

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

Fakultät Agrarwissenschaften und Landschaftsarchitektur (AuL)

Klausurfragen Bio Data Science

für Pflichtmodule

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

(Prüfungsleistung der Wahlpflichtmodule ist eine Portfolioprüfung)

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

Wintersemester 2024/25



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



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

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



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

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



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

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



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

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



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

Erlaubte Hilfsmittel

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

Endnote

_____ von 20 Punkten sind aus den Multiple Choice Aufgaben erreicht.
_____ von 67 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 87 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
83.0 - 87.0	1,0
79.0 - 82.5	1,3
74.5 - 78.5	1,7
70.5 - 74.0	2,0
66.0 - 70.0	2,3
61.5 - 65.5	2,7
57.5 - 61.0	3,0
53.0 - 57.0	3,3
49.0 - 52.5	3,7
43.5 - 48.5	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	8	9	10	10	9	12	9

- Es sind ____ von 67 Punkten erreicht worden.

Multiple Choice Aufgaben

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

ANOVA

1. Aufgabe

(2 Punkte)

Sie führen einen Versuch mit einer Behandlung und drei Faktorleveln durch. Danach rechnen Sie eine einfaktorielle ANOVA und es ergibt sich ein $\eta^2 = 0.78$. Welche Aussage ist richtig?

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

2. Aufgabe

(2 Punkte)

Sie führen ein Feldexperiment durch um das Gewicht von Brokoli zu steigern. Die Pflanzen wachsen unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen. Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.26$. Welche Aussage ist richtig?

- A** ☐ Es werden 26% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen erklärt wird.
- B** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 26% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 74%.
- C** ☐ Es werden 26% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- D** ☐ Es werden 74% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.
- E** ☐ Das η^2 beschreibt den Anteil der Varianz, der durch den Forschenden entsteht. Es gilt die Regel, dass ca. 70% der Varianz eines Versuches durch die Versuchsdurchführung entstehen sollen.

3. Aufgabe

(2 Punkte)

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

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

- D** ☐ Die ANOVA berechnet die F-Statistik aus den SS Behandlung geteilt durch die SS Fehler.
- 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)

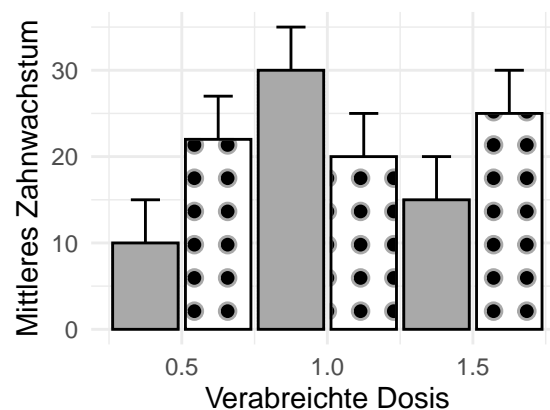
Viele statistische Verfahren nutzen eine Teststatistik um eine Aussage über den Zusammenhang zwischen der Grundgesamtheit und der Stichprobe abzubilden. Ein statistisches Testwerkzeug ist hierbei die ANOVA. Die ANOVA rechnet dabei...

- A** ☐ ... den Unterschied zwischen mehreren Varianzen aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.
- B** ☐ ... den Unterschied zwischen der Varianz ausgelöst durch alle Behandlungsgruppen und der Varianz aus globalen Behandlungsgruppen der Kontrollen. Wenn die ANOVA nicht signifikant ist, muss ein Posthoc-Test ausgeschlossen werden.
- C** ☐ ... 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.
- D** ☐ ... 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.
- E** ☐ ... den Unterschied zwischen der Varianz durch verschiedene Behandlungsgruppen unter der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, kann kein Effekt η^2 bestimmt werden.

5. Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung von Vitamin C auf das Zahnwachstum bei Hasen. Der Versuch wurde an 59 Tieren durchgeführt, wobei jedes Tier eine von drei Vitamin-C-Dosen (0.5, 1 und 1.5 mg/Tag) über eine von zwei Verabreichungsmethoden erhielt. Welche Aussage ist richtig im Bezug auf eine zweifaktorielle ANOVA?



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

Deskriptive Statistik & Explorative Datenanalyse

6. Aufgabe

(2 Punkte)

Berechnen Sie den Mittelwert und Standardabweichung von y mit 5, 13, 8, 18 und 15.

- A** ☐ Es berechnet sich 11.8 ± 27.7
- B** ☐ Sie erhalten 11.8 ± 5.26
- C** ☐ Es ergibt sich 12.8 ± 2.63
- D** ☐ Sie erhalten 11.8 ± 2.29
- E** ☐ Es berechnet sich 12.8 ± 27.7

7. Aufgabe

(2 Punkte)

Wie lautet der Median, das 1st Quartile sowie das 3rd Quartile von y mit 11, 6, 4, 20, 9, 36 und 42.

- A** ☐ Es ergibt sich 11 ± 6
- B** ☐ Es ergibt sich 18 ± 6
- C** ☐ Sie erhalten 11 ± 36
- D** ☐ Sie erhalten 11 [6; 36]
- E** ☐ Sie erhalten 11 [4; 34]

8. Aufgabe

(2 Punkte)

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

- A** ☐ Wir sollten zwei bis fünf Beobachtungen mindestens pro Gruppe vorliegen haben.
- B** ☐ Damit wir hier sauber eine Abbildung von einem
- C** ☐ Die Mindestanzahl liegt bei fünf Beobachtungen.
- D** ☐ Die optimale Anzahl ist größer als hundert Beobachtungen, wobei es gerne sehr viel mehr sein können.
- E** ☐ 10 Beobachtungen.

9. Aufgabe


(2 Punkte)

Um die Standardabweichung zu berechnen müssen wir folgende Rechenoperationen durchführen.

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

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

11. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit zu Maiss 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. 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** ☐ Wenn sich der Mittelwert und der Median nicht unterscheiden, liegen vermutlich keine Outlier in den Daten vor.
- C** ☐ Wenn sich der Mittelwert und der Median unterscheiden, liegen vermutlich keine Outlier in den Daten vor.
- D** ☐ Da sich der Mittelwert und der Median unterscheiden, liegen vermutlich keine Outlier in den Daten vor. Wir verwenden den Datensatz so wie er ist.
- E** ☐ Wenn sich der Mittelwert und der Median unterscheiden, liegen vermutlich Outlier in den Daten vor.

12. Aufgabe

(2 Punkte)

Um zu überprüfen, ob die Daten die Annahme einer Normalverteilung genügen, können wir folgende Visualisierung nutzen. Dabei kommt dann auch die entsprechende Regel zur Abschätzung der Annahme einer Normalverteilung zur Anwendung.

- A** ☐ Nach dem Einlesen der Daten nutzen wir einen Boxplot um zu schauen, ob alle Boxen über alle Behandlungen in etwa gleich groß sind. Damit ist dann auch das IQR in allen Behandlungen in etwa gleich.
- B** ☐ Einen Boxplot. Der Median, dargestellt als Linie, muss in der Mitte des IQR, dargestellt durch die Box, liegen.
- C** ☐ Einen Dotplot. Die Punkte müssen sich wie an einer Perlenschnur anordnen. Eine Abweichung führt zur Ablehnung der Annahme einer Normalverteilung.
- D** ☐ Einen Barplot. Die Mittelwerte müssen alle auf einer Höhe liegen. Die Fehlerbalken haben hier keine Informationen.
- E** ☐ Einen Violinplot. Der Bauch der Violine muss hierbei einen höheren Wert annehmen als der Steg der Violine. Dann kann die Annahme einer Normalverteilung angenommen werden.

13. Aufgabe

(2 Punkte)

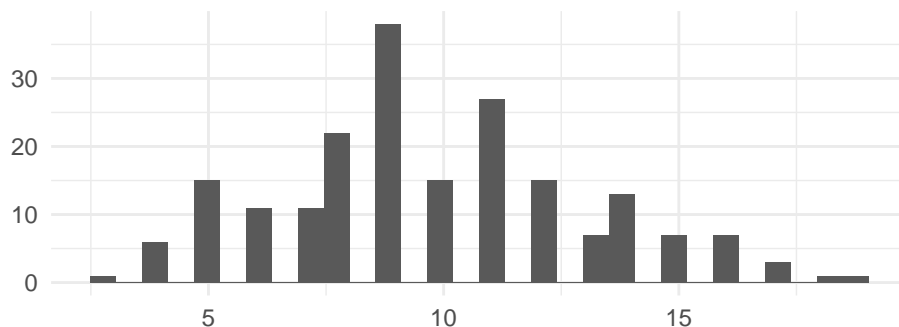
Sie wollen in Ihrer Abschlussarbeit über eine explorative Datenanalyse überprüfen, ob Ihr gemessener Endpunkt einer Normalverteilung folgt. Welche drei Abbildungen eignen sich insbesondere für die Überprüfung?

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

14. Aufgabe

(2 Punkte)

Bevor Sie in Ihrer Abschlussarbeit einen statistischen Test rechnen, wollen Sie einmal betrachten, welcher Verteilung Ihre $n = 184$ geernteten Pflanzen folgen. Welche Verteilung ist abgebildet?



- A ☐ Es handelt sich um eine Normalverteilung.
- B ☐ Es handelt sich um eine Binomial-Verteilung.
- C ☐ Wir haben eine Poisson-Verteilung vorliegen.
- D ☐ Wir haben eine Gammaverteilung vorliegen.
- E ☐ Eine multivariate Normalverteilung.

Lineare Regression & Korrelation

15. Aufgabe

(2 Punkte)

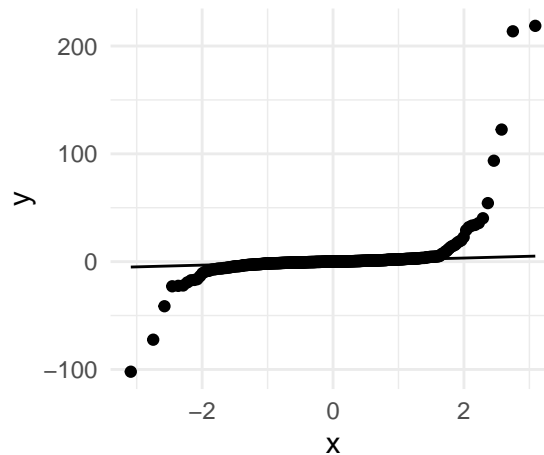
Im Allgemeinen gibt es zwei mögliche Ziele für ein Regressionsmodell. Wir können eine Vorhersagemodell oder ein kausales Modell rechnen. Welche Aussage ist für ein kausales Modell richtig?

- A ☐ Ein kausales Modell schliesst grundsätzlich lineare Modell aus. Es muss ein Graph gefunden werden, der alle Punkte beinhaltet. Erst dann kann das R^2 berechnet werden.
- B ☐ Es wird ein Trainingsdatensatz zum Trainieren des Modells benötigt. Der Testdatensatz dient zur Validierung. Dies gilt insbesondere für ein kausales Modell.
- C ☐ 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 X auf den gemessenen Endpunkt Y aus?
- D ☐ 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.
- E ☐ Wir modellieren den Zusammenhang zwischen X und Y wenn ein kausales Modell rerechnet wird. Dabei kann nicht der gesamte Datensatz genutzt werden. Es wird ein Trainingsdatensatz zum Trainieren des Modells benötigt.

16. Aufgabe

(2 Punkte)

Sie rechnen in eine linearen Regression und erhalten folgenden QQ Plot um die Annahme der normalverteilten Residuen zu überprüfen. Welche Aussage ist richtig?

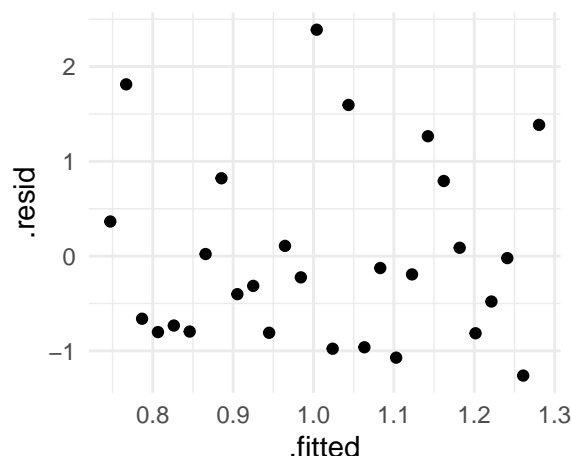


- A** ☐ Wir betrachten insbesondere die beiden Enden der Gerade. Der Rest ist mehr oder minder egal, dann ist die Annahme an die Normalverteilung der Residuen erfüllt.
- B** ☐ Wir betrachten die Punkte auf der Geraden. Wenn die Punkte einigermaßen auf der Geraden liegen, dann gehen wir von normalverteilten Residuen aus. Wir können hier von normalverteilten Residuen ausgehen.
- C** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden und Korrelation ist negativ.
- D** ☐ Wir betrachten die Punkte. Wenn die Punkte einigermaßen gleichmäßig verteilt liegen, dann gehen wir von normalen Residuen aus.
- E** ☐ Wir betrachten die Gerade. Wenn die Punkte einigermaßen gleichmäßig um die Gerade verteilt liegen, dann gehen wir von normalverteilten Residuen aus. Dies ist hier nicht der Fall. Wir haben keine normalverteilten Residuen vorliegen.

17. Aufgabe

(2 Punkte)

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



- A** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Es ist kein Muster zu erkennen.
- B** ☐ Wir betrachten die Nulllinie und alle Punkte sollten ohne Muster gleichmäßig um die Nulllinie liegen. Da dies der Fall ist, gehen wir von keinen Ausreißern aus.

- C** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Diagonalen. Damit ist das Modell erfolgreich geschätzt worden.
- 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 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 .

18. Aufgabe

(2 Punkte)

Welche Aussage über den Korrelationskoeffizienten ρ ist richtig?

- A** ☐ Der Korrelationskoeffizienten ρ zeigt keinen Zusammenhang zwischen zwei Variablen x und y bei einem Wert von 0. Einen maximalen negativen Zusammenhang bei -1 und somit auch einen maximalen positiven Zusammenhang bei 1. Korrelationskoeffizienten ρ ist einheitslos.
- B** ☐ Korrelationskoeffizienten ρ liegt zwischen 0 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als Standardisierung verstanden werden.
- C** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen 0 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos. Eine Signifikanz kann nicht nachgewiesen werden.
- D** ☐ Der Korrelationskoeffizienten ρ 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.
- E** ☐ Der Korrelationskoeffizienten ρ ist eine veraltete Darstellungsform von Effekten in der linearen Regression und wird wie das η^2 aus der ANOVA interpretiert. Der Korrelationskoeffizienten ρ beschreibt den Anteil an erklärter Varianz durch die Regression.

19. Aufgabe


(2 Punkte)

Nach einer simplen linearen Regression zur Untersuchung vom Einfluss der NO_3 -Konzentration in $[\mu g]$ im Wasser auf den absoluten Proteingehalt von Brokkoli in $[kg]$ erhalten Sie einen β_{NO_3} Koeffizienten von 6.9×10^{-7} und einen hoch signifikanten p -Wert mit 0.00051. Warum sehen Sie so einen kleinen Effekt bei einer so deutlichen Signifikanz?

- A** ☐ 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.
- B** ☐ Manchmal ist die Einheit der Einflussvariable X zu groß gewählt, so dass der Anstieg von 1 Einheit in X zu einer zu großen Änderung in y führt. Daher kann der Effekt β_{NO_3} sehr klein wirken, da der p -Wert wird auf einer einheitslosen Teststatistik bestimmt wird.
- C** ☐ 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.
- 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. 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.

20. Aufgabe

(2 Punkte)

Neben der klassischen Regression kann die Funktion $\text{lm}()$ in  auch für welche andere Art von Anwendung genutzt werden?

- A** ☐ 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.
- B** ☐ Ist die Einflussvariable X ein Faktor so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich. Die Funktion `lm()` kann dabei eigentlich weggelassen werden, wird aber traditionell gerechnet.
- C** ☐ Ist die Einflussvariable X ein Faktor so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich. Dennoch muss zuerst ein lineares Modell mit der Funktion `lm()` in **R** gerechnet werden.
- 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** ☐ Ist die Einflussvariable X numerisch so werden die Gruppenmittelwerte geschätzt und eine anschließende ANOVA sowie multipler Gruppenvergleich mit `{emmeans}` ist möglich.

21. Aufgabe

(2 Punkte)

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

- A** ☐ Das GLM ist eine Vereinfachung des LM in R. Mit dem GLM lassen sich polygonale Regressionen rechnen. Somit stehen neben der Normalverteilung noch weitere Verteilungen zu Verfügung.
- B** ☐ Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsgruppen für das X bzw. die Einflussvariablen in einer linearen Regression zu wählen.
- C** ☐ In **R** ist mit dem *generalisierten linearen Modell (GLM)* eine Modellierung implementiert, die neben der klassischen Normalverteilung auch die Poissonverteilung für Zähldaten oder die Binomialverteilung für 0/1-Daten modellieren kann.
- D** ☐ 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.
- E** ☐ Das GLM erlaubt auch nicht normalverteilte Residuen in der Schätzung der Regressionsgrade.

Vermischte Themen

22. Aufgabe


(2 Punkte)







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

- A** ☐ Durch eine Randomisierung können wir nicht von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.
- B** ☐ Randomisierung erlaubt erst die Mittelwerte zu schätzen. Ohne Randomisierung keine Mittelwerte. Ohne Mittelwerte keine Varianz und somit auch kein statistischer Test.
- C** ☐ Randomisierung sorgt für Strukturgleichheit und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- D** ☐ Randomisierung ist die direkte Folge von Strukturgleichheit. Die Strukturgleichheit erlaubt es erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- E** ☐ 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.

23. Aufgabe


(2 Punkte)


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

- A ☐ Die Spracherkennung von  ist nicht in der Lage Deutsch zu verstehen.
- B ☐ Programmiersprachen können nur englische Begriffe verarbeiten. Zusätzliche Pakete können zwar geladen werden, aber meist funktionieren diese Pakete nicht richtig. Deutsch ist International nicht bedeutend genug.
- C ☐ Die  Pakete sind nur in englischer Sprache verfasst. Das ist aber nicht der Hauptgrund, denn  hat wie alle Programmiersprachen Probleme mit Umlauten und Sonderzeichen.
- D ☐  Pakete sind nur in englischer Sprache verfasst. Es macht keinen Sinn  daher in Deutsch zu bedienen.
- E ☐ Programmiersprachen haben Probleme mit Umlauten und Sonderzeichen der deutschen Sprache. Daher ist die Nutzung in Deutsch in den AGBs von  untersagt.

24. Aufgabe

(2 Punkte)

Bei der explorativen Datenanalyse (EDA) in  gibt es eine richtige Abfolge von Prozessschritten, auch *exitCircle of life* genannt. Wie lautet die richtige Reihenfolge für die Erstellung einer EDA?

- A ☐ Die Funktionsreihenfolge ist wie folgt: `read_excel()` -> `mutate()` -> `ggplot()`. Dabei ist bei der Transformation der Daten darauf zu achten, dass keine Faktoren erstellt werden.
- B ☐ Wir lesen die Daten über eine generische Funktion `read()` ein und müssen dann die Funktion `ggplot()` nur noch installieren. Dann haben wir die Abbildungen als `*.png` vorliegen.
- C ☐ Wir lesen die Daten ein und mutieren die Daten. Dabei ist wichtig, dass wir nicht das Paket `tidyverse` nutzen, da dieses Paket veraltet ist. über die Funktion `library(tidyverse)` entfernen wir das Paket von der Analyse.
- D ☐ Für eine explorativen Datenanalyse (EDA) in  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 eigentliche EDA.
- E ☐ Wir transformieren die Spalten über `mutate()` in ein `tibble` und können dann über `ggplot()` uns die Abbildungen erstellen lassen. Dabei beachten wir das wir keine Faktoren in den Daten haben.

25. Aufgabe

(2 Punkte)

Es sei $n_1 < n_2$ in dem Modell $Y \sim X$. Welche Aussage ist richtig?

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

26. Aufgabe

(2 Punkte)

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

- A** ☐ Die Ferkel stammen von der gleichen Sau und sind somit untereinander unabhängig.
- B** ☐ Untereinander stark korreliert. Die Ferkel sind von einer Mutter und somit miteinander korreliert. Dies wird in der Statistik jedoch meist nicht modelliert.
- C** ☐ Die Ferkel stammen von der gleichen Sau und sind somit untereinander abhängig.
- D** ☐ Untereinander unabhängig. Sollten die Mütter verwandt sein, so ist die Varianzstruktur ähnlich und muss modelliert werden.
- E** ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.

27. Aufgabe

(2 Punkte)

In einer Studie wollen Sie den Effektschätzer Risk ratio berechnen. Sie finden in Ihrem Experiment zur Behandlung von Klaueninfektionen bei Kühen in 3 Tieren Erkrankung der Klauen vor. 12 Tiere sind gesund. Welche Aussage ist richtig?

- A** ☐ Es ergibt sich ein Risk ratio von 4, da es sich um ein Anteil handelt.
- B** ☐ Es ergibt sich ein Risk ratio von 0.2, da es sich um ein Anteil handelt. Wir berechnen den Anteil der Kranken.
- C** ☐ Es ergibt sich ein Risk ratio von 0.25, da es sich um ein Anteil handelt.
- D** ☐ Es ergibt sich ein Risk ratio von 0.2, da es sich um eine Chancenverhältnis handelt.
- E** ☐ Es ergibt sich ein Risk ratio von 0.25, da es sich um eine Chancenverhältnis handelt

28. Aufgabe

(2 Punkte)

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

- A** ☐ 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** ☐ Relevanz und Signifikanz haben nichts miteinander zu tun. Daher gibt es auch keinen Zusammenhang zwischen hoher Fallzahl ($n > 10000$) und einem signifikanten Test. Ein Effekt ist immer relevant und somit signifikant.
- D** ☐ Aktuell werden zu grosse Datensätze für die gängige Statistik gemessen. Daher wendet man maschinelle Lernverfahren für kausale Modelle an. Hier ist die Relevanz gleich Signifikanz.
- 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: 0.001, 0.02, 0.01, 0.03, 0.42 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 0.006, 0.12, 0.06, 0.18, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 0.83% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 2e-04, 0.0033, 0.0017, 0.005, 0.07 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 2e-04, 0.0033, 0.0017, 0.005, 0.07 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 0.006, 0.12, 0.06, 0.18, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.006, 0.12, 0.06, 0.18, 2.52 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** ☐ Contrast letter display. Unterschiede in den Behandlungen werden durch den gleichen Buchstaben oder Symbol dargestellt. Die Interpretation des CLD führt häufig in die Irre.
- B** ☐ Compound letter display. Gleichheit in dem Outcomes wird durch den gleichen Buchstaben oder Symbol dargestellt. Teilweise ist die Interpretation des Verbunds (eng. compound) herausfordernd, da wir ja nach dem Unterschied suchen.
- C** ☐ Compact letter display. 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.
- D** ☐ 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.
- E** ☐ 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.

31. Aufgabe

(2 Punkte)

In Ihrer Bachelorarbeit müssen Sie einen Feldversuch auswerten. Nachdem Sie die zweifaktorielle ANOVA gerechnet haben und keine signifikante Interaktion vorliegt, wollen Sie jetzt einen Posthoc-Test rechnen. Welches R Paket nutzen Sie dafür am besten?

- 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 {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.
- C** ☐ 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.
- D** ☐ Das R Paket {emmeans} erlaubt die Durchführung eines multiplen Gruppenvergleichs. Aus einem emmeans Objekt lässt sich leider kein CLD erstellen. Dennoch ist das Paket einfach zu bedienen und wird deshalb genutzt. Die Interpretation der statistischen Auswertung wird über einen Barplot abgebildet.
- E** ☐ Das R Paket {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** Beim multiplen Testen kann es zu einer α -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Welch das bekannteste Verfahren ist.
- ☐ **B** 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
- ☐ **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** Beim multiplen Testen kann es zu Varianzheterogenität kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5%. Daher müssen die p-Werte entsprechend adjustiert werden. Das Verfahren nach Welch, bekannt aus dem t-Test, ist hier häufig anzuwenden.
- ☐ **E** Die Adjustierung der p-Werte nach Bonferroni erlaubt es gegen die β -Inflation vorzugehen, die häufig beim multiplen Testen auftritt. Das globale Powerniveau liegt nicht mehr bei 80% sondern sehr viel niedriger.

33. Aufgabe

(2 Punkte)

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

- ☐ **A** Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ nicht adjustiert werden im Gegensatz zu den p-Werten.
- ☐ **B** 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.
- ☐ **C** 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.
- ☐ **D** 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.
- ☐ **E** Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ adjustiert werden im Gegensatz zu den p-Werten.

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

35. Aufgabe

(2 Punkte)

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

- A** ☐ ... dass Fehlerterme in statistischen Modellen nicht verifiziert werden können.
- B** ☐ ... dass Annahmen an statistische Modelle meist falsch sind.
- C** ☐ ... dass ein minderwertiges Modell durch ein minderwertiges Modell ersetzt wird. Es gilt das Verifikationsprinzip nach Karl Popper.
- D** ☐ ... dass ein schlechtes Modell durch ein weniger schlechtes Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.
- E** ☐ ... dass ein schlechtes Modell durch das Falsifikationsprinzip durch ein noch schlechteres Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.

36. Aufgabe

(2 Punkte)

In fast allen wissenschaftlichen Disziplinen liegt der Grenzwert für das Signifikanzniveau α bei 5%. Wieso wurde dieser Konsens über die Signifikanzschwelle in dieser Form getroffen?

- A** ☐ Der Begründer der modernen Statistik, R. Fischer, hat die Grenze simuliert und berechnet. Dadurch ergibt sich dieser optimale Cut-Off.
- B** ☐ In der Wissenschaft gibt es neben der Naturkonstante, die sich aus der Beobachtung der Welt ergibt, noch die Kulturkonstante, die von einer Gruppe Menschen selbstgewählt wird. Dabei ist $\alpha = 5\%$ eine Kulturkonstante und wurde somit eher zufällig gewählt.
- C** ☐ 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.
- 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** ☐ Da Wissenschaftler eine Schwelle für die statistische Testentscheidung benötigen wurde α in einer großen Konferenz 1945 gewählt. Damit ist $\alpha = 5\%$ eine Kulturkonstante mit einem Rang einer Naturkonstante.

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

38. Aufgabe

(2 Punkte)

Sie versuchen folgende Aussage richtig in die Analogie der statistischen Testtheorie zu setzen. Welche Analogie ist richtig?

H_0 beibehalten obwohl die H_0 falsch ist

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

39. Aufgabe

(2 Punkte)

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

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

40. Aufgabe

(2 Punkte)

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

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

41. Aufgabe

(2 Punkte)

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

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

42. Aufgabe

(2 Punkte)

In Ihrer Forschungsarbeit wollen Sie eine Aussage über die untersuchte Population treffen. Dazu nutzen Sie eine ANOVA als statistischen Test. Erhalten Sie eine valide Aussage aus einem statistischen Test?

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

43. Aufgabe

(2 Punkte)

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

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

44. Aufgabe

(2 Punkte)

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

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

45. Aufgabe

(2 Punkte)

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

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

46. Aufgabe

(2 Punkte)

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

- A** ☐ Mit der Nullhypothese H_A und der Alternativhypothese H_0 gibt es zwei Hypothesen, die aber selten genutzt werden.
- B** ☐ Es gibt - bedingt durch das das Falsifikationsprinzip - ein Set von k Nullhypothesen, die iterative gegen $k - 1$ Alternativhypothesen getestet werden.
- C** ☐ Es gibt ein Hypothesenset bestehend aus k Hypothesen. Meistens wird die Nullhypothese H_0 und die Alternativhypothese H_A verwendet. Wegen des Falsifikationsprinzips ist es wichtig, die bekannte falsche und unbekannte richtige Hypothese mit in das Set zu nehmen.
- D** ☐ Die Hypothesen H_0 und H_A sind rein prosarischer Natur und bilden keinen mathematischen Hintergrund ab. In der Statistik wird die wissenschaftliche Fragestellung getestet. Daher stehen auch die verständlichen Hypothesen im Mittelpunkt der biologischen Interpretation.
- E** ☐ Mit der Nullhypothese H_0 und der Alternativhypothese H_A oder H_1 gibt es zwei Hypothesen.

Statistische Tests für Gruppenvergleiche

47. Aufgabe

(2 Punkte)

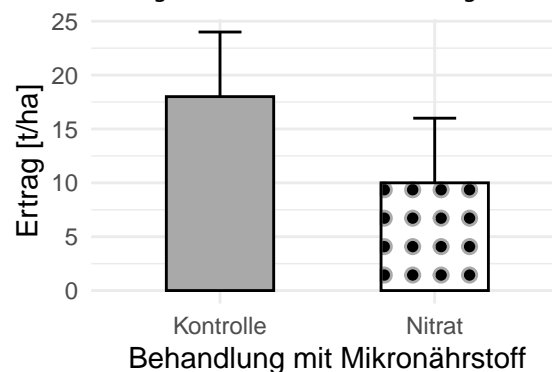
Welche Aussage über den t-Test im Allgemeinen ist richtig? Berücksichtigen Sie den Welch t-Test wie auch den Student t-Test!

- ☐ **A** 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.
- ☐ **B** Der t-Test testet generell zu einem erhöhten α -Niveau von 20%.
- ☐ **C** Der t-Test berechnet die Differenz von zwei Mittelwerten als Effekt und gibt eine Entscheidung, ob sich die beiden Mittelwerte *jeweils* von Null unterscheiden.
- ☐ **D** Der t-Test vergleicht die Mittelwerte von zwei Gruppen.
- ☐ **E** Der t-Test ist ein Vortest der ANOVA und basiert daher auf dem Vergleich von Streuungsparametern

48. Aufgabe

(2 Punkte)

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



- ☐ **A** Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei -8 unter einer groben Abschätzung. Wir müssen aber eine ANOVA rechnen um den Effekt wirklich bestimmen zu können.
- ☐ **B** Der Test deutet auf einen signifikanten Unterschied hin. Der Effekt liegt vermutlich bei -8.
- ☐ **C** Der Effekt und die Signifikanz lassen sich nicht aus Barplots abschätzen. Höchstens der Effekt als relativer Unterschied zwischen der Höhe der Barplots. Standard ist der mediane Unterschied aus Boxplots.
- ☐ **D** Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei -0.8.
- ☐ **E** Die Barplots deuten auf kein signifikanten Unterschied. Der Effekt liegt vermutlich bei -8.

49. Aufgabe

(2 Punkte)

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

- ☐ **A** Abhängige Beobachtungen müssen gesondert in einem gepaarten t-Test modelliert werden. Wenn wiederholt an dem gleichen Tier oder Pflanze gemessen wird, dann bilden wir die Differenz zwischen den beiden Zeitpunkten. Auf den Differenzen rechnen wir den gepaarten t-Test.
- ☐ **B** 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.

- C** ☐ 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.
- D** ☐ Der gepaarte t-Test nutzt die Varianz der Beobachtungen jeweils paarweise und bildet dafür eine verbundene Stichprobe. Dieser Datensatz d dient dann zur Differenzbildung.
- E** ☐ Abhängige Beobachtungen müssen gesondert in einem gepaarten t-Test modelliert werden. Wenn wiederholt an dem gleichen Tier oder Pflanze gemessen wird, dann bilden wir den Quotienten zwischen den beiden Zeitpunkten. Auf den Quotienten rechnen wir den gepaarten t-Test.

50. Aufgabe

(2 Punkte)

Sie führen paarweise t-Tests für alle Vergleiche der verschiedenen Rapssorten in Ihrem Experiment durch. Nach der Adjustierung für multiples Testen ist kein p-Wert unter der α -Schwelle. Ihr Experiment beinhaltet drei Rapssorten und eine ANOVA ergibt $p = 0.049$ für den Ertrag. Sie schauen sich auch die rohen, unadjustierten p-Werte an und finden hier als niedrigsten p-Wert $p_{3-2} = 0.052$. Welche Aussage ist richtig?

- A** ☐ Der Fehler liegt in den t-Tests. Wenn eine ANOVA signifikant ist, dann muss zwangsweise auch ein t-Test signifikant sein.
- B** ☐ 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.
- C** ☐ Die ANOVA testet auf der gesamten Fallzahl. Es wäre besser die ANOVA auf der gleichen Fallzahl wie die einzelnen t-Tests zu rechnen.
- D** ☐ Das Beispiel kann so nicht auftreten, da die ANOVA und die t-Tests algorithmisch miteinander verknüpft sind.
- 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 Unterschied nachweisen. Die p-Werte sind immer etwas kleiner als bei den t-Tests.

Teil I.


Deskriptive Statistik & Explorative Datenanalyse

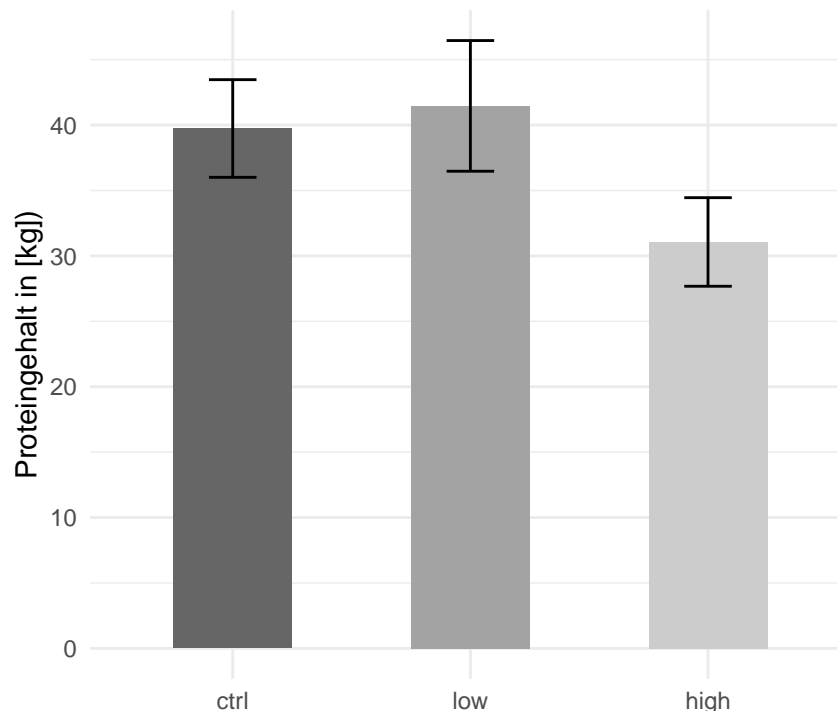
51. Aufgabe


(8 Punkte)


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



Zerforschen des Barplots Jonas steht vor einem ersten Problem, denn wenn es nach seiner Betreuer geht, soll er in einem Gewächshausexperiment Lauch auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Am Ende dann doch besser Stricken. Wunderbar. Eine echte Ablenkung für Jonas. Das heißt erstmal überlegen für Jonas. 'Hm...', Snickers und Iron Maiden. Das ist und bleibt die beste Kombination zum Nachdenken für Jonas. Die Behandlung werden verschiedene Düngestufen (*ctrl*, *low* und *high*) sein. In seiner Exceldatei wird er den Outcome (Y) *Proteingehalt* als *protein* aufnehmen. Vorab soll Jonas aber einmal die folgenden Barplots seiner Betreuer nachbauen, damit er den  Code schonmal für später vorliegen hat. Damit geht das Problem schon los. Wenn die Erschöpfung nicht wäre, ja dann wäre wohl vieles möglich für Jonas! Aber so..



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

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

52. Aufgabe

(8 Punkte)

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



Visualisierung des Barplots Jessica steht vor einem ersten Problem, denn wenn es nach ihrem Betreuer geht, soll sie in einem Freilandversuch Erbsen auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Warhammer. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Jessica denkt gerne über Warhammer nach. Die Behandlung waren verschiedene Lichtstufen (*none*, *200lm* und *600lm*). In ihrer Exceldatei hat sie den Endpunkt (*Y*) *Proteingehalt* als *protein* aufgenommen. Nun soll Jessica die Daten einmal als Barplots in einer Präsentation visualisieren, damit ihrem Betreuer wieder klar wird, was sie eigentlich nochmal gemacht hat und was für ein Ergebnis in einem statistischen Test zu erwarten wäre. Wäre da nicht noch etwas. Eine echte Herausforderung für sie war schon immer der Mangel gewesen. Ein leidiges Lied. Aber egal. Einfach mal raus um Rad zu fahren. Ohne Ziel und Uhr. Das ist was für Jessica.

treatment	protein
200lm	32.2
600lm	30.8
200lm	33.5
600lm	31.5
200lm	36.7
none	39.5
none	48.9
200lm	48.0
none	37.8
none	37.0
none	46.4
600lm	27.7

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

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

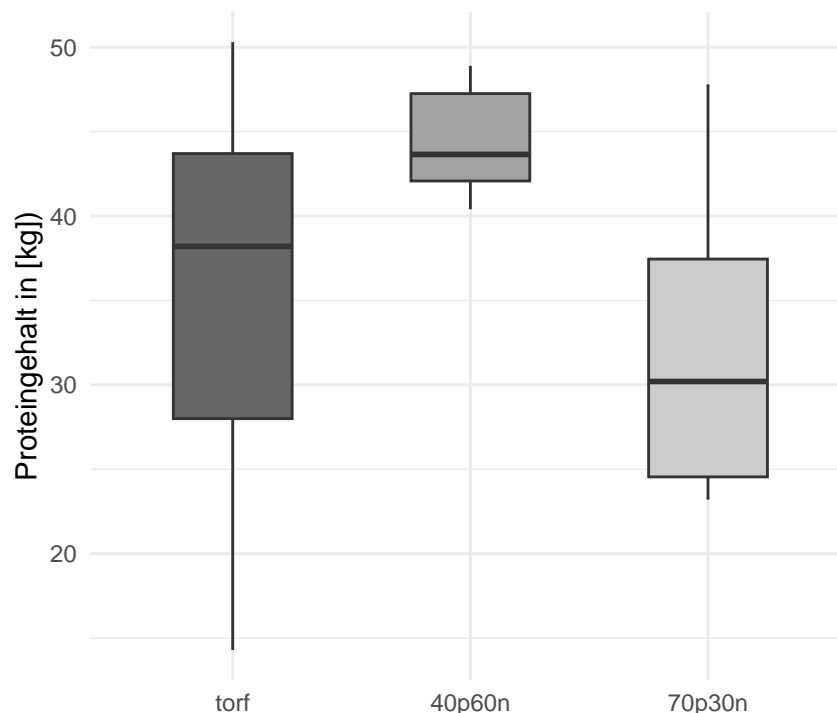
53. Aufgabe


(9 Punkte)


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



Zerforschen des Boxplots Wenn die Erschöpfung nicht wäre, ja dann wäre wohl vieles möglich für Jonas! Aber so.. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Jonas ist es eine Möglichkeit schneller ans Ziel zu gelangen. Jonas soll in seiner Hausarbeit Erdbeeren untersuchen. Die Behandlung in seiner Hausarbeit werden verschiedene Substrattypen (*torf*, *40p60n* und *70p30n*) sein. Erheben wird Jonas als Endpunkt (Y) *Proteingehalt* benannt als *protein* in seiner Exceldatei. Von seiner Betreuerin erhält er nun folgende Abbildung von Boxplots, die er erstmal zur Übung nachbauen soll, bevor er mit dem eigentlichen Versuch beginnt. Aber nur in passender Atmosphäre! Auf seinem Second Screen läuft Mission Impossible und Jonas schaufelt Snickers. Nicht effizient, aber gut.



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

54. Aufgabe

(9 Punkte)

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



Visualisierung des Boxplots Tina steht vor einem ersten Problem, denn wenn es nach ihrem Betreuer geht, soll sie in einem Versuch in einer Klimakammer Erdbeeren auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Astronomie. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Tina denkt gerne über Astronomie nach. Die Behandlung waren verschiedene Lichtstufen (*none* und *600lm*). In ihrer Exceldatei hat sie den Endpunkt (*Y*) *Trockengewicht* als *drymatter* aufgenommen. Nun soll Tina die Daten einmal als Boxplots in einer Präsentation visualisieren, damit ihrem Betreuer wieder klar wird, was sie eigentlich nochmal gemacht hat und was für ein Ergebnis in einem statistischen Test zu erwarten wäre. Anhand von Boxplots lässt sich eine Aussage über die Normalverteilung von *Y* treffen. Wäre da nicht noch etwas. Tina und die Wut, eine unendliche Geschichte mit kniffligen Wendungen. Aber egal. Tina will später nochmal raus um zu Boxen. Druck ablassen, dass muss sie auch.

treatment	drymatter
600lm	44.6
none	28.9
none	31.2
none	28.3
600lm	45.9
600lm	48.2
none	26.2
600lm	35.6
600lm	45.7
600lm	29.4
none	29.8
none	33.1
600lm	33.2
none	28.2
none	37.2
none	28.7

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

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

55. Aufgabe

(8 Punkte)

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



Visualisierung des Histogramm für kategoriale Daten In ihrer Abschlussarbeit möchte Yuki gerne die Daten aus einem Freilandversuch mit Kartoffeln in einem Histogramm darstellen. Das Histogramm erlaubt ihr dabei Rückschlüsse auf die Verteilung über das Outcome (Y) zu treffen. Aus den Boxen wummert London Grammar und ihr Mund ist verklebt von Reese's Peanut Butter Cups. 'Herrlich', denkt Yuki. In seinem Experiment hat Yuki die Knötchen gezählt. Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für sie war schon immer die Faulheit gewesen. Ein leidiges Lied. Yuki nickt im Takt von London Grammar und bemerkt dabei gar nicht was das Minischwein schon wieder anstellt.

Die Knötchen: 11, 4, 4, 4, 3, 1, 2, 3, 4, 5, 9, 7, 4, 3, 6, 6, 7, 2, 4, 5, 2, 3, 3, 4, 5, 5, 6, 3, 4, 3, 8, 7, 3

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

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

56. Aufgabe

(8 Punkte)

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



Visualisierung des Histogramm für kontinuierliche Daten In einem Gespräch mit ihrem Betreuer wird Paula gebeten seine Daten aus einem Feldexperiment mit Spargel in einem Histogramm darzustellen. Paula schmeißt noch eine Handvoll Smarties in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von White Lies. In ihrem Experiment hat er die mittleren Mehлтаuspots erst fotografiert und dann ausgezählt. Laut ihrem Betreuer soll das Histogramm helfen, die Verteilung der die mittleren Mehлтаuspots zu bestimmen. Es wäre einfacher, wenn da nicht noch was wäre. Paula und der Perfektionismus, eine unendliche Geschichte mit kniffligen Wendungen. Wenn White Lies ertönt, dann sucht die Ratte schleunigst Schutz unter dem Sofa. Paula schüttelt den Kopf.

Die mittleren Mehлтаuspots: 7.5, 11, 9.2, 9.5, 10, 9.4, 10.2, 11.4, 8.6, 7.3, 7.3, 9.2, 14.1, 8.3, 9, 9.1, 9.6, 11.7, 12.9, 11.5, 6.9, 11.6, 8.3, 12.5, 9.8

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

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

57. Aufgabe

(10 Punkte)

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



Visualisierung des Scatterplots Aus den Boxen wummert David Bowie und ihr Mund ist verklebt von Schokobons. 'Herrlich', denkt Jessica. Jetzt heißt es aber erstmal auf was anderes konzentrieren. Jessica möchte gerne den Zusammenhang zwischen durchschnittlicher Tagestemperatur [C/d] und Gewichtszuwachs in der 1LW im Kontext von Milchvieh herausfinden. Hierfür hat Jessica einen Leistungssteigerungsversuch im Emsland durchgeführt. Nach einigen unvorgesehenen Ereignissen hat sie es geschafft folgende Datentabelle zu erstellen. Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen. Aber das steht auch nicht im Zentrum. Nun stellt sich die Frage für sie, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Deshalb möchte Jessica als erstes eine explorative Datenanalyse durchführen. Dann was anderes. Das Verrückte ist, dass die Hündin Herr der Ringe wirklich liebt. Das ist Jessica sehr recht, denn sie braucht Entspannung.

Gewichtszuwachs in der 1LW	Durchschnittlicher Tagestemperatur [C/d]
17.3	19.3
20.8	18.8
20.0	16.8
24.0	22.4
18.6	16.2
19.9	17.8
21.2	19.0
18.6	19.9
21.5	18.3
20.5	22.7
18.2	16.3

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

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

58. Aufgabe

(10 Punkte)

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



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

Schlachtgewicht im Zielbereich	Ökologisch	Schlachtgewicht im Zielbereich	Ökologisch
ja	ja	nein	nein
ja	ja	nein	nein
nein	nein	nein	ja
ja	nein	nein	nein
ja	ja	nein	nein
ja	ja	ja	ja
ja	ja	ja	ja
nein	ja	nein	ja
nein	ja	ja	ja
nein	ja	nein	nein
ja	nein	nein	ja
ja	ja	ja	ja
nein	nein	nein	ja
ja	nein	ja	ja
nein	nein	ja	ja
nein	nein	nein	ja
nein	nein	ja	nein

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

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

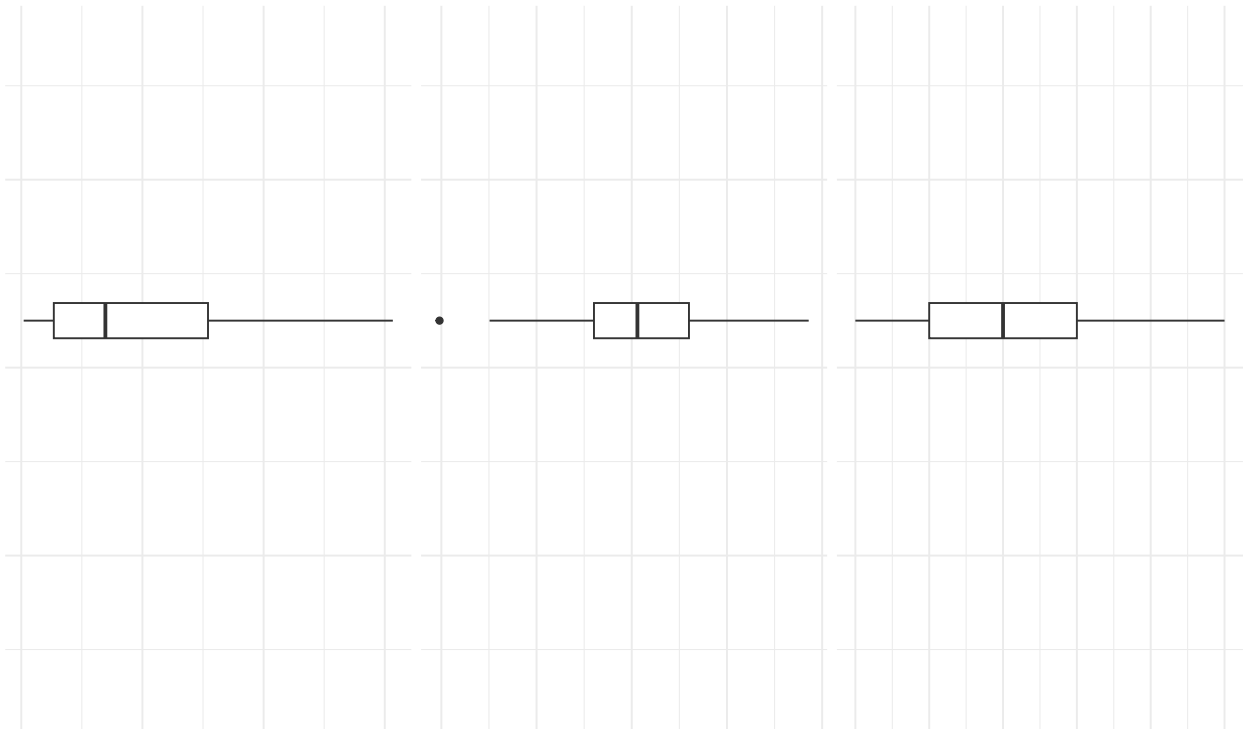
59. Aufgabe

(10 Punkte)

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



Visualisierung von Verteilungen 'Was soll das denn jetzt schon wieder sein? Drei Boxplot, die auf der Seite liegen?', entfährt es Nilufar und schaut dabei Jonas an. 'Keine Ahnung. Es ist bestimmt wieder so ein Lernziel mit der Verteilung und so.', meint Jonas sichtlich genervt und mampft noch ein paar Snickers. 'Du weißt doch wie es heißt, *Frei ist, wer missfallen kann*.¹', merkt Nilufar 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. Jonas und die Erwartung machen die Sache nicht einfacher.



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

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

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

60. Aufgabe

(10 Punkte)

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



Visualisierung der Normalverteilung Jessica und der Mangel machen die Sache mit dem Studium nicht einfacher. Immerhin ist noch Alex zur Hilfe mit dabei. Alex hat Schokobons mitgebracht und David Bowie aufgedreht. Das ist immerhin eine Ablenkung. Nicht so gut wie Warhammer, 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 Jessica. Durch das Mampfen von Alex versteht sie kein Wort der Antwort. Alex lächelt.

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

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

61. Aufgabe

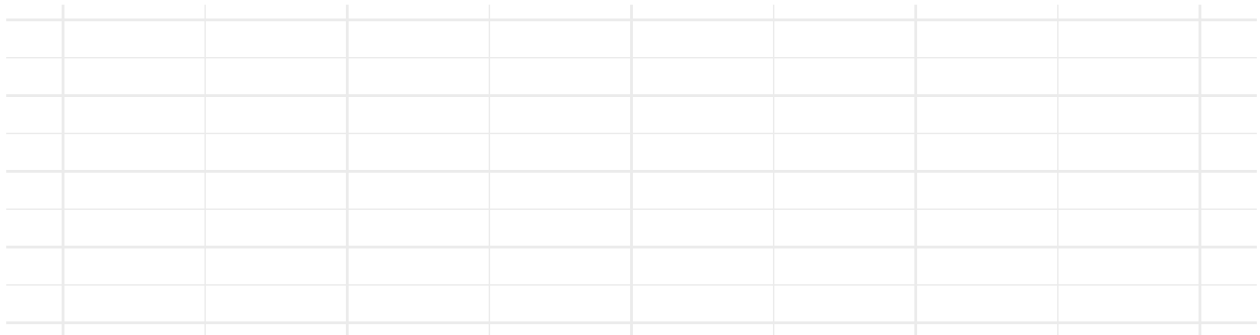
(10 Punkte)

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

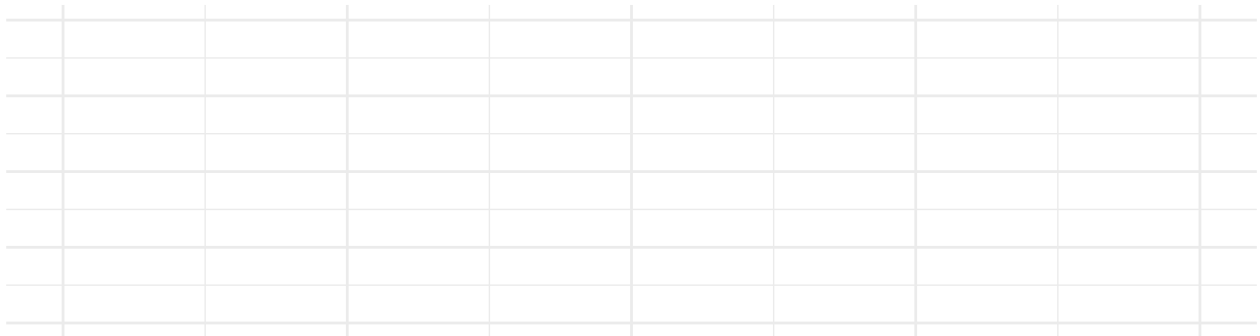


Visualisierung der Normalverteilung und der Poissonverteilung 'Was sollen wir hier dann noch zeichnen?!', entfährt es Paula und schaut dabei Nilufar an. 'Wir sollen eine Normalverteilung mit einem Mittelwert von $\bar{y}_1 = 3$ und einer Standardabweichung von $s_1 = 0.25$ zeichnen. Sowie eine weitere Normalverteilung mit einem Mittelwert von $\bar{y}_2 = 2$ und einer Standardabweichung von $s_2 = 0.25$. Keine Ahnung wie das geht. Darunter sollen dann noch eine Poissonverteilung mit einem Mittelwert von $\lambda_1 = 2$ sowie einer weiteren Poissonverteilung mit einem Mittelwert von $\lambda_2 = 15$ gezeichnet werden.', meint Nilufar sichtlich genervt und mampft noch ein paar Takis Blue Heat. Im Hintergrund spielt leise Deichkind. 'Wirre Geschichte...', merkt Paula nickend an. Die beiden schauen angestrengt auf die leeren Flächen für die Abbildungen. Nilufar und der Perfektionismus machen die Suche nach der Lösung nicht einfacher.

Zwei Normalverteilungen



Zwei Poissonverteilungen



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

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

Teil II.

Statistisches Testen & statistische Testtheorie

62. Aufgabe

(9 Punkte)



Grundgesamtheit und experimentelle Stichprobe foo

Leider kennen sich Nilufar und Tina mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion.

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)



Das Nullritual - Die statistische Testtheorie foo

Leider kennen sich Jonas und Tina mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion.

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)**

H_0 wahr Testentscheidung Richtige Entscheidung α -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!

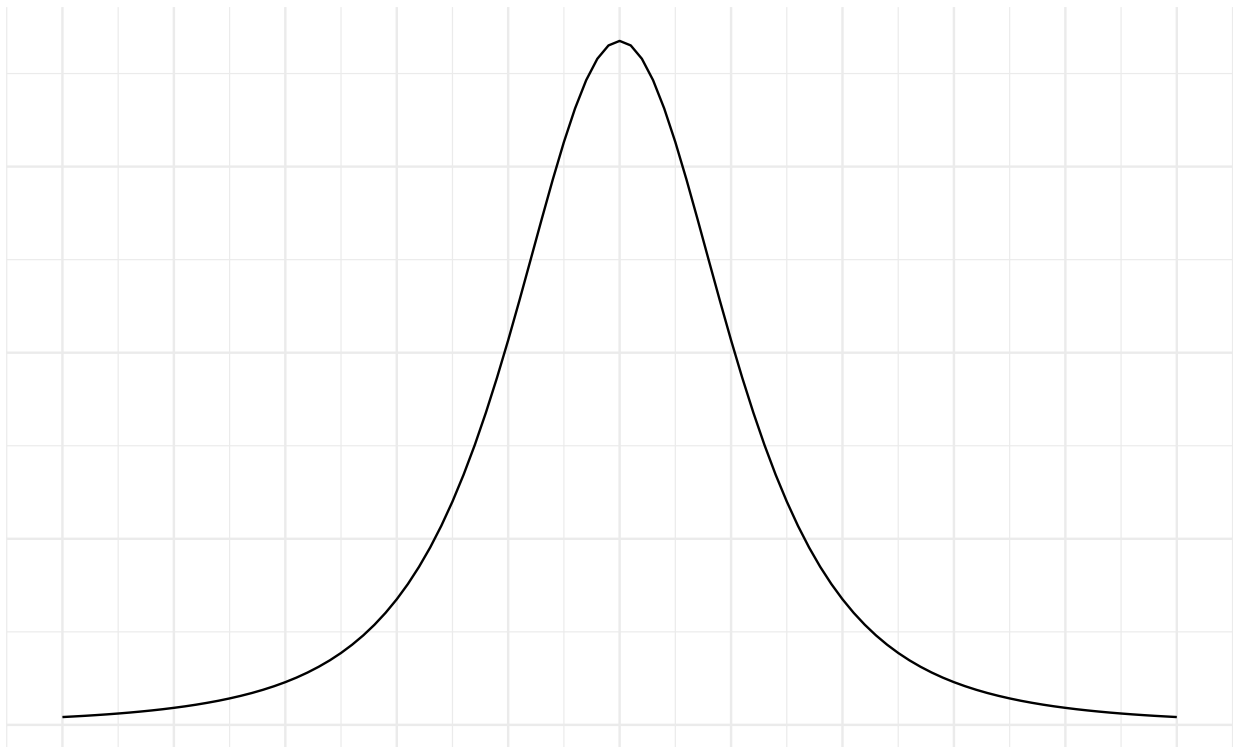


Visualisierung der Teststatistik T_D und dem p-Wert

Leider kennen sich Mark und Nilufar mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion.

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 „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)



Visualisierung des 95% Konfidenzintervalls foo

Leider kennen sich Steffen und Jessica mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion.

Sie rechnen einen t-Test für Gruppenvergleiche der Mittelwerte. Sie schätzen den Unterschied zwischen dem mittleren Trockengewicht nach Düngergabe 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 signifikantes, nicht relevantes 95% Konfidenzintervall
 - (b) Ein 95% Konfidenzintervall mit höherer Varianz s_p in der Stichprobe als der Rest der 95% Konfidenzintervalle
 - (c) Ein signifikantes, relevantes 95% Konfidenzintervall
 - (d) Ein nicht signifikantes, nicht relevantes 95% Konfidenzintervall
 - (e) Ein signifikantes, relevantes 99% 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!



Zusammenhang zwischen dem Effekt, der Streuung sowie der Fallzahl foo

Leider kennen sich Alex und Tina mit der zweifaktoriellen ANOVA überhaupt nicht aus. Geschweige denn mit der Interpretation einer Interaktion.

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 99%-igen Konfidenzintervall! Begründen Sie Ihre Visualisierung anhand der Formel des Konfidenzintervalls des t-Tests mathematisch! **(3 Punkte)**

Teil III.

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

67. Aufgabe

(9 Punkte)

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



Jessica ist in der Uckermark für einen Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) mit Hühnern. Allein diese Tatsache ist für sie eine Erzählung wert. Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen. Für ihre Abschlussarbeit musste sie ein Kreuzungsexperiment mit Hühnern durchführen und das sollte laut ihrem Betreuer an diesem Ort besonders gut gelingen, da man hier gut neue technische Anlagen und Behandlungen fernab der Bevölkerung testen könne. Zeugen gibt es hier jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn ihre Behandlung Ernährungszusatz (*ctrl* und *fedX*) und der Messwert Protein/Fettrate [%/kg] sollen mit einem t-Test ausgewertet werden. Immerhin weiß sie, dass ihr Messwert einer Normalverteilung folgt. Hm..., was entspannendes wäre gut. Um Rad zu fahren geht Jessica dann später nochmal raus. Echte Entspannung.

treatment	weight
dose	23.9
ctrl	18.4
dose	18.4
dose	17.3
ctrl	13.3
ctrl	17.8

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

1. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
2. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ 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 Jessica über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

68. Aufgabe

(12 Punkte)

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



Der Teuteburgerwald, unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer von Tina, die mit ihrer 1 Frau 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 Tina. Für ihren Projektbericht ist Tina ins Nichts gezogen. Wenn die Wut nicht wäre, ja dann wäre wohl vieles möglich für Tina! Aber so.. Was macht sie nun? Tina hat ein Kreuzungsexperiment mit Fleischrindern durchgeführt. Die Behandlung Flüssignahrung (*ctrl* und *fIOW*) wurde an Fleischrindern getestet. Gemessen hat sie dann als ein normalverteiltes Outcome (Y) Gewichtszuwachs in der 1LW. Jetzt soll sie ihrer Betreuerin nach testen, ob die Behandlung Flüssignahrung (*ctrl* und *fIOW*) ein signifikantes Ergebnis liefert. Hm..., was spannendes wäre gut. 'Hm...', Katjes und TocoTronic. Das ist und bleibt die beste Kombination zum Nachdenken für Tina.

Flüssignahrung	Gewichtszuwachs
fIOW	53.8
ctrl	50.6
fIOW	58.2
ctrl	40.3
fIOW	36.0
fIOW	48.2
fIOW	47.8
fIOW	61.3
fIOW	26.9
fIOW	60.7
ctrl	41.6
fIOW	66.6
ctrl	31.6
fIOW	35.2
ctrl	33.6
ctrl	38.7
ctrl	44.4

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 2.68$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *keinen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann die Teststatistik T_D ? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Tina über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

69. Aufgabe

(12 Punkte)

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



Alex ist im Wendland für einen Versuch mit Hühnern. Allein diese Tatsache ist für ihn eine Erzählung wert. Eine echte Herausforderung für ihn war schon immer die Gefälligkeit gewesen. Ein leidiges Lied. Für seinen Projektbericht musste er ein Stallexperiment mit Hühnern durchführen und das sollte laut seiner Betreuerin an diesem Nichtort besonders gut gelingen. Ablenkung gibt es jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn seine Behandlung Lüftungssystem (*keins* und *vorhanden*) und der Messwert Fettgehalt [%/kg] sollen mit einem t-Test ausgewertet werden. Immerhin weiß er, dass sein Messwert einer Normalverteilung folgt. Hm..., was spannendes wäre gut. Alex will später nochmal raus um zu Laufen. Druck ablassen, dass muss er auch.

Lüftungssystem	Fettgehalt
keins	45.5
keins	23.6
vorhanden	31.0
vorhanden	15.7
vorhanden	66.2
vorhanden	27.0
keins	37.3
keins	36.5
vorhanden	67.5
vorhanden	42.7
keins	28.5
vorhanden	40.7
vorhanden	29.4
vorhanden	51.0
keins	36.7
keins	50.9
keins	57.4
keins	49.7
keins	42.5
keins	44.4

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 2.68$ 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 Kartoffeln. Das haben Jonas und Jessica gemeinsam in einer Abschlussarbeit gemacht! Worum ging es aber konkret? Beide haben als einen normalverteilten Endpunkt (Y) Chlorophyllgehalt (SPAD-502Plus) [SPAD] von Kartoffeln bestimmt. Die Daten haben beide zusammen in einem Freilandversuch erhoben. In dem Experiment ging es um eine vorher/nachher Untersuchung an den gleichen Kartoffeln. Als Behandlung wurde Jäten (*Aussaat* und *Ernte*) eingesetzt. Nach der Meinung des Betreuers muss hier ein gepaarter t-Test gerechnet werden. Jonas schaut nachdenklich zu Jessica. Wenn die Erschöpfung nicht wäre, ja dann wäre wohl vieles möglich für Jonas! Aber so... Steffen denkt derweil angestrengt an Schokobons. Im Hintergrund wummert Iron Maiden.

ID	treatment	freshmatter
2	Aussaat	39.9
11	Aussaat	47.6
3	Aussaat	46.0
5	Ernte	20.5
6	Ernte	32.6
7	Aussaat	30.1
1	Ernte	32.2
7	Ernte	25.7
1	Aussaat	30.0
4	Aussaat	28.8
3	Ernte	26.2
6	Aussaat	16.2
2	Ernte	28.3
5	Aussaat	24.5
9	Aussaat	39.7
10	Aussaat	43.9
8	Aussaat	40.5
4	Ernte	31.7

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


1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines gepaarten t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.84$ 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 Jonas über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

71. Aufgabe

(10 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 Yuki sichtlich stolz. Ein paar Mal hat sie schon die Faulheit gehindert weiterzumachen. 'Nach Meinung der Betreuerin soll es aber nur erstmal ein t-Test sein. Und die Ausgabe ist schon wirr genug.', merkt Mark an. Steffen und Mark sind bei Yuki um sich in  helfen zu lassen. Im Hintergrund wummert London Grammar. Mark streichelt zur Beruhigung das Minischwein von Yuki. Die beiden waren 2 Monate im Oldenburger Land um einen Versuch mit Kartoffeln in einem Gewächshausexperiment durchzuführen. Ziel war es das Outcome (Y) Trockengewicht [kg/ha] zu bestimmen. Yuki überlegt, ob er die beiden nicht noch auf den Film *Matrix* einlädt oder dann doch lieber raus geht um zu Boldern? Vielleicht will ja Mark mit. Besser als der Film.

```
##
## Two Sample t-test
##
## data: Trockengewicht by Lüftungssystemen
## t = -4.7094, df = 17, p-value = 0.0002021
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -23.745533 -9.052245
## sample estimates:
## mean in group ctrl mean in group tornado
## 33.39000 49.78889
```

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




1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.11|$ 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!



‘Wir sind uns relativ sicher, dass unser Messwert Gewichtszuwachs in der 1LW ist!’, ruft Mark wild gestikulierend. Mark wäre mehr präsent, wenn es die Unsicherheit nicht gäbe. Als würde sowas die Ausgabe von  interessieren. Mark und Tina sind in einem Café mit Steffen um sich Hilfe von ihm in  zu holen. Während Steffen Kirschstreuselkuchen und Oreos mampft, versuchen die Mark und Tina ihren Versuch im Oldenburger Land mit Puten in einem Kreuzungsexperiment zu erklären. Steffen hofft insgeheim, dass die  Ausgabe des t-Tests ihm mehr Informationen liefert. Eigentlich würde er dann doch lieber raus um zu Ringen vielleicht mit Tina?

```
##  
## Two Sample t-test  
##  
## data: Gewichtszuwachs by Lüftungssystem  
## t = 3.5946, df = 15, p-value = 0.002655  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## 3.050488 11.938401  
## sample estimates:  
## mean in group keins mean in group vorhanden  
## 34.15000 26.65556
```

Helfen Sie Steffen bei der Interpretation des t-Tests! Sonst geht es auch für Mark 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 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!



R ist schon ein tolles Programm, wenn man mit dem Ding umgehen kann. Super umgehen kann damit Steffen. Eine echte Herausforderung für Steffen ist die Romantik, aber das ist noch eine andere Sache. Deshalb sind aber Tina und Alex nicht bei ihm. Sondern um sich bei einem gemeinsamen Projekt helfen zu lassen. Im Hintergrund wummert Taylor Swift. Beide arbeiten gemeinsam an einer Abschlussarbeit. In dem zu beschreibenden Versuch geht es im Teuteburgerwald um einem Leistungssteigerungsversuch mit Hühnern. Dabei ging darum herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Ernährungszusatz (*ctrl* und *fedX*) und dem Messwert Schlachtgewicht [kg] gibt. Da der Messwert Schlachtgewicht [kg] normalverteilt ist kann ein t-Test gerechnet werden. Tina möchte dann später noch mehr über Steffens Hobby Klemmbausteine erfahren.

```
##
## Two Sample t-test
##
## data: Schlachtgewicht by Ernährungszusatz
## t = -0.99065, df = 16, p-value = 0.3366
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -9.880514 3.587008
## sample estimates:
## mean in group ctrl mean in group fedX
##          31.38182          34.52857
```

Helfen Sie Steffen bei der Interpretation des t-Tests! Sonst geht es auch für Tina und Alex 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!



Tina und Jessica haben sich dazu entschieden zusammenzuarbeiten. Das sollte alles etwas einfacher machen. Jeder hat zwar ein getrenntes Themenfeld aber den Hauptversuch machen beide gemeinsam. Das hat sich schonmal als gut Idee soweit herausgestellt. Eine echte Herausforderung für sie war schon immer die Wut gewesen. Ein leidiges Lied. In einem Projektbericht sollen beide herausfinden, ob es einen Zusammenhang zwischen Bestandsdichte (*hoch* und *niedrig*) und Fettgehalt [%/kg] gibt. Die Besonderheit ist hierbei, dass die Messungen an der gleichen Beobachtung stattfinden. Beide messen also zweimal an den gleichen Zandern. Hier muss dann wohl auf einen normalverteilten Messwert (Y) ein gepaarter t-Test gerechnet werden. Leider kennen sich beide nicht sehr gut in  aus. Aber wenigstens haben beide eine Menge an Katjes.

```
##
## Paired t-test
##
## data: Fettgehalt by Bestandsdichte
## t = 5.4148, df = 8, p-value = 0.0006347
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
##  9.938863 24.683359
## sample estimates:
## mean difference
##      17.31111
```

Jetzt brauchen Tina 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 Jonas und Alex schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten visualisieren damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse zu erwarten sind. Die beiden waren im Oldenburger Land um ein Kreuzungsexperiment mit Puten durchzuführen. Dabei haben Jonas und Alex den Messwert Protein/Fettrate [%/kg] unter der Behandlung Flüssignahrung (*ctrl*, *superIn* und *fIOW*) ermittelt. Kennengelernt haben sich die beiden auf einem Konzert von Abba. Später wird noch Alien geguckt. Alex befürwortet das!

Flüssignahrung	Protein/Fettrate
fIOW	46
ctrl	42
fIOW	44
ctrl	42
ctrl	39
superIn	36
fIOW	45
superIn	35
fIOW	44
fIOW	43
fIOW	45
superIn	37
ctrl	42
superIn	35
ctrl	39
superIn	36
ctrl	40
ctrl	39
superIn	36

Leider kennen sich Jonas 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.', Paula schaut Yuki fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Da hilft das Minischwein von Yuki auch nur bedingt. Yuki tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren im Oldenburger Land um ein Freilandversuch mit Erbsen durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Bewässerungstypen (*ctrl*, *low*, *mid* und *high*) und dem Messwert Proteingehalt [g/kg] gibt. Nachher wollen sich beide noch mit dem Hobby Orchideen von Yuki beschäftigen. Kennt Paula noch nicht, klingt aber interessant.

Leider kennen sich Paula und Yuki mit Berechnung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe, das Minischwein reicht als Hilfe nicht aus!

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Bewässerungstypen	3	3845.31			
error	20	1348.52			
Total	23				

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.1$ ab. Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? **(2 Punkte)**

77. Aufgabe

(12 Punkte)

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



Die einfaktoriellen ANOVA und der Student t-Test 'Als erstes bauen wir uns aus unsere Daten die ANOVA Tabelle dann sehen wir schon, ob unser Gruppenvergleich in der ANOVA signifikant ist.', Steffen schaut Nilufar fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Nilufar schmeißt sich noch ein paar Takis Blue Heat in den Rachen. Beide tun sich sehr schwer mit der einfaktoriellen ANOVA. Nun möchte erstmal ihre Betreuung der Arbeit eine ANOVA Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Beide waren im Oldenburger Land um ein Gewächshausexperiment mit Erdbeeren durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Substrattypen (*torf*, *40p60n*, *30p20n* und *70p30n*) und dem Messwert Trockengewicht [kg/ha] gibt. Später wollen die beiden dann noch raus um zu Kicken.

Leider kennen sich Steffen und Nilufar 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)
Substrattypen	3	581.95			
Error	23	722.05			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.03$ 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)

Substrattypen	Fallzahl (n)	Mittelwert	Standardabweichung
torf	7	15.43	7.25
40p60n	6	5.00	3.95
30p20n	6	3.83	6.08
70p30n	8	5.75	4.53

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

78. Aufgabe

(9 Punkte)

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



Die einfaktorielle ANOVA in R Tina schaut entnervt auf und klappt den Laptop zu. Tina dreht Tocotron auf, so dass sich die Nachbarn beschweren werden. Nun möchte ihr Betreuer ihrem Projektbericht erstmal eine ANOVA sehen und *dann* die Ergebnisse präsentiert bekommen bevor es überhaupt mit der Abschlussarbeit weitergeht. Dabei war sie extra im Oldenburger Land um ein Gewächshausexperiment mit Erdbeeren durchzuführen. Und dort was es wirklich nicht schön geschweige denn spannend wie bei ihren Kommilitonen, die in Almería waren. Hätte sie es vorher gewusst, dann hätte sie die Abschlussarbeit bei wem anders geschrieben. Aber gut, jetzt als die ANOVA in R.

```
## Analysis of Variance Table
##
## Response: Chlorophyllgehalt
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Lüftungssysteme  3 326.10   108.70   4.9029 0.009292
## Residuals       22 487.75    22.17
```

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

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

79. Aufgabe

(12 Punkte)

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



Ergebnistabelle der zweifaktoriellen ANOVA Paula steht im Oldenburger Land. Und das ist noch langweiliger als es sich anhört. Wäre es nur so spannend wie bei ihren Kommilitonen, die in Almería waren. Ödnis wohin man nur blickt. Oder eben Hühnern. Die Ratte durchbohrt sie mit leeren Blick. 'Woher zum Teufel!', entfährt es ihr. Aber da ist es schon weg. Ja, darum geht es in ihrer Hausarbeit. Und wäre das nicht noch alles genug, ist ihr Experiment auch noch als Stallexperiment komplex geraten. Es wurde der Messwert Protein/Fettrate [%/kg] mit dem Behandlung Elterlinie (*ctrl*, *Standard*, *Yray* und *Xray*) sowie der Behandlung Ernährungszusatz (*ctrl* und *getIt*) untersucht. 'Hmpf', denkt Paula und ruft 'Und jetzt!?' in die Leere. Und eigentlich wollte Paula doch noch ihrem Hobby nachgehen! Harry Potter. Ein wunderbares Hobby um sich drin zu verlieren und Abstand zu bekommen. Paula denkt gerne über Harry Potter nach.

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

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Elterlinie	3	156.71			
Ernährungszusatz	1	91.94			
Elterlinie:Ernährungszusatz	3	768.05			
Error	18	214.44			

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

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

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

80. Aufgabe

(10 Punkte)

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



Die zweifaktorielle ANOVA in R In einen Versuch in einer Klimakammer wurden Kartoffeln mit der Behandlung Lichtstufen (*none*, 200lm, 400lm und 600lm) sowie der Behandlung Substrattypen (*torf*, 70p30n) untersucht. Es wurde als Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] bestimmt. Jetzt starrt Steffen mit auf die R Ausgabe einer zweifaktoriellen ANOVA. Leider starrt sein Betreuer in der gleichen Art Steffen zurück an. Das wird ein langer Nachmittag, denkt er sich und kreuselt seinen Mund. 'Und was machen wir jetzt?' entfährt es ihm überrascht entnervt. Immerhin war geht es ja um seine Abschlussarbeit. Steffen hätte doch nichts mit Kartoffeln machen sollen. Kartoffeln – was soll das auch bedeutendes sein? Eigentlich wollte Steffen nachher noch einen Film schauen. Wenn Harry Potter läuft, dann ist die Schlange nicht mehr da. Aber jetzt braucht er mal Entspannung!

```
## Analysis of Variance Table
##
## Response: Chlorophyllgehalt
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Lichtstufen    2  208.099   104.049   10.3096 0.001038
## Substrattypen   1   96.053    96.053    9.5172 0.006385
## Lichtstufen:Substrattypen  2  119.169    59.585    5.9038 0.010678
## Residuals     18  181.665    10.093
```

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

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

81. Aufgabe

(12 Punkte)

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



Zusammenhang zwischen der ANOVA und dem t-Test Tina schaut konzentriert auf die Formeln der ANOVA und des t-Tests. In ihrem Experiment wurde als Messwert Proteingehalt [g/kg] bestimmt. Eine echte Herausforderung für sie war schon immer die Wut gewesen. Ein leidiges Lied. Dann wäre es nicht noch komplizierter. In ein Gewächshausexperiment wurden Maiss mit der Behandlung Lüftungssysteme (*ctrl*, *storm*, *thunder* und *tornado*) sowie der Behandlung Genotypen (*AA*, und *BB*) untersucht. Beide Verfahren müssen etwas miteinander zu tun haben und ihr Betreuer möchte das jetzt auch noch verstehen. Im Hintergrund läuft leise Indiana Jones auf ihrem Second Screen. Immerhin hat sie die beiden Formeln vorliegen.

Gegebene Formeln

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

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


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

82. Aufgabe


(11 Punkte)

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



Interaktion in der zweifaktoriellen ANOVA In ein Kreuzungsexperiment wurden Fleischrindern mit der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) sowie der Behandlung Flüssignahrung (*ctrl* und *fLOW*) untersucht. Es wurde als Messwert Gewichtszuwachs in der 1LW bestimmt. Jetzt starrt Jonas mit auf die  Ausgabe einer zweifaktoriellen ANOVA. Leider starrt sein Betreuer in der gleichen Art Jonas zurück an. Es liegt anscheinend eine signifikante Interaktion vor? 'Das wird ein langer Nachmittag.', denkt er sich und kreuselt seinen Mund. 'Und was machen wir jetzt?' entfährt es ihm überrascht entnervt. Immerhin war geht es ja um der Projektbericht. Jonas hätte doch nichts mit Fleischrindern machen sollen. Fleischrindern – was soll das auch bedeutendes sein? Eigentlich wollte Jonas nachher noch einen Film schauen. Irgendwie komisch, wenn er Mission Impossible anmacht, dann ist das Meerschweinchen eigentlich sofort vor dem Bildschirm und starrt hinein.

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


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

83. Aufgabe

(11 Punkte)

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



Zusammenhang zwischen der ANOVA und dem Post-hoc-Test In ein Gewächshausexperiment wurden die Daten D von Spargel mit der Behandlung Bewässerungstypen (*ctrl*, *low*, *mid* und *high*) untersucht. Es wurde als Messwert Trockengewicht [kg/ha] bestimmt. Jetzt startet Jonas mit auf die  Ausgabe einer einfaktoriellen ANOVA. Anscheinend gibt es ein Problem mit der Annahme der Normalverteilung und der Varianzhomogenität der ANOVA in den Daten. 'Wir haben jetzt bei der ANOVA einen p-Wert mit 0.058 raus sowie eine F-Statistik F_D mit 1.78 berechnet. Nach den Boxplots müsste sich eigentlich ein Unterschied zwischen *high* und *low* ergeben. Der Unterschied ist in {emmeans} auch signifikant mit einem p-Wert von 0.049. Wie kann das sein?', meint Jonas ungläubig. Leider startet sein Betreuer in der gleichen Art Jonas zurück an. Jonas hat schon genug Probleme. Wenn die Erschöpfung nicht wäre, dann wäre es einfacher. Das wird ein langer Nachmittag, denkt er sich und kreuselt seinen Mund. 'Und was machen wir jetzt?' entfährt es ihm überrascht entnervt. Immerhin war geht es ja um der Projektbericht. Jonas hätte doch nichts mit Spargel machen sollen. Spargel – was soll das auch bedeutendes sein? Hoffentlich kommt er noch zum Konzert von Iron Maiden. Nervös krault er das Meerschweinchen.

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 Jonas und sein Betreuer mit der Interpretation einer ANOVA überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe und die Zeit wird knapp.

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

Teil V.

Multiple Gruppenvergleiche

84. Aufgabe

(12 Punkte)

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



In ein Stallexperiment mit Hühnern wurde die Behandlung Ernährungszusatz (*ctrl*, *fedX*, *proteinX* und *getIt*) gegen die Ergebnisse einer früheren Studie von Meyer et al. (2021) verglichen. Im Rahmen des Experiments haben Jessica und Yuki verschiedene Student t-Tests für den Mittelwertsvergleich für den Messwert Schlachtgewicht [kg] 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?
0.0300		
0.0020		
0.0012		
0.7600		

Leider kennen sich Jessica 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!



Tina hatte in ihrem Projektbericht ein Feldexperiment durchgeführt. Soweit so gut. Dabei hat sie sich mit Erdbeeren beschäftigt. Angeblich der neueste heiße Kram... aber das ist wiederum was anderes. So richtig mitgenommen hat Tina das Thema dann doch nicht. Hat sie sich doch mit Genotypen (00, AA, AB und BB) und Chlorophyllgehalt (SPAD-502Plus) [SPAD] schon eine Menge an Daten angeschaut. Nach ihrem Betreuer soll sie nun ein CLD bestimmen. Weder weiß sie was ein CLD ist, noch war ihr erster Gedanke mit Köln und die LGBTQ Community richtig...

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

Leider kennen sich Tina 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 hatte in die Projektbericht einen Leistungssteigerungsversuch durchgeführt. Soweit so gut. Dabei hat sie sich mit Puten beschäftigt. Angeblich der neueste heiße Kram... aber das ist wiederum was anderes. So richtig mitgenommen hat Nilufar das Thema dann doch nicht. Hat sie sich doch mit Genotypen (00, AA, AB und BB) und Gewichtszuwachs in der 1LW schon eine Menge an Daten angeschaut. Nach ihre Betreuerin soll sie nun ein CLD bestimmen. Weder weiß sie was ein CLD ist, noch war ihr erster Gedanke mit Köln und die LGBTQ Community richtig... Als erstes solle sie die Gruppen nach absteigender Effektstärke sortieren. Was immer das jetzt bringen soll.

Genotypen	Fallzahl (n)	Mittelwert	Standardabweichung
00	7	9.97	2.72
AA	8	9.25	3.05
AB	9	16.73	1.77
BB	7	5.41	3.50

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!



'Okay, dann nochmal für mich. Ich habe jetzt alles in SPSS gemacht, aber das Wichtigste, was gemacht werden soll, nämlich das CLD, das kann ich nicht in SPSS machen?', Yuki muss sich echt beherrschen. Immerhin betreut ihr Betreuer ja erst nicht seit gestern Abschlussarbeiten und wusste ja was gemacht werden soll! Yuki hatte sich zwei Variablen mit Substrattypen (*kompost*, *torf*, *40p60n* und *70p30n*) und Trockengewicht [kg/ha] in ein Feldexperiment mit Brokkoli angeschaut. Jetzt möchte sie eigentlich fertig werden und nicht nochmal alles neu in **R** und `{emmeans}` machen. Deshalb soll jetzt das CLD per Hand aus der Matrix der *p*-Wert abgeleitet werden. 'Ich glaube ich wechsel nochmal das Thema...', denkt Yuki, verwirft dann aber den Gedanken.

	kompost	torf	40p60n	70p30n
kompost	1.0000000	0.4351585	0.4019395	0.0747364
torf	0.4351585	1.0000000	0.1187149	0.2649014
40p60n	0.4019395	0.1187149	1.0000000	0.0160085
70p30n	0.0747364	0.2649014	0.0160085	1.0000000

Leider kennen sich Yuki 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 der Matrix der *p*-Werte ergebenden Barplots! **(2 Punkte)**
4. Ergänzen Sie das *Compact letter display (CLD)*! Begründen Sie Ihre Antwort! **(4 Punkte)**
5. Interpretieren Sie das *Compact letter display (CLD)* für Yuki und Yuki! **(2 Punkte)**

Teil VI.

Der Chi-Quadrat-Test & Der diagnostische Test

88. Aufgabe

(12 Punkte)

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



Den Chi-Quadrat-Test berechnen Nilufar hat sich ein Herz gefasst und war für ihrem Projektbericht in die Niederlande gegangen. Das war eine super Zeit in der sie viel gelernt hat. Klar gab es auch die ein oder andere Besonderheit, aber das gehört hier eher nicht hin. Dann noch schnell Deichkind auf das Ohr und los gehts. Nilufar ist schon eine ganze Zeit im Büro, da ihre Betreuerin möchte, dass sie jetzt auf ihren Daten mit $n = 112$ Beobachtungen von Fleischrindern einen χ^2 -Test rechnet. Das ginge, da sie als Behandlung *Außenklimakontakt [ja/nein]* bestimmt und zum anderen die Variable *Fettgehalt erreicht [ja/nein]* ermittelt hat. Wie genau, das ist jetzt eine andere Frage. Eigentlich wollte Nilufar nachher noch einen Film schauen. Wenn Star Trek läuft, dann ist das Huhn nicht mehr da. Aber jetzt braucht sie mal Entspannung!

	24	21	
	23	44	

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

89. Aufgabe

(10 Punkte)

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



Der Chi-Quadrat-Test konzeptionell verstehen 'Der χ^2 -Test auf einer 2x2-Kreuztabelle berechnet.', liest Jonas in seiner Mitschrift. So richtig helfen tut ihm das jetzt eherlichweise dann doch nicht. Dann noch schnell Snickers zur Stärkung und los gehts. Jonas hatte sich in ein Kreuzungsexperiment $n = 157$ Beobachtungen von Hühnern angeschaut. Dabei hat er als Behandlung *Automatische Fütterung* [ja/nein] bestimmt und zum anderen die Variable *Protein/Fettrate im Zielbereich* [ja/nein] ermittelt. Am Ende möchte dann seine Betreuerin gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Am Ende des Tages möchte er dann noch sein Hobby Stricken genießen. Das muss auch mal sein!

			87
			70
	79	78	157

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

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

90. Aufgabe

(10 Punkte)

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



Der Chi-Quadrat-Test in R Am Ende hätte Nilufar dann doch einen normalverteilten Endpunkt in ihrem Projektbericht nehmen sollen. Vor ihr liegen jetzt die Daten von zwei Variablen als Kategorien oder wie es in R so schön heißt, als Faktoren. Aber immerhin, hofft sie das was bei den Daten rausgekommen ist. Dann noch schnell Star Trek starten und los gehts mit der Kraft von Takis Blue Heat. Gezählt hat Nilufar einiges mit $n = 157$ Beobachtungen von Brokkoli. Zum einen hat sie als Behandlung *KI-gesteuert* [*ja/nein*] bestimmt und zum anderen die Variable *Frischegewicht über Zielwert* [*ja/nein*] ermittelt. Nun möchte ihr Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Nach ihrem Experiment erhielt sie folgende 2x2 Kreuztabelle aus ihren erhobenen Daten.

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

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: Frischegewicht über Zielwert
## p-value = 0.1811
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.07075797 1.67575062
## sample estimates:
## odds ratio
##  0.3611682
```

Leider kennt sich 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)**

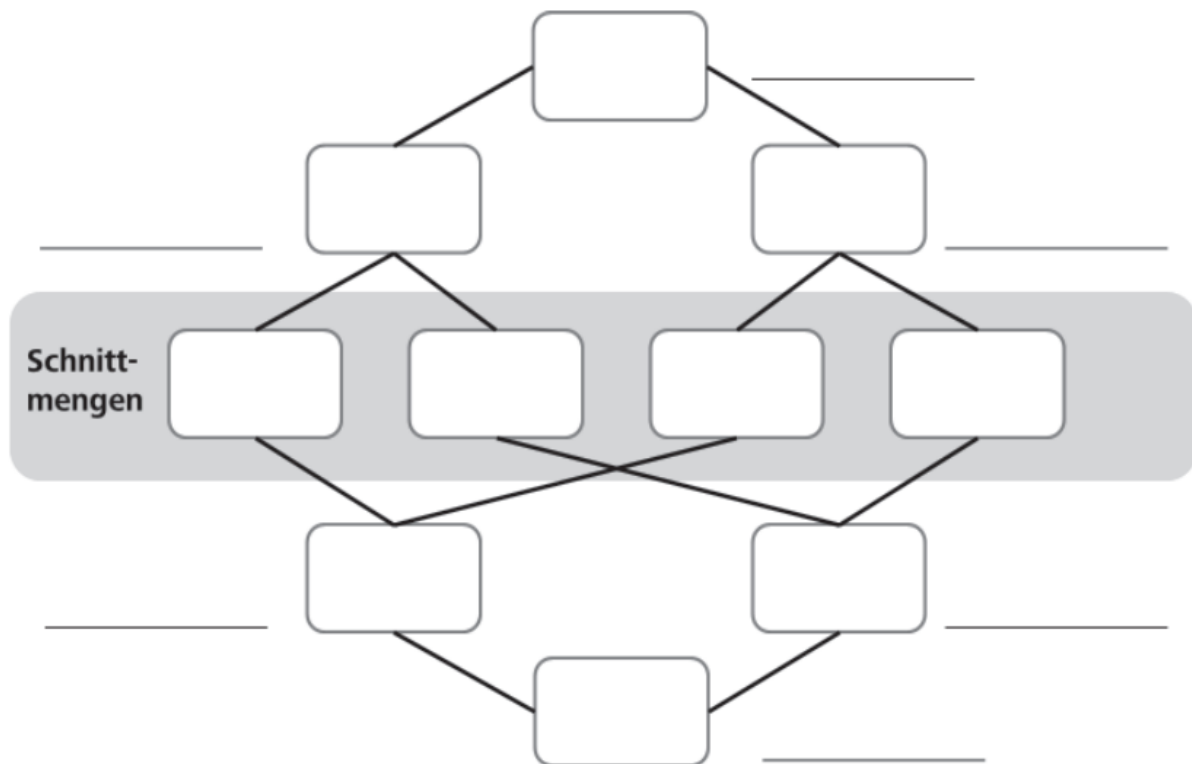
91. Aufgabe

(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 Maiss wird mit 2% angenommen. In 85% der Fälle ist ein Test positiv, wenn das Pflanze erkrankt ist. In 8.0% der Fälle ist ein Test positiv, wenn das Pflanze nicht erkrankt ist und somit gesund ist. Wir führen einen Test auf Gelbmosaikvirus an 2000 Maiss mit einem diagnostischen Test durch.' Jessica klappt die Kinnlade runter. In der Stille dudelt Tocotronic. Tina schaut kompetent und schmeißt sich mit offenen Mund Schokobons an den Kopf vorbei.



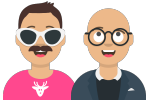
Leider kennen sich Tina und Jessica mit dem diagnostischen Testen überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

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

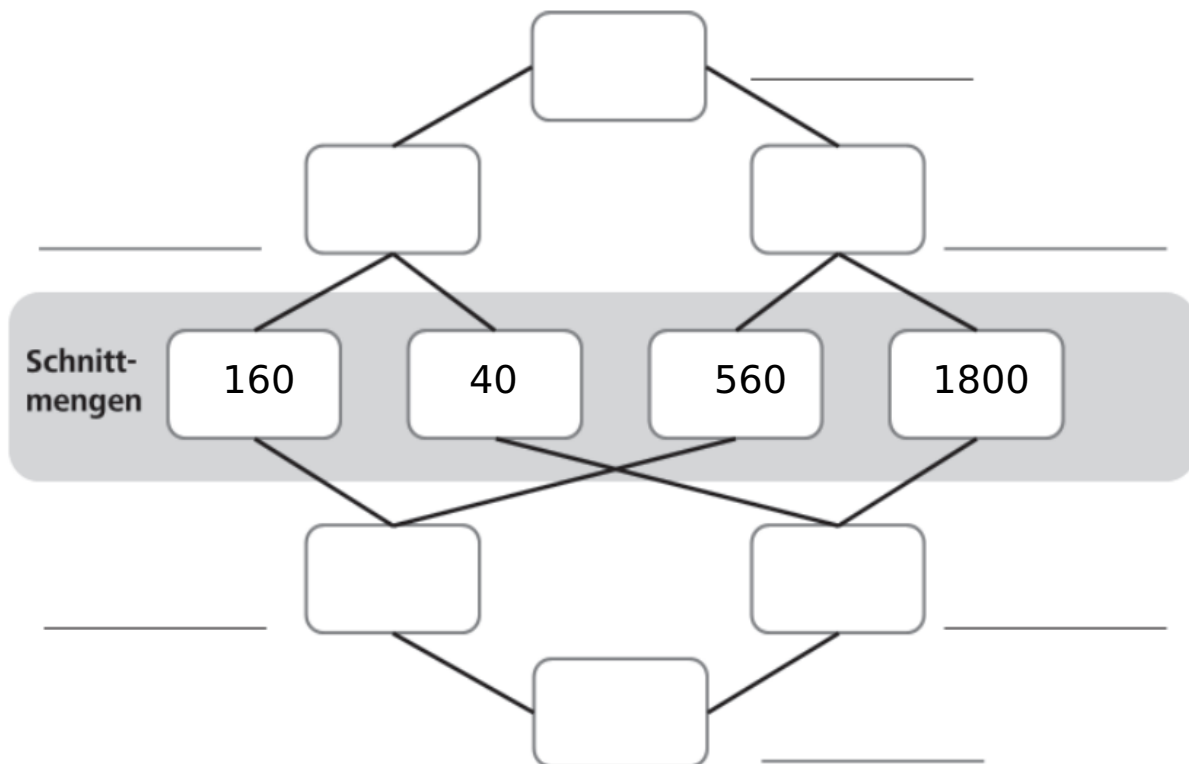
92. Aufgabe

(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 Alex. 'Hm... ich glaube es handelt sich um einen Doppelbaum, den wir beim diagnostischen Testen brauchen.', meint Yuki und dreht Alien 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 die Gefälligkeit von Alex macht es heute auch nicht mehr einfacher. 'Es geht um Escherichia coli (E. coli) an Puten.', stellt Yuki fest. Eigentlich wollte Yuki eher los um zu Boldern. Das wird aber wohl nichts mehr.



Leider kennen sich Alex und Yuki 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 Sensifität 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!



‘Ich glaube du bringst da was durcheinander. Wir nutzen zwar auch für die ANOVA die Funktion $\text{lm}()$ aber hier wollen wir, glaube ich, eine Gerade durch die Punkte zeichnen.’, merkt Steffen an. ‘Ich sehe keine Punkte...’, antwortet Tina sichtlich übernächtigt. ‘Wir müssen die Daten ja auch erst visualisieren!’, spricht Steffen sehr deutlich und langsam. Die beiden hatten einen Versuch in einer Klimakammer in der Uckermark mit Erbsen durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Regenwurmdichte [Anzahl/l] und Frischegewicht [kg/ha]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt.

Durchschnittlicher Regenwurmdichte [Anzahl/l]	Frischegewicht [kg/ha]
23.5	28.4
24.8	30.8
22.7	23.9
25.7	29.6
26.3	26.6
22.7	27.3
27.3	32.2
23.9	27.9
24.1	31.1

Leider kennen sich Steffen und Tina 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!



‘Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen!’, ruft Jessica. ‘Ich sehe nur zwei Zeilen und keine Punkte. Wie soll ich da denn jetzt eine Gerade durchzeichnen?’, fragt Nilufar. Jessica atmet schwer ein und starrt auf die  Ausgabe der Funktion `lm()`. Die beiden hatten ein Stallexperiment im Oldenburger Land mit Milchvieh durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche Tagestemperatur [C/d] und Fettgehalt [%/kg]. Jetzt will die Betreuung von den beiden einmal die Visualisierung der Daten und auch gleich noch die lineare Regression gerechnet bekommen. Das haben beide in  gemacht, aber wie soll das jetzt gehen?

term	estimate	std.error	t statistic	p-value
(Intercept)	2.82	1.97		
Durchschnittliche Tagestemperatur	1.90	0.20		

Leider kennen sich Jessica 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. 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!



'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 Jessica an. 'Ich sehe keine Punkte... das ist doch eine Ausgabe in . Überhaupt, darum geht es doch gar nicht in unserem Versuch. Wir wollen doch keine Gerade zeichnen?.', antwortet Steffen sichtlich übernächtigt. 'Doch wir müssen nur die Koeffizienten der linearen Regression erst richtig interpretieren und vor unserem geistigen Auge erscheint eine Gerade!', spricht Jessica sehr deutlich und langsam. Die beiden hatten ein Feldexperiment in der Uckermark mit Brokkoli durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Anteil an Ton [%/l] und Proteingehalt [g/kg]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der  Ausgabe möglich sein.

```
##
## Call:
## lm(formula = "Proteingehalt ~ Durchschnittlicher_Anteil", data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.86602 -1.03884  0.07407  0.90035  2.10032
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2168     1.3592   0.895   0.376
## Durchschnittlicher_Anteil  1.5874     0.1347  11.784 6.69e-15
##
## Residual standard error: 1.195 on 42 degrees of freedom
## Multiple R-squared:  0.7678, Adjusted R-squared:  0.7623
## F-statistic: 138.9 on 1 and 42 DF, p-value: 6.691e-15
```

Leider kennen sich Jessica und Steffen 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 Jonas laut an. 'Ich sehe keine Punkte... das ist doch eine Ausgabe in . Überhaupt, darum geht es doch gar nicht in meinem Versuch. Ich wollte doch keine Gerade zeichnen?.', antwortet Jonas sich sichtlich übernächtigt selber. Wenn Mission Impossible läuft, dann ist das Meerschweinchen nicht mehr da. Aber jetzt braucht er mal Entspannung! Die Nacht war zu lang und überhaupt. Wenn die Erschöpfung nicht wäre, ja dann wäre wohl vieles möglich für Jonas! Aber so.. Jonas hatte ein Kreuzungsexperiment im Emsland mit Schweinen durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Bewegungsscore [Movement/h] und Gewichtszuwachs in der 1LW. Jetzt will er erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der  Ausgabe möglich sein.

```
##  
## Pearson's correlation  
##  
## data: Durchschnittlicher Bewegungsscore and Gewichtszuwachs  
## t = 2.8034, df = 8, p-value = 0.02308  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
##  0.1335037 0.9240270  
## sample estimates:  
##      cor  
## 0.7039571
```

Leider kennt sich Jonas mit der Korrelationsanalyse in  überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

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

97. Aufgabe

(12 Punkte)

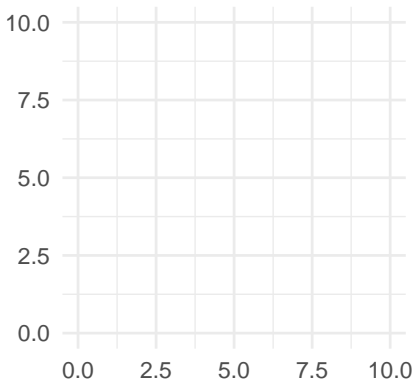
Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



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

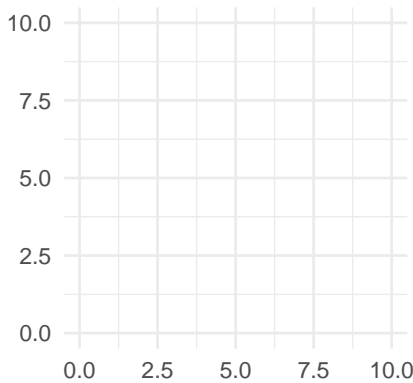
Pearsons $\rho = 0.25$

$R^2 = 1$



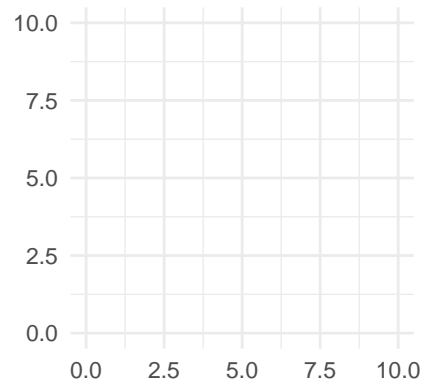
Pearsons $\rho = 0$

$R^2 = 0.75$



Pearsons $\rho = -0.25$

$R^2 = 0.25$



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

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

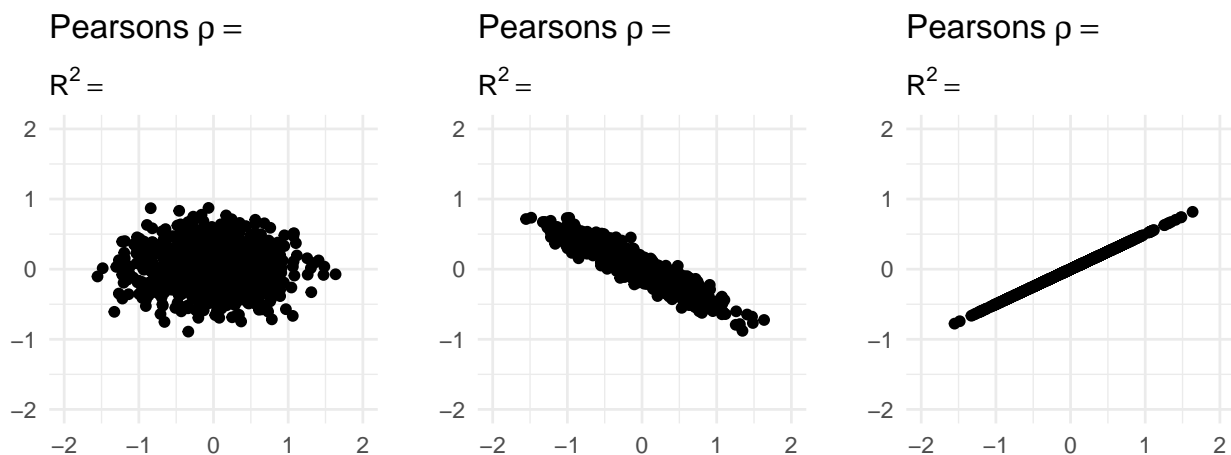
98. Aufgabe

(12 Punkte)

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



Der Bildschirm strahlt blau in das Gesicht von Alex. Es ist schon spät. Und das hat einen Grund. Hm, lecker Gummibärchen und dazu dann im Hintergrund Aliens laufen lassen. Alex überlegt, aber seine Gedanken sind etwas zäh. 'Was soll das hier alles bedeuten?', fragt sich Alex. Irgendwie ist ihm nicht klar wie er p -Werte oder R^2 -Werte abschätzen soll. Alles nicht so einfach. Eine echte Herausforderung für ihn war schon immer die Gefälligkeit gewesen. Ein leidiges Lied.



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

Proteinanteil	Mittlere Eisenkonzentration	\hat{y}	ϵ
31.6	15.8	31.5	
25.0	11.5	26.1	
26.2	10.7	25.0	
35.2	18.3	34.6	
28.2	12.4	27.2	
27.4	11.5	26.0	
20.6	8.1	21.8	
29.0	13.7	28.8	
24.8	10.7	25.1	
30.0	16.2	32.0	

Leider kennt sich Jonas 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!



Jonas hat ein Freilandversuch mit Lauch durchgeführt. Soweit so gut. Dann war er bei seinem Betreuer. Leider war der Schritt nicht so hilfreich. Jonas und die Erschöpfung, eine unendliche Geschichte mit kniffligen Wendungen. Aber es muss ja weitergehen. Jonas hatte dann in seiner Abschlussarbeit einfach zu viele Endpunkte gemessen und ist jetzt vollkommen durcheinander, welche Analyse er nun wie rechnen soll. Naja, dann heißt es jetzt eben Iron Maiden aufdrehen und darüber nachdenken, was hier eigentlich gemacht wurde. Jonas fängt einfach an und nimmt den ersten Endpunkt Anzahl Läsionen auf den Blättern. Dann kann er sich voran arbeiten. Später dann noch raus um zu Schwimmen um mal zu entspannen und vielleicht ist Mark auch da. Wäre toll.

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

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

Teil VIII.

Experimentelles Design




101. Aufgabe

(16 Punkte)



Die Spinne macht mal wieder Randalie in Tinas Zimmer und rennt davon! Jessica und Yuki sind bei Tina in im Wendland wo der neue, bessere Versuch stattfinden soll. Dabei soll in einem Gewächshausexperiment im Wendland mit Erdbeeren durchgeführt werden. Auf dem Tisch stapeln sich Katjes aus Vollkorndinkelmehl. Eine Spezialität der Komune hier. Jessica hasst Vollkorn wie Reese's Peanut Butter Cups geliebt werden. In dem neuen Versuch geht es um den Zusammenhang zwischen der Behandlung Lichtstufen (*none*, 200lm, 400lm und 600lm) und dem Messwert Trockengewicht [kg/ha]. Immerhin ist der Messwert normalverteilt, was einges einfacher macht. Was es nicht so einfacher macht ist, dass Yuki als zusätzliche Herausforderung noch die Faulheit mitgebracht hat. Daher entscheiden sich alle drei für ein *Complete randomized design* (CRD). 'Naja, so viel einfacher ist es dann doch nicht...', merkt Yuki an und sucht die Spinne.

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

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

102. Aufgabe

(20 Punkte)



Steffen und Jessica sind bei Paula um sich Hilfe für eine Versuchsplanung in **R** zu holen. Im Hintergrund läuft viel zu laut Tocotronic. Dabei geht es um den Zusammenhang zwischen der Behandlung Bestandsdichte (*standard*, *eng*, *weit* und *kontakt*) sowie Flüssignahrung (*ctrl* und *flow*) sowie drei Blöcken und dem Messwert Protein/Fettrate [%/kg] in Puten. Der Versuch soll in einem Leistungssteigerungsversuch im Wendland durchgeführt werden. Nach der Dozentin ist der Messwert Protein/Fettrate [%/kg] normalverteilt. Die beiden entschieden sich für ein faktorielles Versuchsdesign. Im ersten Schritt überlegt Paula ein komplexeres experimentelles Design zu probieren. Daher entscheiden sich alle drei für ein *Strip plot design* oder auch *Streifenanlage* mit Berücksichtigung einer Interaktion. Das sollte für den anfang erstmal reichen. 'Und jetzt, was machen wir jetzt?', Jessica schaut die anderen beiden mit großen Augen an. Die zucken mit der Schulter. Alle mampfen Katjes.

Leider kennen sich Paula, Steffen und Jessica 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 **R**! **(2 Punkte)**
5. Erstellen Sie das statistische Modell in der in **R** üblichen Schreibweise für eine ANOVA! Skizzieren Sie die notwendige Funktionen in **R**! **(4 Punkte)**
6. Skizzieren Sie die weitere Datenanalyse hinsichtlich eines multiplen Gruppenvergleiches! **(2 Punkte)**
7. Skizzieren Sie eine mögliche Abbildung im Kontext der wissenschaftlichen Fragestellung! Beschriften Sie die Abbildung! **(3 Punkte)**
8. Ergänzen Sie zu der Abbildung ein mögliches Ergebnis des multiplen Gruppenvergleichs! Welche Annahme hinsichtlich der Modellierung haben Sie getroffen? Begründen Sie Ihre Antwort! **(3 Punkte)**




Teil IX.

Programmieren in R

103. Aufgabe

(9 Punkte)





Jessica muss ihrer Abschlussarbeit mit  arbeiten. Deshalb sitzt sie jetzt mit Ihnen zusammen und hat einige Fragen zu den Grundlagen in  an Sie! Na dann wollen Sie mal helfen. Immerhin will ihre Betreuerin, dass  genutzt wird.

Jessica: *Es gibt ja in R unter anderem `library()` und `Packages`. Was ist der Unterschied und wozu brauche ich die?* **(1 Punkt)**

Sie antworten:

Jessica: *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:

Jessica: *Wie heißen nochmal die beiden  Pakete, die wir fast immer laden, wenn wir  nutzen wollen?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

Jessica: *Was macht überhaupt dieses `c()` hier überall im  Code?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

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

Sie antworten:

Jessica: *Hä? Warum ändert sich nichts an meinen Daten? In R sehe ich doch die Änderungen aber irgendwie speichert R meine Änderungen meines Datensatzes ab. Was ist da los?* **(1 Punkt)**

Sie antworten:

Jessica: *Wir brauchen recht häufig die Tilde (`~`) in R. Wo wird die nochmal angewandt und genutzt?* **(1 Punkt)**

Sie antworten:

104. Aufgabe

(9 Punkte)



Jessica muss ihrer Abschlussarbeit mit **R** arbeiten. Leider ist die Analyse etwas komplexer, so dass es eben in Excel dann nicht mehr geht. Deshalb also gleich alles in **R**. Das ist auch der Grund warum sie jetzt mit Ihnen in der Küche sitzt und einige vertiefende Fragen zu **R** an Sie hat! Na dann wollen Sie mal helfen. Immerhin will ihr Betreuer, dass **R** genutzt wird und die Abgabe ist dann auch schon in gut einem Monat.

Jessica fragt: *Die R Pakete {emmeans}, {ggplot} und {readxl} sind für was nochmal gut? (2 Punkte)*

Sie antworten:

Jessica fragt: *Ich bräuchte mal eine Analogie für das Prinzip von dem r Paket {ggplot}. Wie funktioniert das Paket nochmal? So konzeptionell? (2 Punkte)*

Sie antworten:

Jessica 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:

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

Sie antworten:

Jessica fragt: *Ich möchte in der Funktion emmeans() den Faktor f_1 getrennt in jedem Level des Faktors f_2 auswerten. Was muss ich da in de Funktion emmeans() angeben? (1 Punkt)*

Sie antworten:

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

Sie antworten:

Jessica fragt: *Wenn ich zwei Faktoren f_1 und f_2 habe, wie spezifiziere ich nochmal eine Interaktion in einem Modell? (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., Ols, H. W., Westendarp, H., Appenroth, K. J., & Ulbrich, A. (2022). Influence of light intensity and spectrum on duckweed growth and proteins in a small-scale, re-circulating indoor vertical farm. *Plants*, 11(8), 1010. [[Link](#)]
- Selle, P. H., Cadogan, D. J., Li, X., & Bryden, W. L. (2010). Implications of sorghum in broiler chicken nutrition. *Animal Feed Science and Technology*, 156(3-4), 57-74. [[Link](#)]
- Wu, G., Knabe, D. A., & Kim, S. W. (2004). Arginine nutrition in neonatal pigs. *The Journal of Nutrition*, 134(10), 2783S-2790S. [[Link](#)]

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

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

- bar

105. Aufgabe



(20 Punkte)

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



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

Leider kennt sich Jonas 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!



Unter einem langen Schnaufen starrt Yuki auf den wissenschaftlichen Datensatz *data2* in seinem Laptop. 'Worum geht es denn eigentlich in diesem Datensatz?', fragt er sich kopfschüttelnd und mampft noch ein paar Reese's Peanut Butter Cups. Yuki soll die Datentabelle nutzen um das eigene Experiment zu planen und eine Blaupause zu haben. Als eine Vorlage sozusagen, die er nur noch ausfüllen muss. Daher möchte sein Betreuer, dass er einmal die Daten sinnvoll zusammenfasst. Das sollte dann doch etwas aufwendiger werden. Das wird dann vermutlich heute Abend nichts mehr mit Matrix

Leider kennt sich Yuki mit der Analyse eines wissenschaftlichen Datensatzes überhaupt nicht aus. Deshalb braucht 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 0.9mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 13m in Brusthöhe hatte.

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

Herodot hatte eine Schulterhöhe von 185cm , eine Breite von 85cm sowie eine Länge von 240cm .

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

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

4. Berechnen Sie die Dicke der Eichenwand in cm ! Verdeutlichen Sie Ihre Berechnungen an einer aussagekräftigen Skizze für Pferd und Eiche! **(2 Punkte)**
5. Unter einer Dicke der Eichenwand von 10cm 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 240 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 7cm. Der Kubikmeterpreis für Torf liegt bei 290 EUR.

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

109. Aufgabe

(10 Punkte)

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



Solar- & Biogasanlagen Um die Energiekosten Ihres Betriebes zu senken, wollen Sie eine Solaranlage auf den Schweinestall montieren lassen. Sie messen Ihren Stall und finden folgende Maße wieder. Die vordere Seite des Schweinestall hat eine Höhe h_v von $6m$. Die hintere Seite des Schweinestall hat eine Höhe h_b von $9m$. Der Schweinestall hat eine Tiefe t von $12m$ und eine Breite b von $50m$.

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

Ebenfalls planen Sie eine neue Biogasanlage für Ihren Betrieb. Der neue Methantank hat einen Radius r von $1.5m$. Leider gibt es ein paar bauliche Beschränkungen auf dem Grundstück. Ihr Fundament des zylindrischen Methantanks kann nur ein Gewicht von maximal $14t$ aushalten bevor der Tank wegbricht. Sie rechnen eine Sicherheitstoleranz von 25% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $235kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $290kg/m^3$. Sie betreiben Ihre Anlage bei $-92^\circ C$.

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

110. Aufgabe

(10 Punkte)

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



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

Aligatorenbirnen und Blaubeeren “Sind Sie ein Riesenfaultier oder warum kaufen Sie so viele Aligatorenbirnen?”, spricht es hinter Ihnen. Irritiert drehen Sie sich um und blicken in das puderrote Gesicht von einem Studienrat mit Stock. “Wieso?”, entfährt es Ihnen und Sie bereuen sogleich die Frage. Sofort werden Sie zu einem Whiteboard voller roter Schnüre geschliffen und müssen folgenden mathematischen untermauerten Argumenten im Edeka über sich ergehen lassen. Da kommen Sie nicht mehr raus, also können Sie auch gleich mitmachen. Das Problem liegt in Chile⁴. Tja, die Deutschen und Südamerika.

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

1. Wenn 4 Blaubeerschalen 7.96 Euro kosten, wie viel kosten 5 Schalen? **(2 Punkte)**
2. Wenn Sie die 5 Blaubeerschalen gekauft haben, wie viele Aligatorkbirnen zu je 1.79 EUR können Sie sich dann noch für 200 EUR leisten? **(1 Punkt)**

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

- Ein Kilo Strauchtomaten benötigt 170l Wasser. Eine Strauchtomate wiegt 120 - 125g.
- Ein Kilo Salat benötigt 140l Wasser. Ein Salatkopf wiegt 300 - 510g.
- Ein Kilo Avocado benötigt 1100l Wasser. Eine Avocado wiegt 150 - 400g.
- Ein Kilo Blaubeeren benötigt 820l Wasser. Eine Blaubeere wiegt 3.2 - 3.8g.

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

Chile exportiert im großem Ausmaß Blaubeeren und Avocados. In dem Exportjahr 2024 blieben die Erträge von Blaubeeren mit 7.9×10^4 t in dem prognostizierten Rahmen. Die Menge reduzierte sich um 8.6%. Die Exporte für Avocados stiegen in dem gleichen Zeitraum um 22.1% auf 2.3×10^5 t.

4. Wie viele Liter 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 zwei Prozent des Wassers des Landes für den häuslichen Verbrauch vorgesehen. In den Dörfern der Anbauregionen versorgen Tankwagen die Bevölkerung jede Woche mit Wasser, es gibt etwa 48 Liter Wasser pro Kopf für den täglichen Bedarf. In *Deutschland* liegt der Verbrauch bei 10 - 15 Liter pro Minute Duschen und 35 - 115 Liter pro Waschgang einer Waschmaschine.

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 Studienrat mit Stock 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 66 Millionen Jahren die Dinosaurier - so groß sie auch waren - nicht von der Schwerkraft zu Boden gerissen wurden. Hat der Dampfplauderer etwa recht und war die Schwerkraft vor Millionen von Jahren eine andere? Sind deshalb alle Lebewesen auf der Erde *heutzutage* so viel kleiner, weil die Schwerkraft größer ist als damals? War die Erde kleiner und hatte weniger Masse? Oder ist es nur ein Rechenfehler wie bei der Theorie der Hohlerde von Edmond Halley aus dem 17.-18. Jahrhundert? Müde reiben Sie sich die Augen. So wird es nichts mehr mit dem Schlafen, dann können Sie auch mal etwas 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 5.21 g/cm^3 . Das Gewicht von einem heute lebenden Waldelefanten mit 2.7t liegt bei 6t und das Gewicht von einem Brachiosaurus bei bis zu 30t.

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

Die Distanz zwischen Sonne und Erde entspricht 0.99 astronomische Einheiten (AE). Die Einheit 1 AE wird mit $1.48 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 84% Wasserstoffkernen mit einer molaren Masse von 1.05 g/mol , 9% Heliumkernen mit 4.11 g/mol sowie 7% weiteren Atomkernen mit 89.32 g/mol . Die Teilchendichte bei Eintritt in die Erdatmosphäre liegt zwischen 0.4 bis 100 Teilchen cm^{-3} pro Sekunde mit einer mittleren Teilchendichte von 6 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) walking, (3) wing/leg stretching und (4) sitting.

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	29cm; 21cm; 25.4%	41cm; 19cm; 25.4%	34cm; 18cm; 24.1%
walking	37cm; 31cm; 2.1%	36cm; 32cm; 4.1%	29cm; 19cm; 1.8%
wing/leg stretching	33cm; 17cm; 1.6%	32cm; 21cm; 0.5%	33cm; 27cm; 0.5%
sitting	39cm; 31cm; 5.1%	35cm; 38cm; 1.2%	26cm; 24cm; 5.1%

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 65 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 222 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 $8000\mu\text{g}/150\text{mg}$ Vitamin C. Der Bedarf liegt bei 115mg pro Tag für Männer.

3. Berechnen Sie die notwendige Menge in t an aufzunehmenden wilden Sellerie auf die Concepción für die ununterbrochene Fahrt von drei Monate und 20 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 $2 \times 10^{28} \text{ kg}$. Wenn die Sonne nun am Ende ihrer Lebenszeit zu einem schwarzen Loch mit dem Radius von 2500m 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 $5.974 \cdot 10^{-11} \text{ m}^3 (\text{kg} \cdot \text{s}^2)^{-1}$

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 $2.9 \times 10^8 \text{ m/s}$ aus. Kann das Lichtteilchen der Gravitation des schwarzen Lochs entkommen? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Stellen Sie den Zusammenhang zwischen dem sich verringernden Radius r des schwarzen Lochs bei gleichbleibender Masse m_s und der notwendigen Fluchtgeschwindigkeit v_f in einer Abbildung dar! **(2 Punkte)**
6. Ein Flugzeug 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 vier 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 500 Jahren ist die Replikation abgeschlossen und wiederum vier Sonden werden ausgesendet. Gehen Sie von 5.16 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 2×10^{11} Sterne in der Milchstraße. Nehmen Sie eine Lichtgeschwindigkeit von $2.8 \times 10^8 \text{ m/s}$ an.

1. Skizzieren Sie in einer Abbildung die ersten drei 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 4.6×10^9 Jahren, wie oft war dann eine Sonde einer außerirdischen Zivilisation schon zu Besuch? Korrigieren Sie Ihre Antwort mit dem Wissen, dass sich die Kontinentalplatten einmal alle 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 73 Grad im Vergleich zu den ägyptischen Pyramiden mit 54 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 38 Königsellen. Eine Königselle misst 52.6cm.

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 38 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 3 Sklaven, die Ihnen bei dem Befüllen der Pyramide mit Schutt zu Seite stehen. Leider haben Ihre Sklaven zu allem Überfluss auch noch chronische Schulterschmerzen entwickelt, als sie von der anstehenden Aufgabe erfahren haben. Gehen Sie daher von einer Effizienz der Sklaven von 75% aus. In eine Schubkarre passen 95 Liter.

3. Wie oft müssen Ihre maladen Sklaven die Rampe mit der Schubkarre zur Spitze der Pyramide hochfahren um die Pyramide mit Schutt zu füllen? **(1 Punkt)**
4. Berechnen Sie die Länge der Rampe zur Spitze der Pyramide mit einem Anstellwinkel von 14° ! **(2 Punkte)**
5. Wie weit reicht Ihre Rampe vom Fuß der Pyramide in die oldenburgische Landschaft? **(2 Punkte)**

Bei der Besichtigung der Pyramide teilt Ihnen der leicht übergewichtige Pharao (Nebenberuf *Versicherungsverteter*) 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 6° ändert! **(2 Punkte)**

117. Aufgabe

(12 Punkte)

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



Geocaching – Von Satelliten und Plastikdosen Es ist Wochenende und das Wetter ist *sweet*. Sie schwingen sich auf Ihr Cachermobil um mit 17km/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	GCND3QS	4.0 4.5 Mikro
B	GC9GZVM	1.0 2.5 Normal
C	GCYYIHX	2.5 2.0 Klein
D	GCGBVPP	1.5 3.5 Normal
E	GCKMS2Y	5.0 1.0 Normal

Im Weiteren sind Ihnen folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \overrightarrow{AB} ist 4km. Im Weiteren ist Ihnen der Entfernungsvektor \overrightarrow{CB} mit 7.5km bekannt. Der Entfernungsvektor \overrightarrow{BE} ist das 1.2-fache des Entfernungsvektor \overrightarrow{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 40° südlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 35° östlich. Vom Ort B betrachtet, bilden die Orte C und D einen rechten Winkel am Ort B. Der Ort B liegt auf gerader Linie zwischen den Orten C und E. Somit liegt der Ort E südlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort C 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

$$\text{Suchzeit} = 0.05 + 0.25 \cdot \text{Schwierigkeit}$$

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 7m lang. Erreichen Sie einen Cache in der Höhe von 8.6m? 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

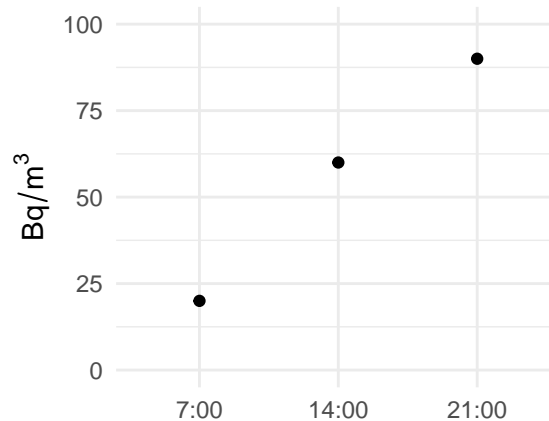
(10 Punkte)

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 21:00 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 $280Bq/m^3$ in Ihrem ungelüfteten Kellerraum erreicht haben? **(2 Punkte)**

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

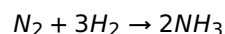
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $280Bq/m^3$ auf unter $90Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	78.1	27.9	
Sauerstoff	20.45	16.2	
Kohlenstoffdioxid	0.035	12.5	

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

Während 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 Aldi die Regale einzuräumen. Dabei ist Ihnen in der Auslage der Sonderangebote das Necronomicon¹² in die Hände gefallen. Nun sind Sie eine Magierin 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 922 n. Chr. für den neuen Lehnsherren Fürsten Arthur. 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 $20mg$ 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 $3.1mm$. Die Löcher im Sieb erlauben ein Tropfendurchmesser von $2.1mm$. Welchen Durchmesser in mm haben Ihre produzierten Bleikugeln? **(3 Punkte)**

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

4. Wie schwer in Kilogramm kg sind die 1.2×10^6 produzierten Bleikugeln, die Sie jetzt auf die Burgmauer transportieren müssen? **(1 Punkt)**

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

5. Wie groß in cm^2 ist Ihr quadratisches Sieb am oberen Ende des Turms, wenn Sie pro Fall ca. 700 Bleikugeln produzieren wollen und die Bleikugel im Fall $0.8cm$ 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 1859 ungefähr 32 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.

$$f(t) = 9 \times 10^9 - 1.2 \times 10^9 \cdot 2.3^{-0.2 \cdot t + 2.4}$$

1. Skizzieren Sie die Sättigungsfunktion *annäherungsweise* in einer Abbildung! **(1 Punkt)**
2. Wie viele Kaninchen können nach der Sättigungsfunktion maximal im westlichen Australien leben? Ergänzen Sie den Wert in Ihrer Abbildung! **(2 Punkte)**
3. Wie viele Millionen Kaninchen leben nach der Sättigungsfunktion nach 12 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 16 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 10 Jahren Wachstum zu durchseuchen? **(1 Punkt)**

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

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

Thomas Austin entließ die Kaninchen im äußersten Norden von Australien. Australien hat eine West-Ost-Ausdehnung von 4100km und eine Nord-Süd-Ausdehnung von knapp 3500km. Die Kaninchen breiten sich radial mit einer Geschwindigkeit von 8.8km 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 13\$ pro Tier und der durchführende Arzt verlangt ca. 35\$ pro Tier.

6. In Ihrem Stall leben 800 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!



Uckermark. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer des Hängebauschweins Frida 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	2.7
01. Apr 2023	5.8

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

Auf dem Weg zu Ihrer Jonagoldplantage wurden Sie mit Ihrem Trecker von einer Gruppe elektrifizierter Rentner abgedrängt. Der Trecker muss wieder aus dem Graben! Frida und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Frida mit 120N . Die elektrifizierten 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 Frida lotrecht über einen Ast oberhalb des Treckers. Die Rentner ziehen in einer geraden Linie über die Böschung hinweg am anderen Seil. Welche Kraft wird aufgebracht? **(2 Punkte)**
6. Im zweiten Versuch ziehen Frida und die Rentner mit einem 50° Winkel mit ihrem Seil an dem Trecker. Welche Kraft wird aufgebracht? **(2 Punkte)**
7. Mit welcher Beschleunigung ziehen Sie den 1.3t 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, 4-mal-gefaltete, 0.6mm, 60-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 60cm und eine Breite von 18cm. 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 90m Zaun zu Verfügung. Sie wollen nun die maximale Fläche des abgeschirmten Vorgartens in Abhängigkeit der Seitenlängen bei der Verwendung von 90m 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

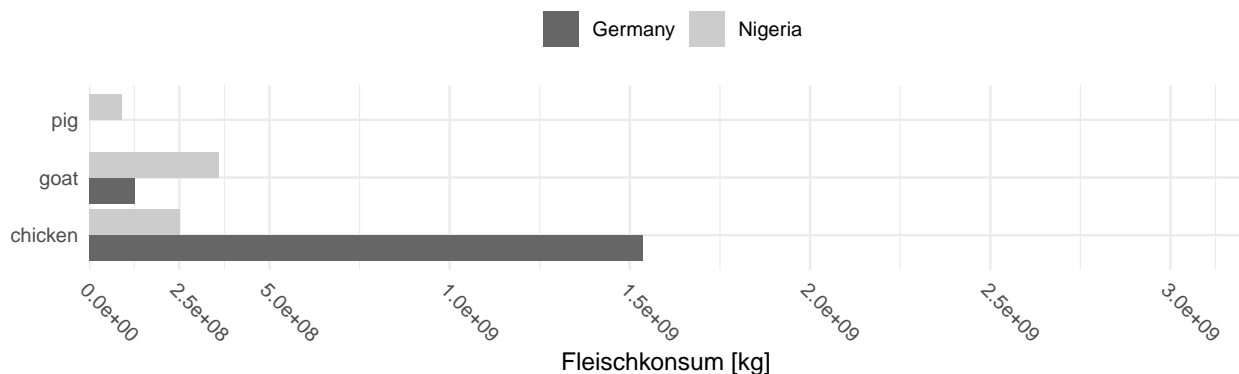
(10 Punkte)

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



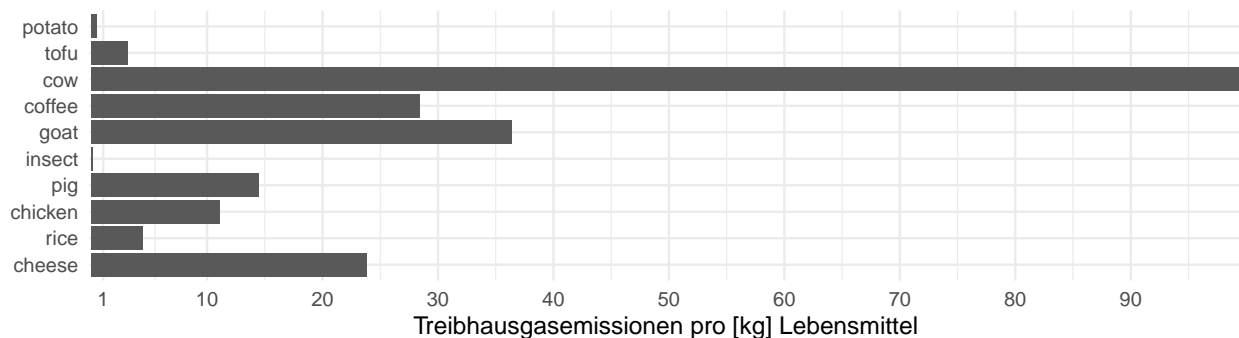
Ein Pfund Insekten, bitte! Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹⁴. Schauen wir uns dazu einmal den Vergleich Deutschland zu Nigeria an. Nigeria hat eine der am schnellsten wachsenden Bevölkerungen der Welt und wird vermutlich im Jahr 2100 zu den Top 5 der bevölkerungsreichsten Länder zählen. Im Jahr 2021 leben ca. 8.4×10^7 Menschen in Deutschland und ca. 1.8×10^8 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 2021 in Deutschland und Nigeria in [kg] einmal dargestellt.



1. Stellen Sie den Fleischkonsum in Deutschland und Nigeria im Jahr 2021 *pro Kopf* in einer aussagekräftigen Tabelle dar! **(2 Punkte)**
2. Ergänzen Sie in der Tabelle eine Spalte in der Sie für den Fleischkonsum in Nigeria auf Deutschland normieren, daher ins Verhältnis Nigeria/Deutschland, setzen! **(1 Punkt)**

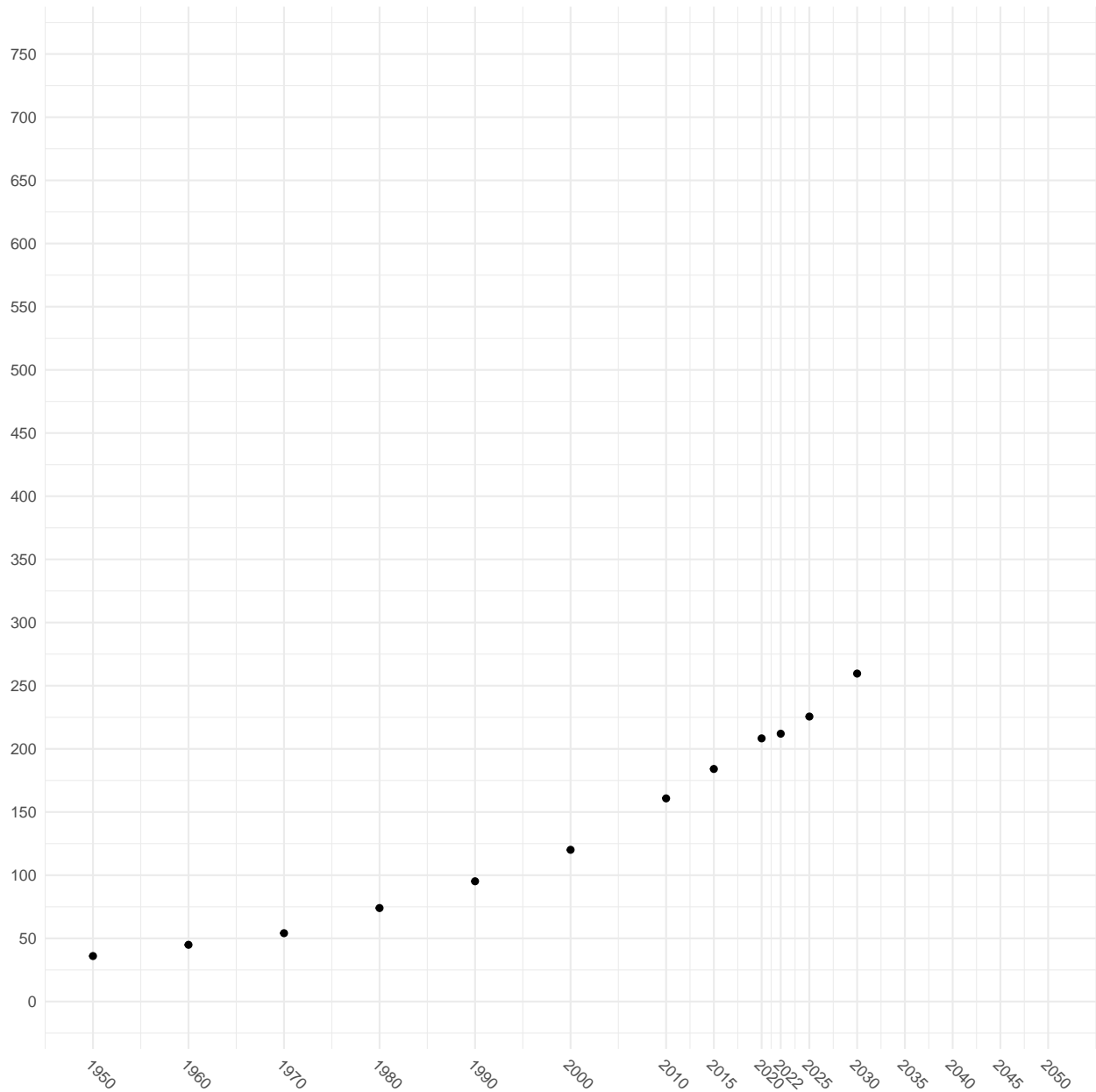
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



3. Stellen Sie in einer Tabelle die Treibhausgasemissionen an CO₂ pro Kopf, die durch den Fleischkonsum in Deutschland und Nigeria im Jahr 2021 entstehen, dar! Ergänzen Sie auch hier das Verhältnis Nigeria zu Deutschland! **(2 Punkte)**

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 80%-iger Angleichung der Lebensbedingungen zu Deutschland im Jahr 2021! **(1 Punkt)**
6. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria im Jahr 2050 im Vergleich zum Jahr 2021, der sich durch den angeglichenen Fleischkonsum ergibt! **(1 Punkt)**
7. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria, wenn die gesamte Proteinaufnahme durch Insekten ersetzt würde! **(1 Punkt)**

124. Aufgabe

(12 Punkte)

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



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

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.8% angenommen. In 90% der Fälle ist ein HIV-Test positiv, wenn der Patient erkrankt ist. In 1% der Fälle ist ein HIV-Test positiv, wenn der Patient *nicht* erkrankt ist und somit gesund ist. Sie stutzen. Wie wahrscheinlich ist es denn eigentlich an AIDS erkrankt zu sein (K^+), wenn Sie einen positiven AIDS-Test vorliegen haben (T^+)? Gehen Sie für die folgenden Berechnungen von $n = 3 \times 10^4$ Patienten mit einem diagnostischen Test für AIDS aus. Sie nehmen sich also einen Kuli und fangen an auf der historischen Ausgabe des Spiegels zu 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 18 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 35%? **(1 Punkt)**

Ihr zu vermarkendes Produkt, hinter dem Sie voll stehen, kostet 75EUR pro Einheit im Direktverkauf. Die direkte Provision für die erste Stufe beträgt 25%. Für die zweite, dritte und vierte Stufe betragen die indirekten Provisionen jeweils 2.5%, 1.5% und 1%. Jeder Ihrer angeworbenen „Partner“ wirbt wiederum drei Partner für sich selbst an. Pro Monat werden im Schnitt vier Einheiten vom Produkt verkauft. Sie wollen nun 2500EUR 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 3375EUR kaufen. Diese Einheiten können Sie nur direkt verkaufen. Leider mussten Sie den Kauf über einen Kredit über 6.2% p.a. über 48 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 4 achtseitige Würfel ($4d8$) zum würfeln in der Hand. Wenn Sie eine 8 würfeln, haben Sie einen Erfolg.

1. Berechnen Sie die Wahrscheinlichkeit *genau* 2 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 zwölfseitigen Würfeln ($2d12$) als Schaden oder das Schwert mit einem zwölfseitigen Würfel plus 4 ($1d12+4$) als Schaden bevorzugen? Begründen Sie Ihre Antwort mathematisch! **(1 Punkt)**

Jetzt wird es immer wilder, da Sie sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass Ihr Rettungswurf gegen den zaubernden Hexer funktioniert. Sie haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.7$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.7$. 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 Helmut und Günther das Team der drei Kandidaten.

Name	$P(\text{win})$	$P(\text{outbid})$
Helmut	0.2	0.05
Günther	0.1	0.11

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 0.12 liegt, mit welcher Wahrscheinlichkeit gewinnt *keiner* die Kaffemaschine von Mitropa? **(1 Punkt)**

Glücksrad für Arme auf der Kirmes! Leider hat es für Maren Gilzer nicht gereicht. Deshalb sind Sie jetzt auf der Kirmes und spielen mit einem Typen in einem Tentakelkostüm um das große Geld. Das Glücksrad hat 24 Felder. Sie drehen das Glücksrad zweimal. Auf 6 Feldern gewinnen Sie 5000EUR 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 6000EUR? **(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.