

Name: _____

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 lieb 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.

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



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

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



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

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



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

Erlaubte Hilfsmittel

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

Endnote

_____ von 20 Punkten sind aus den Multiple Choice Aufgaben erreicht.

_____ von 71 Punkten sind aus den Rechen- und Textaufgaben erreicht.

_____ von 91 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
87.0 - 91.0	1,0
82.5 - 86.5	1,3
78.0 - 82.0	1,7
73.5 - 77.5	2,0
69.0 - 73.0	2,3
64.5 - 68.5	2,7
60.0 - 64.0	3,0
55.5 - 59.5	3,3
51.0 - 55.0	3,7
45.5 - 50.5	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	10	11	9	11	9	12

- Es sind ____ von 71 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)

Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.31$. 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** ☐ Das η^2 wird genutzt um zu erfahren welchen Anteil der Varianz die Behandlungsbedingungen erklären.
- C** ☐ 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.
- D** ☐ Die Berechnung von η^2 ist ein Wert für die Interaktion.
- E** ☐ Das η^2 ist die Korrelation der ANOVA. Mit der Ausnahme, dass 0 der beste Wert ist.

2. Aufgabe

(2 Punkte)

Sie führen ein Feldexperiment durch um das Gewicht von Erbsen zu steigern. Die Pflanzen wachsen unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen. Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.16$. Welche Aussage ist richtig?

- A** ☐ 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.
- B** ☐ Es werden 84% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.
- C** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 16% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 84%.
- D** ☐ Es werden 16% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- E** ☐ Es werden 16% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen erklärt wird.

3. Aufgabe

(2 Punkte)

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

- A** ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 0 annähert kann die Nullhypothese abgelehnt werden.
- B** ☐ Die F-Statistik wird berechnet indem die MS der Behandlung durch die MS des Fehlers geteilt werden. Wenn die F-Statistik sich kaum von der Null unterscheidet kann die Nullhypothese nicht abgelehnt werden.
- C** ☐ Die ANOVA berechnet die T-Statistik indem den Mittelwertsunterschied der Gruppen simultan durch die Standardabweichung der Gruppen teilt. Wenn die T-Statistik höher als 1.96 ist, kann die Nullhypothese abgelehnt werden.

- D** ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 1 annähert kann die Nullhypothese nicht abgelehnt werden.
- E** ☐ Die ANOVA berechnet die T-Statistik aus der Multiplikation der MS Behandlung mit der MS der Fehler. Wenn die F-Statistik genau 0 ist, kann die Nullhypothese nicht abgelehnt werden.

4. Aufgabe

(2 Punkte)

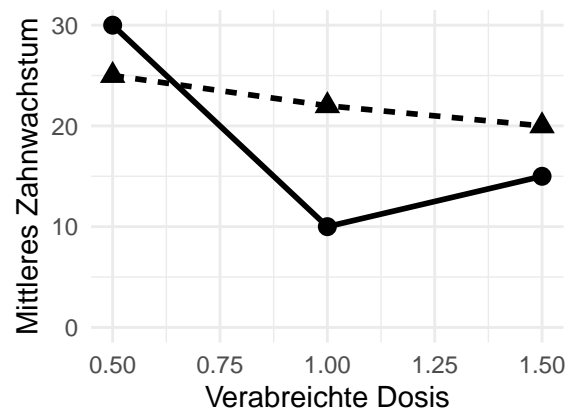
Die ANOVA ist ein statistisches Verfahren welches häufig in den Auswertungen von Experimenten in den Agrarwissenschaften angewendet wird. Dabei wird die ANOVA als ein erstes statistischen Werkzeug für die Übersicht über die Daten benutzt. Eine ANOVA testet dabei...

- A** ☐ ... den Unterschied zwischen zwei paarweisen Mittelwerten aus verschiedenen Behandlungsgruppen. Wenn die signifikant ist, ist daher bekannt welcher Vergleich konkret unterschiedlich ist.
- B** ☐ ... den Unterschied zwischen der Varianz aus verschiedenen Behandlungsgruppen und der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in den Gruppen exakt zu bestimmen.
- C** ☐ ... den Unterschied zwischen der Varianz über alle Behandlungsgruppen oder der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss sich zwischen einem der beiden Varianzquellen entschieden werden.
- D** ☐ ... den Unterschied zwischen der Varianz in den verschiedenen Behandlungsgruppen und der Varianz in einer der Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in einer der Gruppen exakt zu bestimmen.
- E** ☐ ... den Unterschied zwischen der Mittelwerte und der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist bekannt welcher Vergleich konkret unterschiedlich ist.

5. Aufgabe

(2 Punkte)

Ein Versuch wurde an 66 Tieren durchgeführt, wobei jedes Tier eine von drei Vitamin-C-Dosen (0.5, 1 und 1.5 mg/Tag) über eine von zwei Verabreichungsmethoden erhielt. Die folgende Abbildung enthält die Daten aus diesem Versuch zur Bewertung der Wirkung von Vitamin D auf das Zahnwachstum bei Hasen. Welche Aussage ist richtig, wenn Sie eine zweifaktorielle ANOVA rechnen?



- A** ☐ Eine Korrelation liegt vor ($p \leq 0.05$).
- B** ☐ Keine Korrelation liegt vor ($p \geq 0.05$).
- C** ☐ Eine negative Interaktion liegt vor ($p \geq 0.5$).
- D** ☐ Keine Interaktion liegt vor ($p \leq 0.05$).
- E** ☐ Eine mittlere bis starke Interaktion liegt vor ($p \leq 0.05$)

Deskriptive Statistik & Explorative Datenanalyse

6. Aufgabe

(2 Punkte)

Wie lautet der Mittelwert und Standardabweichung von y mit 13, 13, 11, 10 und 10.

- A ☐ Sie erhalten 11.4 +/- 1.23
- B ☐ Es berechnet sich 11.4 +/- 1.52
- C ☐ Sie erhalten 11.4 +/- 0.76
- D ☐ Es berechnet sich 12.4 +/- 2.3
- E ☐ Es ergibt sich 10.4 +/- 1.15

7. Aufgabe

(2 Punkte)

Berechnen Sie den Median, das 1st Quartile sowie das 3rd Quartile von y mit 20, 11, 30, 20, 23, 24, 16, 26, 8, 15 und 42.

- A ☐ Es ergibt sich 20 [15; 26]
- B ☐ Es berechnet sich 21 [16; 27]
- C ☐ Es berechnet sich 21 [16; 25]
- D ☐ Sie erhalten 20 [13; 24]
- E ☐ Es ergibt sich 21 +/- 15

8. Aufgabe

(2 Punkte)

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

- A ☐ Wir brauchen fünf oder mehr Beobachtungen.
- B ☐ erhalten, sollten wir mindestens zwanzig Beobachtungen haben.
- C ☐ Die optimale Anzahl ist größer als hundert Beobachtungen, wobei es gerne sehr viel mehr sein können.
- D ☐ Die untere Grenze liegt bei zwei bis fünf Beobachtungen.
- E ☐ Die untere Grenze liegt bei einer Beobachtung.

9. Aufgabe


(2 Punkte)

Um die Varianz zu berechnen müssen wir folgende Rechenoperationen durchführen.

- A ☐ Den Median berechnen, dann die quadratischen Abstände zum Median aufsummieren, dann die Wurzel ziehen.
- B ☐ Den Mittelwert berechnen und die Abstände quadrieren. Die Summe mit der Fallzahl multiplizieren.
- C ☐ Den Mittelwert berechnen, dann die absoluten Abstände zum Mittelwert aufsummieren
- D ☐ Den Mittelwert berechnen, dann die quadratischen Abstände zum Mittelwert aufsummieren und durch die Fallzahl teilen.
- E ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend teilen wir durch die Fallzahl. Nicht zu vergessen, am Ende dann noch die Wurzel zu ziehen.

10. Aufgabe

(2 Punkte)

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

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

11. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit zu Kartoffel finden Sie auf einmal seltsame Daten. Jedenfalls kommt Ihnen das so vor. Daher berechnen Sie den Mittelwert und den Median. Der Mittelwert \bar{y} und der Median \tilde{y} unterscheiden sich nicht. Welche Aussage ist richtig?

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

12. Aufgabe

(2 Punkte)

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

- A** ☐ Einen Boxplot. Das IQR muss über alle Behandlungen zusammen mit den Whiskers ungefähr gleich aussehen.
- B** ☐ 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.
- C** ☐ Einen Violinplot. Der Bauch der Violine muss hierbei einen höheren Wert annehmen als der Steg der Violine. Dann kann die Annahme einer Varianzhomogenität angenommen werden.
- D** ☐ In einer explorativen Datenanalyse nutzen wir den Violinplot. Dabei sollte der Bauch am Rand liegen. Dann können wir von einer Varianzhomogenität ausgehen.
- E** ☐ Wir erstellen uns für jede Behandlung einen Dotplot und schauen, ob die Dots und damit die Varianz für jede Behandlung gleich groß sind.

13. Aufgabe

(2 Punkte)

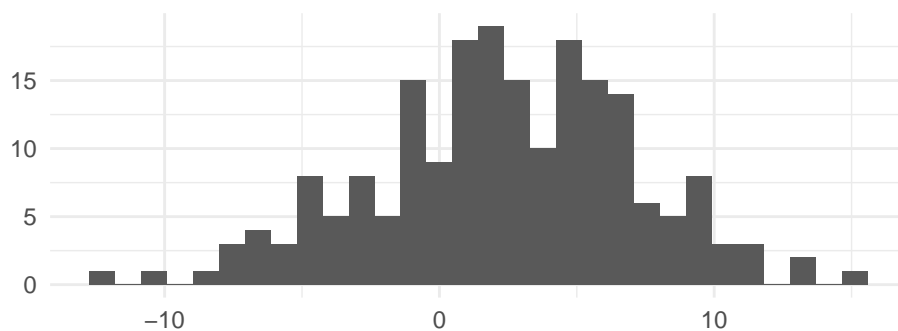
Nach der Durchführung Ihres Feldexperiments wollen Sie eine ANOVA rechnen. Dafür muss aber Ihr Messwert zumindestens approximativ einer Normalverteilung folgen. Welche der drei Abbildungen erlaubt Ihnen abzuschätzen, ob Sie eine Normalverteilung in Ihrem Endpunkt vorliegen haben?

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

14. Aufgabe

(2 Punkte)

Bevor Sie in Ihrer Abschlussarbeit einen statistischen Test rechnen, wollen Sie einmal betrachten, welcher Verteilung Ihre $n = 172$ geernteten Pflanzen folgen. Welche Verteilung ist abgebildet?



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

Lineare Regression & Korrelation

15. Aufgabe

(2 Punkte)

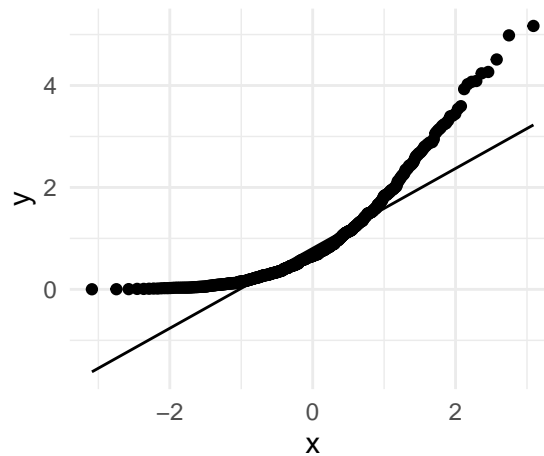
Sie haben das Modell $Y \sim X$ vorliegen und wollen nun ein kausales Modell rechnen. Welche Aussage ist richtig?

- 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 ☐ Ein kausales Modell basiert auf einem Trainingsdatensatz und einem Testdatensatz. Auf dem Trainingsdatensatz wird das Modell trainiert und auf dem Testdatensatz validiert.
- C ☐ Ein kausales Modell benötigt mindestens eine Fallzahl von über 100 Beobachtungen und darf keine fehlenden Werte beinhalten. Die Varianzkomponenten müssen homogen sein.
- D ☐ Wenn ein kausales Modell gerechnet werden soll dann kann dies auf dem gesamten Datensatz geschehen. Das Ziel ist es einen Zusammenhang von X auf Y zu modellieren. Wie wirken sich die Einflussvariablen Y auf die gemessenen Endpunkte $X = x_1, \dots, x_p$ aus?
- E ☐ 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 .

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?

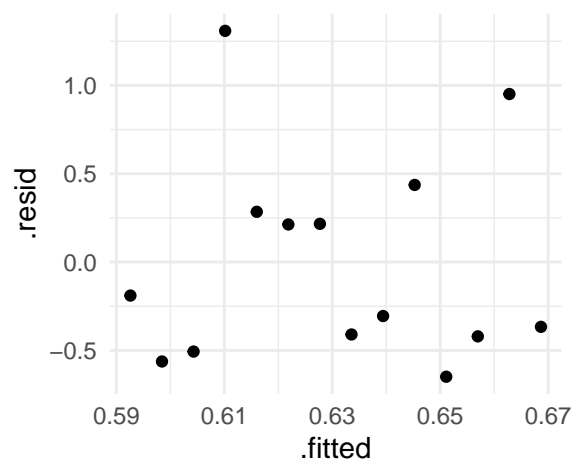


- ☐ A Die Annahme der normalverteilten Residuen ist nicht erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Geraden.
- ☐ B Wir betrachten die Punkte auf der Geraden. Wenn die Punkte einigermaßen auf der Geraden liegen, dann gehen wir von normalverteilten Residuen aus. Wir können hier von normalverteilten Residuen ausgehen.
- ☐ C Wir betrachten die Gerade und dabei insbesondere die beiden Enden der Gerade. Hier sollten die Punkte auf der Geraden liegen, dann ist die Annahme an die Normalverteilung der Residuen erfüllt. Diese Annahme ist nicht erfüllt.
- ☐ D Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.
- ☐ E Wir betrachten die Gerade, die durch die einzelnen Punkte laufen sollte. Wenn die 95% der Punkte von der Geraden getroffen werden, dann gehen wir von normalverteilten Residuen aus.

17. Aufgabe

(2 Punkte)

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



- ☐ A Wenn wir die Nulllinie betrachten so liegen die Punkte nicht gleichmäßig über und unter der Nulllinie. Unser Modell erfüllt nicht die Annahme von normalverteilten Residuen mit einem Mittelwert von 0 und einer Streuung von s^2 .

- B** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Es ist ein Muster zu erkennen und wir können damit auf die Signifikanz von x_1, \dots, x_p schließen.
- C** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Kein Muster ist zu erkennen und keine Outlier zu beobachten.
- D** ☐ Die Punkte müssen gleichmäßig in dem positiven Bereich liegen. Dies ist hier klar nicht der Fall. Einzelne Ausreißer können beobachtet werden. Die Analyse ist gescheitert.
- 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 wird wie das η^2 aus der ANOVA interpretiert. Der Korrelationskoeffizienten ρ beschreibt den Anteil an erklärter Varianz durch die Regression. Dabei gibt er jedoch eine Richtung an und kann auch negativ werden.
- B** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen 0 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos. Eine Signifikanz kann nicht nachgewiesen werden.
- C** ☐ Korrelationskoeffizienten ρ liegt zwischen 0 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als Standardisierung verstanden werden.
- D** ☐ Der Korrelationskoeffizienten ρ liegt zwischen -1 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als standardisierte Steigung verstanden werden.
- E** ☐ 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.

19. Aufgabe

(2 Punkte)

Sie haben ein Feldexperiment mit Erbsen durchgeführt und wollen nun in einer simplen linearen Regression den Einfluss der NO_3 -Konzentration in $[\mu g]$ im Wasser auf den absoluten Proteingehalt in $[kg]$ untersuchen. Sie erhalten einen β_{NO_3} Koeffizienten von 1.1×10^{-5} und einen p -Wert mit 0.00051. Welche Aussage zu der Signifikanz und dem Effekt ist richtig?

- A** ☐ Das Gewicht und die NO_3 -Konzentration korrelieren sehr stark, deshalb wird der β_{NO_3} Koeffizient sehr klein. Mit einer ANOVA kann für die Korrelation korrigiert werden und der Effektschätzer passt dann zum p -Wert.
- B** ☐ Die Einheit der NO_3 -Konzentration ist zu klein gewählt. Die Erhöhung der NO_3 -Konzentration um 1 Einheit führt nur zu einem sehr winzigen Anstieg von β_{NO_3} im Gewicht der Wasserlinsen. Die Einheit $[\mu g]$ muss besser gewählt werden.
- C** ☐ Wenn der Effekt β_{NO_3} 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_{NO_3} in x . Wir müssen daher die Einheit von y entsprechend anpassen.
- D** ☐ Die Fallzahl ist zu hoch angesetzt. Je höher die Fallzahl ist, desto kleiner ist die Teststatistik und damit ist dann auch der p -Wert sehr klein. Es sollte über eine Reduzierung der Fallzahl nachgedacht werden. Dann sollte der Effekt zum p -Wert passen.
- E** ☐ Die Einheit der NO_3 -Konzentration ist zu klein gewählt. Dadurch sehen wir den sehr kleinen p -Wert. Der p -Wert und die Einheit von der NO_3 -Konzentration hängen antiproportional zusammen.

20. Aufgabe

(2 Punkte)

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

- ☐ A Die Funktion `lm()` in R 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()` in R 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.
- ☐ C Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich.
- ☐ D Die Funktion `lm()` in R wird klassischerweise für die nicht-lineare Regression genutzt. Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt.
- ☐ E Neben der klassischen Verwendung der Funktion `lm()` in der linearen Regression kann auch ein Gruppenvergleich gerechnet werden. Dafür müssen aber alle Faktoren aus den Daten entfernt und numerisch umgewandelt werden. Dann kann das R Paket `{emmeans}` genutzt werden um die Korrelation zu berechnen. Eine Adjustierung ist dann nicht mehr notwendig.

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 R ist mit dem *generalisierten linearen Modell (GLM)* eine Modellierung implementiert, die die Poissonverteilung für Zähldaten oder die Binomialverteilung für 0/1-Daten modellieren kann. Weitere Modellierungen sind in R auch mit zusätzlich geladenen Paketen nicht möglich.
- ☐ B Das GLM ist eine allgemeine Erweiterung der linearen Regression auf die Normalverteilung.
- ☐ C Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsfamilien für das Y bzw. das Outcome in einer linearen Regression zu wählen.
- ☐ D Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsgruppen für das X bzw. die Einflussvariablen in einer linearen Regression zu wählen.
- ☐ E Dank dem *generalisierten linearen Modell (GLM)* können auch andere Verteilungsfamilien – außer die Normalverteilung – mit einer linearen Regression modelliert werden. Dafür werden alle Verteilungen in eine Normalverteilung überführt und anschließend standardisiert.

Vermischte Themen

22. Aufgabe

(2 Punkte)

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

- ☐ A Durch eine Randomisierung können wir nicht von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.
- ☐ B Randomisierung erlaubt erst die Varianzen zu schätzen. Ohne eine Randomisierung ist die Berechnung von Mittelwerten und Varianzen nicht möglich. Dadurch lässt sich erst ein Experiment auswerten.
- ☐ C Randomisierung war bis 1952 bedeutend, wurde dann aber in Folge besserer Rechnerleistung nicht mehr verwendet. Aktuelle Statistik nutzt keine Randomisierung mehr.
- ☐ D Strukturgleichheit ist durch Randomisierung gegeben. Leider hilft die Randomisierung noch nicht um von der Stichprobe auf die Grundgesamtheit zu schließen. Deshalb wurde das Falsifikationsprinzip entwickelt.

- ☐ E Randomisierung sorgt für Strukturgleichheit und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.

23. Aufgabe

(2 Punkte)

Wenn Sie einen Datensatz erstellen, dann ist es ratsam die Spalten und die Einträge in englischer Sprache zu verfassen, wenn Sie später die Daten in R auswerten wollen. Welcher Aussage ist richtig?

- ☐ A Im Allgemeinen haben Programmiersprachen Probleme mit Umlauten und Sonderzeichen, die in der deutschen Sprache vorkommen. Eine Nutzung der englischen Sprache umgeht dieses Problem auf einfache Art.
- ☐ B Programmiersprachen haben Probleme mit Umlauten und Sonderzeichen der deutschen Sprache. Daher ist die Nutzung in Deutsch in den AGBs von R untersagt.
- ☐ C Es gibt keinen Grund nicht auch deutsche Wörter zu verwenden. Es ist ein Stilmittel.
- ☐ D R Pakete sind nur in englischer Sprache verfasst. Es macht keinen Sinn R daher in Deutsch zu bedienen.
- ☐ E Alle Funktionen und auch Anwendungen sind in R in englischer Sprache. Die Nutzung von deutschen Wörtern ist nicht schick und das ist zu vermeiden.

24. Aufgabe

(2 Punkte)

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

- ☐ A Für eine explorativen Datenanalyse (EDA) in R müssen wir als erstes die Daten über `read_excel()` einlesen. Danach müssen wir schauen, dass wir die Zeilen richtig über `mutate()` transformiert haben. Insbesondere müssen Variablen mit kontinuierlichen Werten in einen Faktor umgewandelt werden. Am Ende nutzen wir die Funktion `ggplot()` für die eigentlich EDA.
- ☐ B 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.
- ☐ C 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.
- ☐ 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 Für eine explorativen Datenanalyse (EDA) in R müssen wir als erstes die Daten über `read_excel()` einlesen. Danach müssen wir schauen, dass wir die Spalten richtig über `mutate()` transformiert haben. Insbesondere müssen Variablen mit Kategorien in einen Faktor umgewandelt werden. Am Ende nutzen wir die Funktion `ggplot()` für die eigentlich EDA.

25. Aufgabe

(2 Punkte)

Sie haben das abstrakte Modell $Y \sim X$ mit X als Faktor mit zwei Leveln vorliegen. Welche Aussage über $s_1^2 \neq s_2^2$ ist richtig?

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

26. Aufgabe

(2 Punkte)

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

- A** ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.
- B** ☐ Die Ferkel stammen von der gleichen Sau und sind somit untereinander unabhängig.
- C** ☐ Je nach Stallanlage kommt eine andere Analyse in Betracht. Eine allgemeine Aussage über Ferkel und Sauen lässt sich statistisch nicht treffen.
- D** ☐ Untereinander stark korreliert. Die Ferkel sind von einer Mutter und somit miteinander korreliert. Dies wird in der Statistik jedoch meist nicht modelliert.
- E** ☐ Die Ferkel stammen vom gleichen Muttertier und haben vermutlich eine ähnlichere Varianzstruktur als die Ferkel von anderen Sauen. Die Ferkel sind untereinander über die Mutter abhängig.

27. Aufgabe

(2 Punkte)

In einer Studie wollen Sie den Effektschätzer Odds ratio berechnen. Sie finden in Ihrem Experiment zur Behandlung von Klaueninfektionen bei Ziegen in 6 Tieren Erkrankung der Klauen vor. 12 Tiere sind gesund. Welche Aussage ist richtig?

- A** ☐ Das Verhältnis der Chancen Odds ratio ergibt ein Chancenverhältnis von 0.33. Wir sind an der Chance krank zu sein interessiert.
- B** ☐ Es ergibt sich ein Odds ratio von 0.33, da es sich um eine Chancenverhältnis handelt.
- C** ☐ Da es sich um ein Chancenverhältnis handelt ergibt sich ein Odds ratio von 0.5.
- D** ☐ Es ergibt sich ein Odds ratio von 0.33, da es sich um ein Anteil handelt. Wir berechnen den Anteil der Kranken.
- E** ☐ Der Anteil der Gesunden wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Odds ratio von 0.33.

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** ☐ 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.
- B** ☐ Eine große Fallzahl führt zu mehr signifikanten Ergebnissen auch bei kleinen Effekten. Daher werden fast alle Vergleich esignifikant, wenn die Fallzahl nur groß genug wird.
- C** ☐ Aktuell werden immer größere Datensätze erhoben. Dadurch wird auch die Varianz immer höher was automatisch zu mehr signifikanten Ergebnissen führt.
- D** ☐ Relevanz und Signifikanz haben nichts miteinander zu tun. Daher gibt es auch keinen Zusammenhang zwischen hoher Fahlzahl ($n > 10000$) und einem signifikanten Test. Ein Effekt ist immer relevant und somit signifikant.
- 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: 0.34, 0.01, 0.21, 0.001, 0.02 und 0.89. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 1, 0.06, 1, 0.006, 0.12 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.0567, 0.0017, 0.035, 2e-04, 0.0033 und 0.1483. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.0567, 0.0017, 0.035, 2e-04, 0.0033 und 0.1483. Die adjustierten p-Werte werden zu einem α -Niveau von 0.83% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 1, 0.06, 1, 0.006, 0.12 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 0.83% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 2.04, 0.06, 1.26, 0.006, 0.12 und 5.34. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.

30. Aufgabe

(2 Punkte)

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

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

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 {emmeans} erlaubt die Durchführung eines multiplen Gruppenvergleichs. Aus einem {emmeans} Objekt lässt sich recht einfach das CLD erstellen und so über Barplots eine schnelle Interpretation der statistischen Auswertung durchführen.
- B** ☐ Das R Paket {ggplot}. Wir erhalten hier sofort eine Visualisierung der Daten. Anhand der Visualisierung lässt sich eine explorative Datenanalyse durchführen, die gleichwertig zu einem Posthoc-Test ist.
- C** ☐ Das R Paket {emmeans} erlaubt die Durchführung eines multiplen Gruppenvergleichs. Aus einem emmeans Objekt lässt sich leider kein CLD erstellen. Dennoch ist das Paket einfach zu bedienen und wird deshalb genutzt. Die Interpretation der statistischen Auswertung wird über einen Barplot abgebildet.
- D** ☐ Das R Paket {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.
- E** ☐ 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.

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 liegt nicht mehr bei 5% sondern sehr viel niedriger, bei ca. 1%. Es kommt zu einer α -Hyperinflation. Dagegen kann mit der Adjustierung der p-Werte nach Bonferroni vorgegangen werden.
- B** ☐ 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.
- C** ☐ 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.
- D** ☐ 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.
- E** ☐ Beim multiplen Testen kann es zu einer α -Deflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist. Die p-Werte werden durch die Anzahl an Vergleichen geteilt

33. Aufgabe

(2 Punkte)

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

- A** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ nach Bonferroni adjustiert werden. Dafür wird der Effekt mit der Anzahl an Vergleichen k multipliziert. Dies geschieht analog zu den p-Werten.
- B** ☐ Beim multiplen Testen werden die Effekte der paarweisen Vergleiche ignoriert. Der Nachteil des multiplen Testens ist ja auch, dass wir am Ende keine Effekte mehr vorliegen haben. Eine ANOVA liefert hier bessere Informationen.
- C** ☐ Beim multiplen Testen muss der Effekt, hier der Mittelwertsunterschied Δ aus den paarweisen t-Tests, nicht adjustiert werden.
- D** ☐ Beim multiplen Testen kann es zu einer Effektüberschätzung (Δ -Inflation) kommen. Daher müssen die Effekte angepasst werden. Dies geschieht nicht händisch sondern intern in den angewendeten Algorithmen.
- E** ☐ Beim multiplen Testen kann es zu einer Δ -Inflation kommen. Das globale Effektniveau liegt nicht mehr bei 20%. Daher müssen die Effekte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der Effekte nach Bonferroni das bekannteste Verfahren ist.

Statistische Testtheorie

34. Aufgabe

(2 Punkte)

Sie haben den mathematischen Ausdruck $Pr(D|H_0)$ vorliegen, welche Aussage ist richtig?

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

35. Aufgabe

(2 Punkte)

Die Testtheorie hat einen philosophischen Unterbau. Eins der Prinzipien ist das Falsifikationsprinzip. Das Falsifikationsprinzip besagt,

- A** ☐ ... dass ein minderwertiges Modell durch ein minderwertiges Modell ersetzt wird. Es gilt das Verifikationsprinzip nach Karl Popper.
- B** ☐ ... dass ein schlechtes Modell durch ein schlechteres Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.
- C** ☐ ... dass Annahmen an statistische Modelle meist falsch sind.
- D** ☐ ... dass ein schlechtes Modell durch das Falsifikationsprinzip durch ein weniger schlechtes Modell ersetzt wird.
- E** ☐ ... dass in der Wissenschaft immer etwas falsch sein muss. Sonst gebe es keinen Fortschritt.

36. Aufgabe

(2 Punkte)

Das Signifikanzniveau α wird auch Fehler 1. Art genannt und liegt bei 5%. Warum wurde der Grenzwert von 5% als Signifikanzschwelle gewählt?

- A** ☐ Der Begründer der modernen Statistik, R. Fischer, hat die Grenze simuliert und berechnet. Dadurch ergibt sich dieser optimale Cut-Off.
- B** ☐ Der Wert ergab sich aus einer Auswertung von 1042 wissenschaftlichen Veröffentlichungen zwischen 1914 und 1948. Der Wert 5% wurde in 28% der Veröffentlichungen genutzt. Daher legte man sich auf diese Zahl fest.
- C** ☐ Da Wissenschaftler eine Schwelle für die statistische Testentscheidung benötigen wurde α historisch gewählt. Damit ist $\alpha = 5\%$ eine Kulturkonstante.
- 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** ☐ Im Rahmen eines langen Disputs zwischen Neyman und Fischer wurde $\alpha = 5\%$ festgelegt. Leider werden die Randbedingungen und Voraussetzungen an statistische Modelle heute immer wieder ignoriert.

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 $T_D = \frac{\text{signal}}{\text{noise}}$
- D** ☐ Es gilt $T_D = \frac{\text{noise}}{\text{signal}}$
- E** ☐ Es gilt $T_D = \frac{\text{signal}}{\text{noise}^2}$

38. Aufgabe

(2 Punkte)

In der Theorie zur statistischen Testentscheidung kann folgende Aussage in welche richtige Analogie gesetzt werden?

H_0 ablehnen obwohl die H_0 gilt

- A ☐ Fire without alarm, dem β -Fehler als Analogie von Rauch im Haus.
- B ☐ Dem β -Fehler mit der Analogie eines Rauchmelders: Fire without alarm.
- C ☐ In die Analogie eines Rauchmelders: Alarm without fire police, dem α -Fehler.
- D ☐ In die Analogie eines Feuerwehrautos: Car without noise.
- E ☐ Dem α -Fehler in der Analogie eines Rauchmelder: Alarm without fire.

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 ☐ Über das Konfidenzintervall. Das Konfidenzintervall beitet eine Entscheidung über die Signifikanz und zusätzlich kann über die Visualisierung des Konfidenzintervalls eine Relevanzschwelle definiert werden.
- C ☐ Der p-Wert. Durch den Vergleich mit α lässt sich über die Signifikanz entscheiden und der β -Fehler erlaubt über die Power eine Einschätzung der Relevanz.
- D ☐ Das OR. Als Chancenverhältnis gibt es das Verhältnis von Relevanz und Signifikanz wieder.
- E ☐ Die Teststatistik. Durch den Vergleich von T_c zu T_k ist es möglich die H_0 abzulehnen. Die Relevanz ergibt sich aus der Fläche rechts vom dem T_c -Wert.

40. Aufgabe

(2 Punkte)

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

- A ☐ Wir 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.
- B ☐ Wir schauen, ob der p-Wert kleiner ist als das Signifikanzniveau α und vergleichen somit Wahrscheinlichkeiten. Die Wahrscheinlichkeiten werden als Flächen unter der Kurve der Teststatistik dargestellt, wenn die H_0 gilt.
- C ☐ 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.
- D ☐ 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.
- E ☐ Wir machen eine Aussage über die individuelle Wahrscheinlichkeit des Eintretens der Nullhypothese H_0 . Der p-Wert wird mit dem Signifikanzniveau verglichen und bewertet.

41. Aufgabe

(2 Punkte)

Die Ergebnisse der einer statistischen Analyse können in die Analogie einer Wettervorhersage gebracht werden. Welche Analogie für die Ergebnisse eines statistischen Tests trifft am besten zu?

- A** ☐ In der Analogie der Wahrscheinlichkeit für Regen: ein statistischer Test erlaubt die Wahrscheinlichkeit für ein Ereignis abzuschätzen. Die Stärke des Effektes können wir nicht bestimmen.
- B** ☐ In der Analogie der Sonnenscheindauer: Wie lange kann mit einem entsprechenden Effekt gerechnet werden? Die Wahrscheinlichkeit für den Effekt gibt der statistische Test wieder.
- C** ☐ 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.
- D** ☐ In der Analogie des Niederschlags oder Regenmenge: ein statistischer Test gibt die Stärke eines Effektes wieder. Zum Beispiel, wie hoch ist der Mittelwertsunterschied.
- E** ☐ 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.

42. Aufgabe

(2 Punkte)

Sie wollen eine Aussage über ein untersuchtes Individuum treffen. Dazu nutzen Sie einen statistischen Test. Können Sie eine valide Aussage aus einem statistischen Test erhalten?

- A** ☐ Nein, es ist nicht möglich ein untersuchtes Individuum mit einem t-Test auszuwerten. Wir erhalten dann leider keine Aussage zum Individuum.
- B** ☐ Ja, es ist möglich ein untersuchtes Individuum mit einem t-Test auszuwerten. Wir erhalten dann eine Aussage zum Individuum.
- C** ☐ 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.
- D** ☐ Nein, wir erhalten nur eine Aussage zu zwei Individuen. Ein statistischer Test liefert Informationen zu einem Individuum im Vergleich zu einem anderen Individuum.
- E** ☐ Ja, wir können ein untersuchtes Individuum nicht mit einem t-Test auswerten. Wir erhalten keine Aussage zum Individuum. Wir können aber den Effekt als Quelle der Relevanz nutzen.

43. Aufgabe

(2 Punkte)

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

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

44. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit sollen Sie neben den p-Werten auch die Effekte mit angeben. Welche Aussage ist richtig?

- A** ☐ 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.
- B** ☐ 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.
- C** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Damit ist der Effekt direkt mit dem Begriff der Signifikanz verbunden. Die Entscheidung über die Signifikanz trifft der Forschende unabhängig von der Relevanz eines statistischen Tests.
- D** ☐ Der Effekt eines statistischen Tests beschreibt den Output oder die Wiedergabe eines Tests in einem Computer.
- E** ☐ 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.

45. Aufgabe

(2 Punkte)

Welche Aussage über die Entscheidung anhand des p-Wertes gegen die Nullhypothese ist richtig?

- A** ☐ Ist $Pr(D|H_0)$ kleiner als das Signifikanzniveau α gleich 5% dann wird die Nullhypothese H_0 abgelehnt.
- B** ☐ Anhand des p-Wertes lässt sich wie folgt eine Entscheidung treffen. Liegt der Wert über oder gleich dem Signifikanzniveau α dann kann die Nullhypothese abgelehnt werden.
- C** ☐ Anhand des p-Wertes 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)

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

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

Statistische Tests für Gruppenvergleiche

47. Aufgabe

(2 Punkte)

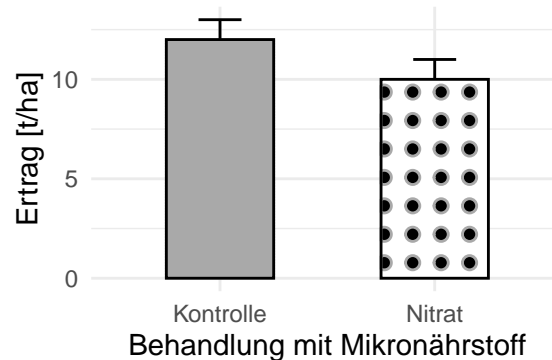
In Ihrer Abschlussarbeit rechnen Sie einen Student t-Test. Welche Aussage ist auch für den Welch t-Test richtig?

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

48. Aufgabe

(2 Punkte)

Ein Versuch wurde in 7 Parzellen pro Gruppe durchgeführt. Die folgende Abbildung enthält die Daten aus diesem Versuch zur Bewertung der Wirkung des Mikronährstoff Nitrat auf den Ertrag in t/ha von Mango im Vergleich zu einer Kontrolle. Welche Aussage ist richtig, wenn Sie einen t-Test rechnen?



- A** ☐ Die Barplots deuten auf kein signifikanten Unterschied. Der Effekt liegt vermutlich bei -2.
- B** ☐ Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei -2 unter einer groben Abschätzung. Wir müssen aber eine ANOVA rechnen um den Effekt wirklich bestimmen zu können.
- C** ☐ Die Barplots deuten auf keinen signifikanten Unterschied. Der Effekt liegt vermutlich bei -2. Wir müssen aber einen Posthoc-Test rechnen um den Effekt wirklich bestimmen zu können.
- D** ☐ 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.
- E** ☐ Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei -2 unter einer groben Abschätzung.

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** ☐ Beim gepaarten t-Test kombinieren wir die Vorteile des Student t-Test für Varianzhomogenität mit den Vorteilen des Welch t-Test für Varianzheterogenität. Wir bilden dafür die Differenz der Einzelbeobachtungen.
- B** ☐ 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.

- C** ☐ Der gepaarte t-Test wird gerechnet, wenn die Beobachtungen abhängig voneinander sind. Wir messen jede Beobachtung nur einmal und berechnen dann die Differenz zu dem Mittel der anderen Beobachtungen.
- 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** ☐ Der gepaarte t-Test nutzt die Varianz der Beobachtungen jeweils paarweise und bildet dafür eine verbundene Stichprobe. Dieser Datensatz d dient dann zur Differenzbildung.

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 fünf Rapssorten und eine ANOVA ergibt $p = 0.049$ für den Ertrag. Sie schauen sich auch die rohen, unadjustierten p-Werte an und finden hier als niedrigsten p-Wert $p_{3-2} = 0.051$. Welche Aussage ist richtig?

- A** ☐ Das ist kein Wunder. Die ANOVA testet nicht auf der gesamten Fallzahl und die paarweisen t-Tests gewinnen immer eine oder mehr Gruppen als Fallzahl dazu. Mit steigender Fallzahl sind mehr signifikante Unterschiede zu erwarten. Die p-Werte unterscheiden sich numerisch auch kaum.
- B** ☐ Es gibt einen Fehler in der Varianzstruktur. Daher kann die ANOVA nicht richtig sein und paarweise t-Tests liefern das richtige Ergebnis.
- C** ☐ Die adjustierten p-Werte deuten in die richtige Richtung. Zusammen mit den nicht signifikanten rohen p-Werten ist von einem Fehler in der ANOVA auszugehen.
- 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 Unterscheid nachweisen.
- E** ☐ Hier kommt der Effekt der steigenden Fallzahl auf die Anzahl an signifikante Ergebnisse zu tragen. Da die ANOVA auf mehr Fallzahl testet als die einzelnen paarweisen t-Tests, kann die ANOVA leichter einen signifikanten Unterscheid nachweisen. Die p-Werte sind immer etwas kleiner als bei den t-Tests.

Teil I.

Deskriptive Statistik & Explorative Datenanalyse

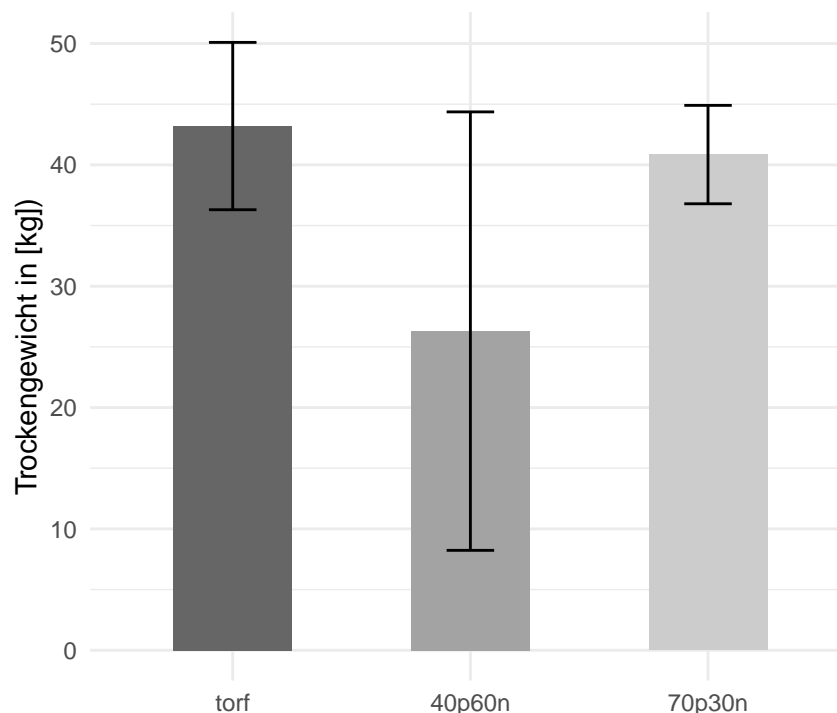
51. Aufgabe

(8 Punkte)

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



Zerforschen des Barplots Jessica steht vor einem ersten Problem, denn wenn es nach ihrer Betreuerin geht, soll sie in einem Gewächshausexperiment Maiss auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Jessica liebt Warhammer. Darin kann sie sich wirklich verlieren und immer wieder neu begeistern. Das heißt erstmal überlegen für Jessica. 'Hm...', Schokobons und David Bowie. Das ist und bleibt die beste Kombination zum Nachdenken für Jessica. Die Behandlung werden verschiedene Substrattypen (*torf*, *40p60n* und *70p30n*) sein. In ihrer Exceldatei wird sie den Outcome (Y) *Trockengewicht* als *drymatter* aufnehmen. Vorab soll Jessica aber einmal die folgenden Barplots ihrer Betreuerin nachbauen, damit sie den R Code schonmal für später vorliegen hat. Damit geht das Problem schon los. Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen.



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

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

52. Aufgabe

(8 Punkte)

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



Visualisierung des Barplots Yuki und die Faulheit, eine unendliche Geschichte mit kniffligen Wendungen. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Yuki ist es eine Möglichkeit schneller ans Ziel zu gelangen. Deshalb hat sich Yuki viele Poster in der Fakultät angeschaut und ist zum Schluß gekommen, dass Barplots eine häufig genutzte Abbildung sind. Yuki soll nun in seinem Projektbericht Brokoli untersuchen. Die Behandlung in seinem Projektbericht sind verschiedene Düngestufen (*ctrl*, *low* und *high*). Erhoben wurden von Yuki als Messwert (*Y*) *Frischegewicht* benannt als *freshmatter* in seiner Exceldatei. Erwartungsgemäß erhält er von seiner Betreuerin den Auftrag die erhobenen Daten als Barplots darzustellen. Dann kann Yuki auch schonmal abschätzen, was bei einem statistischen Test rauskommen könnte. Na dann mal los. Yuki schafft sich die nötige Stimmung. Wenn London Grammar ertönt, dann sucht das Minischwein schleunigst Schutz unter dem Sofa. Yuki schüttelt den Kopf.

treatment	freshmatter
high	34.8
high	36.4
ctrl	43.7
low	30.2
high	35.8
high	34.7
ctrl	39.1
low	25.9
ctrl	49.9
low	30.9
high	32.9

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

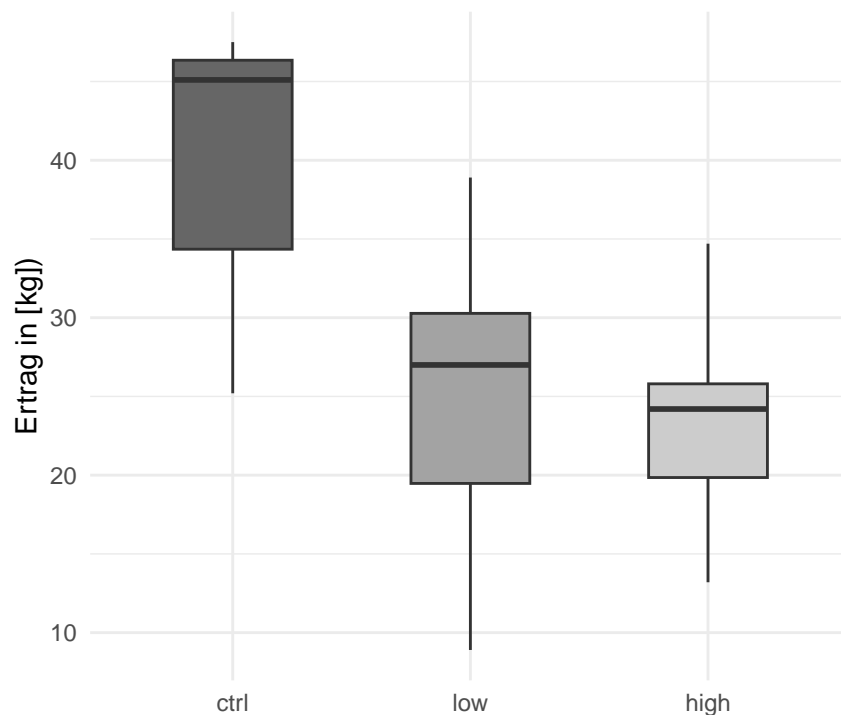
53. Aufgabe


(9 Punkte)


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



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



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

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

54. Aufgabe

(9 Punkte)

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



Visualisierung des Boxplots Boxplots sind bedeutend in der Darstellung von wissenschaftlichen Ergebnissen. Leider hat sich Jessica nicht gemerkt, welche statistischen Maßzahlen für einen Boxplot erhoben werden müssen. Besser wäre was anderes gewesen. Am Ende dann doch besser Warhammer. Wunderbar. Eine echte Ablenkung für Jessica. Das ist in soweit doof, da nach ihrem Betreuer nun Boxplots aus ihren Daten gebaut werden sollen, bevor es mit dem statistischen Testen weitergeht. Anhand von Boxplots lässt sich eine Aussage über die Varianzhomogenität über die Behandlungsgruppen treffen. Die Behandlung für Brokoli waren verschiedene Substrattypen (*torf* und *70p30n*). Erfasst wurde von Jessica als Outcome (Y) *Proteingehalt*. Jessica hat dann *protein* in ihrer Exceldatei eintragen. Aber nur in passender Atmosphäre! Schon dutzende Male gesehen: Herr der Ringe. Aber immer noch großartig zusammen mit Schokobons.

treatment	drymatter
70p30n	52.9
torf	28.1
torf	37.6
torf	19.1
70p30n	39.3
torf	30.1
70p30n	36.5
torf	41.6
70p30n	41.9
70p30n	41.7
70p30n	61.9
70p30n	50.7
torf	24.1
torf	22.6
torf	23.1
70p30n	41.0
70p30n	42.2
torf	29.0
70p30n	42.9

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

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

55. Aufgabe

(8 Punkte)

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



In einem Gespräch mit ihrem Betreuer wird Nilufar gebeten seine Daten aus einem Kreuzungsexperiment mit Lamas in einem Histogramm darzustellen. Aus den Boxen wummert Deichkind und ihr Mund ist verklebt von Takis Blue Heat. 'Herrlich', denkt Nilufar. In ihrem Experiment hat er die Anzahl an gedrehten Haaren/Federn pro cm^2 erst fotografiert und dann ausgezählt. Laut ihrem Betreuer soll das Histogramm helfen, die Verteilung der die Anzahl an gedrehten Haaren/Federn pro cm^2 zu bestimmen. Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für sie war schon immer die Erwartung gewesen. Ein leidiges Lied. Nilufar streichelt liebevoll das Huhn. Der Kopf ist in ihrem Schloß vergraben um den Klang von Deichkind zu dämpfen.

Die Anzahl an gedrehten Haaren/Federn pro cm^2 : 4, 3, 5, 4, 3, 2, 5, 5, 4, 5, 3, 1, 4, 4, 5, 4, 9, 5, 1, 5, 7, 1, 2, 3, 6, 4, 1, 8, 4, 2, 6, 3, 1, 9, 7, 4

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

1. Zeichnen Sie ein Histogramm um die Verteilung der Daten zu visualisieren! (3 Punkte)
2. Beschriften Sie die Achsen der Abbildung! (2 Punkte)
3. Ergänzen Sie die absoluten und relativen Häufigkeiten in der Abbildung! (1 Punkt)
4. Berechnen Sie aus den Daten die *Wahrscheinlichkeit* 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!



'Hm...', Marzipankugeln und Andrea Berg. Das ist und bleibt die beste Kombination zum Nachdenken für Mark. Mark betrachtet die folgenden Daten nach einem Versuch in einer Klimakammer mit Kartoffeln. In dem Experiment wurden die mittleren Mehлтаuspots gezählt. Nach der Meinung seiner Betreuerin muss als erstes geschaut werden, wie diese verteilt sind. Also welcher statistischen Verteilung die mittleren Mehлтаuspots folgen. Dazu soll Mark ein Histogramm verwenden. Dann hätte man auch einen guten Überblick über den Endpunkt (Y). Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für ihn war schon immer die Unsicherheit gewesen. Ein leidiges Lied. Wenn Andrea Berg ertönt, dann sucht der Hamster schleunigst Schutz unter dem Sofa. Mark schüttelt den Kopf.

Die mittleren Mehлтаuspots: 10.6, 10.7, 10.6, 10.8, 5.9, 10.6, 11.1, 14.7, 10.7, 10.2, 10.8, 7.2, 9.2, 8.9, 9.1, 8.1, 7.2, 10.2, 10.9, 9.1, 11.9, 13.8

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

1. Zeichnen Sie ein Histogramm um die Verteilung der Daten zu visualisieren! **(3 Punkte)**
2. 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!



Alex schmeißt noch eine Handvoll Gummibärchen in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Abba. Jetzt heißt es aber erstmal auf was anderes konzentrieren. Alex möchte gerne den Zusammenhang zwischen durchschnittliche UV Einstrahlung [UV/d] und Trockengewicht [kg/ha] im Kontext von Spargel herausfinden. Hierfür hat Alex ein Feldexperiment im Emsland durchgeführt. Nach einigen unvorhergesehenen Ereignissen hat er es geschafft folgende Datentabelle zu erstellen. Eine echte Herausforderung für ihn war schon immer die Gefälligkeit gewesen. Ein leidiges Lied. Aber das steht auch nicht im Zentrum. Nun stellt sich die Frage für ihn, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Deshalb möchte Alex als erstes eine explorative Datenanalyse durchführen. Dann was anderes. Irgendwie komisch, wenn er Alien anmacht, dann ist die Katze eigentlich sofort vor dem Bildschirm und starrt hinein.

Trockengewicht [kg/ha]	Durchschnittliche UV Einstrahlung [UV/d]
8.9	13.0
15.8	15.7
21.8	19.3
18.3	22.3
9.5	13.5
29.7	24.9
16.0	15.9
22.3	22.5
6.4	5.4
14.3	18.1
15.8	18.0

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

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

58. Aufgabe

(10 Punkte)

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



Zwei kategoriale Variablen darzustellen ist nicht so einfach. Jessica hatte erst über einen Mittelwert nachgedacht, dann aber die Idee verworfen. Wäre da nicht noch was anderes. Eine echte Herausforderung für sie war schon immer der Mangel gewesen. Ein leidiges Lied. Dabei hatte sie sich in ein Feldexperiment im Oldenburger Land zum einen die Behandlung Mechanische Bearbeitung [ja/nein] und zum anderen die Messung Chlorophyllgehalt unter Zielwert [ja/nein] im Kontext von Brokkoli angeschaut. Jetzt möchte ihr Betreuer erstmal die langen Tabellen mit ja/nein in einer explorativen Datenanalyse zusammengefasst bekommen. Sonst geht es bei ihrer Abschlussarbeit nicht weiter. Was super nervig ist. Um Rad zu fahren geht Jessica dann später nochmal raus. Echte Entspannung.

Mechanische Bearbeitung	Chlorophyllgehalt unter Zielwert	Mechanische Bearbeitung	Chlorophyllgehalt unter Zielwert
ja	nein	nein	ja
ja	nein	nein	nein
ja	nein	ja	nein
nein	ja	nein	nein
nein	nein	ja	ja
nein	nein	ja	nein
ja	nein	ja	nein
nein	nein	nein	nein
ja	nein	ja	nein
nein	nein	ja	ja
ja	ja	ja	nein
nein	nein	ja	nein
nein	ja	nein	nein
nein	nein	ja	ja
ja	nein	nein	nein
nein	nein	nein	nein
nein	ja	ja	nein
nein	ja	ja	nein

Leider kennt sich Jessica 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 *kein* Effekt von der Behandlung vorliegen würde, wie würde die Tabelle und die Visualisierung aussehen? **(2 Punkt)**

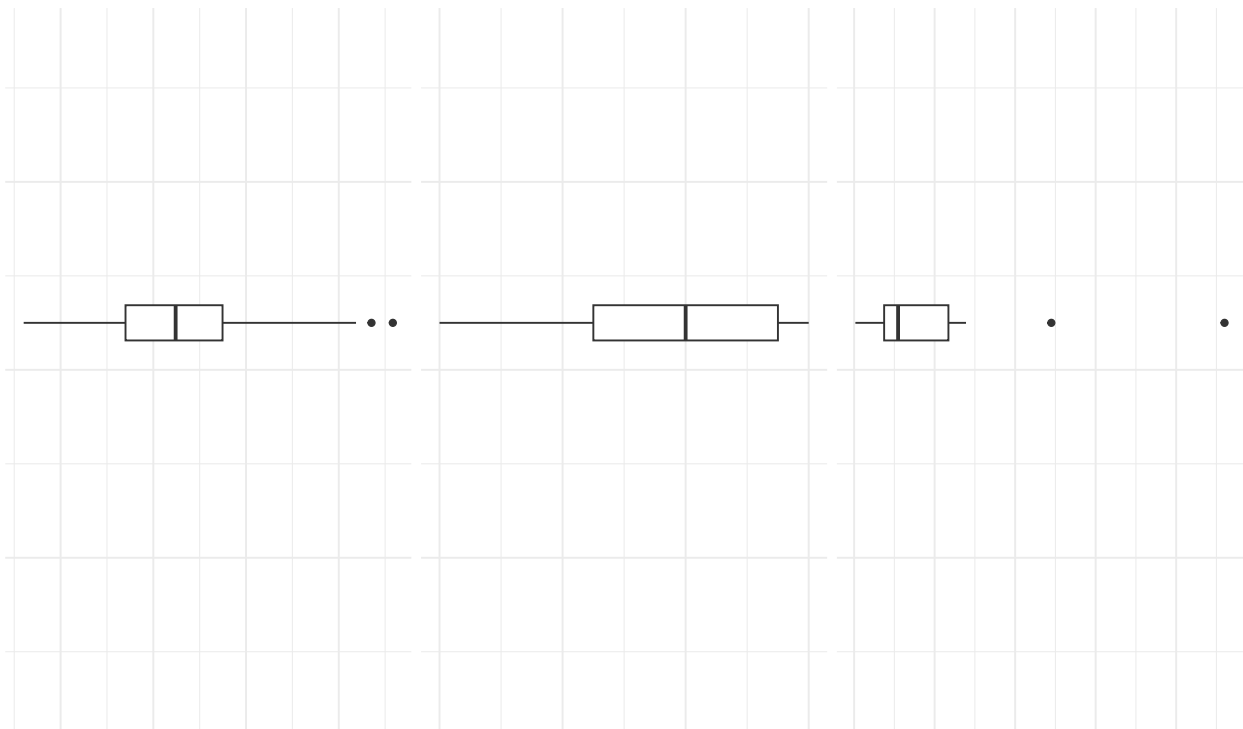
59. Aufgabe

(10 Punkte)

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



‘Was soll das denn jetzt schon wieder sein? Drei Boxplot, die auf der Seite liegen?’, entfährt es Paula und schaut dabei Tina an. ‘Keine Ahnung. Es ist bestimmt wieder so ein Lernziel mit der Verteilung und so.’, meint Tina sichtlich genervt und mampft noch ein paar Katjes. ‘Du weißt doch wie es heißt, *Frei ist, wer missfallen kann*.¹’, merkt Paula nickend an. 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 der Perfektionismus machen die Sache nicht einfacher.



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

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

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

60. Aufgabe

(10 Punkte)

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



Mark und die Unsicherheit machen die Sache mit dem Studium nicht einfacher. Immerhin ist noch Nilufar zur Hilfe mit dabei. Nilufar hat Marzipankugeln mitgebracht und Andrea Berg aufgedreht. Das ist immerhin eine Ablenkung. Nicht so gut wie Geocaching, 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 Mark. Durch das Mampfen von Nilufar versteht er kein Wort der Antwort. Nilufar lächelt.

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

1. Skizzieren Sie vier Normalverteilungen mit $\bar{y}_1 \neq \bar{y}_2 \neq \bar{y}_3 \neq \bar{y}_4$ und $s_1 = s_2 = s_3 = s_4$! **(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!

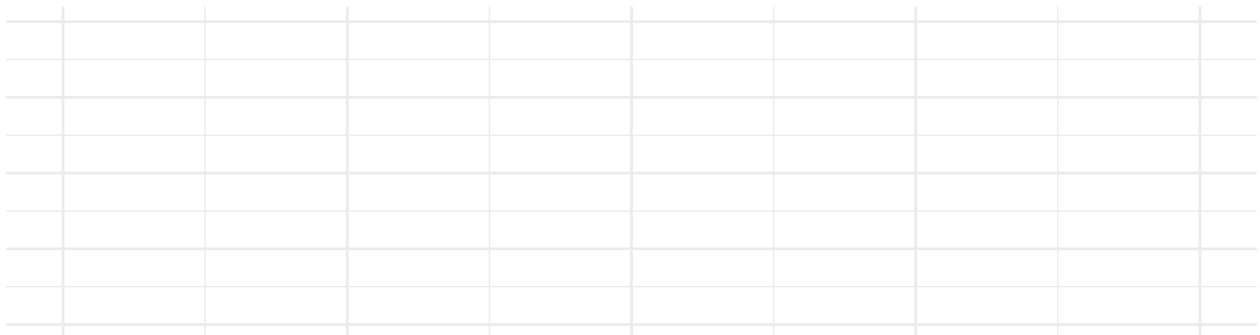


‘Was sollen wir hier dann noch zeichnen?!’, entfährt es Nilufar und schaut dabei Jonas an. ‘Wir sollen eine Normalverteilung mit einem Mittelwert von $\bar{y}_1 = 4$ und einer Standardabweichung von $s_1 = 1$ zeichnen. Sowie eine weitere Normalverteilung mit einem Mittelwert von $\bar{y}_2 = 0$ und einer Standardabweichung von $s_2 = 1$. 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 $\lambda_2 = 1$ gezeichnet werden.’, meint Jonas sichtlich genervt und mampft noch ein paar Snickers. Im Hintergrund spielt leise Iron Maiden. ‘Wirre Geschichte...’, merkt Nilufar nickend an. Die beiden schauen angestrengt auf die leeren Flächen für die Abbildungen. Jonas und die Erwartung machen die Suche nach der Lösung nicht einfacher.

Zwei Normalverteilungen



Zwei Poissonverteilungen



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

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

Teil II.

Statistisches Testen & statistische Testtheorie

62. Aufgabe

(9 Punkte)



Grundlage des statistischen Testen ist das Verständnis von der Grundgesamtheit (eng. *population* oder *ground truth*) und der experimentellen Stichprobe (eng. *sample*).

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)



Für ein besseres Verständnis der statistischen Testtheorie, auch Null-Ritual genannt, kann eine Visualisierung als Kreuztabelle genutzt werden.

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

5% Richtige Entscheidung H_0 beibehalten β -Fehler

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)**

64. Aufgabe

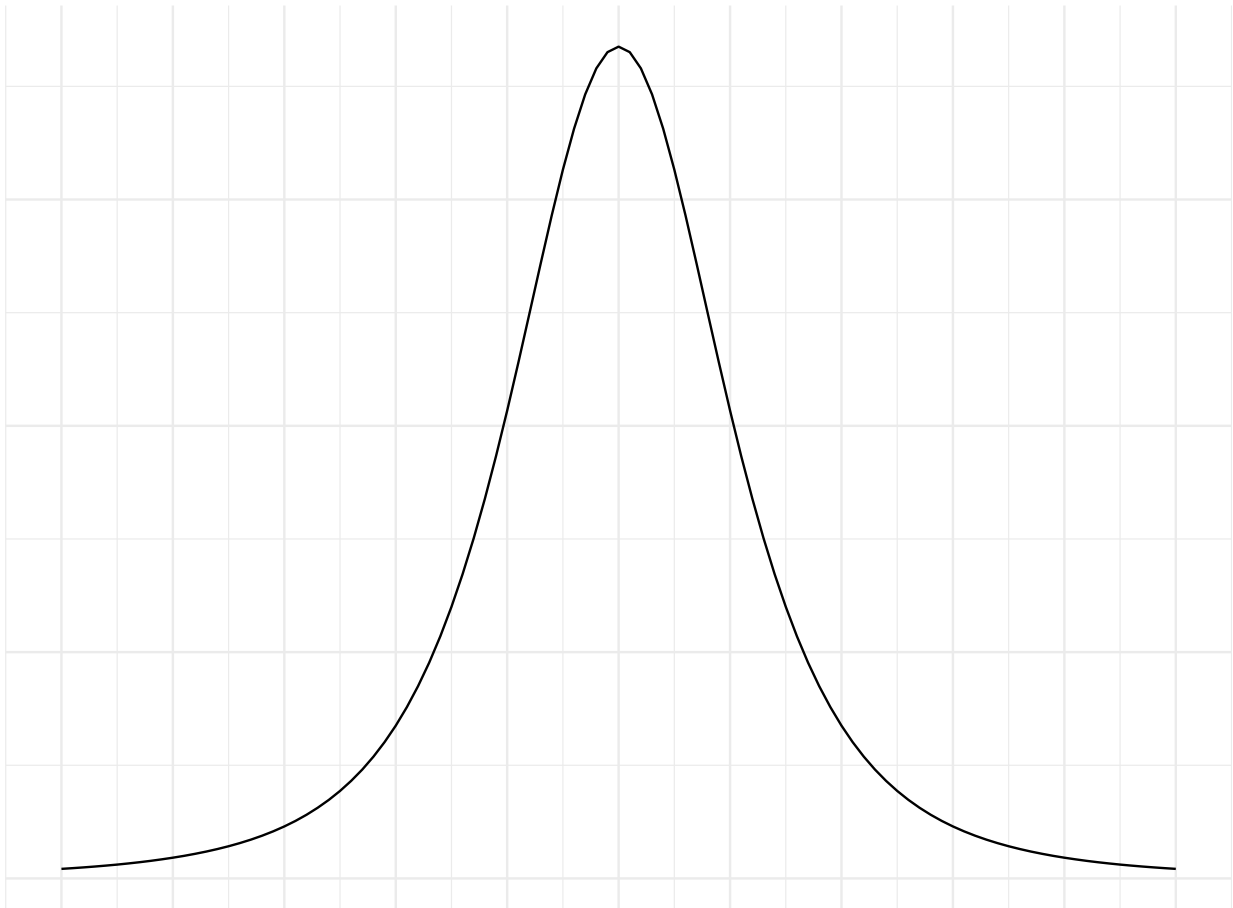
(9 Punkte)

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



Abgebildet ist die t-Verteilung unter der Annahme der Gültigkeit der Nullhypothese. Beachten Sie, dass im Folgenden keine numerisch korrekte Darstellung verlangt wird! Es gilt Erkennbarkeit vor Genauigkeit!

1. Ergänzen Sie eine beschriftete x-Achse! **(1 Punkt)**
2. Ergänzen Sie „ $\bar{y}_1 = \bar{y}_2$ “! **(1 Punkt)**
3. Ergänzen Sie „ $A = 0.95$ “! **(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 signifikant p-Wert in die Abbildung! Begründen Sie Ihre Antwort! **(2 Punkte)**



65. Aufgabe

(10 Punkte)



Sie rechnen einen t-Test für Gruppenvergleiche der Mittelwerte. Sie schätzen den Unterschied zwischen dem mittleren Befall mit Parasiten zu einer unbehandelten Kontrolle.

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



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 $KI_{1-\alpha}$?

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

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

3. Visualisieren Sie ein 95%-iges Konfidenzintervall im Vergleich zu einem 90%-igen Konfidenzintervall! Begründen Sie Ihre Visualisierung anhand der Formel des Konfidenzintervalls des t-Tests mathematisch! **(3 Punkte)**

Teil 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!



Der t-Test. Mark erschauert. Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. Ein mächtiges Werkzeug ist der t-Test in den Händen desjenigen, der einen normalverteilten Endpunkt (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Mark überhaupt aus? 'Hm...', Marzipankugeln und Andrea Berg. Das ist und bleibt die beste Kombination zum Nachdenken für Mark. Mark hat ein Feldexperiment mit Spargel durchgeführt um eine neue technische Versuchsanlage zu testen. Bei dem Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) wurde die Behandlung Genotypen (AA und BB) an den Spargel getestet und dabei wurde geschaut, ob der Versuch überhaupt technisch klappen könnte. Gemessen hat Mark dann als Messwert Proteingehalt [g/kg]. Warum der Versuch in der Uckermark für seinen Projektbericht stattfinden musste, ist ihm bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Proteingehalt [g/kg]?

treatment	weight
dose	31.1
dose	18.0
ctrl	12.7
ctrl	13.0
dose	16.5
ctrl	12.6

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

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

68. Aufgabe

(12 Punkte)

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



Die Uckermark, unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer von Yuki, 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 Yuki. Für seiner Hausarbeit ist Yuki ins Nichts gezogen. Yuki und die Faulheit, eine unendliche Geschichte mit kniffligen Wendungen. Was macht er nun? Yuki hat ein Freilandversuch mit Lauch durchgeführt. Die Behandlung Bewässerungstypen (*low* und *high*) wurde an Lauch getestet. Gemessen hat er dann als einen normalverteilten Messwert (Y) Proteingehalt [g/kg]. Jetzt soll er seinem Betreuer nach testen, ob die Behandlung Bewässerungstypen (*low* und *high*) ein signifikantes Ergebnis liefert. Hm..., was spannendes wäre gut. 'Hm...', Reese's Peanut Butter Cups und London Grammar. Das ist und bleibt die beste Kombination zum Nachdenken für Yuki.

Bewässerungstypen	Proteingehalt
high	41.1
low	22.3
low	24.0
high	42.8
high	43.6
low	25.7
high	36.5
low	25.5
high	34.0
high	36.3
low	23.8
high	35.7
low	22.4
high	39.0
low	22.0
high	50.1
low	27.5
high	45.5
high	47.4

Leider kennt sich Yuki 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 $T_{\alpha=5\%} = 1.84$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *keinen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann der Effekt? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Yuki über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

69. Aufgabe

(12 Punkte)

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



'Der t-Test testet einen normalverteilten Messwert (Y).', liest Alex laut. Das hilft jetzt auch nur bedingt weiter. Alex und die Gefälligkeit, eine unendliche Geschichte mit kniffligen Wendungen. Laut seinem Betreuer ist zwar ihm Messwert Proteingehalt [g/kg] normalverteilt, aber wie rechnet er jetzt einen t-Test? Für seine Abschlussarbeit musste er einen Versuch in einer Klimakammer mit Erdbeeren im Oldenburger Land durchführen. Als wäre das nicht schon anstrengend genug gewesen. Jetzt soll er auch noch testen, ob die Behandlung Bewässerungstypen (*low* und *high*) ein signifikantes Ergebnis liefert. Hm..., was spannendes wäre gut. Schon dutzende Male gesehen: Alien. Aber immer noch großartig zusammen mit Gummibärchen.

Bewässerungstypen	Proteingehalt
low	21.0
high	29.1
high	45.5
high	31.1
low	16.0
low	19.4
high	60.1
high	51.7
low	21.4
low	25.5
high	37.0
high	53.0
high	35.6
high	54.0
high	32.2
low	28.6
low	19.0
low	28.5

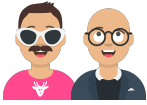
Leider kennt sich Alex mit der Berechnung eines t-Tests überhaupt nicht aus. Deshalb braucht er bei der Berechnung Ihre Hilfe!

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 $T_{\alpha=5\%} = 1.84$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie das 90% 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!



Alles voll mit Lamas. Das haben Alex und Yuki gemeinsam in einem Projektbericht gemacht! Worum ging es aber konkret? Beide haben als ein normalverteiltes Outcome (Y) Schlachtgewicht [kg] von Lamas bestimmt. Die Daten haben beide zusammen in einem Stallexperiment erhoben. In dem Experiment ging es um eine vorher/nachher Untersuchung an den gleichen Lamas. Als Behandlung wurde Flüssignahrung (1l/d und 5l/d) eingesetzt. Nach der Meinung des Betreuers muss hier ein gepaarter t-Test gerechnet werden. Alex schaut nachdenklich zu Yuki. Wenn die Gefälligkeit nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so... Steffen denkt derweil angestrengt an Reese's Peanut Butter Cups. Im Hintergrund wummert Abba.

ID	treatment	freshmatter
2	1l/d	41.4
4	1l/d	28.5
1	1l/d	38.6
3	1l/d	36.3
7	1l/d	42.9
3	5l/d	20.3
7	5l/d	25.1
9	5l/d	18.7
4	5l/d	24.5
5	5l/d	20.5
6	1l/d	42.2
6	5l/d	20.7
5	1l/d	34.0
11	5l/d	17.5
2	5l/d	29.3
10	5l/d	26.4
8	5l/d	25.9
1	5l/d	27.1

Leider kennen sich Alex und Yuki 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 Alex über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

71. Aufgabe

(10 Punkte)

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



Jonas und Tina sind bei Nilufar um sich Hilfe in  zu holen. Im Hintergrund wummert Deichkind. Die beiden hatten zwar schon erste Kontakte mit  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 Tina nicht das Problem hätte die Erschöpfung zu händeln. In einer Abschlussarbeit haben beide zusammen Erbsen untersucht. Dabei ging es um den Zusammenhang zwischen der Behandlung Genotypen (AA und BB) und dem Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Der Versuch wurde in einem Versuch in einer Klimakammer im Teuteburgerwald durchgeführt. Nach der Betreuung ist der Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] 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: Chlorophyllgehalt by Genotypen
## t = 0.71999, df = 15, p-value = 0.4826
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -6.433844 12.997733
## sample estimates:
## mean in group AA mean in group BB
## 36.74444 33.46250
```

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.13|$ 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!



Almería. Spanien. Sonne und Strand. Paula und Yuki haben ihren gemeinsamen Auslandsaufenthalt sichtlich genossen. Dann hatte sich auch noch angeboten ihre Abschlussarbeit gemeinsam in Almería durchzuführen. Es hätte sogar noch besser funktioniert, wenn Nilufar nicht die Erwartung ein paar Mal im Weg gestanden hätte und Paula nicht das Problem gehabt hätte die Faulheit zu händeln. Nun müssen jetzt alle Daten in **R** ausgewertet werden, da **R** international der Standard in der Datenauswertung ist und die Betreuer in Spanien nur **R** können. Während beide Nilufar Oliven mit Takis Blue Heat füttern, hoffen Paula und Yuki mehr Informationen von Nilufar über die seltsame **R** Ausgabe des t-Tests. Immerhin erinnern beide sich an die Behandlung Substrattypen (*torf* und *70p30n*) und das es um Erdbeeren ging. Im Hintergrund wummert Deichkind und Fotos zeigen Nilufar mit dem Hobby Hip Hop.

```
##  
## Two Sample t-test  
##  
## data: Proteingehalt by Substrattypen  
## t = 2.4893, df = 17, p-value = 0.02346  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## 1.108114 13.429664  
## sample estimates:  
## mean in group torf mean in group 70p30n  
## 45.78889 38.52000
```

Helfen Sie Nilufar bei der Interpretation des t-Tests! Sonst geht es auch für Paula und Yuki 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!



'Mit dem R Paket {emmeans} können wir gleich die Gruppenvergleiche rechnen und uns das *compact letter displac*' ausgeben lassen!', verkündet Jessica sichtlich stolz. Ein paar Mal hat sie schon der Mangel gehindert weiterzumachen. 'Nach Meinung der Betreuerin soll es aber nur erstmal ein t-Test sein. Und die Ausgabe ist schon wirr genug.', merkt Paula an. Paula und Tina sind bei Jessica um sich in  helfen zu lassen. Im Hintergrund wummert David Bowie. Tina streichelt zur Beruhigung die Hündin von Jessica. Die beiden waren 3 Monate im Teuteburgerwald um einen Versuch mit Puten in einem Stallexperiment durchzuführen. Ziel war es das Outcome (Y) Schlachtgewicht [kg] zu bestimmen. Jessica überlegt, ob sie die beiden nicht noch auf den Film *Herr der Ringe* einlädt oder dann doch lieber raus geht um Rad zu fahren? Vielleicht will ja Tina mit. Besser als der Film.

```
##
## Two Sample t-test
##
## data: Schlachtgewicht by Ernährungszusatz
## t = -2.1969, df = 18, p-value = 0.04136
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -16.745871 -0.374129
## sample estimates:
## mean in group ctrl mean in group fedX
##                33.96                42.52
```

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


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

74. Aufgabe


(10 Punkte)

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



Es gibt ja immer die Möglichkeit sich Hilfe zu holen. Das geht natürlich auch immer in einem Projektbericht. Deshalb arbeiten Mark und Jessica gemeinsam an einem Projektbericht. Das macht dann auch die Analyse ihres Hauptversuches einfacher. Zwar hat jeder von ihnen noch ein Subthema, aber auch da kann man sich ja helfen. Das hilft dann teilweise nur bedingt. Eine echte Herausforderung für ihn war schon immer die Unsicherheit gewesen. Ein leidiges Lied. In dem Hauptversuch wurde Folgendes von den beiden gemacht. Mark und Jessica haben sich Milchvieh angeschaut. Dabei geht um Zusammenhang zwischen Genotypisierung (0d und 14d) und Protein/Fettrate [%/kg]. Jetzt sollen beide einen gepaarten t-Test rechnen. Leider kennen sich beide nicht sehr gut in  aus. Aber wenigstens haben beide eine Menge an Marzipankugeln und in der Wohnung wummert Andrea Berg.

```
##
## Paired t-test
##
## data: Protein/Fettrate by Genotypisierung
## t = -2.6261, df = 8, p-value = 0.03036
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -13.793715 -0.895174
## sample estimates:
## mean difference
## -7.344444
```

Jetzt brauchen Mark und Jessica Ihre Hilfe bei der Berechnung eines gepaarten t-Tests in  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?', Nilufar schaut Alex fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Die beiden hatten sich auf einem Konzert von Abba kennengelernt. Alex tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren im Teutoburgerwald um ein Stallexperiment mit Milchvieh 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 Alien geguckt. Alex befürwortet das!

Ernährungszusatz	Gewichtszuwachs
getIt	49
ctrl	35
getIt	45
ctrl	35
fedX	25
getIt	46
getIt	46
ctrl	35
fedX	25
fedX	23
ctrl	35
fedX	24
getIt	44
ctrl	35
fedX	23
fedX	27
getIt	43
getIt	45
ctrl	35
fedX	26
ctrl	35

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

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

5. Schätzen Sie die Effekte der Behandlungsstufen! **(2 Punkte)**

76. Aufgabe

(9 Punkte)

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



Ergebnistabelle der einfaktoriellen ANOVA 'Als erstes bauen wir uns aus unsere Daten die ANOVA Tabelle dann sehen wir schon, ob unser Gruppenvergleich in der ANOVA signifikant ist.', Jonas schaut Tina fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Da hilft die Spinne von Tina auch nur bedingt. Tina tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren im Teuteburgerwald um ein Kreuzungsexperiment mit Fleischrindern durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Elterlinie (*ctrl*, *Standard*, *Yray* und *Xray*) und dem Messwert Protein/Fettrate [%/kg] gibt. Nachher wollen sich beide noch mit dem Hobby Astronomie von Tina beschäftigen. Kennt Jonas noch nicht, klingt aber interessant.

Leider kennen sich Jonas und Tina mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe, die Spinne 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)
Elterlinie	3				
error	24	483.37			
Total	27	5215			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.01$ ab. Begründen Sie Ihre Antwort! (2 Punkte)
5. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? (2 Punkte)

77. Aufgabe

(12 Punkte)

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



Die einfaktoriellen ANOVA und der Student t-Test Jonas und Paula schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten mit einer einfaktoriellen ANOVA auswerten damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse in den multiplen Gruppenvergleichen zu erwarten sind. Deshalb erstmal Smarties mampfen, die Paula mitgebracht hat. Nun möchte erstmal ihre Betreuung der Arbeit eine ANOVA Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Jonas schaut Paula sehen erstmal gar nichts. Die beiden waren im Teuteburgerwald um einen Versuch in einer Klimakammer mit Erdbeeren durchzuführen. Dabei haben Jonas und Paula den Messwert Frischegewicht [kg/ha] unter der Behandlung Lichtstufen (*none*, *200lm*, *400lm* und *600lm*) ermittelt. Später wollen die beiden dann noch raus um zu Fechten.

Leider kennen sich Jonas 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)
Lichtstufen	3	5245.23			
Error	24	1229.49			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.01$ 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)**

Lichtstufen	Fallzahl (n)	Mittelwert	Standardabweichung
none	5	7.40	2.30
200lm	7	4.71	5.41
400lm	7	3.14	1.95
600lm	9	34.00	11.24

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

78. Aufgabe

(9 Punkte)

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



Die einfaktorielle ANOVA in R 'Uff... die einfaktorielle ANOVA und R. Nicht so einfach... Was sagt mir jetzt die Ausgabe der ANOVA und wo sehe ich, ob da was signifikant ist?', denkt Nilufar und hebt die Augenbraue. Nilufar hatte sich ein Freilandversuch mit Spargel angeschaut. Als wäre das nicht alles schon schwer genug. Wenn die Erwartung nicht wäre, ja dann wäre wohl vieles möglich für Nilufar! Aber so.. Dabei ging es beim Experiment herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lichtstufen (*none*, 200lm, 400lm und 600lm) und dem Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] gibt. Nun möchte ihr Betreuer ihrer Abschlussarbeit erstmal eine ANOVA sehen und die Ergebnisse präsentiert bekommen. Und eigentlich will sie ja was anderes... Hm, lecker Takis Blue Heat und dazu dann im Hintergrund Star Trek laufen lassen.

```
## Analysis of Variance Table
##
## Response: Chlorophyllgehalt
##           Df Sum Sq Mean Sq F value Pr(>F)
## Lichtstufen  3 236.85   78.950   3.8258 0.02204
## Residuals    25 515.91   20.636
```

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

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

79. Aufgabe

(12 Punkte)

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



Ergebnistabelle der zweifaktoriellen ANOVA Wie absolut ärgerlich. Jetzt stellt sich tatsächlich heraus, dass seiner Betreuerin keine Ahnung von der zweifaktoriellen ANOVA hat. Woher soll Steffen jetzt das Wissen nehmen? Steffen mampft aus Frust noch eine Handvoll Oreos. Immerhin muss er ja noch mit seiner Hausarbeit dieses Jahr fertig werden. In ein Kreuzungsexperiment hatte er Schweinen mit der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) sowie der Behandlung Flüssignahrung (*ctrl* und *flOw*) im Teutoburgerwald untersucht. Es wurde als Messwert Gewichtszuwachs in der 1LW bestimmt. Jetzt muss er erstmal die zweifaktorielle ANOVA verstehen. Und eigentlich wollte Steffen doch noch zum Sport! Um zu Ringen geht Steffen dann später nochmal raus. Echte Entspannung.

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

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Lüftungssystem	3	318.22			
Flüssignahrung	1	379.85			
Lüftungssystem:Flüssignahrung	3	16.75			
Error	18	245.39			

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

	$F_{\alpha=5\%}$
Lüftungssystem	4.26
Flüssignahrung	3.40
Lüftungssystem:Flüssignahrung	5.23

5. Was bedeutet ein signifikantes Ergebnis in einer zweifaktoriellen ANOVA? (2 Punkte)
6. Was sagt der Term *Lüftungssystem:Flüssignahrung* 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 'Mit der zweifaktoriellen ANOVA lässt sich die Interaktion zwischen den beiden Behandlungen nachweisen!', ihre Betreuerin scheint die zweifaktoriellen ANOVA zu verstehen. Warum jetzt sie jetzt nochmal alles wiederkauen muss, wird Nilufar echt nicht so klar. Wenn es doch so klar ist? Nilufar war im Teutoburgerwald und hatte dort ein Gewächshausexperiment mit Spargel durchgeführt. Die Komune wo sie untergekommen war, war cool gewesen. Nur jetzt muss eben das Experiment fertig ausgewertet werden. Nilufar hatte zwei Behandlungen auf Spargel angewendet. Einmal Genotypen (*AA*, *AB* und *BB*) sowie als zweite Behandlung Bewässerungstypen (*ctrl*, und *high*). Gemessen wurde der Messwert (*Y*) Proteingehalt [g/kg]. Jetzt muss das hier zu einem Ende kommen! Eigentlich wollte Nilufar nachher noch einen Film schauen. Irgendwie komisch, wenn sie Star Trek anmacht, dann ist das Huhn eigentlich sofort vor dem Bildschirm und starrt hinein.

```
## Analysis of Variance Table
##
## Response: Proteingehalt
##
##          Df Sum Sq Mean Sq F value    Pr(>F)
## Genotypen    2   0.56    0.28  0.0137    0.9864
## Bewässerungstypen 1 49.78   49.78  2.4569    0.1344
## Genotypen:Bewässerungstypen 2 651.92  325.96 16.0894 9.835e-05
## Residuals   18 364.67   20.26
```

Leider kennt 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. 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 Alex schaut konzentriert auf die Formeln der ANOVA und des t-Tests. In seinem Experiment wurde als Messwert Fettgehalt [%/kg] bestimmt. Eine echte Herausforderung für ihn war schon immer die Gefälligkeit gewesen. Ein leidiges Lied. Dann wäre es nicht noch komplizierter. In ein Kreuzungsexperiment wurden Milchvieh mit der Behandlung Flüssignahrung (*ctrl*, *superIn* und *fLOW*) sowie der Behandlung Elterlinie (*ctrl*, und *Xray*) untersucht. Beide Verfahren müssen etwas miteinander zu tun haben und seine Betreuerin möchte das jetzt auch noch verstehen. Im Hintergrund läuft leise Alien auf seinem Second Screen. Immerhin hat er die beiden Formeln vorliegen.

Gegebene Formeln

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

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

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

82. Aufgabe


(11 Punkte)

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



Interaktion in der zweifaktoriellen ANOVA Es ist schon kurz nach fünf und Mark wird langsam nervös. Mark wollte heute Abend noch seine E-Sport Qualifikation schauen und dann zum Sport. Stattdessen versucht seine Betreuerin die Ausgabe der zweifaktoriellen ANOVA zu visualisieren und zu überprüfen, ob es mit der Visualisierung der Daten als Boxplots zusammenpasst. Es liegt anscheinend eine signifikante Interaktion vor? Mark hatte im Teuteburgerwald ein Feldexperiment mit Erbsen durchgeführt. Es gab dabei zwei Behandlungen. Einmal Düngestufen (*ctrl*, *low*, *mid* und *high*) sowie als zweite Behandlung Bewässerungstypen (*ctrl*, und *high*). Gemessen wurde der Messwert (Y) Trockengewicht [kg/ha]. So kompliziert kann das jetzt doch nicht sein! Eigentlich wollte Mark nachher noch zum Sport. Einfach mal raus um zu Reiten. Ohne Ziel und Uhr. Das ist was für Mark.

Leider kennen sich Mark und seine Betreuerin 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 Geocaching 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 Es ist schon kurz nach fünf und Jessica wird langsam nervös. Jessica wollte heute Abend noch ihre E-Sport Qualifikation schauen. Hoffentlich kommt sie noch rechtzeitig zum Streamen. Angestrengend krault sie die Hündin. Stattdessen versucht ihre Betreuerin die Ausgabe der einfaktoriellen ANOVA zu visualisieren und zu überprüfen, ob es mit der Visualisierung der Daten als Boxplots zusammenpasst. Anscheinend gibt es ein Problem mit der Annahme der Normalverteilung und der Varianzhomogenität der ANOVA in den Daten. 'Wir haben jetzt bei der ANOVA einen p-Wert mit 0.058 raus sowie eine F-Statistik F_D mit 1.2 berechnet. Nach den Boxplots müsste sich eigentlich ein Unterschied zwischen *fLow* und *ctrl* ergeben. Der Unterschied ist in {emmeans} auch signifikant mit einem p-Wert von 0.036. Wie kann das sein?', grummelt ihre Betreuerin. Jessica hatte im Teuteburgerwald ein Stallexperiment mit Milchvieh durchgeführt. Dabei wurden die Daten D erhoben. Es gab dabei eine Behandlungen Flüssignahrung (*ctrl*, *superIn* und *fLow*). Gemessen wurde der Messwert (Y) Gewichtszuwachs in der 1LW. So kompliziert kann das jetzt doch nicht sein! Jessica hat schon genug Probleme. Wenn der Mangel 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 Jessica und ihre Betreuerin mit der Interpretation einer ANOVA überhaupt nicht aus. Deshalb braucht sie bei der Erstellung Ihre Hilfe und die Zeit wird knapp.

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

Teil V.

Multiple Gruppenvergleiche

84. Aufgabe

(12 Punkte)

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



In ein Gewächshausexperiment mit Kartoffeln wurde die Behandlung Substrattypen (*kompost*, *torf*, *40p60n*, *30p20n* und *70p30n*) gegen die Ergebnisse einer früheren Studie von Meyer et al. (2021) verglichen. Im Rahmen des Experiments haben Nilufar und Yuki 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 Meyer et al. (2021). Jetzt sollen die beiden einmal schauen, was in den Daten so drin ist.

Rohen p -Werte	Adjustierte p -Werte	Nullhypothese ablehnen?
0.230		
0.020		
0.080		
0.002		
0.340		

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

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

85. Aufgabe

(10 Punkte)

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



Jessica betrachtet in sich gekehrt die Poster vor dem Büro von ihrem Betreuer. Viele der explorativen Abbildungen sagen ihr etwas. Die Barplots und die Boxplots könnte sie dann schon nachbauen. Das macht sie dann zuversichtlich die Abschlussarbeit auch hinzukriegen. Etwas komischer sind die seltsamen Buchstaben über den Barplots. Jessica betrachtet ein Poster das sich mit Lamas beschäftigt. Bestandsdichte (*effizient*, *standard*, *eng*, *weit* und *kontakt*) und Fettgehalt [%/kg] wurden dort bestimmt. So richtig schlau, wird sie daraus nicht.

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

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich 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!



Nilufar betrachtet in sich gekehrt die Poster vor dem Büro von ihre Betreuerin. Viele der explorativen Abbildungen sagen ihr etwas. Die Barplots und die Boxplots könnte sie dann schon nachbauen. Das macht sie dann zuversichtlich die Abschlussarbeit auch hinzukriegen. Etwas komischer sind die seltsamen Buchstaben über den Barplots. Nilufar betrachtet ein Poster das sich mit Kartoffeln beschäftigt. Substrattypen (*kompost*, *torf*, *40p60n* und *70p30n*) und Trockengewicht [kg/ha] wurden dort bestimmt. So richtig schlau, wird sie daraus nicht. Als erstes müsse müsse man die Gruppen nach absteigender Effektstärke sortieren, liest Nilufar im Methodenteil und ist dann noch verwirrter als vorher schon.

Substrattypen	Fallzahl (n)	Mittelwert	Standardabweichung
kompost	8	10.12	3.19
torf	8	14.62	2.00
40p60n	8	17.11	3.04
70p30n	9	15.92	2.29

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

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

87. Aufgabe

(10 Punkte)

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



'Oh, nee!', ruft Mark aus und rollt entnervt mit seinen Augen. Mark hatte seine gesamte Analyse in SPSS gerechnet. Das war ja auch alles in Ordnung. Abbildungen haben geklappt und auch die statistischen Tests gingen dann irgendwie doch. Aber das CLD nicht. Mark findet einfach keine Möglichkeit ein CLD in SPSS zu erhalten. Aber seine Betreuerin 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 Elterlinie (*ctrl*, *Standard*, *SLOW*, und *Xray*) und Protein/Fettrate [%/kg] in ein Kreuzungsexperiment mit Fleischrindern angeschaut. Wo kriegt er jetzt ein CLD her? Dann eben per Hand aus der Matrix der *p*-Wert. Mark stöhnt...

	ctrl	Standard	SLOW	Xray
ctrl	1.0000000	0.6238086	0.1135440	0.0225907
Standard	0.6238086	1.0000000	0.0512722	0.0099146
SLOW	0.1135440	0.0512722	1.0000000	0.4208090
Xray	0.0225907	0.0099146	0.4208090	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 Yuki! **(2 Punkte)**

Teil VI.


Der Chi-Quadrat-Test & Der diagnostische Test

88. Aufgabe

(12 Punkte)

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



Den Chi-Quadrat-Test berechnen Am Ende hätte Mark 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  so schön heißt, als Faktoren. Aber immerhin, hofft er das was bei den Daten rausgekommen ist. Dann noch schnell Andrea Berg auf das Ohr und los gehts. Gezählt hat Mark einiges mit $n = 110$ Beobachtungen von Erdbeeren. Zum einen hat er als Behandlung *KI-gesteuert* [ja/nein] bestimmt und zum anderen die Variable *Chlorophyllgehalt unter Zielwert* [ja/nein] ermittelt. Nun möchte sein Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Eigentlich wollte Mark nachher noch einen Film schauen. Irgendwie komisch, wenn er Columbo anmacht, dann ist der Hamster eigentlich sofort vor dem Bildschirm und starrt hinein.

	38	11	
	23	38	

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. 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 war es für Mark in seiner Hausarbeit dann doch kein normalverteiltes Outcome. Das was jetzt etwas doof, da er sich auf eine ANOVA gefreut hatte. Dann noch schnell Marzipankugeln zur Stärkung und los gehts. Prinzipiell ginge das auch irgendwie, aber nun möchte sein Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Mark hatte sich in ein Kreuzungsexperiment $n = 121$ Beobachtungen von Zandern angeschaut. Dabei hat er als Behandlung *Ökologisch* [ja/nein] bestimmt und zum anderen die Variable *Fettgehalt erreicht* [ja/nein] ermittelt. Jetzt muss Mark mal schauen, wie er das jetzt rechnet. Am Ende des Tages möchte er dann noch sein Hobby Geocaching genießen. Das muss auch mal sein!

			59
			62
	79	42	121

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

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

90. Aufgabe

(10 Punkte)

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



Der Chi-Quadrat-Test in R Am Ende war es für Nilufar in ihrer Hausarbeit dann doch kein normalverteiltes Outcome. Das was jetzt etwas doof, da er sich auf eine ANOVA gefreut hatte. Dann noch schnell Star Trek starten und los gehts mit der Kraft von Takis Blue Heat. Prinzipiell ginge das auch irgendwie, aber nun möchte ihr Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Nilufar hatte sich in einen Leistungssteigerungsversuch $n = 121$ Beobachtungen von Schweinen angeschaut. Dabei hat sie als Behandlung *Außenklimakontakt* [ja/nein] bestimmt und zum anderen die Variable *Gewichtszuwachs erreicht* [ja/nein] ermittelt. Jetzt muss Nilufar mal schauen, wie sie das jetzt rechnet. Nach ihrem Experiment erhielt sie folgende 2x2 Kreuztabelle aus ihren erhobenen Daten.

```
##                               Außenklimakontakt
## Gewichtszuwachs erreicht ja nein
##                               ja    6    8
##                               nein  6    4
```

Dann rechnete Nilufar 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 = 0.6802
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.06993249 3.40089546
## sample estimates:
## odds ratio
##  0.5149174
```

Leider kennt sich Nilufar 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. 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)**

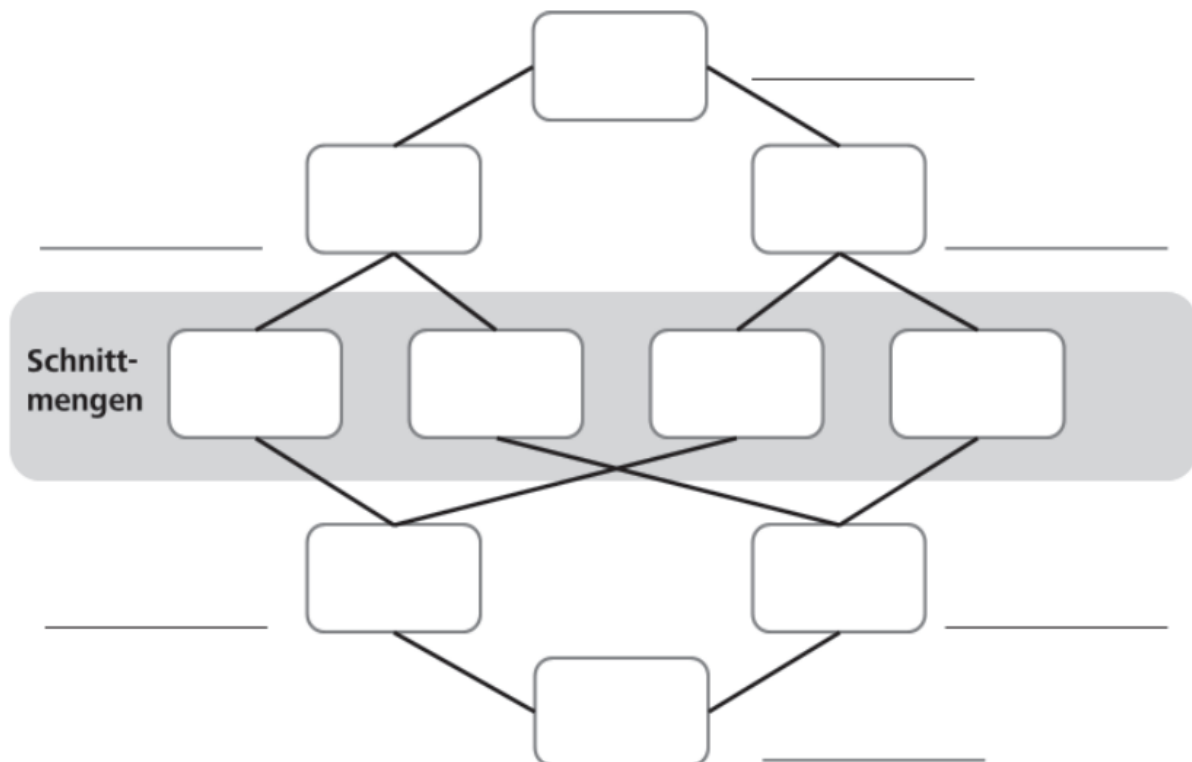
91. Aufgabe

(11 Punkte)

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



Den diagnostische Test am Doppelbaum berechnen Tina liest laut vor. 'Die Prävalenz von Klauenseuche bei Erdbeeren wird mit 4% angenommen. In 85% der Fälle ist ein Test positiv, wenn das Pflanze erkrankt ist. In 8.5% der Fälle ist ein Test positiv, wenn das Pflanze nicht erkrankt ist und somit gesund ist. Wir führen einen Test auf Gelbmosaikvirus an 4000 Erdbeeren mit einem diagnostischen Test durch.' Paula klappt die Kinnlade runter. In der Stille dudelt Tocotronic. Tina schaut kompetent und schmeißt sich mit offenen Mund Smarties an den Kopf vorbei.



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

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

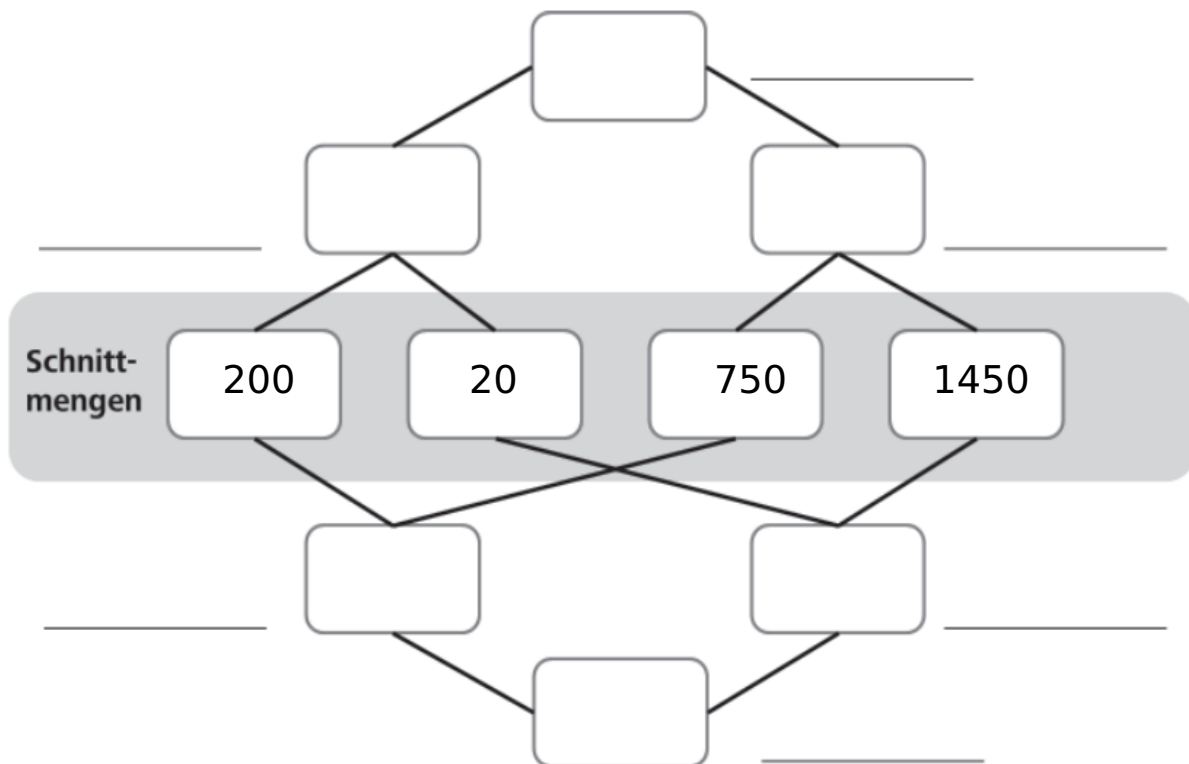
92. Aufgabe

(11 Punkte)

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



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



Leider kennen sich Jessica und Alex 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!



‘Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen!’, ruft Yuki. ‘Ich sehe nur eine Zahlen und keine Punkte. Wie soll ich da denn jetzt eine Gerade durchzeichnen?’, fragt Mark. Yuki atmet schwer ein. Die beiden hatten ein Kreuzungsexperiment im Oldenburger Land mit Zandern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: mittlerer Anzahl an weißen Blutkörperchen [LEU/ml] 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]	Mittlerer Anzahl an weißen Blutkörperchen [LEU/ml]
15.5	23.7
12.0	20.3
12.9	22.8
16.3	24.7
20.2	23.4
17.6	23.1
17.3	23.2
26.3	32.7
19.8	26.0
19.7	23.2
12.1	19.0
21.6	31.2

Leider kennen sich Yuki und Mark 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!



'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 Steffen an. 'Ich sehe keine Punkte... ich sehe nur zwei Zeilen einer Tabelle und ich glaube du hast gerade was gelöscht.', antwortet Mark sichtlich übernächtigt. 'Wir müssen die Koeffizienten der linearen Regression ja auch erst interpretieren!', spricht Steffen sehr deutlich und langsam. Die beiden hatten ein Gewächshausexperiment im Emsland mit Erdbeeren durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche UV-Einstrahlung [UV/d] und Frischegewicht [kg/ha]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der  Ausgabe möglich sein.

term	estimate	std.error	t statistic	p-value
(Intercept)	4.70	1.92		
Durchschnittliche UV-Einstrahlung	0.72	0.20		

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



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

95. Aufgabe


(11 Punkte)

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



'Hä? Was ist denn das? Das wird ja immer wilder! Hatten wir das als Aufgabe eine lineare Regression zu rechnen? Wir bauen aus kontinuierlichen Daten eine Abbildung und interpretieren diese dann?', fragt Steffen. Nilufar schaut fragend zurück. 'Keine Ahnung... das ist jetzt jedenfalls keine Abbildung von irgendwas sondern eine  Ausgabe mit ganz wilden Bezeichnungen...', antwortet Nilufar leicht angespannt. Die beiden hatten ein Stallexperiment im Teuteburgerwald mit Schweinen durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche Tagestemperatur [C/d] und Schlachtgewicht [kg]. Jetzt haben die beiden eigentlich alles zusammen. *Eigentlich...*, denn mit der  Ausgabe haben beide jetzt ein Problem.

```
##
## Call:
## Schlachtgewicht ~ Durchschnittliche_Tagestemperatur
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.3251 -0.6083  0.0847  0.7635  2.6016
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   4.0792     1.9310   2.112   0.0413
## Durchschnittliche_Tagestemperatur 0.2891     0.1941   1.489   0.1447
##
## Residual standard error: 1.384 on 38 degrees of freedom
## Multiple R-squared:  0.05514, Adjusted R-squared:  0.03028
## F-statistic: 2.218 on 1 and 38 DF,  p-value: 0.1447
```

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. 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!



'Ich glaube ich bringe da was durcheinander. Ich möchte eine Gerade durch die Punkte zeichnen oder doch eine Korrelation berechnen?', merkt Tina laut an. 'Ich sehe keine Punkte... das ist doch eine Ausgabe in R. Überhaupt, darum geht es doch gar nicht in meinem Versuch. Ich wollte doch keine Gerade zeichnen?.', antwortet Tina sich sichtlich übernächtigt selber. Irgendwie komisch, wenn sie Indiana Jones anmacht, dann ist die Spinne eigentlich sofort vor dem Bildschirm und starrt hinein. Die Nacht war zu lang und überhaupt. Wenn die Wut nicht wäre, ja dann wäre wohl vieles möglich für Tina! Aber so.. Tina hatte ein Kreuzungsexperiment im Teutoburgerwald mit Lamas durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche Tagestemperatur [C/d] und Schlachtgewicht [kg]. Jetzt will sie erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der R Ausgabe möglich sein.

```
##  
## Pearson's correlation  
##  
## data: Durchschnittliche Tagestemperatur and Schlachtgewicht  
## t = 5.8879, df = 8, p-value = 0.0003667  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
##  0.6284096 0.9767005  
## sample estimates:  
##      cor  
## 0.9013893
```

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

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

97. Aufgabe

(12 Punkte)

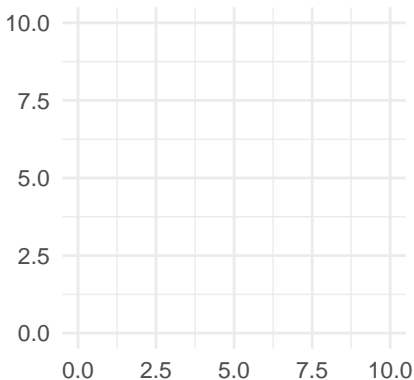
Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



'Hm..., drei leere Abbildungen. Was soll ich da hetzt machen?', fragt sich Yuki. Yuki kennt sich nur begrenzt bis gar nicht mit der linearen Regression und Korrelation aus. Dafür mit etwas anderem. Yuki liebt Orchideen. Darin kann sie sich wirklich verlieren und immer wieder neu begeistern. Aber das hilft hier auch nur so halb, die Aufgabe zu lösen und mehr über den Korrelationskoeffizienten zu erfahren.

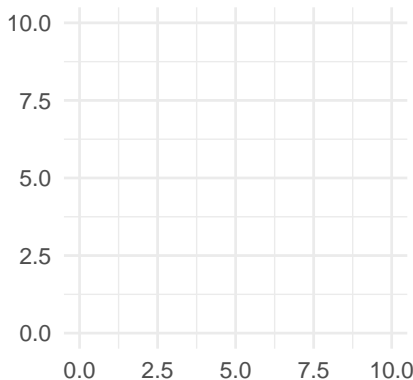
Pearsons $\rho = 1$

$R^2 = 0.75$



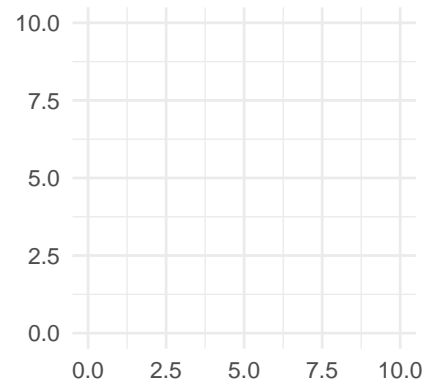
Pearsons $\rho = 0.25$

$R^2 = 0.25$



Pearsons $\rho = 0.75$

$R^2 = 0.5$



Leider kennt sich Yuki 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)**

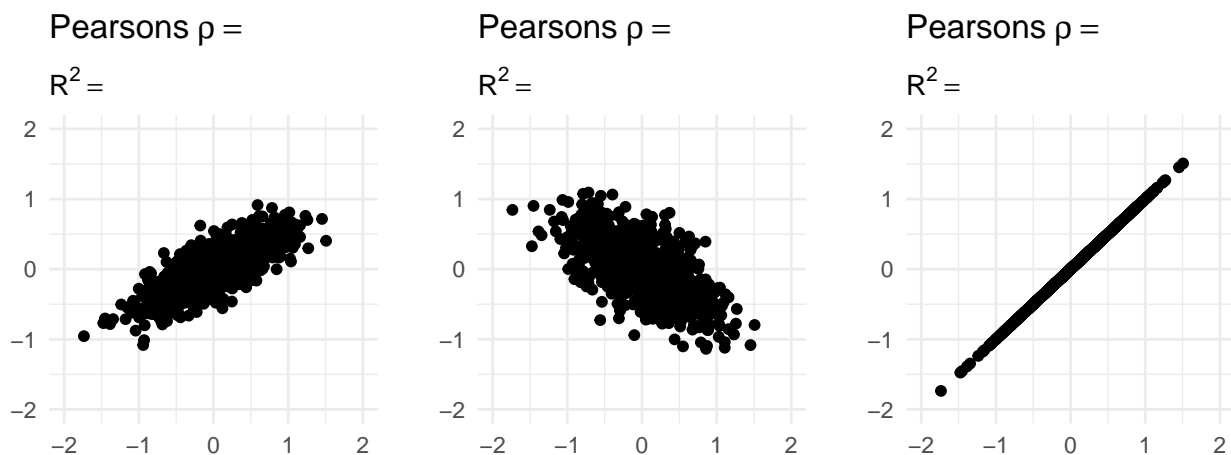
98. Aufgabe

(12 Punkte)

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



Der Bildschirm strahlt blau in das Gesicht von Jonas. Es ist schon spät. Und das hat einen Grund. Schon dutzende Male gesehen: Mission Impossible. Aber immer noch großartig zusammen mit Snickers. . Jonas überlegt, aber seine Gedanken sind etwas zäh. 'Was soll das hier alles bedeuten?', fragt sich Jonas. Irgendwie ist ihm nicht klar wie er ρ -Werte oder R^2 -Werte abschätzen soll. Alles nicht so einfach. Wenn die Erschöpfung nicht wäre, ja dann wäre wohl vieles möglich für Jonas! Aber so..



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

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

99. Aufgabe

(11 Punkte)

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



'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 Steffen stolz. Leider hat Steffen vergessen wie der `R` Code für den Residuenplot geht. Steffen hatte anderes im Kopf. Schon dutzende Male gesehen: Harry Potter. Aber immer noch großartig zusammen mit Oreos. Aber sowas hilft ihm natürlich hier nicht. Da schmeißt sich Steffen noch ein paar Oreos in den Mund und kaut los.

Durchschnittlicher Anteil	Proteingehalt	\hat{y}	ϵ
28.4	10.8	26.4	
23.6	9.2	24.5	
27.9	12.0	27.9	
21.5	10.6	26.3	
32.4	15.5	32.3	
23.4	8.2	23.2	
26.7	12.3	28.3	
31.8	13.3	29.6	
22.3	6.3	20.7	
27.2	10.7	26.4	
22.6	8.1	23.1	
28.8	11.8	27.7	

Leider kennt sich Steffen 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!



Yuki hat einen Versuch in einer Klimakammer mit Brokkoli durchgeführt. Soweit so gut. Dann war sie bei ihrer Betreuerin. Leider war der Schritt nicht so hilfreich. Yuki und die Faulheit, eine unendliche Geschichte mit kniffligen Wendungen. Aber es muss ja weitergehen. Yuki hatte dann in ihrer Abschlussarbeit einfach zu viele Endpunkte gemessen und ist jetzt vollkommen durcheinander, welche Analyse sie nun wie rechnen soll. Naja, dann heißt es jetzt eben London Grammar aufdrehen und darüber nachdenken, was hier eigentlich gemacht wurde. Yuki fängt einfach an und nimmt den ersten Endpunkt Chlorophyllgehalt unter Zielwert [ja/nein]. Dann kann sie sich voran arbeiten. Später dann noch raus um zu Boldern um mal zu entspannen und vielleicht ist Nilufar auch da. Wäre toll.

Leider kennt sich Yuki mit dem Kontext der linearen Regression überhaupt nicht aus. Deshalb braucht sie 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)



Die Schlange macht mal wieder Randalen in Steffens Zimmer und rennt davon! Jessica und Mark sind bei Steffen in im Emsland wo der neue, bessere Versuch stattfinden soll. Dabei soll in einem Stallexperiment im Emsland mit Schweinen durchgeführt werden. Auf dem Tisch stapeln sich Oreos aus Vollkorndinkelmehl. Eine Spezialität der Komune hier. Jessica hasst Vollkorn wie Marzipankugeln geliebt werden. In dem neuen Versuch geht es um den Zusammenhang zwischen der Behandlung Bestandsdichte (*standard*, *eng*, *weit* und *kontakt*) und dem Messwert Fettgehalt [%/kg]. Immerhin ist der Messwert normalverteilt, was einges einfacher macht. Was es nicht so einfacher macht ist, dass Mark als zusätzliche Herausforderung noch die Unsicherheit mitgebracht hat. Daher entscheiden sich alle drei für ein *Latin square design*. 'Naja, so viel einfacher ist es dann doch nicht...', merkt Mark an und sucht die Schlange.

Leider kennen sich Steffen, Jessica und Mark mit dem *Latin square design* überhaupt nicht aus. Deshalb brauchen die Drei bei der Erstellung Ihre Hilfe!

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




102. Aufgabe

(20 Punkte)



Die Schlange macht mal wieder Randalen in Nilufars Zimmer und rennt davon! Mark und Steffen sind bei Nilufar in im Emsland wo der neue, bessere Versuch stattfinden soll. Dabei soll in einem Stallexperiment im Emsland mit Fleischrindern durchgeführt werden. Auf dem Tisch stapeln sich Oreos aus Vollkorndinkelmehl. Eine Spezialität der Komune hier. Mark hasst Vollkorn wie Marzipankugeln geliebt werden. In dem neuen Versuch geht es um den Zusammenhang zwischen der Behandlung Genotypen (*AA*, *AB* und *BB*) sowie Bestandsdichte (*standard* und *kontakt*) sowie zwei Blöcken und dem Messwert Protein/Fettrate [%/kg]. Immerhin ist der Messwert normalverteilt, was einiges einfacher macht. Was es nicht so einfacher macht ist, dass Steffen als zusätzliche Herausforderung noch die Unsicherheit mitgebracht hat. Daher entscheiden sich alle drei für ein *Strip plot design* oder auch *Streifenanlage*. 'Naja, so viel einfacher ist es dann doch nicht...', merkt Steffen an und sucht die Schlange.

Leider kennen sich Nilufar, Mark und Steffen mit dem *Strip plot design* oder auch *Streifenanlage* 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)



'Hm.  ist eigentlich gar nicht so schwer, wenn man die Grundlagen kann.', meint Mark ganz zuversichtlich. Nur leider kennt er sich überhaupt nicht mit  aus! Das heißt, Sie müssen hier einmal Rede und Antwort stehen und helfen.

Mark: *Jetzt lese ich hier von einem Faktor. Was ist ein Faktor in  ?* (1 Punkt)

Sie antworten:

Mark: *Was ist der Unterschied zwischen `library()` und `Packages` und warum brauche ich sowas in  ?* (1 Punkt)

Sie antworten:

Mark: *Wie war nochmal der Name der Funktion in dem wir in R Daten intern abspeichern? Was waren da nochmal die Vorteile?* (1 Punkt)

Sie antworten:

Mark: *Warum brauche ich eigentlich das RStudio und R? Wo ist denn da der Unterschied?* (1 Punkt)

Sie antworten:

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



Sie antworten:

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

Sie antworten:

Mark: *Was war eigentlich nochmal ein Vorteil von der Nutzung von  ?* (1 Punkt)

Sie antworten:

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

Sie antworten:

Mark: *Ich habe doch die Spalte mutiert und geändert. Warum sehe ich das in R aber mein Datensatz ändert sich nicht?* (1 Punkt)

Sie antworten:

104. Aufgabe

(9 Punkte)



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

Tina fragt: *Hm... wenn ich in R Daten mit Gruppen eingelesen habe, was muss ich dann nochmal als erstes machen? Und warum?* **(2 Punkte)**

Sie antworten:

Tina fragt: *Welche Funktionen brauche ich in welcher Reihenfolge um eine ANOVA zu rechnen?* **(1 Punkt)**

Sie antworten:

Tina fragt: *Nach der EDA zu urteilen, habe ich Varianzheterogenität zwischen den Gruppen vorliegen. Wie kann ich da die Funktion `emmeans()` umstellen.* **(1 Punkt)**

Sie antworten:

Tina 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:

Tina fragt: *Es gibt ja nur ein richtiges Format für die Eingabe eines Datums. Wie lautet das Format?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

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

Sie antworten:

Teil X.

Forschendes Lernen

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

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

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

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

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

- bar

105. Aufgabe



(20 Punkte)

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



Mark hält die wissenschaftliche Veröffentlichung *Selle, P. H., et al. (2010). Implications of sorghum in broiler chicken nutrition* unter einem Schnaufen in die Luft. 'Worum geht es denn eigentlich in dieser Arbeit?', fragt er stirnrunzelnd und wirft die Arme in die Luft, da hilft dann auch nicht mehr die beruhigende Wirkung von Andrea Berg. Mark soll die Veröffentlichung nutzen um das eigene Experiment zu planen. Als eine Vorlage sozusagen. Daher möchte seine Betreuerin, dass er einmal die Veröffentlichung sinnvoll zusammenfasst. Das sollte dann doch etwas aufwendiger werden. Das wird dann vermutlich heute Abend nichts mehr mit seinem Hobby Geocaching. Der Hamster schaut mitleidig.

Leider kennt sich Mark 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)² **(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)**

²Schimel, 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!



Vor dem Start der eigenen Arbeit möchte sein Betreuer, dass Jonas einmal die wissenschaftlichen Daten *data4* sinnvoll zusammenfasst. Dann würde die eigene Arbeit auch leichter von der Hand gehen und Jonas 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 Snickers. Das wird dann vermutlich heute Abend nichts mehr mit Mission Impossible

Leider kennt sich Jonas 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 Während der Besetzung Mecklenburgs durch die Franzosen kamen Napoleon die Geschichten des berühmten Apfelschimmels Herodot aus Ivenack zu Gehör. Herodot lief zwar niemals Rennen, war aber eines der berühmtesten Pferde dieser Zeit. Napoleon selbst gab den Auftrag, diesen Schimmel durch die Armee nach Frankreich zu bringen. Der Legende nach sollen Arbeiter den Schimmel im hohlen Stamm einer 1000-jährigen Eiche aus Ivenack vor den Franzosen versteckt haben. Doch Herodot verriet sein Versteck durch lautes Wiehern, woraufhin die französische Armee den Schimmel beschlagnahmte und nach Frankreich führte³.

Forschungsfrage: "Konnten die Ivenacker den Apfelschimmel Herodot vor dem Zugriff von Napoleon in der 1000-jährigen Eiche verstecken?"

Gehen Sie von einem radialen Wachstum der 1000-jährigen Eiche von 1mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 12m in 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 195cm , eine Breite von 85cm sowie eine Länge von 230cm .

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)**

³Die 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 In einem Experiment wollen Sie die Wuchshöhe von 120 Stockrosen bestimmen. Bevor Sie 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 müssen Sie die Stockrosen auch bewegen und in ein Gewächshaus platzieren. Die Töpfe für die Keimung haben einen Durchmesser von 8cm und eine Höhe von 10cm. Der Kubikmeterpreis für Torf liegt bei 290 EUR.

1. Skizzieren Sie den Versuchsplan auf *zwei* Tischen im Gewächshaus! **(2 Punkte)**
2. Berechnen Sie die benötigte Anzahl an Pflanztöpfen, wenn Sie Randpflanzen mit berücksichtigen wollen! **(1 Punkt)**
3. 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 Um die Energiekosten Ihres Betriebes zu senken, wollen Sie eine Solaranlage auf den Rinderstall montieren lassen. Sie messen Ihren Stall und finden folgende Maße wieder. Die vordere Seite des Rinderstall hat eine Höhe h_v von $6.5m$. Die hintere Seite des Rinderstall hat eine Höhe h_b von $8.5m$. Der Rinderstall hat eine Tiefe t von $14m$ und eine Breite b von $70m$.

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

Ebenfalls planen Sie eine neue Biogasanlage für Ihren Betrieb. Der neue Methantank hat einen Radius r von $1m$. 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. Sie rechnen eine Sicherheitstoleranz von 15% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $220kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $290kg/m^3$. Sie betreiben Ihre Anlage bei $-90^\circ C$.

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 puderrote Gesicht von einem Mädchen mit Zöpfen und Zahnsperre. “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 Rewe über sich ergehen lassen. Da kommen Sie nicht mehr raus, also können Sie auch gleich mitmachen. Das Problem liegt in Chile⁴. Tja, die Deutschen und Südamerika.

Zuerst werden Ihre Fähigkeiten getestet, der Mathematik folgen zu können. Oder berechnen Sie gerade den Einkauf von einem Mädchen mit Zöpfen und Zahnsperre?

1. Wenn 5 Blaubeerschalen 7.95 Euro kosten, wie viel kosten 6 Schalen? **(2 Punkte)**
2. Wenn Sie die 6 Blaubeerschalen gekauft haben, wie viele Aligatorenbirnen zu je 2.59 EUR können Sie sich dann noch für 100 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 Rewe über Gemüse erfährt.

- Ein Kilo Strauchtomaten benötigt 170l Wasser. Eine Strauchtomate wiegt 90 - 120g.
- Ein Kilo Salat benötigt 140l Wasser. Ein Salatkopf wiegt 320 - 490g.
- Ein Kilo Avocado benötigt 950l Wasser. Eine Avocado wiegt 140 - 410g.
- Ein Kilo Blaubeeren benötigt 820l Wasser. Eine Blaubeere wiegt 3.1 - 3.6g.

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

Chile exportiert im großem Ausmaß Blaubeeren und Avocados. In dem Exportjahr 2024 blieben die Erträge von Blaubeeren mit 7.9×10^4 t in dem prognostizierten Rahmen. Die Menge reduzierte sich um 6.8%. Die Exporte für Avocados fielen in dem gleichen Zeitraum um 21.2% auf 2×10^5 t.

4. Wie viele Kubikmeter Wasser hat Chile in dem Exportjahr 2023 exportiert? **(2 Punkte)**

Chile ist eines der wenigen Länder der Welt, die ihr Wasser komplett privatisiert haben. Derzeit sind nur drei Prozent des Wassers des Landes für den häuslichen Verbrauch vorgesehen. In den Dörfern der Anbauregionen versorgen Tankwagen die Bevölkerung jede Woche mit Wasser, es gibt etwa 48 Liter Wasser pro Kopf für den täglichen Bedarf. In *Deutschland* liegt der Verbrauch bei 8 - 17 Liter pro Spülmaschinenlauf und 10 - 15 Liter pro Minute Duschen.

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

Das alles hätten Sie nicht von einem Mädchen mit Zöpfen und Zahnsperre 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)**

⁴Die Quelle der Inspiration für die Aufgabe waren folgende Reportagen: “Bis zum letzten Tropfen” in AMNESTY – Magazin der Menschenrechte vom August 2021 und “Wasserknappheit in Chile: Eine Folge der Privatisierung?” in Die Welternährung dem Fachjournal der Welthungerhilfe vom April 2022.

111. Aufgabe

(12 Punkte)

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



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

Die Dampfnudelerde “Was für einen Unsinn!“, rufen Sie. Jetzt haben Sie 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 65 Millionen Jahren die Dinosaurier - so groß sie auch waren - nicht von der Schwerkraft zu Boden gerissen wurden. Hat der Dampfplauderer etwa recht und war die Schwerkraft vor Millionen von Jahren eine andere? Sind deshalb alle Lebewesen auf der Erde *heutzutage* so viel kleiner, weil die Schwerkraft größer ist als damals? War die Erde kleiner und hatte weniger Masse? Oder ist es nur ein Rechenfehler wie bei der Theorie der Hohlerde von Edmond Halley aus dem 17.-18. Jahrhundert? Müde reiben Sie sich die Augen. So wird es nichts mehr mit dem Schlafen, dann können Sie auch mal etwas rechnen⁵.

Betrachten wir die Schwerkraft oder Gewichtskraft, die auf Lebewesen damals und heute gewirkt haben soll. Nehmen Sie für die Fallbeschleunigung g der Erde *heutzutage* einen Wert von 9.81 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.1956 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.44 g/cm^3 . Das Gewicht von einem heute lebenden asiatischen Elefanten liegt bei 3t bis 5t und das Gewicht von einem Brachiosaurus bei bis zu 30t.

1. Welchen Durchmesser müsste die Erde vor 65 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 65 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 65 Millionen Jahren! **(2 Punkte)**
2. Beantworten Sie die Eingangsfrage mit 1-2 Antwortsätzen! **(1 Punkt)**

Die Distanz zwischen Sonne und Erde entspricht 1.05 astronomische Einheiten (AE). Die Einheit 1 AE wird mit $1.55 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 85% Wasserstoffkernen mit einer molaren Masse von 1.08 g/mol , 9% Heliumkernen mit 4.32 g/mol sowie 6% weiteren Atomkernen mit 152.01 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)**

⁵Die 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 Enten in Mastställen sagen. Das ist natürlich etwas ungünstig, den dann kommt es zu Picken und Kannibalismus. Denn wenn der Nachbar nervt, dann muss zu Maßnahmen gegriffen werden. Kennt jeder aus einer mittelmäßigen Wohngemeinschaft. Das wollen wir aber als vorsorgliche Enten-Halter:innen nicht⁶. Betrachten wir also einmal das Platzangebot (eng. *space allowance*, abk. *SA*) der Enten 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 Ente plus dem benötigten Radius für das Verhalten *i*.
- *R_i* dem notwendigen Abstand zu den Nachbarn für das Verhalten *i*.
- *i* dem Verhalten: (1) dustbathing, (2) sitting, (3) walking und (4) foraging incl. scratching.

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)	Jabcocks et al. (2019)
dustbathing	28cm; 19cm; 50.1%	42cm; 32cm; 50.1%	31cm; 22cm; 50.1%
sitting	29cm; 21cm; 4.1%	34cm; 35cm; 5.2%	44cm; 20cm; 2.1%
walking	32cm; 27cm; 0.6%	33cm; 37cm; 0.6%	37cm; 25cm; 0.4%
foraging incl. scratching	41cm; 30cm; 8.1%	38cm; 21cm; 1.2%	38cm; 27cm; 7.6%

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 Enten für ein Verhalten *i*. Nutzen Sie hierfür vereinfachte geometrische Formen! **(2 Punkte)**
5. Sie entnehmen der Literatur folgende Aussage zur Verteilung der Enten 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 α , die ein Tier einnimmt. Welche Annahmen haben Sie für die Berechnung der Körperfläche getroffen? **(2 Punkte)**

⁶Die 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 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 40 Tagen die ersten Symptome ein; die ersten Toten sind nach 70 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 $4000\mu\text{g}/50\text{mg}$ Vitamin C. Der Bedarf liegt bei 110mg pro Tag für Männer.

3. Berechnen Sie die notwendige Menge in kg 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 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 3000m kollabiert, wird die Sonne 45% der aktuellen Masse verloren haben. Ein Lichtteilchen mit der Masse m_f und der Fluchtgeschwindigkeit v_f will dem schwarzen Loch entkommen. Sie haben folgende Formeln für die kinetische Energie des Lichtteilchens E_{kin} und der Gravitationsenergie des schwarzen Lochs E_{grav} gegeben⁷.

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

mit

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

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

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

⁷Die 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!



Das Fermi Paradoxon Der 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? Wie lange würde eine außerirdische Zivilisation benötigen um die gesamte Milchstraße zu besuchen, wenn das maximale Reisetempo die Geschwindigkeit der Voyager 1 Sonde wäre?⁸

Wir treffen folgende Annahmen. Eine außerirdische Zivilisation schickt zwei Voyager 1 ähnliche Sonden mit der Geschwindigkeit von $6.3587 \times 10^4 \text{ km/h}$ los um sich auf den erreichten Planeten selbst zu replizieren. Nach 750 Jahren ist die Replikation abgeschlossen und wiederum zwei Sonden werden ausgesendet. Gehen Sie von 7.81 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 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 annäherungsweise benötigt um die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
4. Bei einem vermuteten Alter der Erde von 4.5×10^9 Jahren, wie oft war dann eine Sonde einer außerirdischen Zivilisation schon zu Besuch? Korrigieren Sie Ihre Antwort mit dem Wissen, dass sich die Kontinentalplatten einmal alle 1.2×10^8 Jahre vollständig im Erdinneren umgewandelt haben! **(2 Punkte)**

⁸Die 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 oldenburgischen Pyramidentage an und Sie 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 74 Grad im Vergleich zu den ägyptischen Pyramiden mit 55 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 bei dem Befüllen der Pyramide mit Schutt zu Seite stehen. Leider haben Ihre Sklaven zu allem Überfluss auch noch chronische Knieschmerzen entwickelt, als sie von der anstehenden Aufgabe erfahren haben. Gehen Sie daher von einer Effizienz der Sklaven von 80% aus. In eine Schubkarre passen 90 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 11° ! **(2 Punkte)**
5. Wie weit reicht Ihre Rampe vom Fuß der Pyramide in die oldenburgische Landschaft? **(2 Punkte)**

Bei der Besichtigung der Pyramide teilt Ihnen der leicht übergewichtige Pharao (Nebenberuf *Chefredakteur*) mit, das die Pyramide zu flach sei und somit nicht in die oldenburgische 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 7° 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*. Sie schwingen sich auf Ihr Cachermobil um mit 16km/h, geleitet von modernster Satellitentechnologie und einem Supercomputer aus dem Jahr 2000 in Ihren Händen, Plastikdosen in der Natur und an sehenswerten Orten zu finden. Sie wollen diesmal endlich die aufwärts Schwierigkeitschallenge durchführen. Die Reihenfolge der Caches nach Schwierigkeitswertung gibt daher die von Ihnen 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 Ihnen für Ihre Planung der Route zu Verfügung⁹.

Ort	Cache	Wertung (S T G)
A	GCRQ6EU	3.5 1.5 Mikro
B	GC26BPM	2.0 4.0 Klein
C	GCRZCIE	4.5 4.5 Mikro
D	GCQACHL	5.0 3.5 Normal
E	GCQN0TQ	4.0 2.0 Klein

Im Weiteren sind Ihnen folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \vec{AC} ist 4km. Im Weiteren ist Ihnen der Entfernungsvektor \vec{CB} mit 7.5km bekannt. Der Entfernungsvektor \vec{BE} ist das 1.2-fache des Entfernungsvektor \vec{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 40° nördlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 60° östlich. Vom Ort B betrachtet, bilden die Orte C und D einen rechten Winkel am Ort B. Der Ort B liegt auf gerader Linie zwischen den Orten C und E. Somit liegt der Ort E nördlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort D Ihre Cachertour.

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 Schwierigkeitschallenge 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

$$\text{Suchzeit} = 0.05 + 0.15 \cdot \text{Schwierigkeit}$$

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

4. An der höchsten Schwierigkeit müssen Sie angeln. Ihre Angel ist ausgefahren 5m lang. Erreichen Sie einen Cache in der Höhe von 6.6m? Berechnen Sie dazu Ihre maximale mögliche Angelhöhe! Welche Annahmen mussten Sie treffen um die Aufgabe zu lösen? **(2 Punkte)**

⁹Die Quelle der Inspiration für die Aufgabe war folgende Tätigkeit: [Geocaching – Mach mit bei der weltweit größten Schatzsuche](#).

118. Aufgabe

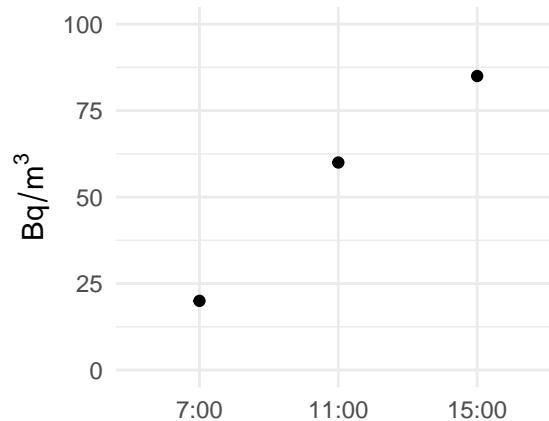
(10 Punkte)

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 Kellerkind vom Dorf wollen das Ausmaß der Radonbelastung in ihrem Kellerzimmer bestimmen und lüften daher nicht. Spart dann auch Energie und lüften wird sowieso überschätzt. Während einer Messperiode von 7:00 Uhr bis 15:00 bestimmen Sie dreimal automatisch die Radonbelastung in Ihrem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung¹⁰.



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

Radon zerfällt mit einer Halbwertszeit von 3.5d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 143d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

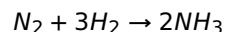
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $300 Bq/m^3$ auf unter $80 Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	77.1	27.9	
Sauerstoff	21.3	16.2	
Kohlenstoffdioxid	0.029	11.8	

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 Sie Ihr etwas pappiges Toastbrot mampfen kommt Ihnen die Dokumentation über Brot aus Luft in den Sinn. Sie denken darüber ein wenig nach. Für die Umwandlung von Stickstoff N_2 mit Wasserstoff H_2 zu Ammoniak NH_3 gilt folgende Reaktionsgleichung¹¹:



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)**

¹⁰Die Quelle der Inspiration für die Aufgabe war folgender Artikel: [Atmende Wand](#)

¹¹Die 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 Ihr Studentenjob war nach Ladenschluss bei Penny die Regale einzuräumen. Dabei ist Ihnen in der Auslage der Sonderangebote das Necronomicon¹² in die Hände gefallen. Nun sind Sie ein Magier der Zeichen geworden! Also eigentlich können Sie nur Mathe und das dämliche Necronomicon hat Sie in die Vergangenheit geschleudert... aber gut, was tut man nicht alles im Jahr 148 n. Chr. für den neuen Lehnsherren Henry dem Roten. Sie bauen natürlich einen Schrottkugelturm um sich den Horden der Finsternis mit genug Schrott erwehren zu können! Ihnen stehen zwei mächtige magische Formeln zur Unterstützung zu Verfügung.

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

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 $10mg$ zu gleichförmigen Bleikugeln bei einer Geschwindigkeit von $13m/s$ bilden.

1. Wie hoch müssen Sie den Schrottkugelturm bauen lassen, damit sich runde Bleikugeln durch die Fallgeschwindigkeit von $13m/s$ bilden? **(3 Punkte)**

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

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

Sie haben jetzt die 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. 900 Bleikugeln produzieren wollen und die Bleikugel im Fall $1.4cm$ Abstand haben müssen? **(1 Punkt)**

¹²Ein 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 mit Ihrer Faultierpension in Down Under nicht geklappt. War vielleicht auch nicht so die beste Idee... aber dafür haben Sie eine Neue! Oder wie es Mike Tyson zugeschrieben wird: „Ich wurde nie niedergeschlagen, ich war immer am Aufstehen!“. Daher machen Sie jetzt einen Großhandel mit Kaninchenfleisch und damit dem teuersten Fleisch in Australien auf. Moment, hopsen hier nicht, seit Thomas Austin im Jahr 1875 ungefähr 26 Kaninchen entlassen hat, Millionen von Kaninchen rum? Wieso ist das Kaninchenfleisch dann so exklusiv? Dem wollen wir mal mathematisch nachgehen!¹³

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

$$f(t) = 1.2 \times 10^{10} - 1.1 \times 10^9 \cdot 2.3^{-0.3 \cdot t + 4.1}$$

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 2 nach folgender Formel.

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

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

Das Myxoma Virus und das RHDV töten 98.5% 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 Osten von Australien. Australien hat eine West-Ost-Ausdehnung von 4000km und eine Nord-Süd-Ausdehnung von knapp 3400km. Die Kaninchen breiten sich radial mit einer Geschwindigkeit von 9.8km pro Jahr aus.

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

Eine jährliche Impfung gegen das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) kosten 12\$ pro Tier und der durchführende Arzt verlangt ca. 35\$ pro Tier.

6. In Ihrem Stall leben 1000 Mastkaninchen. Mit welchen jährlichen Zusatzkosten für die Impfungen der Kaninchen müssen Sie daher kalkulieren? **(1 Punkt)**

¹³Die Quelle der Inspiration für die Aufgabe war der folgendes YouTube Video: [Incredible Stories – Why don't they eat wild rabbits in Australia? They have millions of them! The reason is surprising...](#)

121. Aufgabe

(12 Punkte)

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



Ostfriesland. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer des Schafs Fridolin und Ihnen. Grünes Gras unter Ihren Füßen und ein strammer Wind im Gesicht, egal wohin Sie schauen. Ein schmatzendes Geräusch ertönt unter Ihnen. Sie sinnieren, sollten Sie Ihre weiten Graslandschaften jetzt schon düngen? Dafür benötigen Sie die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Sie sehen nicht ein, Geld für einen Agrarmetrologen zu bezahlen. Also rechnen Sie mit folgenden Informationen zu Monatsmultiplikatoren des GLT-Wertes: Januar mit $0.6\times$, Februar mit $0.8\times$ und März mit $1.2\times$. Sie haben noch im letzten Jahr folgende Temperaturen gemessen.

Datum	C°
01. Jan 2023	0.4
01. Feb 2023	1.5
01. Mrz 2023	3.1
01. Apr 2023	6.1

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 Ihrer Boskoopplantage wurden Sie mit Ihrem Trecker von einer Gruppe elektrifizierter Rentner abgedrängt. Der Trecker muss wieder aus dem Graben! Fridolin und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Fridolin mit 180N . Die elektrifizierten Rentner bringen eine Kraft von 190N auf.

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

5. Im ersten Versuch legen Sie das Seil für Fridolin lotrecht über einen Ast oberhalb des Treckers. Die Rentner ziehen in einer geraden Linie über die Böschung hinweg am anderen Seil. Welche Kraft wird aufgebracht? **(2 Punkte)**
6. Im zweiten Versuch ziehen Fridolin und die Rentner mit einem 50° Winkel mit ihrem Seil an dem Trecker. Welche Kraft wird aufgebracht? **(2 Punkte)**
7. Mit welcher Beschleunigung ziehen Sie den 1.2t schweren Trecker *jeweils* aus dem Graben, wenn $F = m \cdot a$ gilt? **(1 Punkt)**

122. Aufgabe

(10 Punkte)

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



In der Kartonagenfabrik Wenn man sich zu spät anmeldet, dann ist die Exkursion nicht so toll. Also geht es mit Rektor Skinner und Mrs. Krabappel in die Kartonagenfabrik. Wie schon im vorherigen Semester... In der Kartonagenfabrik angekommen erfahren Sie, dass die Kartons zum Versand von Nägeln nicht hier zusammengebaut werden sondern das sich die Endfertigung in Flint, Michigan befindet. Unter anderem wird dort der berühmte *Doppelt gewellte, 6-mal-gefaltete, 0.8mm, 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 Sie wenig Zuversicht zu erkennen.

Jetzt heißt es Kartons optimieren. Der nun zu optimierende, flache Karton hat eine Länge von 40cm und eine Breite von 20cm. 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. 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

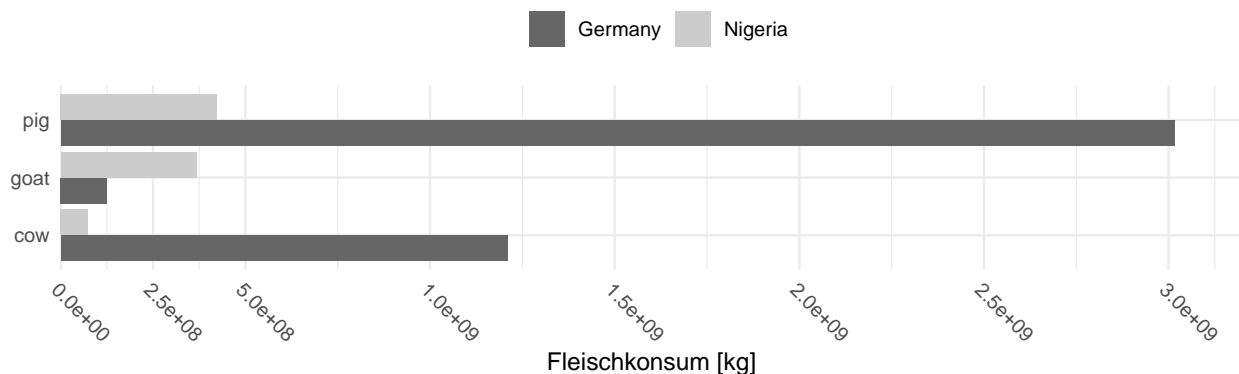
(10 Punkte)

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



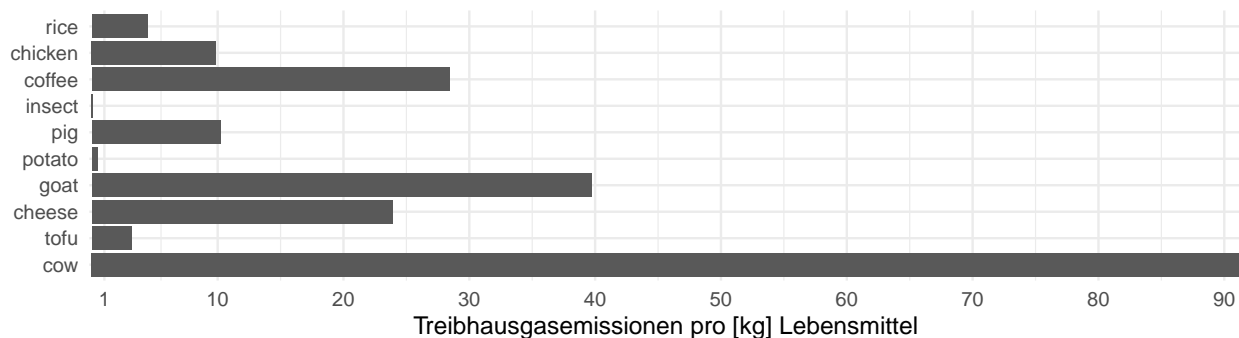
Ein Pfund Insekten, bitte! Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹⁴. Schauen wir uns dazu einmal den Vergleich Deutschland zu Nigeria an. Nigeria hat eine der am schnellsten wachsenden Bevölkerungen der Welt und wird vermutlich im Jahr 2100 zu den Top 5 der bevölkerungsreichsten Länder zählen. Im Jahr 2023 leben ca. 8.4×10^7 Menschen in Deutschland und ca. 1.84×10^8 Menschen in Nigeria. Mit den Informationen wollen wir anfangen und dann eine Prognose für den Fleischkonsum im Jahr 2050 zu treffen.

Im folgenden ist Abbildung des Fleischkonsums im Jahr 2023 in Deutschland und Nigeria in [kg] einmal dargestellt.



1. Stellen Sie den Fleischkonsum in Deutschland und Nigeria im Jahr 2023 *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)**

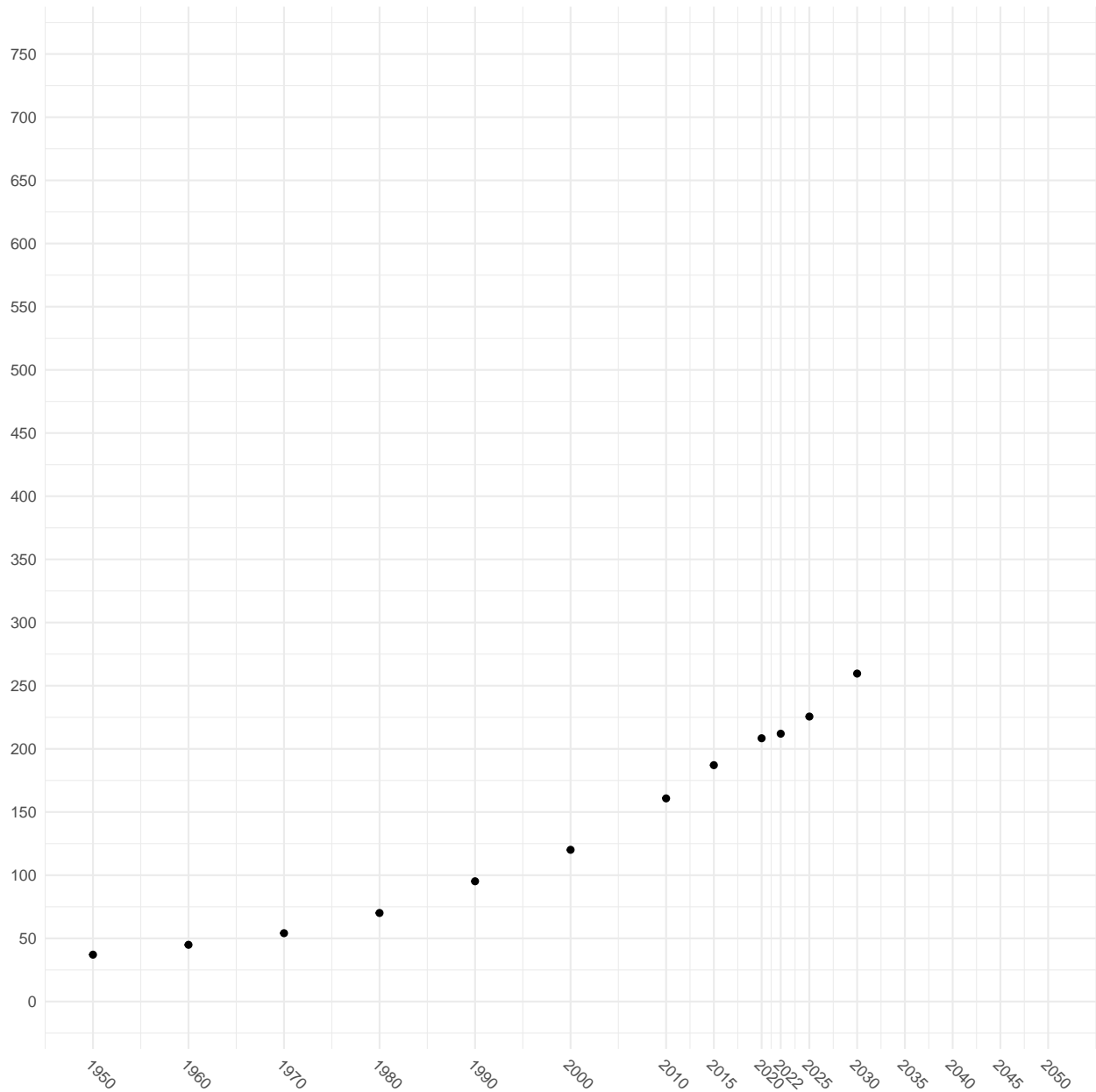
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



3. Stellen Sie in einer Tabelle die Treibhausgasemissionen an CO₂ pro Kopf, die durch den Fleischkonsum in Deutschland und Nigeria im Jahr 2023 entstehen, dar! Ergänzen Sie auch hier das Verhältnis Nigeria zu Deutschland! **(2 Punkte)**

¹⁴Die 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 70%-iger Angleichung der Lebensbedingungen zu Deutschland im Jahr 2023! **(1 Punkt)**
6. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria im Jahr 2050 im Vergleich zum Jahr 2023, der sich durch den angeglichenen Fleischkonsum ergibt! **(1 Punkt)**
7. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria, wenn die gesamte Proteinaufnahme durch Insekten ersetzt würde! **(1 Punkt)**

124. Aufgabe

(12 Punkte)

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



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

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.5% angenommen. In 92% der Fälle ist ein HIV-Test positiv, wenn der Patient erkrankt ist. In 3% der Fälle ist ein HIV-Test positiv, wenn der Patient *nicht* erkrankt ist und somit gesund ist. Sie stutzen. Wie wahrscheinlich ist es denn eigentlich an AIDS erkrankt zu sein (K^+), wenn Sie einen positiven AIDS-Test vorliegen haben (T^+)? Gehen Sie für die folgenden Berechnungen von $n = 4 \times 10^4$ Patienten mit einem diagnostischen Test für AIDS aus. Sie nehmen sich also einen Kuli und fangen an auf der historischen Ausgabe des Spiegels zu rechnen¹⁵.

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

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

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

¹⁵Die Quelle der Inspiration für die Aufgabe war der folgende wissenschaftlicher Artikel: Binder et al. (2022) Von Baumdiagrammen über Doppelbäume zu Häufigkeitsnetzen – kognitive Überlastung oder didaktische Unterstützung? Journal für Mathematik-Didaktik, 1-33

125. Aufgabe

(12 Punkte)

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



Network-Marketing oder Schneeballschlacht! Eine Möglichkeit, leicht Geld zu verdienen, ist es anderen Menschen für Geld zu versprechen, wie man leicht reich werden kann. Am besten natürlich ohne viel Aufwand und ortsunabhängig. Dann wollen wir mal loslegen. Schnell ein YouTube-Werbevideo gedreht und auf geht es mit unserem Network-Marketing. Aber Moment, wie funktioniert Network-Marketing eigentlich und was hat das alles mit einer Schneeballschlacht zu tun? Wir wollen hier einmal in die Untiefen des „passiven Einkommens“ abtauchen¹⁶.

Das Jahr 2022 war das erfolgreichste Jahr in der Geschichte von Direct Finanzanlagen Left/Right (D-FL/R). Das Unternehmen steigerte den Umsatz um rund 24 Prozent von 260 Millionen Euro im Jahr 2021. Doch wie viel kommt bei den Partnern an? Laut D-FL/R habe das Unternehmen 2.8×10^5 aktive Partner.

1. Berechnen Sie zuerst den Umsatz der Firma D-FL/R im Jahr 2022! **(1 Punkt)**
2. Wie viel von dem Umsatz im Jahr 2022 wird im Durchschnitt von jedem aktiven Partner erwirtschaftet? **(1 Punkt)**
3. Welche *monatlicher* Umsatz ergibt sich dadurch im Durchschnitt für jeden aktiven Partner bei einer direkten Provision von 20%? **(1 Punkt)**

Ihr zu vermarktendes Produkt, hinter dem Sie voll stehen, kostet 200EUR pro Einheit im Direktverkauf. Die direkte Provision für die erste Stufe beträgt 30%. Für die zweite, dritte und vierte Stufe betragen die indirekten Provisionen jeweils 2%, 1% und 0.5%. Jeder Ihrer angeworbenen „Partner“ wirbt wiederum drei Partner für sich selbst an. Pro Monat werden im Schnitt zwei Einheiten vom Produkt verkauft. Sie wollen nun 1800EUR im Monat *passiv* – also durch indirekte Provisionen – erwirtschaften.

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

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

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

Sie mussten zum Einstieg bei D-FL/R Einheiten des Produkts für 6000EUR kaufen. Diese Einheiten können Sie nur direkt verkaufen. Leider mussten Sie den Kauf über einen Kredit über 6% p.a. über 36 Monate finanzieren.

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

¹⁶Die Quellen der Inspiration für die Aufgabe waren folgendes YouTube Video: [Simplicissimus – Die meistgesuchte Betrügerin der Welt](#) und der Artikel: [Deutschlandfunk Kultur – Die Illusion, schnell reich zu werden](#)

126. Aufgabe

(10 Punkte)

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



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

In dem Spiel haben Sie nun auf einmal 5 sechseitige Würfel ($5d6$) zum würfeln in der Hand. Wenn Sie eine 6 würfeln, haben Sie 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)**

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

3. Würden Sie die Axt mit zwei vierseitigen Würfeln ($2d4$) als Schaden oder das Schwert mit einem zwölfseitigen Würfel plus 5 ($1d12+5$) als Schaden bevorzugen? Begründen Sie Ihre Antwort mathematisch! **(1 Punkt)**

Jetzt wird es immer wilder, da Sie sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass Ihr Rettungswurf gegen den zaubernden Hexer funktioniert. Sie haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.6$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.75$. Sie haben mitgezählt und festgestellt, dass in 40 von 100 Fällen Ihr Rettungswurf bei einem erfolgreichen Zauber funktioniert hat.

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

127. Aufgabe

(12 Punkte)

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



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

Name	$P(\text{win})$	$P(\text{outbid})$
Frida	0.3	0.08
Elke	0.2	0.043

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

Glücksrad für Arme auf der Kirmes! Leider hat es für Maren Gilzer nicht gereicht. Deshalb sind Sie jetzt auf der Kirmes und spielen mit Catwoman um das große Geld. Das Glücksrad hat 20 Felder. Sie drehen das Glücksrad zweimal. Auf 8 Feldern gewinnen Sie 4000EUR sonst 1000EUR. Ganz schön viel Geld und ganz schön zwielichtig hier...

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

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

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

Teil XII.

Angewandte Nutztier- und Pflanzenwissenschaften (M.Sc.)

128. Aufgabe

(6 Punkte)

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

129. Aufgabe

(8 Punkte)

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

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

mit

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

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

130. Aufgabe

(6 Punkte)

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