha von Mango im Vergleich zu einer Kontrolle. Der Versuch wurde in 10 Parzellen pro Gruppe durchgeführt. Welche Aussage im Bezug auf eine statistische Auswertung ist richtig?



- ☐ A Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei -8. Wir müssen aber einen Posthoc-Test rechnen um den Effekt wirklich bestimmen zu können.
- ☐ B Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt kann nicht bei einem t-Test aus Barplots bestimmt werden.
- ☐ C 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.
- ☐ D Der Test deutet auf einen signifikanten Unterschied hin. Der Effekt liegt vermutlich bei -8.
- ☐ E Der Test deutet auf keinen signifikanten Unterschied hin. Der Effekt liegt vermutlich bei -8.

49. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit betrachten Sie die Effekte von einer Behandlung vor und nach der Gabe eines Vitamins. Sie müssen einen gepaarten t-Test rechnen. Welche Aussage ist richtig?

- ☐ A Der gepaarte t-Test nutzt die Varianz der Beobachtungen jeweils paarweise und bildet dafür eine verbundene Stichprobe. Dieser Datensatz d dient dann zur Differenzbildung.
- ☐ B Der gepaarte t-Test wird genutzt, wenn die Differenzen der Beobachtungen verbunden sind und wir dadurch die Unabhängigkeit nicht mehr vorliegen haben.
- ☐ C Wenn die Beobachtungen nicht unabhängig voneinander sind, rechnen wir einen gepaarten t-Test. Messen wir wiederholt an dem gleichen Tier oder Pflanze dann bilden wir die Differenz zwischen den zwei Messpunkten.
- ☐ D Wenn die Beobachtungen unabhängig voneinander sind, rechnen wir einen gepaarten t-Test. Messen wir wiederholt an dem gleichen Tier oder Pflanze dann bilden wir das Produkt zwischen den zwei Messpunkten.
- ☐ E 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.

50. Aufgabe

(2 Punkte)

Sie führen paarweise t-Tests für alle Vergleiche der verschiedenen Rapssorten in Ihrem Experiment durch. Nach der Adjustierung für multiples Testen ist kein p-Wert unter der α -Schwelle. Ihr Experiment beinhaltet drei Rapssorten und eine ANOVA ergibt $p = 0.048$ für den Ertrag. Sie schauen sich auch die rohen, unadjustierten p-Werte an und finden hier als niedrigsten p-Wert $p_{3-2} = 0.052$. Welche Aussage ist richtig?

- ☐ A Hier kommt der Effekt der steigenden Fallzahl auf die Anzahl an signifikante Ergebnisse zu tragen. Da die ANOVA auf mehr Fallzahl testet als die einzelnen paarweisen t-Tests, kann die ANOVA leichter einen signifikanten Unterschied nachweisen. Die p-Werte sind immer etwas kleiner als bei den t-Tests.

- B** ☐ Es gibt einen Fehler in der Varianzstruktur. Daher kann die ANOVA nicht richtig sein und paarweise t-Tests liefern das richtige Ergebnis.
- C** ☐ Der Fehler liegt in den t-Tests. Wenn eine ANOVA signifikant ist, dann muss zwangsweise auch ein t-Test signifikant sein.
- D** ☐ Hier kommt der Effekt der steigenden Fallzahl auf die Anzahl an signifikante Ergebnisse zu tragen. Da die ANOVA auf weniger Fallzahl testet als die paarweisen t-Tests, kann die ANOVA schwerer einen signifikanten Unterschied nachweisen.
- E** ☐ Das Beispiel kann so nicht auftreten, da die ANOVA und die t-Tests algorithmisch miteinander ver-schränkt sind.

Teil I.

Deskriptive Statistik & Explorative Datenanalyse

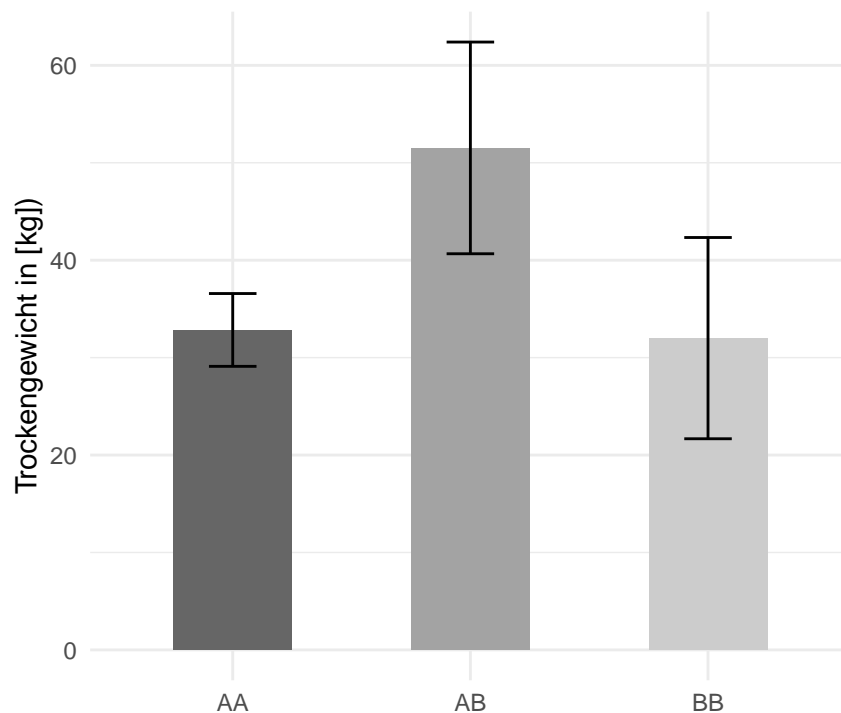
51. Aufgabe


(8 Punkte)


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



Zerforschen des Barplots Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Mark ist es eine Möglichkeit schneller ans Ziel zu gelangen. Mark soll in seiner Hausarbeit Erbsen untersuchen. Die Behandlung in seiner Hausarbeit werden verschiedene Genotypen (AA, AB und BB) sein. Erheben wird Mark als Outcome (Y) *Trockengewicht* benannt als *drymatter* in seiner Exceldatei. Von seiner Betreuerin erhält er nun folgende Abbildung von Barplots, die er erstmal zur Übung nachbauen soll, bevor er mit dem eigentlichen Versuch beginnt.



Leider kennt sich Mark 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 Mark einen Unterschied zwischen den Behandlungen erwarten? Begründen Sie Ihre Antwort! **(2 Punkte)**

52. Aufgabe

(8 Punkte)

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



Visualisierung des Barplots Jonas steht vor einem ersten Problem, denn wenn es nach seinem Betreuer geht, soll er in einem Freilandversuch Kartoffeln auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Jonas liebt Stricken. Darin kann er sich wirklich verlieren und immer wieder neu begeistern. Die Behandlung waren verschiedene Substrattypen (*torf*, *40p60n* und *70p30n*). In seiner Exceldatei hat er den Endpunkt (*Y*) *Proteingehalt* als *protein* aufgenommen. Nun soll Jonas die Daten einmal als Barplots in einer Präsentation visualisieren, damit seinem Betreuer 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. Wenn die Erschöpfung nicht wäre, ja dann wäre wohl vieles möglich für Jonas! Aber so.. Aber egal. Einfach mal raus um zu Schwimmen. Ohne Ziel und Uhr. Das ist was für Jonas.

treatment	protein
torf	26.0
70p30n	34.2
torf	24.7
40p60n	29.1
40p60n	29.9
torf	29.2
70p30n	33.9
torf	35.6
70p30n	29.4
torf	29.4
40p60n	32.3
40p60n	29.0

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

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

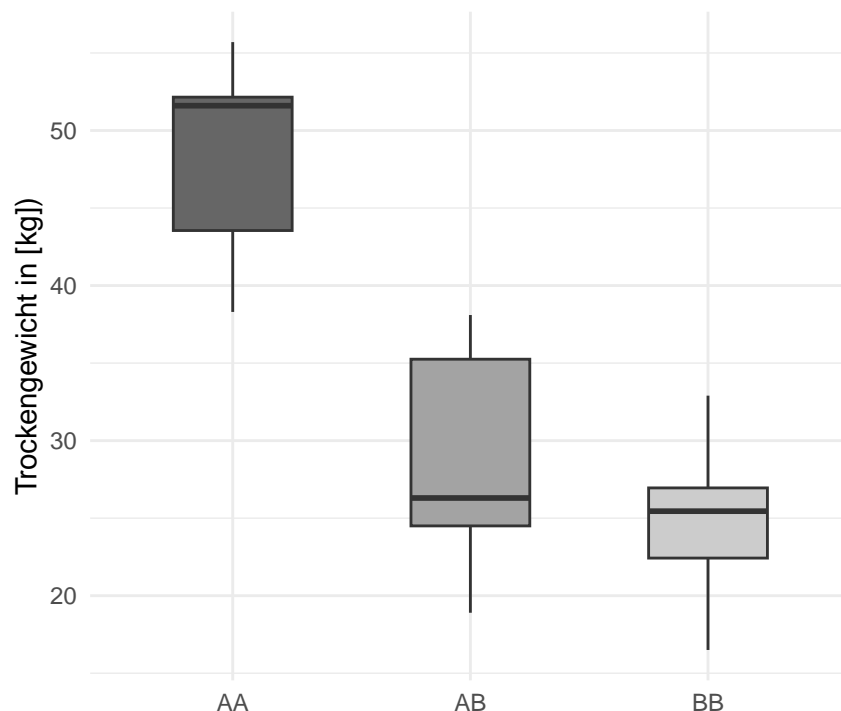
53. Aufgabe

(9 Punkte)

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



Zerforschen des Boxplots Eine echte Herausforderung für sie war schon immer der Perfektionismus gewesen. Ein leidiges Lied. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Paula ist es eine Möglichkeit schneller ans Ziel zu gelangen. Paula soll in ihrer Abschlussarbeit Maiss untersuchen. Die Behandlung in ihrer Abschlussarbeit werden verschiedene Genotypen (AA, AB und BB) sein. Erheben wird Paula als Endpunkt (Y) *Trockengewicht* benannt als *drymatter* in ihrer Exceldatei. Von ihrem Betreuer erhält sie nun folgende Abbildung von Boxplots, die sie erstmal zur Übung nachbauen soll, bevor sie mit dem eigentlichen Versuch beginnt. Aber nur in passender Atmosphäre! Hm, lecker Smarties und dazu dann im Hintergrund Jagd auf roter Oktober laufen lassen.



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

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

54. Aufgabe

(9 Punkte)

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



Visualisierung des Boxplots Boxplots sind bedeutend in der Darstellung von wissenschaftlichen Ergebnissen. Leider hat sich Steffen nicht gemerkt, welche statistischen Maßzahlen für einen Boxplot erhoben werden müssen. Besser wäre was anderes gewesen. Klemmbausteine. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Steffen denkt gerne über Klemmbausteine nach. Das ist in soweit doof, da nach seiner Betreuerin nun Boxplots aus seinen 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 Erbsen waren verschiedene Substrattypen (*torf* und *70p30n*). Erfasst wurde von Steffen als Outcome (Y) *Proteingehalt*. Steffen hat dann *protein* in seiner Exceldatei eintragen. Aber nur in passender Atmosphäre! Hm, lecker Oreos und dazu dann im Hintergrund Harry Potter laufen lassen.

treatment	drymatter
70p30n	34.6
70p30n	33.0
70p30n	41.3
torf	23.2
torf	23.1
70p30n	24.0
torf	12.4
torf	31.3
70p30n	35.6
torf	29.1
torf	28.5
torf	16.6
70p30n	34.0
torf	10.8
70p30n	38.0
torf	25.7
70p30n	33.4
70p30n	33.7
torf	15.5

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

1. Zeichnen Sie in *einer* Abbildung die beiden Boxplots für die zwei Behandlungen von Erbsen! 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 Erbsen erwarten würden, wie sehen dann die beiden Boxplots aus? Antworten Sie mit einer Skizze der Boxplots! **(1 Punkt)**

55. Aufgabe

(8 Punkte)

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



Visualisierung des Histogramm für kategoriale Daten Alex schmeißt noch eine Handvoll Gummibärchen in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Abba. Alex betrachtet die folgenden Daten nach einem Feldexperiment mit Brokkoli. In dem Experiment wurden die Läsionen auf den Blättern gezählt. Nach der Meinung seiner Betreuerin muss als erstes geschaut werden, wie diese verteilt sind. Also welcher statistischen Verteilung die Läsionen auf den Blättern folgen. Dazu soll Alex ein Histogramm verwenden. Dann hätte man auch einen guten Überblick über den Messwert (Y). Es wäre einfacher, wenn da nicht noch was wäre. Wenn die Gefälligkeit nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so.. Wenn Abba ertönt, dann sucht die Katze schleunigst Schutz unter dem Sofa. Alex schüttelt den Kopf.

Die Läsionen auf den Blättern: 5, 3, 3, 6, 3, 5, 3, 3, 8, 2, 5, 1, 3, 5, 4, 3, 2, 4, 0, 6, 2, 6, 6, 2, 4, 7, 2, 3, 4, 1, 2, 0, 3, 6, 7, 4

Leider kennt sich Alex 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)**

56. Aufgabe

(8 Punkte)

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



Visualisierung des Histogramm für kontinuierliche Daten In seiner Hausarbeit möchte Jonas gerne die Daten aus einem Leistungssteigerungsversuch mit Schweinen in einem Histogramm darstellen. Das Histogramm erlaubt ihm dabei Rückschlüsse auf die Verteilung über den Messwert (Y) zu treffen. Jonas schmeißt noch eine Handvoll Snickers in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Iron Maiden. In seinem Experiment hat Jonas die mittleren auffälligen Hautflecken gezählt. Es wäre einfacher, wenn da nicht noch was wäre. Jonas und die Erschöpfung, eine unendliche Geschichte mit kniffligen Wendungen. Wenn Iron Maiden ertönt, dann sucht das Meerschweinchen schleunigst Schutz unter dem Sofa. Jonas schüttelt den Kopf.

Die mittleren auffälligen Hautflecken: 10.1, 8.8, 11.3, 4.8, 10.6, 8.2, 8.5, 9.3, 12.3, 10.5, 10, 11.9, 7.1, 8.4, 11.6, 12.1, 11.2, 11.9, 7.8, 13.5, 11.9, 11.6, 7.8, 10.7, 12.9

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

57. Aufgabe

(10 Punkte)

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



Visualisierung des Scatterplots Mark liest laut: 'Wenn zwei kontinuierliche Variablen vorliegen, können diese in einer explorativen Datenanalyse...'. Mark stoppt. Mark schmeißt noch eine Handvoll Marzipankugeln in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Andrea Berg. Was waren noch gleich kontinuierliche Variablen? In seiner Hausarbeit hatte er ein Stallexperiment im Teuteburgerwald durchgeführt. Dabei ging es um den Zusammenhang zwischen Fettgehalt [%/kg] und mittlerer Anzahl an weißen Blutkörperchen [LEU/ml] im groben Kontext von Zandern. Nun stellt sich die Frage für ihn, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Dafür war eine explorative Datenanalyse gut! Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. Dann was anderes. Wenn Columbo läuft, dann ist der Hamster nicht mehr da. Aber jetzt braucht er mal Entspannung!

Fettgehalt [%/kg]	Mittlerer Anzahl an weißen Blutkörperchen [LEU/ml]
25.2	15.3
17.0	11.3
23.5	12.9
28.2	19.6
24.6	17.3
24.7	18.7
25.5	16.6
28.5	26.7
27.2	16.2
31.6	18.3

Leider kennt sich Mark 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 *kein* Effekt von x auf y vorhanden wäre, wie würde die Gerade verlaufen und welche Werte würden die statistischen Maßzahlen annehmen? **(2 Punkt)**

58. Aufgabe

(10 Punkte)

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



Visualisierung des Mosaicplots Das Verrückte ist, dass das Huhn Star Trek wirklich liebt. Das ist Nilufar sehr recht, denn sie braucht Entspannung. Aber Ablenkung hilft nur begrenzt. 'Uff!', denkt sich Nilufar. Jetzt hat sie doch tatsächlich zwei kategoriale Variablen in ihrem Projektbericht gemessen. Zum einen die Behandlung Außenklimakontakt [ja/nein] und zum anderen die Messung Protein/Fettrate im Zielbereich [ja/nein] im Kontext von Zandern. Hierfür hat sie ein Kreuzungsexperiment im Wendland durchgeführt. Jetzt möchte Nilufar die Daten einmal in einer explorativen Datenanalyse darstellen. Danach kann sie dann über den passenden statistischen Test nachdenken. Dabei unterstützt ihre Betreuerin diesen Ansatz bevor es in der Datenanalyse weiter geht. So schön wie so gut. Eine echte Herausforderung für sie war schon immer die Erwartung gewesen. Ein leidiges Lied.

Außenklimakontakt	Protein/Fettrate im Zielbereich	Außenklimakontakt	Protein/Fettrate im Zielbereich
ja	ja	ja	nein
nein	nein	ja	ja
ja	nein	nein	nein
nein	nein	ja	nein
nein	nein	ja	ja
nein	nein	nein	nein
ja	nein	nein	nein
nein	nein	ja	ja
nein	nein	ja	nein
nein	nein	ja	ja
nein	ja	ja	ja
ja	ja	ja	nein
nein	nein	nein	nein
ja	ja	nein	nein
nein	nein	ja	ja
ja	ja	ja	ja
ja	nein	nein	nein
nein	nein	nein	nein

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

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

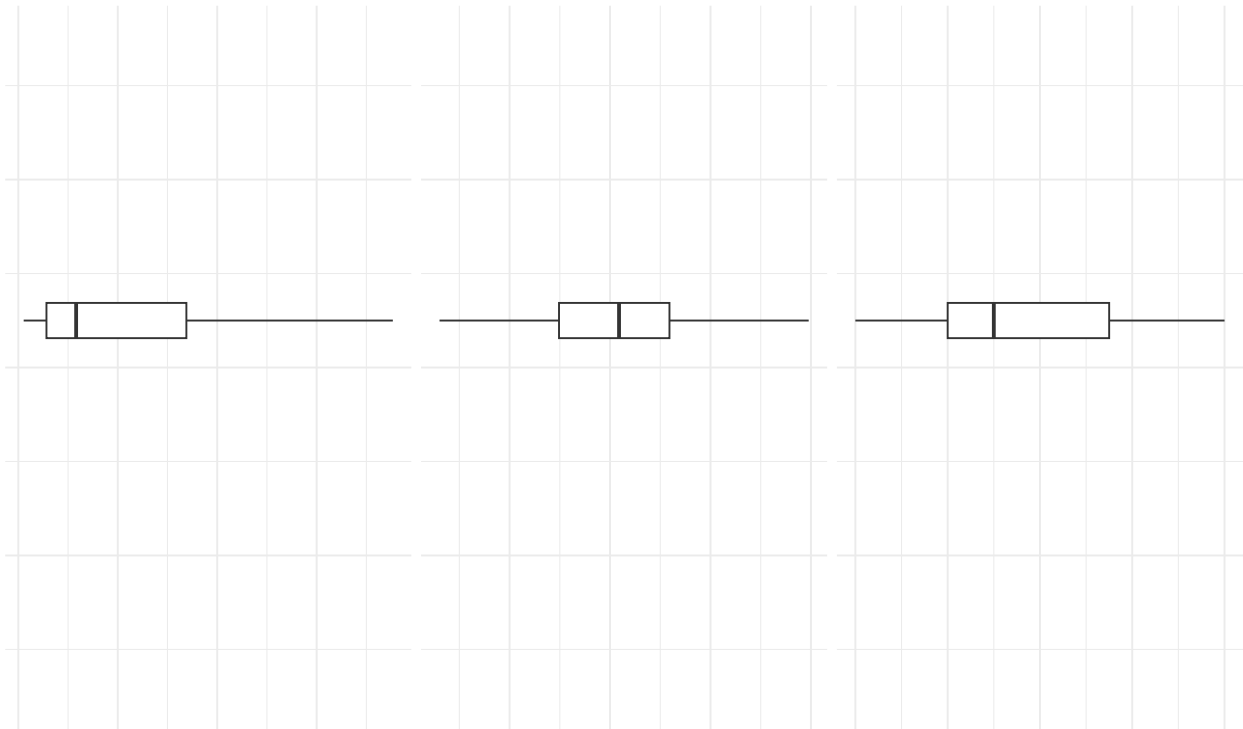
59. Aufgabe

(10 Punkte)

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



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

60. Aufgabe

(10 Punkte)

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



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

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

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

61. Aufgabe

(10 Punkte)

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

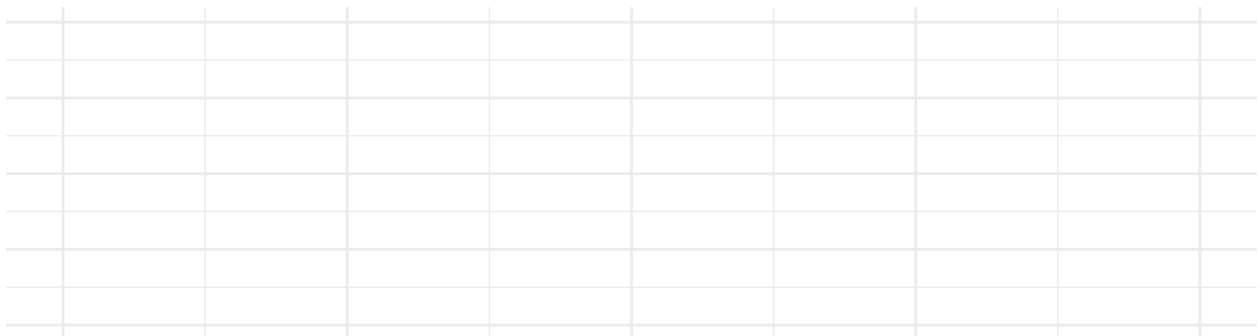


Visualisierung der Normalverteilung und der Poissonverteilung 'Wer hat sich denn sowas ausgedacht? Verteilungen?', entfährt es Jonas und schaut dabei Yuki an. In seiner Hand zerdrückt er Reese's Peanut Butter Cups von Yuki. 'Wir sollen eine Normalverteilung mit einem Mittelwert von $\bar{y}_1 = 4$ und einer Standardabweichung von $s_1 = 0.25$ zeichnen. Sowie eine weitere Normalverteilung mit einem Mittelwert von $\bar{y}_2 = 1$ und einer Standardabweichung von $s_2 = 0.25$. Keine Ahnung wie das geht. Darunter sollen dann noch eine Poissonverteilung mit einem Mittelwert von $\lambda_1 = 20$ sowie einer weiteren Poissonverteilung mit einem Mittelwert von $\lambda_2 = 2$ gezeichnet werden.', meint Yuki sichtlich eingeschüchtert und rettet noch ein paar Reese's Peanut Butter Cups in seinem Mund. Jonas und die Erschöpfung machen die Suche nach der Lösung nicht einfacher. Im Hintergrund spielt viel zu laut London Grammar, die diesmal Jonas ausgewählt hat und das Meerschweinchen schon in die Küche vertrieben hat, wo es den Biomüll mampft.

Zwei Normalverteilungen



Zwei Poissonverteilungen



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

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

Teil II.

Statistisches Testen & statistische Testtheorie

62. Aufgabe

(9 Punkte)



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

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

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

63. Aufgabe

(9 Punkte)



Das Nullritual - Die statistische Testtheorie 'Das Känguruh hat mir gerade zugezwinkert. Das macht mir Angst', bemerkt Tina. Es reicht ja schon ein Problem. Wenn die Wut nicht wäre, ja dann wäre wohl vieles möglich für Tina! 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 Katjes von dir. Bäh!', antwortet Jonas. Beide sind im Zoo und wollen sich von der statistische Testtheorie ablenken lassen. Eigentlich wollte ja Tina stoppen wie lange Tiere pinkeln², scheiterte aber an einer Oma mit Stock, die die beiden beschimpfte.

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

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

α -Fehler 20% H_0 wahr Testentscheidung

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

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

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

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

64. Aufgabe

(9 Punkte)

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

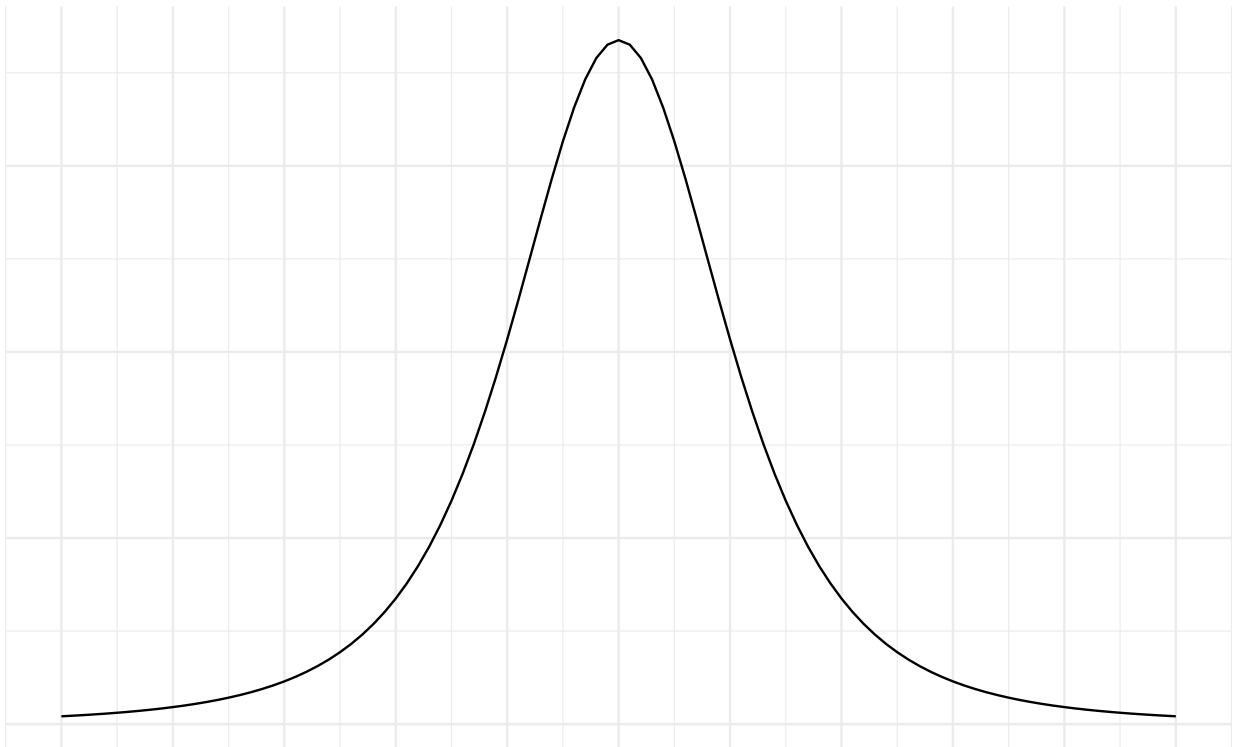


Visualisierung der Teststatistik T_D und dem p-Wert 'Kannst du mir nochmal an einer Visualisierung erklären, wie der Zusammenhang zwischen der Teststatistik aus den Daten T_D und dem p-Wert ist? Ich habe hier zig Fachbegriffe, kriege die aber nicht zusammen...', fragt Paula nachdrücklich Steffen. Das hilft aber nur bedingt, denn Steffen hat wenig geschlafen und träumt zu den Klängen von White Lies. Paula hatte den ganzen Abend mit Steffen über den Perfektionismus diskutiert und nun sind beide voll neben der Spur. So wird es nichts mit der Klausur. Paula mampft noch ein paar Smarties und nickt ein. Jetzt brauchen die beiden gesondert Hilfe!

Leider kennen sich Paula und Steffen 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 „ $\alpha = 95\%$ “! **(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 signifikanten 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 Tina und schaut in das leere Gesicht von Alex. 'Keine Ahnung. Irgendwas mit Relevanz und Effekt oder Signifikanz. Da kannst du irgendwie was verbinden. Keine Ahnung warum', entgegnet Alex. 'Wir haben doch als Messwert *Wasserverbrauch der Bewässerung* erhoben.', stellt Tina fest. Jetzt haben beide das Problem, die möglichen 95% Konfidenzintervalle zu interpretieren.

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

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



66. Aufgabe

(10 Punkte)

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



Zusammenhang zwischen dem Effekt, der Streuung sowie der Fallzahl Es regnet. Wie immer. Aber dafür sind Jessica und Steffen ja auch in Regenbrück zum Lernen verabredet. Gibt es dafür ein besseres Wetter? Eine große Kanne Kaffee und Berge von Schokobons liegen bereit und warten darauf gegessen zu werden. Jessica 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 $KI_{1-\alpha}$?

Steffen hebt die Augenbraue. 'Mir ist kalt und es zieht bei dir. Ich bleibe dabei. Wir sollten erstmal Harry Potter schauen, bis dein Backofen hier mal die Küche geheizt hat. Den Film habe ich doch extra mitgebracht! Genauso wie die Pizza!' Jessica ist der Idee nicht abgeneigt und auch die Hündin kommt in die Küche um sich zu wärmen.

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

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

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

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

Teil III.

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

67. Aufgabe

(9 Punkte)

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



Berechnung des Student t-Test oder Welch t-Test Mark ist im Wendland für einen Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) mit Hühnern. Allein diese Tatsache ist für ihn eine Erzählung wert. Mark und die Unsicherheit, eine unendliche Geschichte mit kniffligen Wendungen. Für seiner Hausarbeit musste er ein Stallexperiment mit Hühnern durchführen und das sollte laut seinem Betreuer 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 Lüftungssystem (*keins* und *vorhanden*) und der Messwert Fettgehalt [%/kg] sollen mit einem t-Test ausgewertet werden. Immerhin weiß er, dass sein Messwert einer Normalverteilung folgt. Hm..., was spannendes wäre gut. Mark will später nochmal raus um zu Reiten. Druck ablassen, dass muss er auch.

treatment	weight
ctrl	17.8
dose	13.8
ctrl	19.4
ctrl	25.1
dose	27.1
dose	18.2

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. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 1.84$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des Student 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 Das Emsland, unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer von Jonas, der mit seiner 1 Mann starken Besatzung 12 Wochen lang unterwegs ist, um neue Welten zu erforschen, neues Leben und neue Zivilisationen. 'Oder nennen wir es Ödnis und Verzweiflung', denkt Jonas. Für seine Abschlussarbeit ist Jonas ins Nichts gezogen. Jonas und die Erschöpfung, eine unendliche Geschichte mit kniffligen Wendungen. Was macht er nun? Jonas hat ein Freilandversuch mit Erbsen durchgeführt. Die Behandlung Lüftungssystemen und Folientunneln (*ctrl* und *tornado*) wurde an Erbsen getestet. Gemessen hat er dann als einen normalverteilten Endpunkt (Y) Frischegewicht [kg/ha]. Jetzt soll er seinem Betreuer nach testen, ob die Behandlung Lüftungssystemen und Folientunneln (*ctrl* und *tornado*) ein signifikantes Ergebnis liefert. Hm..., was entspannendes wäre gut. Aus den Boxen wummert Iron Maiden und sein Mund ist verklebt von Snickers. 'Herrlich', denkt Jonas.

Lüftungssystemen	Frischegewicht
ctrl	47.2
tornado	48.3
ctrl	29.2
tornado	37.8
tornado	34.3
ctrl	39.9
tornado	38.1
ctrl	26.8
tornado	46.4
ctrl	19.5
tornado	26.4
ctrl	49.2
ctrl	47.1
tornado	32.9
tornado	23.8
tornado	51.2
tornado	32.9
ctrl	25.7
tornado	58.3

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *einen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann die Teststatistik T_D ? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Jonas 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 ein normalverteiltes Outcome (Y).', liest Alex laut. Das hilft jetzt auch nur bedingt weiter. Wenn die Gefälligkeit nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so.. Laut seiner Betreuerin ist zwar ihm Messwert Gewichtszuwachs in der 1LW normalverteilt, aber wie rechnet er jetzt einen t-Test? Für seine Abschlussarbeit musste er einen Leistungssteigerungsversuch mit Schweinen im Wendland durchführen. Als wäre das nicht schon anstrengend genug gewesen. Jetzt soll er auch noch testen, ob die Behandlung Elterlinie (*Standard* und *Xray*) ein signifikantes Ergebnis liefert. Hm..., was entspannendes wäre gut. Hm, lecker Gummibärchen und dazu dann im Hintergrund Alien laufen lassen.

Elterlinie	Gewichtszuwachs
Standard	25.6
Xray	33.3
Standard	35.0
Standard	38.6
Xray	38.0
Xray	23.1
Standard	39.0
Xray	24.2
Standard	27.1
Xray	36.2
Xray	38.3
Xray	34.5
Standard	37.7
Standard	45.9
Xray	33.4
Xray	34.3
Xray	33.4
Xray	41.0
Standard	36.4

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie das 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 Alex über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

70. Aufgabe

(11 Punkte)

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



Berechnung des gepaarten t-Test Yuki und Jonas haben sich dazu entschieden zusammenzuarbeiten. Das sollte alles etwas einfacher machen. Jeder hat zwar ein getrenntes Themenfeld aber den Hauptversuch machen beide gemeinsam. Das hat sich schonmal als gut Idee soweit herausgestellt. In einer Hausarbeit sollen beide herausfinden, ob es einen Zusammenhang zwischen Ausgeizen (*ctrl* und *28d*) und Frischgewicht [kg/ha] gibt. Die Besonderheit ist hierbei, dass die Messungen an der gleichen Beobachtung stattfinden. Beide messen also zweimal an den gleichen Erbsen. Hier muss dann wohl auf ein normalverteiltes Outcome (Y) ein gepaarter t-Test gerechnet werden. Yuki schaut etwas flehentlich zu Jonas. Wenn die Faulheit nicht wäre, ja dann wäre wohl vieles möglich für Yuki! Aber so... Steffen denkt derweil angestrengt an Iron Maiden und wippt leicht mit dem Fuß.

ID	treatment	freshmatter
7	28d	26.7
11	28d	33.1
5	28d	21.0
8	28d	31.7
7	ctrl	45.1
4	ctrl	44.5
6	28d	14.9
2	ctrl	57.7
4	28d	18.0
6	ctrl	48.6
1	ctrl	49.5
1	28d	31.5
2	28d	28.0
10	28d	28.1
9	28d	41.2
5	ctrl	40.2
3	28d	29.7
3	ctrl	37.6

Leider kennen sich Yuki und Jonas 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 Yuki 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 Steffen und Mark sind bei Nilufar um sich Hilfe in R zu holen. Im Hintergrund wummert Deichkind. Die beiden hatten zwar schon erste Kontakte mit R sind sich aber unsicher bei der Interpretierung der Ausgabe eines t-Tests für ihren gemeinsamen Versuch. Es würde auch besser funktionieren, wenn Nilufar nicht die Erwartung im Weg stehen würde und Mark nicht das Problem hätte die Romantik zu händeln. In einem Projektbericht haben beide zusammen Brokkoli untersucht. Dabei ging es um den Zusammenhang zwischen der Behandlung Bewässerungstypen (*low* und *high*) und dem Messwert Proteingehalt [g/kg]. Der Versuch wurde in einem Gewächshausexperiment im Emsland durchgeführt. Nach des Betreuers ist der Messwert Proteingehalt [g/kg] normalverteilt und ein t-Test passt daher. Das wird jetzt nicht mehr angezweifelt...Nilufar überlegt, ob sie die beiden nicht noch auf den Film *Star Trek* einlädt.

```
##
## Two Sample t-test
##
## data: Proteingehalt by Bewässerungstypen
## t = 1.9627, df = 14, p-value = 0.06987
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -0.6305603 14.2242111
## sample estimates:
## mean in group low mean in group high
## 33.38571 26.58889
```

Helfen Sie Nilufar bei der Interpretation des t-Tests! Sonst geht es auch für Steffen 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 $T_{\alpha=5\%} = |2.14|$ einzeichnen! **(4 Punkte)**
5. Beschriften Sie die Abbildung! **(1 Punkt)**
6. Berechnen Sie den Effekt des t-Tests! **(1 Punkt)**

72. Aufgabe

(10 Punkte)

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



Interpretation des t-Tests in R - das 95% Konfidenzintervall Yuki und Mark sind bei Alex um sich Hilfe in R zu holen. Im Hintergrund wummert Abba. Die beiden hatten zwar schon erste Kontakte mit R sind sich aber unsicher bei der Interpretierung der Ausgabe eines t-Tests für ihren gemeinsamen Versuch. Es würde auch besser funktionieren, wenn Alex nicht die Gefälligkeit im Weg stehen würde und Mark nicht das Problem hätte die Faulheit zu händeln. In einer Hausarbeit haben beide zusammen Lamas untersucht. Dabei ging es um den Zusammenhang zwischen der Behandlung Ernährungszusatz (*ctrl* und *fedX*) und dem Messwert Protein/Fettrate [%/kg]. Der Versuch wurde in einem Kreuzungsexperiment im Oldenburger Land durchgeführt. Nach des Betreuers ist der Messwert Protein/Fettrate [%/kg] normalverteilt und ein t-Test passt daher. Das wird jetzt nicht mehr angezweifelt...Alex überlegt, ob er die beiden nicht noch auf den Film *Alien* einlädt.

```
##
## Two Sample t-test
##
## data: Protein/Fettrate by Ernährungszusatz
## t = -1.0358, df = 18, p-value = 0.314
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -6.844163 2.324163
## sample estimates:
## mean in group ctrl mean in group fedX
##                29.32                31.58
```

Helfen Sie Alex bei der Interpretation des t-Tests! Sonst geht es auch für Yuki 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 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 'Wir waren in der Uckermark um Erbsen in einem Feldexperiment zu messen.', Mark legt das Dokument auf den Tisch und schaut Paula und Jonas fragend an. Beide schauen fragend zurück. Gäbe es die Erschöpfung nicht, dann wäre es für Jonas irgendwie einfacher hier zu helfen. Echt unangenehm. Die beiden sind zu Mark gekommen, da sie sich nicht mit R auskennen und daher Hilfe bei der Interpretation des t-Tests brauchen. Im Hintergrund wummert Andrea Berg und leere Marzipankugeln Packungen stapeln sich auf dem Boden. 'Kein Problem', sagt Mark und streichelt langsam der Hamster. 'Aber worum es in dem Versuch geht, lässt sich nur aus dem Text in seiner Hand erraten.' merkt er an. Vielleicht hilft da ja die Ausgabe des t-Tests in R weiter. Draußen geht blutrot die Sonne unter.

```
##  
## Two Sample t-test  
##  
## data: Frischegewicht by Lüftungssystemen  
## t = -4.8542, df = 17, p-value = 0.0001488  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -27.65753 -10.89929  
## sample estimates:  
## mean in group ctrl mean in group tornado  
## 26.81250 46.09091
```

Helfen Sie Mark bei der Interpretation des t-Tests! Sonst geht es auch für Paula und Jonas 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 Zandern. Aber das haben Steffen und Tina eben gemeinsam in einer Hausarbeit gemacht! Worum ging es aber konkret? Beide haben als ein normalverteiltes Outcome (Y) Schlachtgewicht [kg] von Zandern bestimmt. Die Daten haben beide zusammen in einem Stallexperiment erhoben. In dem Experiment ging es um eine vorher/nachher Untersuchung an den gleichen Zandern. Als Behandlung wurde Genotypisierung (0d und 14d) eingesetzt. Nach der Meinung der Betreuerin muss hier ein gepaarter t-Test gerechnet werden. Leider kennen sich beide nicht sehr gut in R aus. Steffen hat einiges an Oreos geholt, so dass beide die Zeit gut durchbringen werden. Dann geht Steffen nochmal zum Sport. Um zu Ringen geht Steffen dann später nochmal raus. Echte Entspannung.

```
##
## Paired t-test
##
## data: Schlachtgewicht by Genotypisierung
## t = 0.3055, df = 8, p-value = 0.7678
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -6.693878  8.738323
## sample estimates:
## mean difference
## 1.022222
```

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

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

Teil IV.

Die einfaktorielle & zweifaktorielle ANOVA

75. Aufgabe

(11 Punkte)

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



Visualisierung der einfaktoriellen ANOVA 'Als erstes visualisieren wir unsere Daten und dann können wir schon abschätzen, ob unser Gruppenvergleich in der ANOVA signifikant werden würde?', Mark schaut Yuki fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Die beiden hatten sich auf einem Konzert von London Grammar kennengelernt. Yuki tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren in der Uckermark um ein Kreuzungsexperiment mit Fleischrindern durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Ernährungszusatz (*ctrl*, *fedX* und *getIt*) und dem Messwert Gewichtszuwachs in der 1LW gibt. Später wird noch Matrix geguckt. Yuki befürwortet das!

Ernährungszusatz	Gewichtszuwachs
fedX	38
getIt	40
getIt	40
ctrl	30
fedX	49
ctrl	30
getIt	39
fedX	47
ctrl	31
getIt	39
fedX	43
ctrl	32
ctrl	32
fedX	43
fedX	48
getIt	40
ctrl	29

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

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

76. Aufgabe

(9 Punkte)

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



Ergebnistabelle der einfaktoriellen ANOVA Yuki und Jonas 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. Da hilft das Meerschweinchen von Jonas auch nur bedingt. Die beiden waren in der Uckermark um ein Feldexperiment mit Maiss durchzuführen. Dabei haben Yuki und Jonas den Messwert Trockengewicht [kg/ha] unter der Behandlung Substrattypen (*torf*, *40p60n*, *30p20n* und *70p30n*) ermittelt. Nachher wollen sich beide noch mit dem Hobby Stricken von Jonas beschäftigen. Kennt Yuki noch nicht, klingt aber interessant.

Leider kennen sich Yuki und Jonas mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe, das Meerschweinchen 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)
Substrattypen	3	3966.45			
error	21	1123.39			
Total	24				

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.07$ ab. Begründen Sie Ihre Antwort! (2 Punkte)
5. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? (2 Punkte)

77. Aufgabe

(12 Punkte)

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



Die einfaktoriellen ANOVA und der Student t-Test 'Als erstes bauen wir uns aus unsere Daten die ANOVA Tabelle dann sehen wir schon, ob unser Gruppenvergleich in der ANOVA signifikant ist.', Yuki schaut Paula fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Paula schmeißt sich noch ein paar Smarties in den Rachen. Beide tun sich sehr schwer mit der einfaktoriellen ANOVA. Nun möchte erstmal ihre Betreuung der Arbeit eine ANOVA Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Beide waren in der Uckermark um ein Gewächshausexperiment mit Spargel durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Genotypen (AA, AB und BB) und dem Messwert Proteingehalt [g/kg] gibt. Später wollen die beiden dann noch raus um zu Fechten.

Leider kennen sich Yuki 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)
Genotypen	2	259.25			
Error	19	666.07			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.52$ 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)

Genotypen	Fallzahl (n)	Mittelwert	Standardabweichung
AA	7	5.86	7.38
AB	8	6.25	4.56
BB	7	13.43	5.68

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 Tina schaut sich fragend in der Bibliothek um. Tina hatte gehofft, dass jemand hier sein würde, den sie kennt und sich mit R auskennt. Wird aber enttäuscht. Tina war in der Uckermark um ein Kreuzungsexperiment mit Milchvieh durchzuführen. Nun möchte ihr Betreuer ihrem Projektbericht erstmal eine ANOVA sehen und die Ergebnisse präsentiert bekommen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Elterlinie (*ctrl*, *Standard*, *Yray* und *Xray*) und dem Messwert Protein/Fettrate [%/kg] gibt. Und eigentlich wollte Tina doch noch zum Sport! Einfach mal raus um zu Boxen. Ohne Ziel und Uhr. Das ist was für Tina.

```
## Analysis of Variance Table
##
## Response: Protein/Fettrate
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Elterlinie  3 5421.9 1807.29   85.173 2.422e-15
## Residuals 32  679.0   21.22
```

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

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

79. Aufgabe

(12 Punkte)

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



Ergebnistabelle der zweifaktoriellen ANOVA Nilufar steht in der Uckermark. Und das ist noch langweiliger als es sich anhört. Wäre es nur so spannend wie bei ihren Kommilitonen, die in Almería waren. Ödnis wohin man nur blickt. Oder eben Milchvieh. Das Huhn durchbohrt sie mit leeren Blick. 'Woher zum Teufel!', entfährt es ihr. Aber da ist es schon weg. Ja, darum geht es in ihrer Abschlussarbeit. Und wäre das nicht noch alles genug, ist ihr Experiment auch noch als einen Leistungssteigerungsversuch komplex geraten. Es wurde der Messwert Fettgehalt [%/kg] mit dem Behandlung Genotypen (AA, AB und BB) sowie der Behandlung Elterlinie (*ctrl*, und *Xray*) untersucht. 'Hmpf', denkt Nilufar und ruft 'Und jetzt!?' in die Leere. Und eigentlich wollte Nilufar doch noch ihrem Hobby nachgehen! Hip Hop. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Nilufar denkt gerne über Hip Hop nach.

Leider kennen sich Nilufar 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. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Genotypen	3	297.4			
Elterlinie	1	11.13			
Genotypen:Elterlinie	3	70.06			
Error	18	444			

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

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

5. Was bedeutet ein signifikantes Ergebnis in einer zweifaktoriellen ANOVA? (2 Punkte)
6. Was sagt der Term *Genotypen:Elterlinie* 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 Steffen wird langsam nervös. Steffen wollte heute Abend noch seine E-Sport Qualifikation schauen. Stattdessen versucht seine Betreuerin die Ausgabe der zweifaktoriellen ANOVA zu visualisieren und zu überprüfen, ob es mit der Visualisierung der Daten als Boxplots zusammenpasst. Steffen hatte in der Uckermark einen Versuch in einer Klimakammer mit Kartoffeln durchgeführt. Es gab dabei zwei Behandlungen. Einmal Düngestufen (*ctrl*, *low*, *mid* und *high*) sowie als zweite Behandlung Substrattypen (*torf*, *70p30n*). Gemessen wurde der Messwert (Y) Chlorophyllgehalt (SPAD-502Plus) [SPAD]. So kompliziert kann das jetzt doch nicht sein! Eigentlich wollte Steffen nachher noch zum Sport. Steffen will später nochmal raus um zu Ringen. Druck ablassen, dass muss er auch.

```
## Analysis of Variance Table
##
## Response: Chlorophyllgehalt
##
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Düngestufen      2  17.04      8.52  0.4690 0.6330562
## Substrattypen     1 357.34    357.34 19.6743 0.0003193
## Düngestufen:Substrattypen  2 342.00    171.00  9.4147 0.0015911
## Residuals       18 326.93     18.16
```

Leider kennt 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. 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 Yuki schaut konzentriert auf die Formeln der ANOVA und des t-Tests. In ihrem Experiment wurde als Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] bestimmt. Wenn die Faulheit nicht wäre, ja dann wäre wohl vieles möglich für Yuki! Aber so.. Dann wäre es nicht noch komplizierter. In ein Gewächshausexperiment wurden Maiss mit der Behandlung Genotypen (AA, AB und BB) sowie der Behandlung Lichtstufen (*none*, und 600lm) untersucht. Beide Verfahren müssen etwas miteinander zu tun haben und ihre Betreuerin möchte das jetzt auch noch verstehen. Im Hintergrund läuft leise Matrix auf ihrem Second Screen. Immerhin hat sie die beiden Formeln vorliegen.

Gegebene Formeln

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

Leider kennen sich Yuki 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 'Mit der zweifaktoriellen ANOVA lässt sich die Interaktion zwischen den beiden Behandlungen nachweisen!', sein Betreuer scheint die zweifaktoriellen ANOVA zu verstehen. Warum jetzt er jetzt nochmal alles wiederkauen muss, wird Steffen echt nicht so klar. Wenn es doch so klar ist? Steffen war in der Uckermark und hatte dort einen Leistungssteigerungsversuch mit Fleischrindern durchgeführt. Die Komune wo er untergekommen war, war cool gewesen. Nur jetzt muss eben das Experiment fertig ausgewertet werden. Es liegt anscheinend eine signifikante Interaktion vor? Steffen hatte zwei Behandlungen auf Fleischrindern angewendet. Einmal Ernährungszusatz (*ctrl*, *fedX* und *getIt*) sowie als zweite Behandlung Elterlinie (*ctrl*, und *Xray*). Gemessen wurde der Messwert (*Y*) Gewichtszuwachs in der 1LW. Jetzt muss das hier zu einem Ende kommen! Eigentlich wollte Steffen nachher noch einen Film schauen. Das Verrückte ist, dass die Schlange Harry Potter wirklich liebt. Das ist Steffen sehr recht, denn er braucht Entspannung.

Leider kennen sich Steffen und sein Betreuer mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion. Deshalb braucht er bei der Erstellung Ihre Hilfe, sonst wird es heute Abend mit seinem Hobby Klemmbausteine 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!', seine Betreuerin scheint die einfaktoriellen ANOVA zu verstehen. Warum jetzt er jetzt nochmal alles wiederkauen muss, wird Mark echt nicht so klar. Wenn es doch so klar ist? 'Wir haben jetzt bei der ANOVA einen p-Wert mit 0.061 raus sowie eine F-Statistik F_D mit 1.2 berechnet. Nach den Boxplots müsste sich eigentlich ein Unterschied zwischen *kontakt* und *eng* ergeben. Der Unterschied ist in {emmeans} auch signifikant mit einem p-Wert von 0.049. Wie kann das sein?', fragt Mark etwas provokant und dreht Andrea Berg leiser. Mark war in der Uckermark und hatte dort einen Leistungssteigerungsversuch mit Puten durchgeführt. Die Komune wo er untergekommen war, war cool gewesen. Dort gab es selbstgemachte Marzipankugeln aus Vollkorn! Nur jetzt muss eben das Experiment fertig ausgewertet werden. Mark hatte eine Behandlungen Bestandsdichte (*standard*, *eng*, *weit* und *kontakt*) auf Puten angewendet. Gemessen wurde der Messwert (Y) Gewichtszuwachs in der 1LW. Dabei wurden die Daten D erhoben. Jetzt muss das hier zu einem Ende kommen! Mark hat schon genug Probleme. Wenn die Unsicherheit 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 Mark und seine Betreuerin mit der Interpretation einer ANOVA überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe und die Zeit wird knapp.

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

Teil V.

Multiple Gruppenvergleiche

84. Aufgabe

(12 Punkte)

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



Adjustierung multipler Vergleiche In ein Freilandversuch mit Brokkoli wurde die Behandlung Genotypen (00, AA, AB und BB) gegen die Ergebnisse einer früheren Studie von Qui et al. (2017) verglichen. Im Rahmen des Experiments haben Tina und Paula verschiedene Student t-Tests für den Mittelwertsvergleich für den Messwert Frischgewicht [kg/ha] gerechnet. Es ergab sich dann die folgende Tabelle der rohen p -Werte für die Vergleiche zu Qui et al. (2017). Jetzt sollen die beiden einmal schauen, was in den Daten so drin ist.

Rohen p -Werte	Adjustierte p -Werte	Nullhypothese ablehnen?
0.23		
0.76		
0.06		
0.08		

Leider kennen sich Tina und Paula 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 sitzt schon etwas länger bei ihrer Betreuerin. So langsam macht Paula sich Gedanken, ob sie nicht doch mal anmerken sollte, dass sie von CLD noch nie was gehört hat. Aber noch kann gelauscht werden, ein Ende ist erstmal nicht in Sicht! Paula hatte in ihrer Hausarbeit einen Leistungssteigerungsversuch durchgeführt. Deshalb sitzt sie hier. Also eigentlich nein, deshalb nicht. Paula will fertig werden. Hat sie sich doch mit Bestandsdichte (*effizient*, *standard*, *eng*, *weit* und *kontakt*) und Protein/Fettrate [%/kg] schon eine Menge angeschaut. Paula beugt sich leicht nach vorne. Nein, doch keine Pause. Weiter warten auf eine Lücke im Fluss...

Behandlung	Compact letter display
effizient	a
standard	a
eng	a
weit	a
kontakt	a

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 Alex hatte in der Projektbericht ein Kreuzungsexperiment durchgeführt. Soweit so gut. Dabei hat er sich mit Zandern beschäftigt. Angeblich der neueste heiße Kram... aber das ist wiederum was anderes. So richtig mitgenommen hat Alex das Thema dann doch nicht. Hat er sich doch mit Ernährungszusatz (*ctrl*, *fedX*, *proteinX* und *getIt*) und Gewichtszuwachs in der 1LW schon eine Menge an Daten angeschaut. Nach sein Betreuer soll er nun ein CLD bestimmen. Weder weiß er was ein CLD ist, noch war sein erster Gedanke mit Köln und die LGBTQ Community richtig... Als erstes solle er die Gruppen nach absteigender Effektstärke sortieren. Was immer das jetzt bringen soll.

Ernährungszusatz	Fallzahl (n)	Mittelwert	Standardabweichung
ctrl	9	14.54	2.13
fedX	7	15.86	1.73
proteinX	9	15.03	2.56
getIt	9	10.21	3.00

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich ergebenden Barplots! **(1 Punkt)**
4. Berechnen Sie die Matrix der *p*-Werte anhand von Student t-Tests! **(4 Punkte)**
5. Ergänzen Sie das *Compact letter display (CLD)* zu den gezeichneten Barplots! Begründen Sie Ihre Antwort! **(4 Punkte)**
6. Interpretieren Sie das *Compact letter display (CLD)* für Alex und Paula! **(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 Excel 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 Excel 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 Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) und Schlachtgewicht [kg] in ein Kreuzungsexperiment mit Lamas angeschaut. Wo kriegt er jetzt ein CLD her? Dann eben per Hand aus der Matrix der *p*-Wert. Mark stöhnt...

	keins	storm	tornado	thunder
keins	1.0000000	0.1740523	0.4410091	0.0041091
storm	0.1740523	1.0000000	0.0418082	0.0834537
tornado	0.4410091	0.0418082	1.0000000	0.0006892
thunder	0.0041091	0.0834537	0.0006892	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 Paula! **(2 Punkte)**

Teil VI.

Der Chi-Quadrat-Test & Der diagnostische Test

88. Aufgabe

(12 Punkte)

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



Den Chi-Quadrat-Test berechnen Am Ende hätte Tina dann doch einen normalverteilten Endpunkt in ihrer Abschlussarbeit nehmen sollen. Vor ihr liegen jetzt die Daten von zwei Variablen als Kategorien oder wie es in R so schön heißt, als Faktoren. Aber immerhin, hofft sie das was bei den Daten rausgekommen ist. Dann noch schnell Tocotronic auf das Ohr und los gehts. Gezählt hat Tina einiges mit $n = 110$ Beobachtungen von Brokkoli. Zum einen hat sie als Behandlung *Mechanische Bearbeitung* [ja/nein] bestimmt und zum anderen die Variable *Frischegewicht über Zielwert* [ja/nein] ermittelt. Nun möchte ihre Betreuerin gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Eigentlich wollte Tina nachher noch einen Film schauen. Das Verrückte ist, dass die Spinne Indiana Jones wirklich liebt. Das ist Tina sehr recht, denn sie braucht Entspannung.

	24	21	
	27	38	

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

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

89. Aufgabe

(10 Punkte)

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



Der Chi-Quadrat-Test konzeptionell verstehen Am Ende hätte Jessica dann doch einen normalverteilten Endpunkt in ihrer Abschlussarbeit nehmen sollen. Dann noch schnell Schokobons zur Stärkung und los gehts. Vor ihr liegen jetzt die Daten von zwei Variablen als Kategorien oder wie es in R so schön heißt, als Faktoren. Aber immerhin, hofft sie das was bei den Daten rausgekommen ist. Gezählt hat Jessica einiges mit $n = 144$ Beobachtungen von Hühnern. Zum einen hat sie als Behandlung *Außenklimakontakt* [ja/nein] bestimmt und zum anderen die Variable *Gewichtszuwachs erreicht* [ja/nein] ermittelt. Nun möchte ihr Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Am Ende des Tages möchte sie dann noch ihr Hobby Warhammer genießen. Das muss auch mal sein!

			61
			83
	76	68	144

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

90. Aufgabe

(10 Punkte)

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



Der Chi-Quadrat-Test in R Am Ende hätte Yuki dann doch einen normalverteilten Endpunkt in seinem Projektbericht nehmen sollen. 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. Dann noch schnell Matrix starten und los gehts mit der Kraft von Reese's Peanut Butter Cups. Gezählt hat Yuki einiges mit $n = 144$ Beobachtungen von Hühnern. Zum einen hat er als Behandlung *Automatische Fütterung* [ja/nein] bestimmt und zum anderen die Variable *Gewichtszuwachs erreicht* [ja/nein] ermittelt. Nun möchte sein Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Nach seinem Experiment erhielt er folgende 2x2 Kreuztabelle aus seinen erhobenen Daten.

```
##               Automatische Fütterung
## Gewichtszuwachs erreicht ja nein
##               ja    13    5
##               nein   7    18
```

Dann rechnete Yuki 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: Gewichtszuwachs erreicht
## p-value = 0.005898
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.462677 32.500828
## sample estimates:
## odds ratio
##  6.352594
```

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

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

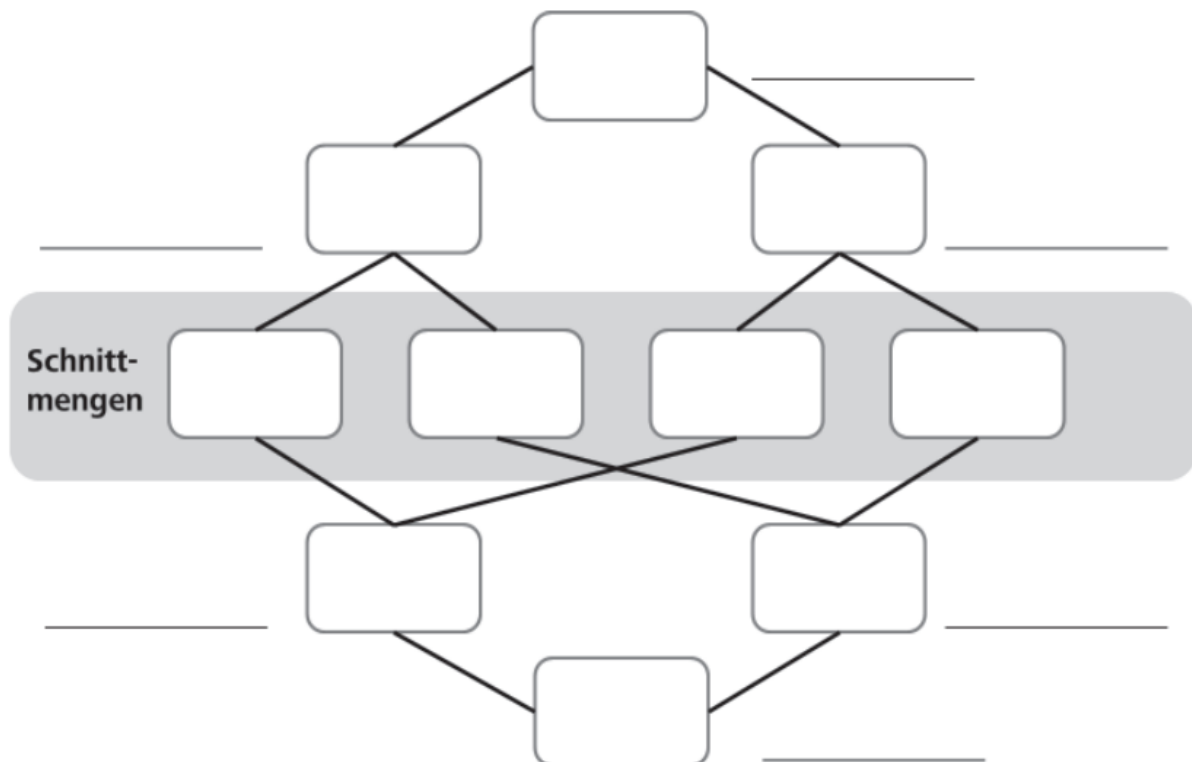
91. Aufgabe

(11 Punkte)

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



Den diagnostische Test am Doppelbaum berechnen Steffen liest laut vor. 'Die Prävalenz von Klauenseuche bei Erdbeeren wird mit 4% angenommen. In 75% der Fälle ist ein Test positiv, wenn das Pflanze erkrankt ist. In 7.5% der Fälle ist ein Test positiv, wenn das Pflanze nicht erkrankt ist und somit gesund ist. Wir führen einen Test auf Kräuselkrankheit an 1000 Erdbeeren mit einem diagnostischen Test durch.' Yuki klappt die Kinnlade runter. In der Stille dudzelt Taylor Swift. Steffen schaut kompetent und schmeißt sich mit offenen Mund Reese's Peanut Butter Cups an den Kopf vorbei.



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

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

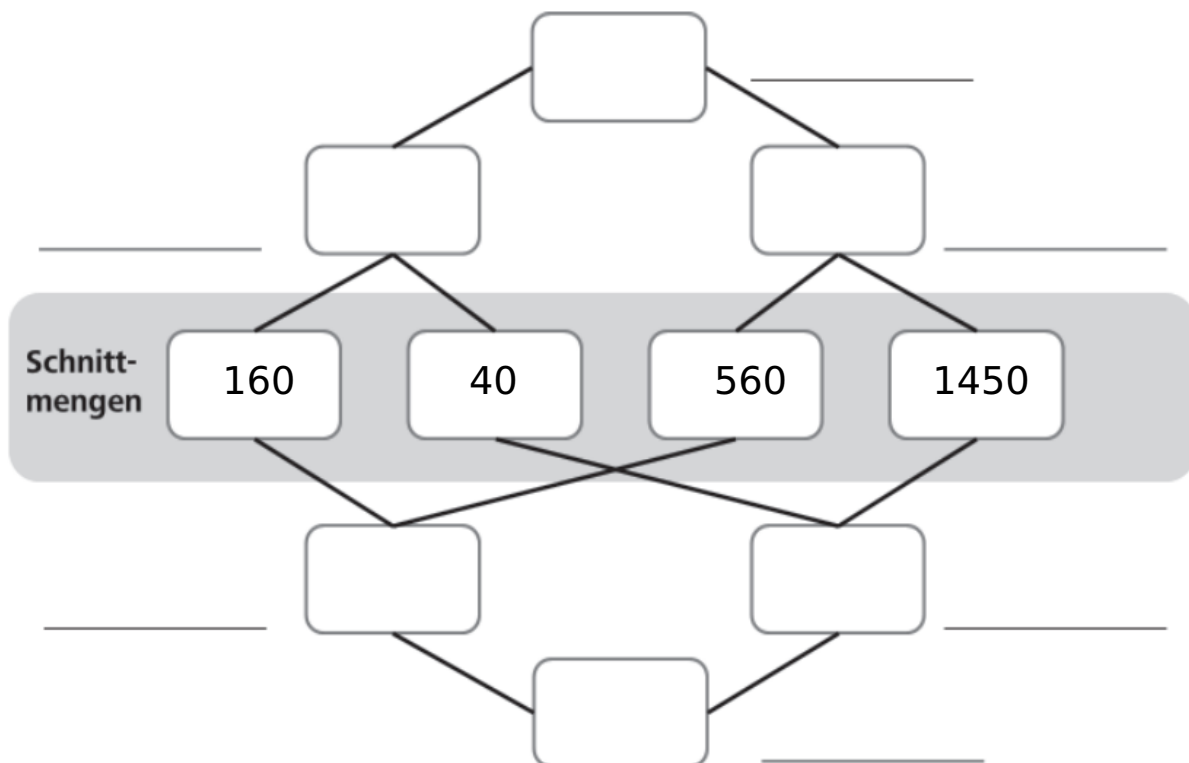
92. Aufgabe

(11 Punkte)

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



Der diagnostische Test und statistische Maßzahlen 'Was ist denn das?', entfährt es Paula. 'Hm... ich glaube es handelt sich um einen Doppelbaum, den wir beim diagnostischen Testen brauchen.', meint Mark und dreht Jagd auf roter Oktober auf dem Second Screen etwas leiser. Was jetzt beide von einem diagnostischen Test haben, ist ihnen auch nicht klar. Es ist ja schon alles komplex genug und der Perfektionismus von Paula macht es heute auch nicht mehr einfacher. 'Es geht um Kräuselkrankheit an Erbsen.', stellt Mark fest. Eigentlich wollte Mark eher los um zu Reiten. Das wird aber wohl nichts mehr.



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

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

Teil VII.

Lineare Regression & Korrelation

93. Aufgabe

(10 Punkte)

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



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

Schlachtgewicht [kg]	Durchschnittlicher Tagestemperatur [C/d]
18.7	27.5
19.0	23.3
18.1	21.5
12.7	15.0
13.3	17.6
13.4	20.3
17.1	21.6
14.5	22.7
12.7	17.6
7.3	9.9

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

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

94. Aufgabe

(12 Punkte)

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



Interpretation der Ergebnisse einer linearen Regression 'Ich glaube du bringst da was durcheinander. Wir nutzen zwar auch für die ANOVA die Funktion `lm()` aber hier wollen wir, glaube ich, eine Gerade durch die Punkte zeichnen.', merkt Mark an. 'Ich sehe keine Punkte... ich sehe nur zwei Zeilen einer Tabelle und ich glaube du hast gerade was gelöscht.', antwortet Nilufar sichtlich übernächtigt. 'Wir müssen die Koeffizienten der linearen Regression ja auch erst interpretieren!', spricht Mark sehr deutlich und langsam. Die beiden hatten ein Gewächshausexperiment im Wendland mit Lauch durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche UV-Einstrahlung [UV/d] und Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der `R` Ausgabe möglich sein.

term	estimate	std.error	t statistic	p-value
(Intercept)	-2.69	2.68		
Durchschnittliche UV-Einstrahlung	0.56	0.26		

Leider kennen sich Mark und Nilufar 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)**

95. Aufgabe

(11 Punkte)

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



Interpretation der Ergebnisse einer linearen Regression in R 'Ich glaube du bringst da was durch-einander. Wir nutzen zwar auch für die ANOVA die Funktion `lm()` aber hier wollen wir, glaube ich, eine Gerade durch die Punkte zeichnen.', merkt Jonas an. 'Ich sehe keine Punkte... das ist doch eine Ausgabe in R. Überhaupt, darum geht es doch gar nicht in unserem Versuch. Wir wollen doch keine Gerade zeichnen?.', antwortet Nilufar sichtlich übernächtigt. 'Doch wir müssen nur die Koeffizienten der linearen Regression erst richtig interpretieren und vor unserem geistigen Auge erscheint eine Gerade!', spricht Jonas sehr deutlich und langsam. Die beiden hatten ein Gewächshausexperiment im Wendland mit Erbsen durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche UV-Einstrahlung [UV/d] und Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der R Ausgabe möglich sein.

```
##
## Call:
## Chlorophyllgehalt ~ Durchschnittliche_UV_Einstrahlung
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6237 -0.7874 -0.2143  0.8516  2.3160
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.1223     1.5052   0.081   0.936
## Durchschnittliche_UV_Einstrahlung  2.6842     0.1479  18.154 <2e-16
##
## Residual standard error: 1.267 on 35 degrees of freedom
## Multiple R-squared:  0.904, Adjusted R-squared:  0.9013
## F-statistic: 329.6 on 1 and 35 DF, p-value: < 2.2e-16
```

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

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

96. Aufgabe

(10 Punkte)

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



Interpretation der Ergebnisse einer Korrelationsanalyse in R 'Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen...', denkt Steffen. 'Ich sehe nur Kauderwelsch und keine Punkte. Ich glaube das war jetzt doch eine Korrelation, die ich rechnen sollte. Und warum überhaupt? War das unsere Fragestellung?', denkt sich Steffen. Steffen atmet schwer ein und starrt auf die R Ausgabe der Funktion `cor.test()`. Das hilft alles nur begrenzt. Wenn Harry Potter läuft, dann ist die Schlange nicht mehr da. Aber jetzt braucht er mal Entspannung! Steffen hatte einen Versuch in einer Klimakammer im Emsland mit Erdbeeren durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche Regenwurmdichte [Anzahl/l] und Trockengewicht [kg/ha]. Jetzt will die Betreuung von ihm die Interpretierung der Daten in Form einer Korrelation berechnet bekommen. Das hat Steffen in R gemacht, aber wie soll das jetzt gehen? Das mit der Interpretation? Wenn die Romantik nicht wäre, ja dann wäre wohl vieles möglich für Steffen! Aber so..

```
##  
## Pearson's correlation  
##  
## data: Durchschnittliche Regenwurmdichte and Trockengewicht  
## t = -9.4511, df = 8, p-value = 1.292e-05  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.9903013 -0.8275875  
## sample estimates:  
## cor  
## -0.9580184
```

Leider kennt sich Steffen mit der Korrelationsanalyse in R 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. Erstellen Sie eine Visualisierung für den Korrelationskoeffizienten! Beschriften Sie die Abbildung! **(2 Punkte)**
4. Nennen Sie die zwei Eigenschaften des Korrelationskoeffizienten! **(2 Punkte)**
5. Interpretieren Sie den Korrelationskoeffizienten hinsichtlich des Effekts und der Signifikanz! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Visualisieren Sie das 95% Konfidenzintervall! Beschriften Sie die Abbildung! **(2 Punkte)**

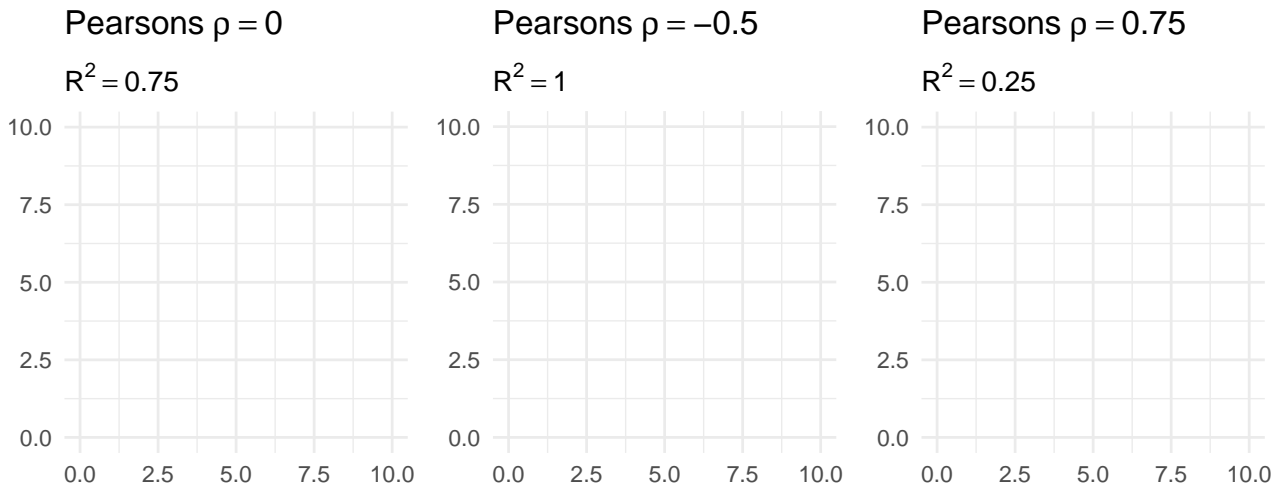
97. Aufgabe

(12 Punkte)

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



Visualisierung der Korrelation und des Bestimmtheitsmaßes Irgendwie komisch, wenn sie Star Trek anmacht, dann ist das Huhn eigentlich sofort vor dem Bildschirm und starrt hinein. Da hilft dann die Aufgabe auch nur bedingt. 'Hm..., drei leere Abbildungen. Was soll ich da jetzt machen?', fragt sich Nilufar und mampft noch ein paar Takis Blue Heat in sich hinein. Nilufar kennt sich nur begrenzt bis gar nicht mit der linearen Regression und Korrelation aus.



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

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

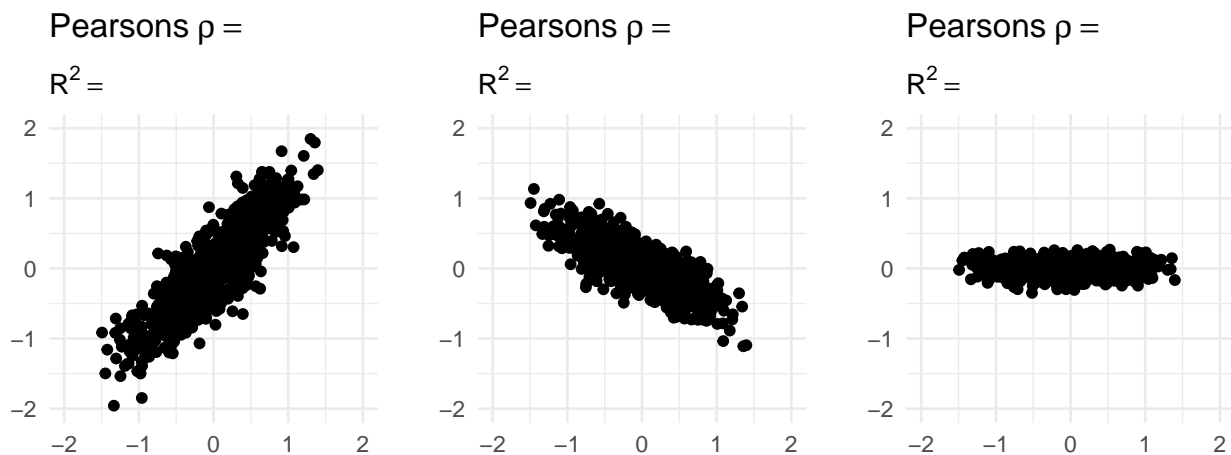
98. Aufgabe

(12 Punkte)

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



Schätzen der Korrelation und des Bestimmtheitsmaßes Der Bildschirm strahlt blau in das Gesicht von Jessica. Es ist schon spät. Und das hat einen Grund. Auf seinem Second Screen läuft Herr der Ringe und Jessica schaufelt Schokobons. Nicht effizient, aber gut. . Jessica überlegt, aber ihre Gedanken sind etwas zäh. 'Was soll das hier alles bedeuten?', fragt sich Jessica. Irgendwie ist ihr nicht klar wie sie ρ -Werte oder R^2 -Werte abschätzen soll. Alles nicht so einfach. Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen.



Leider kennt sich Jessica 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)**

99. Aufgabe

(11 Punkte)

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



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

Mittlere Eisenkonzentration	Proteinanteil	\hat{y}	ϵ
5.8	-2.3	5.6	
24.8	7.9	21.5	
26.0	10.3	25.3	
22.5	8.9	23.2	
20.2	7.5	21.0	
20.4	7.9	21.5	
30.2	13.9	30.9	
19.0	7.6	21.1	
28.9	11.8	27.7	

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

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

100. Aufgabe


(12 Punkte)

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



Visualisierung des Regressionskreuzes Mark hat ein Gewächshausexperiment mit Erdbeeren 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 Unsicherheit gewesen. Ein leidiges Lied. Aber es muss ja weitergehen. Mark hatte dann in seiner Abschlussarbeit einfach zu viele Endpunkte gemessen und ist jetzt vollkommen durcheinander, welche Analyse er nun wie rechnen soll. Naja, dann heißt es jetzt eben Andrea Berg aufdrehen und darüber nachdenken, was hier eigentlich gemacht wurde. Mark fängt einfach an und nimmt den ersten Endpunkt Frischegewicht [kg/ha]. Dann kann er sich voran arbeiten. Später dann noch raus um zu Reiten um mal zu entspannen und vielleicht ist Jonas auch da. Wäre toll.

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

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

Teil VIII.

Experimentelles Design

101. Aufgabe

(16 Punkte)



Einfache experimentelle Designs Jonas und Paula sind bei Yuki um sich Hilfe für eine Versuchsplanung in **R** zu holen. Im Hintergrund läuft viel zu laut London Grammar. Dabei geht es um den Zusammenhang zwischen der Behandlung Flüssignahrung (*ctrl*, *superIn* und *floW*) und dem Messwert Protein/Fettrate [%/kg] in Schweinen. Der Versuch soll in einem Leistungssteigerungsversuch im Oldenburger Land durchgeführt werden. Nach der Dozentin ist der Messwert Protein/Fettrate [%/kg] normalverteilt. Die beiden entschieden sich für ein faktorielles Versuchsdesign. Im ersten Schritt überlegt Yuki ein einfaches experimentelles Design zu probieren. Daher entscheiden sich alle drei für ein *Randomized complete block design (RCBD)* mit zwei *Blöcken*. Das sollte für den Anfang erstmal reichen. 'Und jetzt, was machen wir jetzt?', Paula schaut die anderen beiden mit großen Augen an. Die zucken mit der Schulter. Alle mampfen Reese's Peanut Butter Cups.

Leider kennen sich Yuki, Jonas und Paula mit dem *Randomized complete block design (RCBD)* mit zwei *Blöcken* 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 **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**! **(3 Punkte)**
6. Skizzieren Sie die weitere Datenanalyse hinsichtlich eines multiplen Gruppenvergleiches! **(2 Punkte)**
7. Skizzieren Sie eine mögliche Abbildung im Kontext der wissenschaftlichen Fragestellung! Beschriften Sie die Abbildung! **(2 Punkte)**
8. Ergänzen Sie zu der Abbildung ein mögliches Ergebnis des multiplen Gruppenvergleichs! Begründen Sie Ihre Antwort! **(2 Punkte)**




102. Aufgabe

(20 Punkte)



Fortgeschrittene experimentelle Designs Das Minischwein macht mal wieder Randalen in Marks Zimmer und rennt davon! Yuki und Paula sind bei Mark im Wendland wo der neue, bessere Versuch stattfinden soll. Dabei soll in einem Versuch in einer Klimakammer im Wendland mit Maiss durchgeführt werden. Auf dem Tisch stapeln sich Reese's Peanut Butter Cups aus Vollkorndinkelmehl. Eine Spezialität der Komune hier. Yuki hasst Vollkorn wie Smarties geliebt werden. In dem neuen Versuch geht es um den Zusammenhang zwischen der Behandlung Lüftungssysteme (*ctrl*, *storm*, *thunder* und *tornado*) sowie Lichtstufen (*none*, und *600lm*) sowie drei Blöcken und dem Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Immerhin ist der Messwert normalverteilt, was einiges einfacher macht. Was es nicht so einfacher macht ist, dass Paula als zusätzliche Herausforderung noch der Perfektionismus mitgebracht hat. Daher entscheiden sich alle drei für ein *Split plot design* oder auch *Spaltanlage*. 'Naja, so viel einfacher ist es dann doch nicht...', merkt Paula an und sucht das Minischwein.

Leider kennen sich Mark, Yuki und Paula 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 ! **(2 Punkte)**
5. Erstellen Sie das statistische Modell in der in  üblichen Schreibweise für eine ANOVA! Skizzieren Sie die notwendige Funktionen in ! **(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 IX.

Programmieren in R

103. Aufgabe

(9 Punkte)



Grundlegende Kenntnisse der Programmierung in R Steffen muss seiner Hausarbeit mit R arbeiten. Deshalb sitzt er jetzt mit Ihnen zusammen und hat einige Fragen zu den Grundlagen in R an Sie! Na dann wollen Sie mal helfen. Immerhin will sein Betreuer, dass R genutzt wird.

Steffen: *Ich verstehe den Pipe-Operator nicht. Wie sieht der aus und was macht der? Gebe mal ein Beispiel!* **(1 Punkt)**

Sie antworten:

Steffen: *Gibt es einen Vorteil von der Nutzung von R?* **(1 Punkt)**

Sie antworten:

Steffen: *Was ist der Unterschied zwischen dem RStudio und R?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

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

Sie antworten:

Steffen: *Warum gibt es eigentlich Objekte, Wörter und Funktionen in R? Wie unterscheiden sich diese überhaupt?* **(1 Punkt)**

Sie antworten:

Steffen: *Wie speichern wir in R intern Daten ab? Ich brauche nochmal den Namen der Funktion. Was sind den die Vorteile von dieser Art der Speicherung?* **(1 Punkt)**

Sie antworten:

Steffen: *Der Zuweisungs-Operator wird sehr häufig genutzt. Wie sieht der aus und wie funktioniert der an einem Beispiel?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

104. 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 Alex 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 Alex dann in seiner Hausarbeit nichts mit der Auswertung und Abgabe. Das kann auch keine Lösung für Alex und Sie sein. Immerhin haben Sie schon so viel gelernt.

Alex fragt: *Ich baue mir ja meinen Datensatz in Excel. Was muss ich da im Bezug auf die Namen der Spalten beachten?* **(1 Punkt)**

Sie antworten:

Alex fragt: *Das Dateiformat in R hat einen Namen. Wie heißt der und gerne mit Beispiel!* **(1 Punkt)**

Sie antworten:

Alex fragt: *Ich will eine ANOVA in R rechnen. Dazu brauche ich zwei Funktionen. Welche waren das noch gleich und wie war die Reihenfolge?* **(1 Punkt)**

Sie antworten:

Alex fragt: *Wie verbindet {ggplot} die einzelnen Ebenen einer Abbildung?* **(1 Punkt)**

Sie antworten:

Alex fragt: *Wozu nutze ich die Funktion mutate() hauptsächlich?* **(1 Punkt)**

Sie antworten:

Alex fragt: *Warum wurden jeweils die R Pakete {emmeans}, {ggplot} und {readxl} geladen?* **(2 Punkte)**

Sie antworten:

Alex fragt: *Ich will das R Paket {ggplot} nutzen, da war so eine Analogie an die ich mich nicht erinnern kann. Was war noch gleich das Prinzip von {ggplot}? Wie funktioniert {ggplot} konzeptionell?* **(2 Punkte)**

Sie antworten:

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 vorge setzt. 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 vorge setzt. 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 Vor dem Start der eigenen Arbeit möchte seine Betreuerin, dass Alex einmal die wissenschaftliche Veröffentlichung *Wu, G., et al. (2004). Arginine nutrition in neonatal pigs* sinnvoll zusammenfasst. Dann würde die eigene Arbeit auch leichter von der Hand gehen und Alex hätte dann schon eine Vorlage. 'Das ist jetzt aber umfangreicher als gedacht!', schnauft er und runzelt die Stirn. Im Hintergrund spielt viel zu leise Abba. Das wird dann vermutlich heute Abend nichts mehr mit seinem Hobby Starcraft. Die Katze schaut mitleidig.

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

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

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

106. Aufgabe


(20 Punkte)

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



Zerforschen eines wissenschaftlichen Datensatzes Vor dem Start der eigenen Arbeit möchte seine Betreuerin, dass Steffen einmal die wissenschaftlichen Daten *data1* sinnvoll zusammenfasst. Dann würde die eigene Arbeit auch leichter von der Hand gehen und Steffen hätte dann schon eine Vorlage um die eigenen erhobenen Daten in eine Tabelle eintragen zu können. 'Das ist jetzt aber umfangreicher als gedacht!', schnauft er und runzelt die Stirn als er in seinen Laptop starrt. Dabei isst er noch ein paar Oreos. Das wird dann vermutlich heute Abend nichts mehr mit Harry Potter

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

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

Teil XI.

Mathematik

107. Aufgabe

(10 Punkte)

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



Herodot – der Schimmel aus Ivenack Die Lerngruppe *Die Pantoffeltieren* bestehend aus Alex, Yuki, Jessica und Nilufar waren auf Exkursion in Mecklenburg-Vorpommern und haben dort Folgendes erarbeitet. Während der Besetzung Mecklenburgs durch die Franzosen kamen Napoleon die Geschichten des berühmten Apfelschimmels Herodot aus Ivenack zu Gehör. Herodot lief zwar niemals Rennen, war aber eines der berühmtesten Pferde dieser Zeit. Napoleon selbst gab den Auftrag, diesen Schimmel durch die Armee nach Frankreich zu bringen. Der Legende nach sollen Arbeiter den Schimmel im hohlen Stamm einer 1000-jährigen Eiche aus Ivenack vor den Franzosen versteckt haben. Doch Herodot verrät 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 1.1mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 11m in Brusthöhe hatte.

1. Wie groß war der Durchmesser in m der Eiche im Jahr 1815 als Herodot in der Eiche versteckt werden sollte? **(2 Punkte)**
2. Skizzieren Sie in einer Abbildung einen linearen Zusammenhang und einen exponentiellen Zusammenhang für das Wachstum der 1000-jährigen Eiche. Erklären Sie die Auswirkungen der Entscheidung für linear oder exponentiell auf Ihre Berechnungen! **(2 Punkte)**

Herodot hatte eine Schulterhöhe von 190cm, eine Breite von 85cm sowie eine Länge von 240cm.

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

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

4. Berechnen Sie die Dicke der Eichenwand in cm ! Verdeutlichen Sie Ihre Berechnungen an einer aussagekräftigen Skizze für Pferd und Eiche! **(2 Punkte)**
5. Unter einer Dicke der Eichenwand von 30cm 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 *E* bestehend aus Tina, Mark, Alex und Paula hat sich zusammengefunden um den ersten Versuch zu planen. In einem Experiment wollen sie die Wuchshöhe von 120 Stockrosen bestimmen. Bevor die Vier überhaupt mit dem Experiment beginnen können, gibt es aber ein paar Abschätzungen über die Kosten und den Aufwand zu treffen. Zum einen müssen sie die Stockrosen einpflanzen und müssen dafür Substrat bestellen. Zum anderen muss die Projektgruppe die Stockrosen auch bewegen und in ein Gewächshaus platzieren. Die Töpfe für die Keimung haben einen Durchmesser von 8.5cm und eine Höhe von 7cm. Der Kubikmeterpreis für Torf liegt bei 270 EUR.

Helfen Sie der Projektgruppe *E* bei der Planung des Versuches!

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

109. Aufgabe

(10 Punkte)

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



Solar- & Biogasanlagen Nilufar bringt ein neues, tolles Projekt mit in die Lerngruppe *Die Blattläuse* bestehend aus ihr, Steffen, Paula sowie Yuki. Um die Energiekosten ihres Betriebes zu senken, will sie eine Solaranlage auf den Hühnerstall montieren lassen. Dafür hat sie ihren Stall ausgemessen und findet folgende Maße wieder. Die vordere Seite des Hühnerstall hat eine Höhe h_v von $6.5m$. Die hintere Seite des Hühnerstall hat eine Höhe h_b von $8.5m$. Der Hühnerstall hat eine Tiefe t von $14m$ und eine Breite b von $70m$. 'Sag mal Nilufar, ist das eine Matheaufgabe oder rechnen wir hier gerade für dich kostenlos als menschliche Computer Sachen für deinen Betrieb?', fragt Paula mit erhobenen Augenbrauen. Yuki und Steffen nicken zustimmend.

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

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

Ebenfalls plant Nilufar eine neue Biogasanlage für ihren Betrieb. Der neue Methantank hat einen Radius r von $1.2m$. Leider gibt es ein paar bauliche Beschränkungen auf dem Grundstück. Ihr Fundament des zylindrischen Methantanks kann nur ein Gewicht von maximal $15t$ aushalten bevor der Tank wegbricht. Nilufar rechnen eine Sicherheitstoleranz von 10% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $235kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $270kg/m^3$. Nilufar betreibt ihre Anlage bei $-88^\circ C$.

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

110. Aufgabe

(10 Punkte)

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



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

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

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

1. Wenn 4 Blaubeerschalen 7.96 Euro kosten, wie viel kosten 7 Schalen? **(2 Punkte)**
2. Wenn Sie die 7 Blaubeerschalen gekauft haben, wie viele Aligatorbirnen zu je 2.89 EUR können Sie sich dann noch für 200 EUR leisten? **(1 Punkt)**

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

- Ein Kilo Strauchtomaten benötigt 190l Wasser. Eine Strauchtomate wiegt 110 - 125g.
- Ein Kilo Salat benötigt 130l Wasser. Ein Salatkopf wiegt 320 - 520g.
- Ein Kilo Avocado benötigt 1050l Wasser. Eine Avocado wiegt 130 - 410g.
- Ein Kilo Blaubeeren benötigt 830l Wasser. Eine Blaubeere wiegt 3 - 3.8g.

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 2022 blieben die Erträge von Blaubeeren mit 7.9×10^4 t in dem prognostizierten Rahmen. Die Menge steigerte sich um 6.1%. Die Exporte für Avocados stiegen in dem gleichen Zeitraum um 22.1% auf 2×10^5 t.

4. Wie viele Tonnen Wasser hat Chile in dem Exportjahr 2021 exportiert? **(2 Punkte)**

Chile ist eines der wenigen Länder der Welt, die ihr Wasser komplett privatisiert haben. Derzeit sind nur ein 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 61 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 Anbaugebieten können Sie in *Deutschland* wie oft Ihren Bedarf stillen? **(1 Punkt)**

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

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

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

111. Aufgabe

(12 Punkte)

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



Stichworte: Kardaschow-Skala • Dyson-Sphäre • 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 Steffen kostbaren Schlaf prokrastiniert um einem Ernährungswissenschaftler auf YouTube über die Erde als Dampfnudel zu lauschen. Irgendwie passt es dann doch mit der Analogie. Übermüdet müssen Sie darüber nachdenken, warum vor 66 Millionen Jahren die Dinosaurier - so groß sie auch waren - nicht von der Schwerkraft zu Boden gerissen wurden. Hat der Dampfplauderer etwa recht und war die Schwerkraft vor Millionen von Jahren eine andere? Sind deshalb alle Lebewesen auf der Erde *heutzutage* so viel kleiner, weil die Schwerkraft größer ist als damals? War die Erde kleiner und hatte weniger Masse? Oder ist es nur ein Rechenfehler wie bei der Theorie der 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 9.78 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.2742 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.21 g/cm^3 . Das Gewicht von einem heute lebenden afrikanischen Elefanten liegt bei 5t bis 7t und das Gewicht von einem Tyrannosaurus rex (T. rex) bei 4.5t bis 8t.

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

Die Distanz zwischen Sonne und Erde entspricht 1.01 astronomische Einheiten (AE). Die Einheit 1 AE wird mit $1.55 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 87% Wasserstoffkernen mit einer molaren Masse von 1.05 g/mol , 12% Heliumkernen mit 4.32 g/mol sowie 1% weiteren Atomkernen mit 89.32 g/mol . Die Teilchendichte bei Eintritt in die Erdatmosphäre liegt zwischen 0.4 bis 100 Teilchen cm^{-3} pro Sekunde mit einer mittleren Teilchendichte von 5 cm^{-3} pro Sekunde.

Lösen Sie den folgenden Aufgabenteil mit einer aussagekräftigen Skizze!

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

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

112. Aufgabe

(10 Punkte)

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



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

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

mit

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

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

	Aldridge et al. (2021)	Baxter et al. (2022)	Jacobs et al. (2019)
preening	40cm; 22cm; 25.4%	38cm; 28cm; 50.1%	38cm; 13cm; 25.4%
sitting	38cm; 25cm; 4.3%	32cm; 21cm; 5.2%	31cm; 36cm; 1.8%
walking	39cm; 30cm; 3.6%	32cm; 23cm; 3.6%	41cm; 19cm; 3.6%
wing/leg stretching	37cm; 25cm; 6.2%	30cm; 26cm; 3.2%	37cm; 23cm; 3.2%

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

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

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

113. Aufgabe

(8 Punkte)

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



Nelken von den Molukken Alex und Jessica 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 Alex ziemlich plötzlich, so dass Jessica die Schokobons aus dem Mund fallen. 'Nein, und ehrlich gesagt bin ich auch ziemlich müde...'. Das ist jetzt aber Alex egal, denn er möchte folgende Sachlage diskutieren. Und Alex 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 60 Tagen zu beklagen; nach 105 Tagen rafft die Skorbut eine ganze Schiffsbesatzung dahin«

Ferdinand Magellan stach im Jahre 1519 in See um eine Passage durch den südamerikanischen Kontinent zu finden. Zu seiner Flotte gehörten fünf Schiffe - das Flaggschiff *Trinidad*, die *San Antonio*, die *Victoria*, die *Concepción* und die *Santiago* - mit einer Besatzung von insgesamt 237 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 95 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 $5000\mu\text{g}/50\text{mg}$ 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 24 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 Jessica ist bei Nilufar um gemeinsam *Event Horizon – Am Rande des Universums* zu streamen. Das war jetzt nicht die beste Idee. Denn Jessica kann Horror überhaupt nicht ab. Deshalb flüchtet sie sich in Logik um ihre Emotionen zu bändigen. Nilufar mampft ungerührt Takis Blue Heat. Folgenden Gedankengang nutzt Jessica um dem Film zu entkommen. Die Sonne hat eine aktuelle, angenommene Masse von $2 \times 10^{29} \text{ kg}$. Wenn die Sonne nun am Ende ihrer Lebenszeit zu einem schwarzen Loch mit dem Radius von 2500m kollabiert, wird die Sonne 35% 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 Jessica für die kinetische Energie des Lichtteilchens E_{kin} und der Graviationsenergie 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 $6.674 \cdot 10^{-11} \text{ m}^3 (\text{kg} \cdot \text{s}^2)^{-1}$

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

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

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

115. Aufgabe

(10 Punkte)

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



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

Das Fermi Paradoxon Jessica und Alex 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?'. Alex schaut sie irritiert und interessiert an. Die beiden hat das Problem gepackt. Deshalb wollen Jessica und Alex das Paradoxon mal mathematisch untersuchen! Wie lange würde eine außerirdische Zivilisation benötigen um die gesamte Milchstraße zu besuchen, wenn das maximale Reisetempo die Geschwindigkeit der Voyager 1 Sonde wäre?⁹

Die beiden treffen folgende Annahmen. Eine außerirdische Zivilisation schickt *drei* Voyager 1 ähnliche Sonden mit der Geschwindigkeit von $6.1198 \times 10^4 \text{ km/h}$ los um sich auf den erreichten Planeten selbst zu replizieren. Nach 500 Jahren ist die Replikation abgeschlossen und wiederum *drei* Sonden werden ausgesendet. Gehen Sie von 7.81 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 1.5×10^{11} Sterne in der Milchstraße. Nehmen Sie eine Lichtgeschwindigkeit von $2.9 \times 10^8 \text{ m/s}$ 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 4.6×10^9 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 8×10^7 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 mecklemburgischen Pyramidentage! Sie und Jessica 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 72 Grad im Vergleich zu den ägyptischen Pyramiden mit 60 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 44 Königsellen. Eine Königselle misst 52.2cm.

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

1. Bei der Königspyramide von Meroe soll eine Seitenlänge der Grundfläche 44 Königsellen lang sein. Welche Höhe der Königspyramide in m ergibt sich? **(1 Punkt)**
2. Die Außenflächen der Pyramide soll begrünt werden. Für die Bepflanzung muss eine 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 Jessica 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 75% 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 14° ! **(2 Punkte)**
5. Wie weit reicht Ihre Rampe vom Fuß der Pyramide in die mecklemburgische Landschaft? **(2 Punkte)**

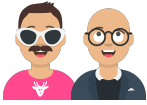
Bei der Besichtigung der Pyramide teilt Ihnen der leicht übergewichtige Pharao (Nebenberuf *Finanzbeamter*) mit, das die Pyramide zu flach sei und somit nicht in die mecklemburgische Landschaft passen würde. Sie müssen nochmal ran.

6. Die Grundfläche der Pyramide ändert sich nicht. Berechnen Sie die Änderung der Höhe in Königsellen, wenn sich der Anstellwinkel der Pyramide um 5° ändert! **(2 Punkte)**

117. Aufgabe

(12 Punkte)

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



Geocaching – Von Satelliten und Plastikdosen Es ist Wochenende und das Wetter ist sweet. Yuki und Alex schwingen sich auf ihre Cachermobile um mit 17km/h, geleitet von modernster Satellitentechnologie und einem Supercompter aus dem Jahr 2000 in den Händen, Plastikdosen in der Natur und an sehenswerten Orten zu finden. Yuki und Alex wollen diesmal endlich die aufwärts 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 Yuki und Alex für die Planung der Route zu Verfügung¹⁰.

Ort	Cache	Wertung (S T G)
A	GCRC52Q	1.0 1.0 Normal
B	GCW53OT	3.5 3.5 Mikro
C	GCNTXDP	3.0 2.0 Klein
D	GC9MOPV	2.0 4.0 Klein
E	GCYZL27	2.5 3.0 Mikro

Im Weiteren sind den beiden folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \overrightarrow{AC} ist 6km. Im Weiteren ist Ihnen der Entfernungsvektor \overrightarrow{CB} mit 6.5km bekannt. Der Entfernungsvektor \overrightarrow{BE} ist das 1.2-fache des Entfernungsvektor \overrightarrow{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 20° südlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 55° 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 südlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort E Ihre Cachertour.

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

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

$$\text{Suchzeit} = 0.15 + 0.13 \cdot \text{Schwierigkeit}$$

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

4. An der höchsten Schwierigkeit müssen Sie angeln. Ihre Angel ist ausgefahren 8m lang. Erreichen Sie einen Cache in der Höhe von 9.8m? Berechnen Sie dazu Ihre maximale mögliche Angelhöhe! Welche Annahmen mussten Sie treffen um die Aufgabe zu lösen? **(2 Punkte)**

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

118. Aufgabe

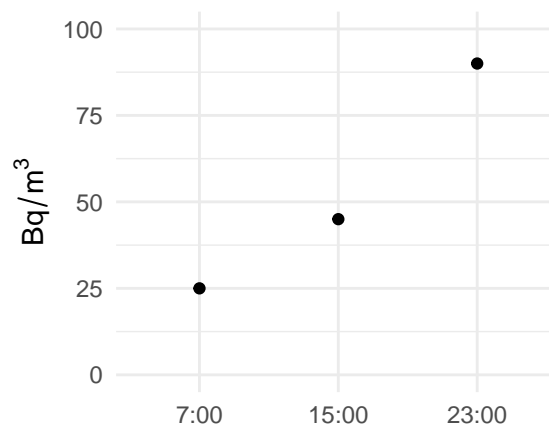
(10 Punkte)

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



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

Die atmende Wand und Brot aus Luft Als Kellerkind¹¹ vom Dorf will Yuki 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 23:00 bestimmt er dreimal automatisch die Radonbelastung in seinem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung¹². Leider helfen die Messwerte Yuki überhaupt nicht weiter. Sie müssen also helfen!



1. Wie lange dauert es in Stunden bis Sie eine kritische Belastung von $320 Bq/m^3$ in Ihrem ungelüfteten Kellerraum erreicht haben? **(2 Punkte)**

Radon zerfällt mit einer Halbwertszeit von 4.1d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 143d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

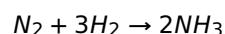
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $320 Bq/m^3$ auf unter $100 Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	78.1	28.4	
Sauerstoff	20.45	15.8	
Kohlenstoffdioxid	0.035	12.1	

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

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



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

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

¹¹Tocotronic - Electric Guitar als passende Untermalung 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 Steffen war nach Ladenschluss bei Penny die Regale einzuräumen. Dabei ist Steffen in der Auslage der Sonderangebote das Necronomicon¹⁴ in die Hände gefallen. Nun ist er eine Magierin der Zeichen geworden! Also eigentlich kann Steffen nur Mathe und das dämliche Necronomicon hat ihn in die Vergangenheit geschleudert... aber gut, was tut man nicht alles im Jahr 1173 n. Chr. für den neuen Lehnsherren Henry dem Roten. Steffen baut natürlich einen Schrottkugelturm um sich den Horden der Finsternis mit genug Schrott erwehren zu können! Steffen 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 $10m/s$ bilden.

1. Wie hoch müssen Sie den Schrottkugelturm bauen lassen, damit sich runde Bleikugeln durch die Fallgeschwindigkeit von $10m/s$ 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 $2.8mm$. Die Löcher im Sieb erlauben ein Tropfendurchmesser von $2mm$. Welchen Durchmesser in mm haben Ihre produzierten Bleikugeln? **(3 Punkte)**

Sie haben jetzt die 2.3×10^5 Bleikugeln zusammen. Blei hat eine Dichte von $15.1g/cm^3$.

4. Wie schwer in Kilogramm kg sind die 2.3×10^5 produzierten Bleikugeln, die Sie jetzt auf die Burgmauer transportieren müssen? **(1 Punkt)**

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

5. Wie groß in cm^2 ist Ihr quadratisches Sieb am oberen Ende des Turms, wenn Sie pro Fall ca. 700 Bleikugeln produzieren wollen und die Bleikugel im Fall $0.5cm$ 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 Yuki mit der Koalakuschule in Down Under nicht geklappt. War vielleicht auch nicht so die beste Idee... aber dafür hat Yuki eine neue Eingebung! Oder wie es Mike Tyson zugeschrieben wird: »Ich wurde nie niedergeschlagen, ich war immer am Aufstehen!«. Daher macht Yuki jetzt einen Großhandel mit Kaninchenfleisch und damit dem teuersten Fleisch in Australien auf. Moment, hopsen hier nicht, seit Thomas Austin im Jahr 1875 ungefähr 30 Kaninchen entlassen hat, Millionen von Kaninchen rum? Wieso ist das Kaninchenfleisch dann so exklusiv? Yuki wird stutzig und frag Sie, dem mal mathematisch nachzugehen!¹⁵

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

$$f(t) = 9 \times 10^9 - 9 \times 10^8 \cdot 1.7^{-0.3 \cdot t + 3.1}$$

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 1.8 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 töteten 99.7% der Kaninchenpopulation innerhalb weniger Wochen.

4. Wie lange in Jahren dauert es bis eine Kaninchenpopulation nach einer Viruspandemie wieder auf 60% der gesättigten Kaninchenpopulation angewachsen ist? **(2 Punkte)**

Thomas Austin entließ die Kaninchen im äußersten Süden von Australien. Australien hat eine West-Ost-Ausdehnung von 4000km und eine Nord-Süd-Ausdehnung von knapp 3500km. Die Kaninchen breiten sich radial mit einer Geschwindigkeit von 7.3km 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 10\$ pro Tier und der durchführende Arzt verlangt ca. 42\$ pro Tier.

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

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

121. Aufgabe

(12 Punkte)

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



Lüneburger Heide. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer des Esels Fridolin und Nilufar. Grünes Gras unter Nilufars Füßen und ein strammer Wind im Gesicht, egal wohin sie schaut. Ein schmatzendes Geräusch ertönt unter Nilufar. Nilufar sinniert, sollte sie ihre weiten Graslandschaften jetzt schon düngen? Dafür benötigt Nilufar die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Nilufar 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 $0.4\times$, Februar mit $0.75\times$ und März mit $1.05\times$. Sie haben noch im letzten Jahr folgende Temperaturen gemessen.

Datum	C°
01. Jan 2023	0.4
01. Feb 2023	1.1
01. Mrz 2023	3.1
01. Apr 2023	4.3

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 190°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 Nilufars Pink Lady Plantage werden Sie beide auf dem Trecker von einer Gruppe elektrifizierter Renter abgedrängt. Der Trecker muss wieder aus dem Graben! Fridolin und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Fridolin mit 180N . Die elektrifizierten Renter bringen eine Kraft von 140N 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 50° Winkel mit ihrem Seil an dem Trecker. Welche Kraft wird aufgebracht? **(2 Punkte)**
7. Mit welcher Beschleunigung ziehen Sie den 1.5t schweren Trecker *jeweils* aus dem Graben, wenn $F = m \cdot a$ gilt? **(1 Punkt)**

122. Aufgabe

(10 Punkte)

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



In der Kartonagenfabrik Mark, Jonas, Alex und Nilufar sitzen im Bus. Wenn man sich zu spät anmeldet, dann ist die Exkursion nicht so toll. Mark hatte den Anderen in der Lerngruppe zu spät Bescheid gesagt. 'Was denn, bin ich eure Nanny oder was?!', entfährt es Mark nachdem die vorwurfsvollen Blicke schon eine Weile auf ihm lasten. Also geht es eben mit Rektor Skinner und Mrs. Krabappel in die Kartonagenfabrik. Wie schon im vorherigen Semester... In der Kartonagenfabrik angekommen erfahren die Vier, dass die Kartons zum Versand von Nägeln nicht hier zusammengebaut werden sondern dass sich die Endfertigung in Flint, Michigan befindet. Unter anderem wird dort der berühmte *Doppelt gewellte, 6-mal-gefaltete, 0.7mm, 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 130m 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 130m 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

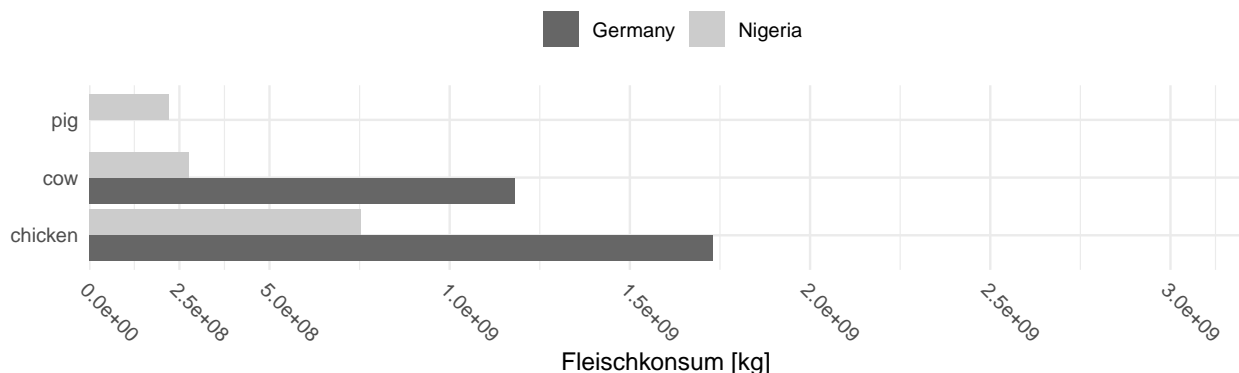
(10 Punkte)

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



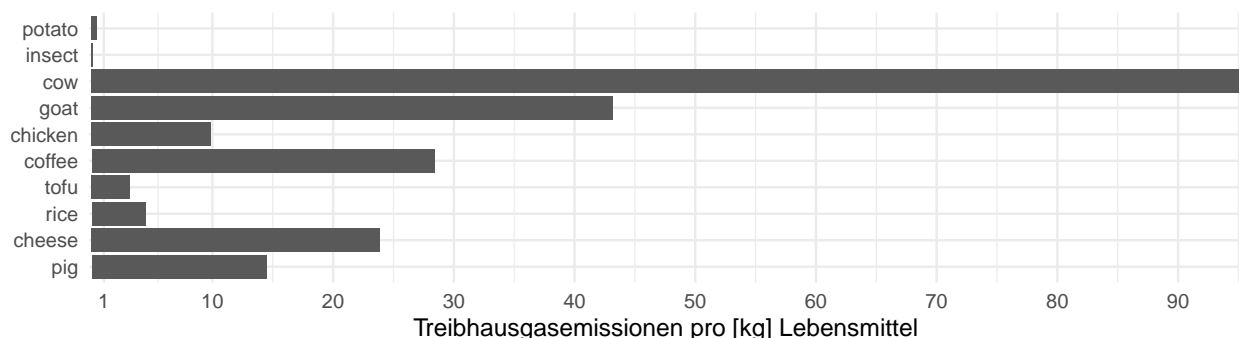
Ein Pfund Insekten, bitte! 'Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹⁶.', merkt Alex an. Die Lerngruppe um Paula, Yuki und Jonas sind bei Alex um mal was außergewöhnliches zu essen. Um den Sinn der Nahrungsumstellung zu verdeutlichen, vergleicht Alex 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 2021 leben ca. 8.2×10^7 Menschen in Deutschland und ca. 1.84×10^8 Menschen in Nigeria. Mit den Informationen wollen Sie und Alex mit der Überzeugungsarbeit anfangen und dann eine Prognose für den Fleischkonsum im Jahr 2050 zu treffen.

Im Folgenden ist Abbildung des Fleischkonsums im Jahr 2021 in Deutschland und Nigeria in [kg] einmal dargestellt.



1. Stellen Sie den Fleischkonsum in Deutschland und Nigeria im Jahr 2021 *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)**

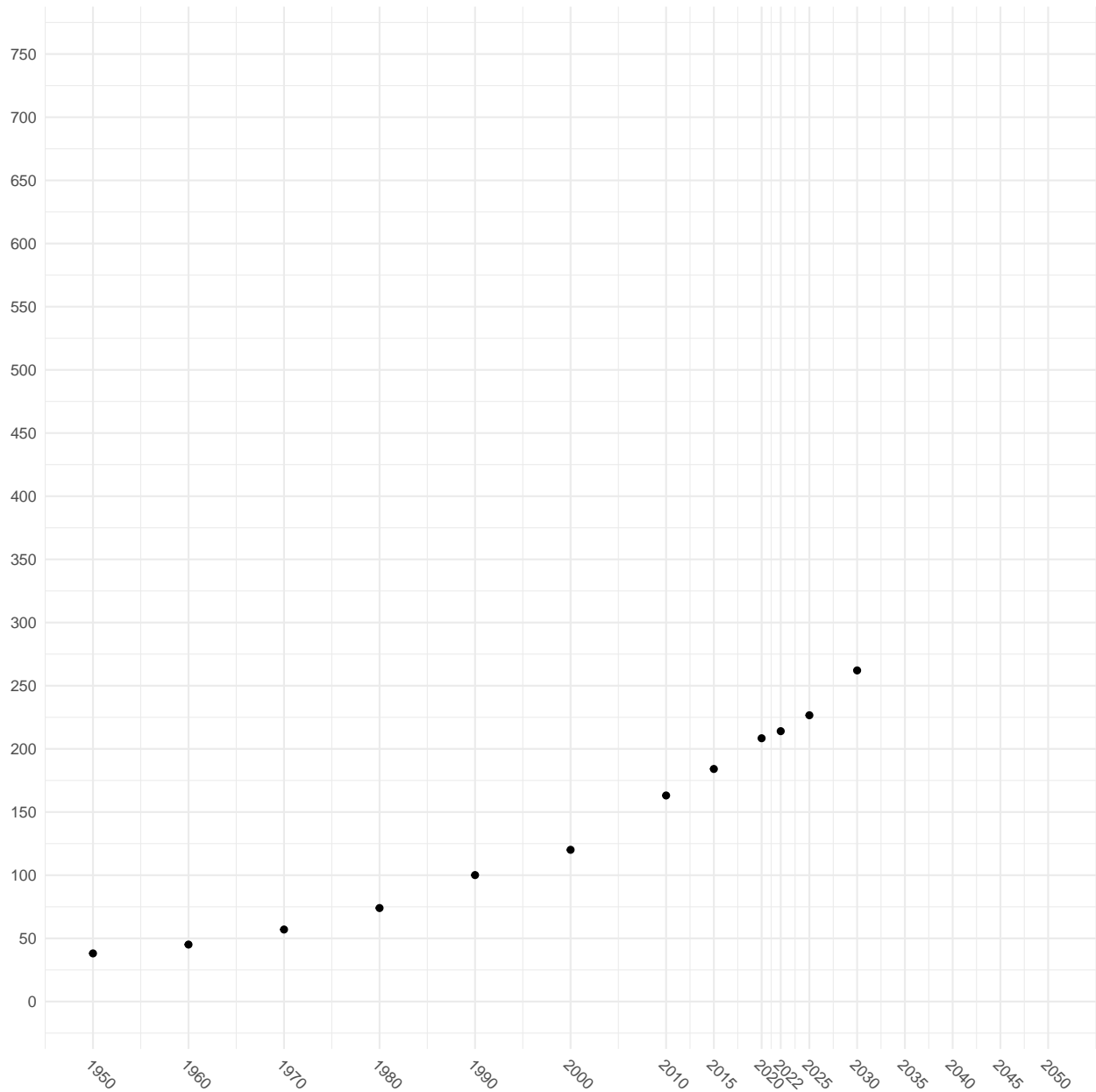
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



3. Stellen Sie in einer Tabelle die Treibhausgasemissionen an CO₂ pro Kopf, die durch den Fleischkonsum in Deutschland und Nigeria im Jahr 2021 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 2021! **(1 Punkt)**
6. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria im Jahr 2050 im Vergleich zum Jahr 2021, 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 Mark die historische Ausgabe des Spiegels aus den 80zigern beiseite. Mark und Tina sind bei ihrem Orthopäden 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 Mark und Tina nun aber über AIDS und dem diagnostischen AIDS-Test, den die beiden nun machen werden? Leider zu wenig. Da brauchen dann Mark und Tina mal wieder Ihre Hilfe bei der Interpretation eines diagnostischen Tests!

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.1% angenommen. In 92% der Fälle ist ein HIV-Test positiv, wenn der Patient erkrankt ist. In 0.5% 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 $n = 10^4$ 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 Mark und Tina, 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 Mark und Tina 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, Steffen und Jessica sitzen bei Alex und hören sich etwas über Network-Marketing an. Alex 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 KH Gesund und Schön Components (KH-GSC). Das Unternehmen steigerte den Umsatz um rund 10 Prozent von 310 Millionen Euro im Jahr 2021. Doch wie viel kommt bei den Partnern an? Laut KH-GSC habe das Unternehmen 3.8×10^5 aktive Partner.

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

Das von Alex zu vermarkende Produkt, hinter dem Alex voll steht, kostet 150EUR pro Einheit im Direktverkauf. Die direkte Provision für die erste Stufe beträgt 25%. Für die zweite, dritte und vierte Stufe betragen die indirekten Provisionen jeweils 1.75%, 0.75% und 0.25%. Jeder von Alex angeworbener »Partner« wirbt wiederum drei Partner für sich selbst an. Pro Monat werden im Schnitt vier Einheiten vom Produkt verkauft. Alex will nun 1800EUR 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)**

Alex musste zum Einstieg bei KH-GSC Einheiten des Produkts für 6750EUR kaufen. Diese Einheiten kann Alex nur direkt verkaufen. Das ganze Wohnzimmer ist voll davon. Leider musste Alex den Kauf über einen Kredit über 5% p.a. über 60 Monate finanzieren. Sie schütteln den Kopf und klären Alex 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 Paula, Nilufar und Mark sitzen bei Alex nachdem sich alle begeistert in der Serie *Stranger Thinks* verloren haben. Alle drei wollen jetzt einmal bei Alex *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 Alex nicht so alleine ist.

In dem Spiel hat Paula nun auf einmal 4 zwölfseitige Würfel ($4d12$) zum würfeln in der Hand. Wenn Paula eine 12 würfelt, hat Paula einen Erfolg.

1. Berechnen Sie die Wahrscheinlichkeit *genau* 3 Erfolge zu erzielen! **(2 Punkte)**
2. Berechnen Sie die Wahrscheinlichkeit keinen Erfolg zu erzielen! **(1 Punkt)**

Nilufar betrachtet nun aufmerksam die ausufernden Ausrüstungstabellen. Nilufar wird aber geholfen und muss sich jetzt nur zwischen der Axt oder dem Schwert entscheiden.

3. Würden Sie die Axt mit zwei zwölfseitigen Würfeln ($2d12$) als Schaden oder das Schwert mit einem sechseitigen Würfel plus 5 ($1d6+5$) als Schaden bevorzugen? Begründen Sie Ihre Antwort mathematisch! **(1 Punkt)**

Jetzt wird es immer wilder, da Nilufar und Mark sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass der Rettungswurf gegen den zaubernden Hexer funktioniert. Nilufar und Mark haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.6$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.75$. Sie haben aber mitgezählt und festgestellt, dass in 40 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 Mark gestern. Die Lerngruppe mit Alex und Paula ging viel zu lang. Während er wegdämmert, kommen in ihm 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 Mark fragt sich, ob er nicht doch lieber bezahlter Gast bei Barbara Salesch hätten sein sollten. Aber Mark 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 Alex und Paula das Team der drei Kandidaten. Mark braucht dringend Ihre Hilfe in seinen Wahnträumen. Sie gehen wie in *Interception* rein!¹⁹

Name	$P(\text{win})$	$P(\text{outbid})$
Alex	0.2	0.12
Paula	0.3	0.02

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 0.08 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 Mark auf der Kirmes und spielen mit einem Typen in einem Tentakelkostüm um das große Geld. Das Glücksrad hat 22 Felder. Sie beide drehen das Glücksrad zweimal. Auf 6 Feldern gewinnen Mark und Sie 3000EUR sonst 2000EUR. Ganz schön viel Geld und ganz schön zwielichtig hier...

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

Im Fiebertraum von Mark reisen sie beide im Zug nach Köln um bei »Geh aufs Ganze!« mitzuspielen. Mark 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.