

Name: _____

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

Fakultät Agrarwissenschaften und Landschaftsarchitektur (AuL)

Klausurfragen Bio Data Science

für Pflichtmodule

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

(Prüfungsleistung der Wahlpflichtmodule ist die Portfolioprüfung)

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

»Wintersemester 2025/26«



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



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

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



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

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



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

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



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

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



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

Erlaubte Hilfsmittel

- Normaler Taschenrechner ohne Möglichkeit der Kommunikation mit anderen Geräten! Ausdrücklich kein Handy!
- Eine DIN A4-Seite als beidseitig, selbstgeschriebene, handschriftliche Formelsammlung. Keine digitalen Ausdrücke!
- **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 81 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 101 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
96.5 - 101	1,0
91.5 - 96.0	1,3
86.5 - 91.0	1,7
81.5 - 86.0	2,0
76.5 - 81.0	2,3
71.5 - 76.0	2,7
66.5 - 71.0	3,0
61.5 - 66.0	3,3
56.5 - 61.0	3,7
50.5 - 56.0	4,0

Es ergibt sich eine Endnote von _____.

Multiple Choice Aufgaben

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

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

- Es sind ____ von 20 Punkten erreicht worden.

Rechen- und Textaufgaben

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

- Es sind ____ von 81 Punkten erreicht worden.

Multiple Choice Aufgaben

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

Programmieren in R

1. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Statistik und Versuchswesen • Biostatistik

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

- ☐ A Die **R** Pakete sind nur in englischer Sprache verfasst. Das ist aber nicht der Hauptgrund, denn **R** hat wie alle Programmiersprachen Probleme mit Umlauten und Sonderzeichen.
- ☐ B **R** Pakete sind nur in englischer Sprache verfasst. Es macht keinen Sinn **R** daher in Deutsch zu bedienen.
- ☐ C Alle Funktionen und auch Anwendungen sind in **R** in englischer Sprache. Die Nutzung von deutschen Wörtern ist nicht schick und das ist zu vermeiden.
- ☐ D Es gibt keinen Grund nicht auch deutsche Wörter zu verwenden. Es ist ein Stilmittel.
- ☐ E 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.

2. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Statistik und Versuchswesen • Biostatistik

Wenn wir ein Experiment in **R** auswerten wollen, dann müssen die Daten in Excel in einem bestimmten Format angelegt werden. Wir sprechen auch vom Long-Format. Welche Aussage zum Long-Format ist richtig?

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

3. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Statistik und Versuchswesen • Biostatistik

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

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

Deskriptive Statistik & Explorative Datenanalyse

4. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik

Wie lautet der Mittelwert und Standardabweichung von y mit 10, 14, 16, 11 und 15.

- A ☐ Es ergibt sich 12.2 +/- 3.35
- B ☐ Sie erhalten 13.2 +/- 1.61
- C ☐ Es berechnet sich 14.2 +/- 6.7
- D ☐ Es ergibt sich 13.2 +/- 2.59
- E ☐ Sie erhalten 13.2 +/- 1.295

5. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik

Berechnen Sie den Median, das 1st Quartile sowie das 3rd Quartile von y mit 12, 24, 25, 35, 18, 20 und 63.

- A ☐ Es ergibt sich 24 [18; 35]
- B ☐ Es ergibt sich 24 +/- 18
- C ☐ Es ergibt sich 28 +/- 18
- D ☐ Es berechnet sich 28 [19; 36]
- E ☐ Sie erhalten 24 [16; 33]

6. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik

Sie wollen nach einem Feldversuch die Standardabweichung berechnen. Welche der folgenden Rechenoperationen müssen durchgeführt werden?

- A ☐ Den Mittelwert berechnen, dann die absoluten Abstände zum Mittelwert aufsummieren. Die Fallzahl ($n - 1$) entsprechend gewichten.
- B ☐ Den Mittelwert berechnen, dann die quadratischen Abstände zum Mittelwert aufsummieren und durch die Fallzahl ($n - 1$) teilen.
- C ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend subtrahieren wir die Fallzahl ($n - 1$).
- D ☐ Wir berechnen erst den Mittelwert und dann die quadratischen Abstände zu dem Mittelwert. Diese quadratischen Abstände summieren wir auf und teilen am Ende durch die Fallzahl ($n - 1$). Als letzten Schritt ziehen wir die quadratische Wurzel.
- E ☐ Den Mittelwert berechnen und die Abstände quadrieren. Die Summe mit der Fallzahl ($n - 1$) multiplizieren.

7. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

- A ☐ Den Median und die Standardabweichung.
- B ☐ Den Median und die Quartile.
- C ☐ Der Boxplot stellt die Mittelwerte und die Standardabweichung dar.
- D ☐ Der Boxplot stellt den Median und die Streuung dar.
- E ☐ Den Mittelwert und die Varianz.

8. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Nachdem Sie in einem Feldexperiment zu Leistungssteigerung von Kartoffel durchgeführt haben, berechnen Sie den Mittelwert und den Median. Der Mittelwert ist \bar{y} gleich 21.2 t/ha und als Median ergibt sich \tilde{y} gleich 9.2 t/ha. Welche Aussage ist richtig?

- A ☐ Wenn sich der Mittelwert und der Median nicht unterscheiden, liegen vermutlich Outlier in den Daten vor.
- B ☐ Nach der Betrachtung der Werte unterscheiden sich Mittelwert und Median nicht. Die Daten können so verwendet werden wie sie vorliegen. Es gibt keinen Outlier.
- C ☐ Nach der Betrachtung der Werte unterscheiden sich Mittelwert und Median nicht. Sie haben Varianzhomogenität vorliegen. Sie können künstlich Outlier zufügen um die Daten auszuwerten.
- 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 ☐ Da sich der Mittelwert und der Median unterscheiden, liegen vermutlich Outlier in den Daten vor. Wir untersuchen den Datensatz nach auffälligen Beobachtungen.

9. Aufgabe

(2 Punkte)

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Im Folgenden sehen Sie ein Normalverteilung dargestellt. Welche Aussage zu der Normalverteilung und der Standardabweichung σ ist richtig?



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

10. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

- A ☐ Die optimale Anzahl ist größer als hundert Beobachtungen, wobei es gerne sehr viel mehr sein können.
- B ☐ Wir sollten eine Beobachtung mindestens pro Gruppe vorliegen haben.
- C ☐ Wir brauchen fünf oder mehr Beobachtungen.
- D ☐ 1 Beobachtung.
- E ☐ Die untere Grenze liegt bei zwei bis fünf Beobachtungen.

11. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

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

12. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

- A ☐ In einer explorativen Datenanalyse nutzen wir den Boxplot. Dabei sollte der Median als dicke Linie in der Mitte der Box liegen. Dann können wir von einer Varianzhomogenität ausgehen.
- B ☐ Einen Violinplot. Der Bauch der Violine muss hierbei einen höheren Wert annehmen als der Steg der Violine. Dann kann die Annahme einer Varianzhomogenität angenommen werden.
- C ☐ In einer explorativen Datenanalyse nutzen wir den Violinplot. Dabei sollte der Bauch am Rand liegen. Dann können wir von einer Varianzhomogenität ausgehen.
- D ☐ Einen Barplot. Die Mittelwerte müssen alle auf einer Höhe liegen. Die Fehlerbalken haben hier keine Informationen.
- E ☐ Einen Boxplot. Das IQR muss über alle Behandlungen zusammen mit den Whiskers ungefähr gleich aussehen.

Statistische Testtheorie

13. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

- A ☐ Die Wahrscheinlichkeit für die Nullhypothese, wenn die Daten wahr sind.
- B ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit nicht die Daten D zu beobachten sondern die Nullhypothese, wenn diese wahr ist.
- C ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit der Alternativehypothese und somit $1 - Pr(H_A)$
- D ☐ $Pr(D|H_0)$ beschreibt die Wahrscheinlichkeit die Teststatistik T_D aus den Daten D zu beobachten, wenn die Nullhypothese wahr ist.
- E ☐ $Pr(D|H_0)$ stellt die Wahrscheinlichkeit die Teststatistik T zu beobachten dar, wenn die Nullhypothese falsch ist.

14. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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 ☐ 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 Rank einer Naturkonstante.
- B ☐ Der Begründer der modernen Statistik, R. Fischer, hat die Grenze simuliert und berechnet. Dadurch ergibt sich dieser optimale Cut-Off.

- C** ☐ Im Rahmen eines langen Disputes zwischen Neyman und Fischer wurde $\alpha = 5\%$ festgelegt. Leider werden die Randbedingungen und Voraussetzungen an statistische Modelle heute immer wieder ignoriert.
- D** ☐ Da Wissenschaftler eine Schwelle für die statistische Testentscheidung benötigen wurde α historisch gewählt. Damit ist $\alpha = 5\%$ eine Kulturkonstante.
- E** ☐ Der Wert ergab sich aus einer Auswertung von 1042 wissenschaftlichen Veröffentlichungen zwischen 1914 und 1948. Der Wert 5% wurde in 28% der Veröffentlichungen genutzt. Daher legte man sich auf diese Zahl fest.

15. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Das statistische Testen basiert auf dem Falsifikationsprinzip. Es besagt,

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

16. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

- A** ☐ Bei der Berechnung der Teststatistik T_D gewichten wir den Effekt *signal* mit der Varianz *noise*. Wir können verallgemeinert $T_D = \text{signal} \cdot \text{noise}$ schreiben.
- B** ☐ Wir gewichten den Effekt *signal* mit der Varianz *noise* und erhalten $T_D = \text{signal}/\text{noise}$.
- C** ☐ Bei der Berechnung der Teststatistik T_D gewichten wir den Effekt *signal* mit der Varianz *noise*. Wir können verallgemeinert $T_D = \text{signal}/\text{noise}^2$ schreiben.
- D** ☐ Es gilt $T_D = \text{signal} \cdot \text{noise}$. Der Effekt *signal* wird mit der Varianz *noise* gewichtet.
- E** ☐ Es gilt $T_D = \text{signal}/\text{noise}$. Der Effekt *noise* wird mit der Varianz *signal* gewichtet.

17. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Sie haben ein Signifikanzniveau α gleich 5% vorliegen. Welche Aussage zusammen mit dem p -Wert ist richtig?

- A** ☐ Wir vergleichen mit dem p -Wert und dem Signifikanzniveau α absolute Werte auf einem Zahlenstrahl und damit den Unterschied der Teststatistiken, wenn die H_0 gilt.
- B** ☐ Wir schauen, ob der p -Wert kleiner ist als das Signifikanzniveau α . Wir sprechen von einem signifikanten Ergebnis. Dabei vergleichen wir somit Wahrscheinlichkeiten. Die Wahrscheinlichkeiten werden als Flächen unter der Kurve der Teststaistik dargestellt, wenn die H_0 gilt.
- C** ☐ Wir vergleichen mit dem p -Wert und dem Signifikanzniveau α Wahrscheinlichkeiten und damit die absoluten Werte auf einem Zahlenstrahl, wenn die H_0 gilt.
- D** ☐ Wir vergleichen die Effekte des p -Wertes mit den Effekten der Signifikanzschwelle unter der Annahme der Nullhypothese. Dabei gilt, dass wir die Nullhypothese nur ablehnen können anhand des Falsifikationsprinzips.
- E** ☐ Wir machen ein Aussage über die Flächen unter der Kurve der Teststatistik der Hypothese H_0 und über die Flächen unter den Kurve der Teststatistik der Hypothese H_A . Dabei werden Wahrscheinlichkeiten verglichen, die durch die Flächen unter der Kurve der beiden Testverteilungen repräsentiert werden.

18. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

- A ☐ In der Analogie der Regenwahrscheinlichkeit in einem bestimmten Gebiet: ein statistischer Test gibt die Wahrscheinlichkeit für ein Ereignis in einem Experiment mit den Daten D wieder und lässt sich kaum verallgemeinern.
- B ☐ In der Analogie der Maximaltemperatur: Was ist der maximale Unterschied zwischen zwei Gruppen. Wir erhalten hier eine Aussage über die Spannweite und den maximalen Effekt.
- C ☐ In der Analogie der Durchschnittstemperatur: Wie oft tritt ein Effekt durchschnittlich ein? Wir erhalten eine Wahrscheinlichkeit für die Effekte. Zum Beispiel, wie hoch ist die Wahrscheinlichkeit für einen Mittelwert als Durchschnitt.
- D ☐ In der Analogie des Niederschlags oder Regenmenge: ein statistischer Test gibt die Stärke eines Effektes wieder. Zum Beispiel, wie hoch ist der Mittelwertsunterschied.
- E ☐ In der Analogie der Wahrscheinlichkeit für Regen: ein statistischer Test erlaubt die Wahrscheinlichkeit für ein Ereignis abzuschätzen. Die Stärke des Effektes können wir nicht bestimmen.

19. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

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

20. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Welche Aussage über den Effekt eines statistischen Tests ist richtig?

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

21. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Roland Fischer entwickelte Anfang des letzten Jahrhunderts als Grundlage für das experimentelle Design in der Statistik die Randomisierung. Warum ist die Randomisierung für die Entscheidung anhand einer statistischen Auswertung so wichtig?

- ☐ A Randomisierung war bis 1952 bedeutend, wurde dann aber in Folge besserer Rechnerleistung nicht mehr verwendet. Aktuelle Statistik nutzt keine Randomisierung mehr.
- ☐ B Randomisierung erlaubt erst die Varianzen zu schätzen. Ohne eine Randomisierung ist die Berechnung von Mittelwerten und Varianzen nicht möglich. Dadurch lässt sich erst ein Experiment auswerten.
- ☐ C Randomisierung erlaubt erst die Mittelwerte zu schätzen. Ohne Randomisierung keine Mittelwerte. Ohne Mittelwerte keine Varianz und somit auch kein statistischer Test.
- ☐ D Durch eine Randomisierung können wir von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.
- ☐ E Durch eine Randomisierung können wir nicht von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.

22. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

- ☐ A Es gibt ein 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.
- ☐ B Mit der Nullhypothese H_A und der Alternativhypothese H_0 gibt es zwei Hypothesen, die aber selten genutzt werden.
- ☐ C Mit der Nullhypothese H_0 und der Alternativhypothese H_A oder H_1 gibt es zwei Hypothesen.
- ☐ D Es gibt - bedingt durch das das Falsifikationsprinzip - ein Set von k Nullhypothesen, die iterative gegen $k - 1$ Alternativhypothesen getestet werden.
- ☐ E Ein statistisches Hypothesenpaare gibt es. Zum einen die Nullhypothese und zum anderen die Alternativhypothese. Es ist aber nur notwendig die Alternative anzugeben, da die Nullhypothese nicht beim Testen benötigt wird.

23. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

H_0 ablehnen obwohl die H_0 gilt

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

24. Aufgabe

(2 Punkte)

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

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

25. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

Sie werten in Ihrer Abschlussarbeit einen sehr großen Datensatz aus einer öffentlichen Datenbank aus. Nun stellen Sie fest, dass Sie ein Problem mit der Bewertung Ihrer Ergebnisse anhand der Signifikanz bekommen. Wie Sie herausfinden, scheint dies ein häufiges Problem in der Bio Data Science zu sein. Welche Aussage ist richtig?

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

26. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

- A ☐ Untereinander unabhängig. Sollten die Mütter verwandt sein, so ist die Varianzstruktur ähnlich und muss modelliert werden.
- 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 ☐ Abhängig von der Stallanlage und des Experiments können die Ferkel abhängig oder unabhängig sein. Allgemein gilt, dass Ferkel von unterschiedlichen Sauen näher miteinander verwandt sind als Ferkel von gleichen Sauen. Das Fisher-Axiom.
- E ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.

27. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

Sie rechnen einen statistischen Test und wollen anhand des 95%-Konfidenzintervalls eine Entscheidung gegen die Nullhypothese treffen. Welche Aussage ist richtig?

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

28. Aufgabe

(2 Punkte)

Inhalt folgender Module: Biostatistik

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

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

Statistische Tests für Gruppenvergleiche

29. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik

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

- A ☐ Der t-Test testet generell zu einem erhöhten α -Niveau von 20%.
- B ☐ Der t-Test vergleicht die Varianzen von mindestens zwei oder mehr Gruppen
- C ☐ Der t-Test vergleicht zwei Gruppen indem die Mittelwerte miteinander verglichen werden.
- D ☐ Der t-Test vergleicht zwei oder mehr Gruppen indem die Mittelwerte miteinander verglichen werden.
- E ☐ Der t-Test vergleicht die Mittelwerte von zwei Gruppen unter der strikten Annahme von Varianzhomogenität. Sollte keine Varianzhomogenität vorliegen, so gibt es keine Möglichkeit den t-Test in einer Variante anzuwenden.

30. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

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

31. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung des Mikronährstoff Eisen auf den Ertrag in t/ha von Weizen im Vergleich zu einer Kontrolle. Der Versuch wurde in 8 Parzellen pro Gruppe durchgeführt. Welche Aussage im Bezug auf eine statistische Auswertung ist richtig?



- A ☐ Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei 0.1.
- B ☐ Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt kann nicht bei einem t-Test aus Barplots bestimmt werden.
- C ☐ Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei 1.
- D ☐ Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt liegt bei 1.
- E ☐ Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei 1. Wir müssen aber einen Posthoc-Test rechnen um den Effekt wirklich bestimmen zu können.

32. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

In Ihrer Abschlussarbeit passen die Ergebnisse einer ANOVA und eines multiplen Vergleiches nicht zusammen. Nach einem Experiment mit vier Maissorten ergibt eine ANOVA ($p = 0.045$). Sie führen anschließend die paarweisen t-Tests für alle Vergleiche durch. Nach der Adjustierung für multiples Testen ist kein p-Wert unter der α -Schwelle. Sie schauen sich auch die rohen, unadjustierten p-Werte an und finden hier als niedrigsten p-Wert $p_{3-2} = 0.053$. Welche Aussage ist richtig?

- A ☐ Die ANOVA testet auf der gesamten Fallzahl. Es wäre besser die ANOVA auf der gleichen Fallzahl wie die einzelnen t-Tests zu rechnen.
- B ☐ Der Fehler liegt in den t-Tests. Wenn eine ANOVA signifikant ist, dann muss zwangsweise auch ein t-Test signifikant sein.
- C ☐ Hier kommt der Effekt der steigenden Fallzahl auf die Anzahl an signifikante Ergebnisse zu tragen. Da die ANOVA auf mehr Fallzahl testet als die einzelnen paarweisen t-Tests, kann die ANOVA leichter einen signifikanten Unterscheid nachweisen. Die p-Werte sind immer etwas kleiner als bei den t-Tests.
- D ☐ Hier kommt der Effekt der steigenden Fallzahl auf die Anzahl an signifikante Ergebnisse zu tragen. Da die ANOVA auf weniger Fallzahl testet als die paarweisen t-Tests, kann die ANOVA schwerer einen signifikanten Unterscheid nachweisen.
- E ☐ 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.

ANOVA

33. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.31$. Welche Aussage ist richtig?

- A ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird.

- B** ☐ Das η^2 beschreibt den Anteil der globalen Varianz, der von den Umweltbedingungen erklärt wird.
- C** ☐ Das η^2 ist die Korrelation der ANOVA. Mit der Ausnahme, dass $\eta^2 = 0$ der beste Wert ist.
- D** ☐ Der Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird, wird durch das $1 - \eta^2$ beschrieben.
- E** ☐ Das η^2 ist ein Wert für die Güte der ANOVA. Je kleiner desto besser. Ein η^2 von 0 bedeutet ein perfektes Modell mit keiner Abweichung. Die Varianz ist null.

34. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

Nach einem Feldexperiment erhalten Sie ein $\eta^2 = 0.23$ aus Ihrer einfaktoriellen ANOVA. Sie haben sich den Chlorophyllgehalt in Kartoffel angeschaut und wollen nun das Ergebnis anhand des η^2 bewerten. Welche Aussage ist richtig?

- A** ☐ Mit dem η^2 lässt sich auf die Qualität der Randomisierung und damit der Strukturgleichheit zwischen der Grundgesamtheit und der Stichprobe schließen. Es gilt dabei die Regel, dass ein η^2 -Wert von 1 zu bevorzugen ist.
- B** ☐ Es werden 23% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen erklärt wird.
- C** ☐ Es werden 23% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- D** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 23% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 77%.
- E** ☐ Es werden 77% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.

35. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik


Die einfaktorielle ANOVA ist ein Standardverfahren in der agrarwissenschaftlichen Forschung wenn es um den Vergleich von Behandlungsgruppen geht. Welche der folgenden Aussage zu der Berechnung der Teststatistik der einfaktoriellen 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** ☐ 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.
- 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 F-Statistik indem die MS der Behandlung durch die MS des Fehlers geteilt werden. Wenn die F-Statistik sich der 0 annähert kann die Nullhypothese nicht abgelehnt werden.

36. Aufgabe

(2 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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

- A** ☐ ... den Unterschied zwischen der Varianz über alle Behandlungsgruppen oder der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss sich zwischen einem der beiden Varianzquellen entschieden werden.
- B** ☐ ... den Unterschied zwischen der F-Statistik anhand der Varianz der Gruppen. Wenn die F-Statistik exakt 0 ist, kann die Nullhypothese abgelehnt werden.
- C** ☐ ... den Unterschied zwischen der Varianz über alle Behandlungsgruppen und der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss ein Posthoc-Test angeschlossen werden.

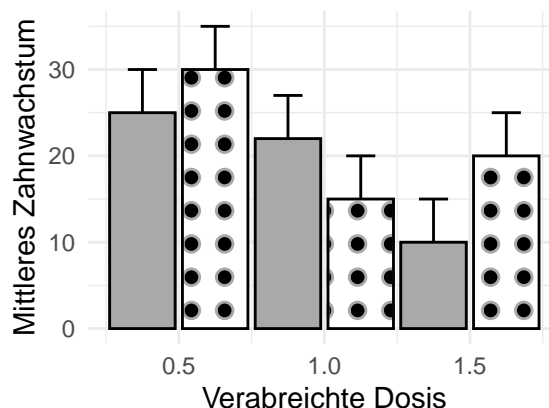
- D** ☐ ... den Unterschied zwischen zwei paarweisen Mittelwerten aus verschiedenen Behandlungsguppen. Wenn die signifikant ist, ist daher bekannt welcher Vergleich konkret unterschiedlich ist.
- E** ☐ ... den Unterschied zwischen der Varianz in den verschiedenen Behandlungsguppen und der Varianz in einer der Behandlungsguppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in einer der Gruppen exakt zu bestimmen.

37. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

Ein Versuch wurde an 50 Tieren durchgeführt, wobei jedes Tier eine von drei Vitamin-C-Dosen (0.5, 1 und 1.5 mg/Tag) über eine von zwei Verabreichungsmethoden erhielt. Die folgende Abbildung enthält die Daten aus diesem Versuch zur Bewertung der Wirkung von Vitamin C auf das Zahnwachstum bei Kanarienvögel. Welche Aussage ist richtig, wenn Sie eine zweifaktorielle ANOVA rechnen?



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

Multiple Gruppenvergleiche

38. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

Sie haben folgende unadjustierten p-Werte gegeben: 0.42, 0.03, 0.21, 0.02 und 0.01. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 2.1, 0.15, 1.05, 0.1 und 0.05. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 1, 0.15, 1, 0.1 und 0.05. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.
- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 1, 0.15, 1, 0.1 und 0.05. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.084, 0.006, 0.042, 0.004 und 0.002. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.084, 0.006, 0.042, 0.004 und 0.002. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.

39. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

- ☐ **A** Compact letter display. Gleiche Buchstaben bedeuten, dass sich die Behandlungen unterscheiden. Daher ist das CLD sehr unintuitiv. Es wäre besser, wenn gleiche Buchstaben Gleichheit anzeigen würden. Dies ist aber leider in der statistischen Testtheorie nicht möglich.
- ☐ **B** 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.
- ☐ **C** 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.
- ☐ **D** Contrast letter display. Unterschiede in den Behandlungen werden durch den gleichen Buchstaben oder Symbol dargestellt. Die Interpretation des CLD führt häufig in die Irre.
- ☐ **E** Compact letter display. Teilweise ist die Interpretation des CLD schwierig, da wir ja nach Unterschieden suchen aber nur Gleichheit in den Buchstaben sehen. Die Gleichheit der Behandlungen wird durch gleiche Buchstaben dargestellt.

40. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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 {lm}. Das Paket {lm} erstellt selbstständig Konfidenzintervalle und entsprechende p-Werte. Da wir in dem Paket nicht adjustieren müssen, ist es bei Anwendern sehr beliebt.
- ☐ **B** Das R Paket {hmisc} erlaubt die Durchführung eines multiplen Gruppenvergleichs aus verschiedenen Modellen heraus. Aus einem hmisc Objekt lässt sich recht einfach das CLD erstellen und so über Barplots eine schnelle Interpretation der statistischen Auswertung durchführen.
- ☐ **C** Das R Paket {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 recht einfach das CLD erstellen und so über Barplots eine schnelle Interpretation der statistischen Auswertung durchführen.
- ☐ **E** 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.

41. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

- ☐ **A** Beim multiplen Testen kann es zu einer α -Deflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist. Die p-Werte werden durch die Anzahl an Vergleichen geteilt.
- ☐ **B** Beim multiplen Testen kann es zu Varianzheterogenität kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5%. Daher müssen die p-Werte entsprechend adjustiert werden. Das Verfahren nach Welch, bekannt aus dem t-Test, ist hier häufig anzuwenden.
- ☐ **C** Das globale Signifikanzniveau explodiert und erreicht Werte größer als Eins. Es kommt zu einer α -Inflation. Dagegen kann mit der Adjustierung der α -Werte nach Bonferroni vorgegangen werden.

- D** ☐ Beim multiplen Testen kann es zu einer α -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Welch das bekannteste Verfahren ist.
- E** ☐ Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern sehr viel höher. Es kommt zu einer α -Inflation. Dagegen kann mit der Adjustierung der p-Werte nach Bonferroni vorgegangen werden.

42. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

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

Lineare Regression & Korrelation

43. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

- A** ☐ Es wird ein Trainingsdatensatz zum Modellieren des Trainingsmodells benötigt. Der Testdatensatz dient rein zur Visualisierung. Dies gilt vor allem für ein kausales Modell.
- B** ☐ Es wird ein Trainingsdatensatz zum Trainieren des Modells benötigt. Der Testdatensatz dient zur Validierung. Dies gilt insbesondere für ein kausales Modell.
- C** ☐ 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.
- D** ☐ 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.
- E** ☐ Wir modellieren den Zusammenhang zwischen X und Y wenn ein kausales Modell rerechnet wird. Dabei kann der gesamte Datensatz genutzt werden. Eine Aufteilung wie in einem prädiktiven Modell ist nicht notwendig.

44. Aufgabe

(2 Punkte)

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

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



- A** ☐ Wir betrachten die Gerade, die durch die einzelnen Punkte laufen sollte. Wenn die 95% der Punkte von der Geraden getroffen werden, dann gehen wir von normalverteilten Residuen aus.
- B** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Geraden.
- 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.

45. Aufgabe

(2 Punkte)

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik

In den Humanwissenschaften wird der Korrelationskoeffizienten ρ sehr häufig verwendet. Daher ist es auch wichtig für andere Forschende den Korrelationskoeffizienten ρ zu verstehen. Welche Aussage zu dem Korrelationskoeffizienten ρ ist richtig?

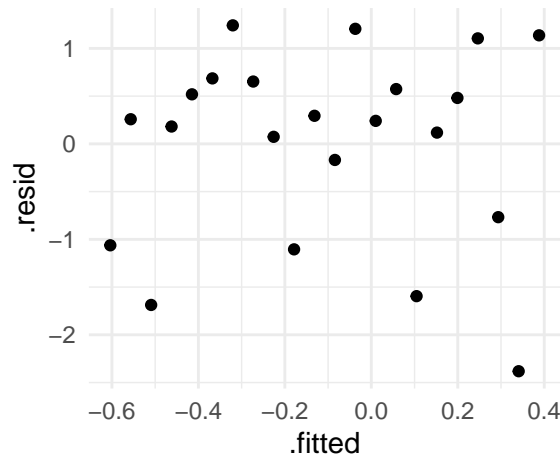
- A** ☐ Der Korrelationskoeffizienten ρ liegt zwischen -1 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als standardisierte Steigung verstanden werden.
- B** ☐ Der Korrelationskoeffizienten ρ zeigt keinen Zusammenhang zwischen zwei Variablen x und y bei einem Wert von 0. Einen negativen Zusammenhang Richtung -1 und somit auch einen positiven Zusammenhang Richtung 1. Je größer die Zahl allgemein, desto stärker der Effekt.
- C** ☐ Der Korrelationskoeffizienten ρ liegt zwischen -1 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ als standardisierte Steigung zu verstehen, wenn eine Standardisierung durchgeführt wurde. Diese Adjustierung nach Fischer muss am Anschluß der Berechnung der Korrelation durchgeführt werden.
- D** ☐ Der Korrelationskoeffizienten ρ ist eine veraltete Darstellungsform von Effekten in der linearen Regression und wird wie das η^2 aus der ANOVA interpretiert. Der Korrelationskoeffizienten ρ beschreibt den Anteil an erklärter Varianz durch die Regression.
- E** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen 0 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos. Eine Signifikanz kann nicht nachgewiesen werden.

46. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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



- A** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Kein Muster ist zu erkennen und keine Outlier zu beobachten.
- B** ☐ Die Punkte müssen gleichmäßig in dem negativen Bereich liegen. Dies ist hier klar nicht der Fall. Einzelne Ausreißer können beobachtet werden. Die Analyse ist gescheitert.
- C** ☐ 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 .
- D** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Diagonalen. Damit ist das Modell erfolgreich geschätzt worden.
- E** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Es ist ein Muster zu erkennen und wir können damit auf die Signifikanz von x_1, \dots, x_p schließen.

47. Aufgabe

(2 Punkte)

Inhalt folgender Module: Biostatistik

In einer linearen Regression kann es vorkommen, dass der Effekt repräsentiert durch den β Koeffizienten nicht so richtig von der Größenordnung zu dem p-Wert passen will. So liefert eine Untersuchung des Einflusses von der PO_2 -Konzentration in $[\mu g]$ im Wasser auf das Trockengewicht in $[kg]$ an Erbsen folgende Effekte und p-Werte: $1e-04$ als p-Wert und einen β_{PO_2} Koeffizienten von 7.4×10^{-6} . Welche Aussage ist richtig?

- A** ☐ Wenn der Effekt β_{PO_2} winzig ist, dann kann es an einer falsch gewählten Einheit liegen. Der Anstieg von einer Einheit in X führt ja zu einer Änderung von β_{PO_2} in x . Wir müssen daher die Einheit von y entsprechend anpassen.
- B** ☐ Die Einheit der PO_2 -Konzentration ist zu klein gewählt. Dadurch sehen wir den sehr kleinen p-Wert. Der p-Wert und die Einheit von der PO_2 -Konzentration hängen antiproportional zusammen.
- C** ☐ Das Gewicht und die PO_2 -Konzentration korrelieren sehr stark, deshalb wird der β_{PO_2} Koeffizient sehr klein. Mit einer ANOVA kann für die Korrelation korrigiert werden und der Effektschätzer passt dann zum p-Wert.
- D** ☐ Die Fallzahl ist zu klein angesetzt. Je kleiner die Fallzahl ist, desto höher ist die Teststatistik und damit auch der p-Wert kleiner. Wir brauchen also mehr Fallzahl um den geringen Effekt noch signifikant zu kriegen.
- E** ☐ Die Einheit der PO_2 -Konzentration ist zu klein gewählt. Die Erhöhung der PO_2 -Konzentration um 1 Einheit führt nur zu einem sehr winzigen Anstieg von β_{PO_2} im Gewicht der Wasserlinsen. Die Einheit $[\mu g]$ muss besser gewählt werden.

48. Aufgabe

(2 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik

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

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

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

49. Aufgabe

(2 Punkte)

Inhalt folgender Module: Biostatistik

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 erlaubt auch nicht normalverteilte Residuen in der Schätzung der Regressionsgrade.
- B** ☐ Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsfamilien für das Y bzw. das Outcome in einer linearen Regression zu wählen.
- C** ☒ 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.
- D** ☐ Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsgruppen für das X bzw. die Einflussvariablen in einer linearen Regression zu wählen.
- E** ☐ Das GLM ist ein faktisch maschineller Lernalgorithmus, der selbstständig die Verteilungsfamilie für Y wählt.

50. Aufgabe

(2 Punkte)

Inhalt folgender Module: Biostatistik

In einer Studie wollen Sie den Effektschätzer Risk ratio berechnen. Sie finden in Ihrem Experiment zur Behandlung von Klaueninfektionen bei Ziegen in 4 Tieren Erkrankung der Klauen vor. 12 Tiere sind gesund. Welche Aussage ist richtig?

- A** ☐ Es ergibt sich ein Risk ratio von 0.25, da es sich um eine Chancenverhältnis handelt.
- B** ☐ Es ergibt sich ein Risk ratio von 0.33, da es sich um ein Anteil handelt.
- C** ☐ Das Verhältnis von Chancen Risk ratio ergibt ein Chancenverhältnis von 0.33.
- D** ☐ Das Verhältnis der Anteile Risk ratio ergibt ein Anteilsverhältnis von 0.25. Wir sind am Anteil der Kranken interessiert.
- E** ☐ Der Anteil der Gesunden wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Risk ratio von 0.25.

Teil I.

Programmieren in R

51. Aufgabe

(9 Punkte)

Inhalt folgender Module: Mathematik & Statistik • Modellierung landwirtschaftlicher Daten • Statistik



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

Paula: *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:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

Paula: *Ich habe die Namen der beiden R Pakete vergessen, die wir eigentlich immer laden. Wie heißen die noch gleich? (1 Punkt)*

Sie antworten:

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

Sie antworten:

52. Aufgabe

(9 Punkte)

Inhalt folgender Module: Modellierung landwirtschaftlicher Daten • Angewandte Statistik und Versuchswesen • Biostatistik



Fortgeschrittene Kenntnisse der Programmierung in R Jonas muss seiner 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 er jetzt mit Ihnen in der Küche sitzt und einige vertiefende Fragen zu R an Sie hat! Na dann wollen Sie mal helfen. Immerhin will sein Betreuer, dass R genutzt wird und die Abgabe ist dann auch schon in gut einem Monat.

Jonas fragt: *Wozu war nochmal die Funktion `mutate()` gut?* **(1 Punkt)**

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

Teil II.

Deskriptive Statistik & Explorative Datenanalyse

53. Aufgabe

(9 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen
Modellierung landwirtschaftlicher Daten • Biostatistik



Zerforschen des Barplots Nilufar steht vor einem ersten Problem, denn wenn es nach ihrer Betreuerin geht, soll sie in einem Gewächshausexperiment Erdbeeren auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Am Ende dann doch besser Hip Hop. Wunderbar. Eine echte Ablenkung für Nilufar. Das heißt erstmal überlegen für Nilufar. Aus den Boxen wummert Deichkind und ihr Mund ist verklebt von Takis Blue Heat. 'Herrlich', denkt Nilufar. Die Behandlung werden verschiedene Genotypen (AA, AB und BB) sein. In ihrer Exceldatei wird sie den Endpunkt (Y) *Frischegewicht* als *freshmatter* aufnehmen. Vorab soll Nilufar aber einmal die folgenden Barplots ihrer Betreuerin nachbauen, damit sie den R Code schonmal für später vorliegen hat. Damit geht das Problem schon los. Eine echte Herausforderung für sie war schon immer die Erwartung gewesen. Ein leidiges Lied.



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

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

54. Aufgabe

(8 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Modellierung landwirtschaftlicher Daten • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Visualisierung des Barplots Tina steht vor einem ersten Problem, denn wenn es nach ihrer Betreuerin geht, soll sie in einem Feldexperiment Maiss auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Tina liebt Astronomie. Darin kann sie sich wirklich verlieren und immer wieder neu begeistern. Die Behandlung waren verschiedene Lichtstufen (*none*, *200lm* und *600lm*). In ihrer Exceldatei hat sie den Endpunkt (Y) *Proteingehalt* als *protein* aufgenommen. Nun soll Tina die Daten einmal als Barplots in einer Präsentation visualisieren, damit ihrer Betreuerin 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. Tina und die Wut, eine unendliche Geschichte mit kniffligen Wendungen. Aber egal. Einfach mal raus um zu Boxen. Ohne Ziel und Uhr. Das ist was für Tina.

treatment	protein
200lm	49.6
none	7.4
200lm	49.8
200lm	35.0
600lm	39.1
none	29.9
none	16.6
600lm	27.7
600lm	37.2
none	20.4
600lm	36.2
none	18.8
600lm	52.5

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

55. Aufgabe

(9 Punkte)

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


Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen
Modellierung landwirtschaftlicher Daten • Biostatistik



Zerforschen des Boxplots Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Mark ist es eine Möglichkeit schneller ans Ziel zu gelangen. Mark soll in seiner Hausarbeit Erdbeeren untersuchen. Die Behandlung in seiner Hausarbeit werden verschiedene Lichtstufen (*none*, *200lm* und *600lm*) sein. Erheben wird Mark als Messwert (*Y*) *Proteingehalt* benannt als *protein* in seiner Exceldatei. Von seinem Betreuer 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! Hm, lecker Marzipankugeln und dazu dann im Hintergrund Columbo laufen lassen.



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

56. Aufgabe

(9 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Modellierung landwirtschaftlicher Daten • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Visualisierung des Boxplots Boxplots sind bedeutend in der Darstellung von wissenschaftlichen Ergebnissen. Leider hat sich Alex nicht gemerkt, welche statistischen Maßzahlen für einen Boxplot erhoben werden müssen. Besser wäre was anderes gewesen. Alex liebt Starcraft. Darin kann er sich wirklich verlieren und immer wieder neu begeistern. Das ist in soweit doof, da nach seinem Betreuer nun Boxplots aus seinen Daten gebaut werden sollen, bevor es mit dem statistischen Testen weitergeht. Anhand von Boxplots lässt sich eine Aussage über die Normalverteilung von Y treffen. Die Behandlung für Erdbeeren waren verschiedene Bewässerungstypen (*low* und *high*). Erfasst wurde von Alex als Messwert (Y) *Proteingehalt*. Alex hat dann *protein* in seiner Exceldatei eingetragen. Aber nur in passender Atmosphäre! Auf seinem Second Screen läuft Alien und Alex schaufelt Gummibärchen. Nicht effizient, aber gut.

treatment	drymatter
high	34.5
high	18.4
low	33.9
high	23.8
low	35.5
low	33.5
high	27.9
high	24.8
low	38.1
high	33.1
high	39.6
low	31.6
low	38.2
high	30.3
low	37.1
low	29.8
low	29.2
high	34.3
low	28.0
high	33.9

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Zeichnen Sie in *einer* Abbildung die beiden Boxplots für die zwei Behandlungen von Erdbeeren! Beschriften Sie die Achsen entsprechend! **(4 Punkte)**
3. Wie ist Ihr Vorgehen, wenn Sie eine *gerade* Anzahl an Beobachtungen pro Gruppe haben? **(1 Punkt)**
4. Beschriften Sie *einen* der beiden Boxplots mit den gängigen statistischen Maßzahlen! **(2 Punkte)**
5. 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)**

57. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Visualisierung des Scatterplots Alex liest laut: 'Wenn zwei kontinuierliche Variablen vorliegen, können diese in einer explorativen Datenanalyse...'. Alex stoppt. Alex schmeißt noch eine Handvoll Gummibärchen in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Abba. Was waren noch gleich kontinuierliche Variablen? In seiner Abschlussarbeit hatte er ein Stallexperiment im Oldenburger Land durchgeführt. Dabei ging es um den Zusammenhang zwischen Gewichtszuwachs in der 1LW und durchschnittlichen Bewegungsscore [Movement/h] im groben Kontext von Hühnern. Nun stellt sich die Frage für ihn, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Dafür war eine explorative Datenanalyse gut! Wenn die Gefälligkeit nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so.. Dann was anderes. Irgendwie komisch, wenn er Alien anmacht, dann ist die Katze eigentlich sofort vor dem Bildschirm und starrt hinein.

Durchschnittlichen Bewegungsscore [Movement/h]	Gewichtszuwachs in der 1LW
25.2	22.0
25.1	19.8
23.9	21.3
27.5	22.9
25.6	19.6
28.3	21.4
25.7	18.9
23.7	19.1
24.7	21.1
26.1	21.0
22.7	21.1

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

1. Erstellen Sie eine Visualisierung für die Datentabelle. Beschriften Sie die Achsen entsprechend! **(4 Punkte)**
2. Schätzen Sie eine Gerade durch die Punkte! **(1 Punkt)**
3. Beschriften Sie die Gerade mit den gängigen statistischen Maßzahlen! Geben Sie die numerischen Zahlenwerte mit an! **(3 Punkte)**
4. Wenn *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!



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Visualisierung des Mosaicplots Das Verrückte ist, dass die Katze Alien wirklich liebt. Das ist Alex sehr recht, denn er braucht Entspannung. Aber Ablenkung hilft nur begrenzt. 'Uff!', denkt sich Alex. Jetzt hat er doch tatsächlich zwei kategoriale Variablen in seiner Abschlussarbeit gemessen. Zum einen die Behandlung Pestizideinsatz [ja/nein] und zum anderen die Messung Trockengewicht über Zielwert [ja/nein] im Kontext von Erbsen. Hierfür hat er ein Freilandversuch in der Uckermark durchgeführt. Jetzt möchte Alex die Daten einmal in einer explorativen Datenanalyse darstellen. Danach kann er dann über den passenden statistischen Test nachdenken. Dabei unterstützt seine Betreuerin diesen Ansatz bevor es in der Datenanalyse weiter geht. So schön wie so gut. Wenn die Gefälligkeit nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so..

Pestizideinsatz	Trockengewicht über Zielwert	Pestizideinsatz	Trockengewicht über Zielwert
nein	ja	nein	ja
ja	nein	ja	ja
nein	ja	nein	nein
nein	ja	nein	ja
ja	ja	ja	nein
nein	ja	nein	nein
nein	nein	nein	nein
ja	ja	ja	ja
ja	nein	ja	ja
ja	nein	nein	ja
nein	ja	ja	ja
ja	nein	nein	ja
ja	ja	ja	ja
ja	ja	ja	ja
ja	ja	nein	ja
nein	ja	nein	nein

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

59. Aufgabe

(8 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Visualisierung des Histogramm für kategoriale Daten In einem Gespräch mit ihrer Betreuerin wird Tina gebeten seine Daten aus einem Leistungssteigerungsversuch mit Schweinen in einem Histogramm darzustellen. 'Hm...', Katjes und Tocotronic. Das ist und bleibt die beste Kombination zum Nachdenken für Tina. In ihrem Experiment hat er die dunklen Pigmentstörungen erst fotografiert und dann ausgezählt. Laut ihrer Betreuerin soll das Histogramm helfen, die Verteilung der die dunklen Pigmentstörungen zu bestimmen. Es wäre einfacher, wenn da nicht noch was wäre. Tina und die Wut, eine unendliche Geschichte mit kniffligen Wendungen. Tina streichelt liebevoll die Spinne. Der Kopf ist in ihrem Schloß vergraben um den Klang von Tocotronic zu dämpfen.

Die dunklen Pigmentstörungen: 6, 1, 5, 1, 3, 2, 3, 3, 1, 3, 4, 3, 4, 4, 5, 6, 5, 5, 3, 5, 5, 3, 5, 1, 5, 5, 5, 6, 8, 9, 6, 6, 5, 7, 5, 2

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

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

60. Aufgabe

(8 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Visualisierung des Histogramm für kontinuierliche Daten Nilufar schmeißt noch eine Handvoll Takis Blue Heat in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Deichkind. Nilufar betrachtet die folgenden Daten nach einem Stallexperiment mit Schweinen. In dem Experiment wurden die mittlere Anzahl an weißen Blutkörperchen gezählt. Nach der Meinung ihrer Betreuerin muss als erstes geschaut werden, wie diese verteilt sind. Also welcher statistischen Verteilung die mittlere Anzahl an weißen Blutkörperchen folgen. Dazu soll Nilufar ein Histogramm verwenden. Dann hätte man auch einen guten Überblick über den Messwert (Y). Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für sie war schon immer die Erwartung gewesen. Ein leidiges Lied. Nilufar streichelt liebevoll das Huhn. Der Kopf ist in ihrem Schloß vergraben um den Klang von Deichkind zu dämpfen.

Die mittlere Anzahl an weißen Blutkörperchen: 11.4, 7.4, 8.3, 9.5, 9.8, 8.5, 9.8, 7.3, 9.7, 11.2, 10.6, 10.4, 12.9, 13.1, 12.2, 10.1, 9.8, 10.4, 8.8, 10.3, 9.3, 10.3, 12.6, 10.1, 9.3, 12.7, 10.4, 11.1

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

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

61. Aufgabe

(10 Punkte)

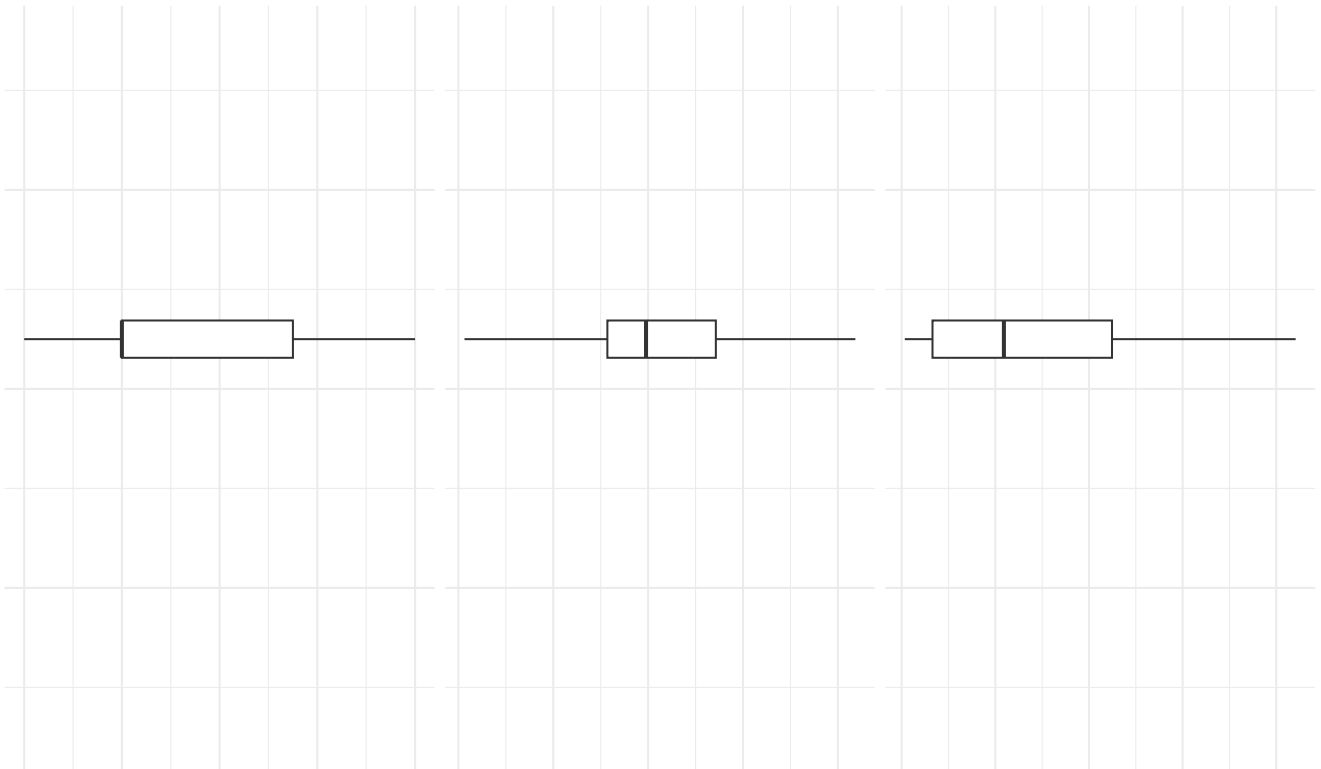
Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



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



Jetzt brauchen Alex und Paula 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 Striche! **(3 Punkte)**
3. Wie viel Prozent der Beobachtungen fallen in das IQR? Ergänzen Sie die Abbildung entsprechend um den Bereich! **(2 Punkte)**
4. Wie viel Prozent der Beobachtungen fallen in $\bar{y} \pm 1s$ und $\bar{y} \pm 2s$ unter der Annahme einer Normalverteilung? **(2 Punkte)**

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

Teil III.

Statistisches Testen & statistische Testtheorie

62. Aufgabe

(9 Punkte)



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



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

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

1. Erklären Sie den Zusammenhang zwischen Stichprobe und Grundgesamtheit an einem Schaubild! Beschriften Sie das Schaubild entsprechend! *Nutzen Sie hierfür als Veranschaulichung ein aussagekräftiges Beispiel!* **(3 Punkte)**
 - a) Nennen Sie das statistische Verfahren um von einer Grundgesamtheit auf eine Stichprobe zu gelangen! **(1 Punkt)**
 - b) Nennen Sie ein konkretes Beispiel zur Durchführung um von einer Grundgesamtheit auf eine Stichprobe zu gelangen! **(1 Punkt)**
 - c) Benennen Sie die statistische Eigenschaft, die zwischen Grundgesamtheit und Stichprobe vorliegen muss! **(1 Punkt)**
2. Erweitern Sie das Schaubild um die Entstehung von $Pr(D|H_0)$! *Nutzen Sie hierfür als Veranschaulichung ein aussagekräftiges Beispiel!* **(3 Punkte)**

63. Aufgabe

(9 Punkte)

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Das Nullritual - Die statistische Testtheorie 'Alien ist der beste Film, den es gibt.', meint Alex. Nilufar entgegnet, 'Ich empfehle ja immer allen Star Trek.' Die beiden sind im Zoo und diskutieren, ob Pinguine Knie haben. Eigentlich wollten beide nochmal die statistische Testtheorie durchgehen, sind dann aber auf abenteuerlichen Wege im Zoo gelandet. Alex starrt wie hypnotisiert auf einen strullenden Elefanten und stoppt die Zeit.² 'Du bist so peinlich.', entfährt es Nilufar und schmeißt sich noch ein paar überteuerte Takis Blue Heat rein.

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

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

H_0 abgelehnt 5% H_0 falsch Richtige Entscheidung

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

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

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

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

64. Aufgabe

(9 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



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

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

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

1. Ergänzen Sie eine beschriftete x-Achse! **(1 Punkt)**
2. Ergänzen Sie „ $\bar{y}_1 = \bar{y}_2$ “! **(1 Punkt)**
3. Ergänzen Sie „ $A = 0.95$ “! **(1 Punkt)**
4. Zeichnen Sie $T_{\alpha=5\%}$ in die Abbildung! **(1 Punkt)**
5. Zeichnen Sie das Signifikanzniveau α in die Abbildung! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Zeichnen Sie $-T_D$ in die Abbildung! **(1 Punkt)**
7. Zeichnen Sie einen signifikant p-Wert in die Abbildung! Begründen Sie Ihre Antwort! **(2 Punkte)**



65. Aufgabe

(10 Punkte)



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



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

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

1. Beschriften Sie die untenstehende Abbildung mit der Signifikanzschwelle! Begründen Sie Ihre Antwort! **(2 Punkte)**
2. Ergänzen Sie eine *in den Kontext passende* Relevanzschwelle! Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Skizzieren Sie in die untenstehende Abbildung sechs einzelne Konfidenzintervalle (a-f) mit den jeweiligen Eigenschaften! **(6 Punkte)**
 - (a) Ein signifikantes, relevantes 99% Konfidenzintervall.
 - (b) Ein 95% Konfidenzintervall mit niedriger Fallzahl n in der Stichprobe als der Rest 95% der Konfidenzintervalle
 - (c) Ein signifikantes, nicht relevantes 95% Konfidenzintervall
 - (d) Ein nicht signifikantes, nicht relevantes 95% Konfidenzintervall
 - (e) Ein signifikantes, relevantes 95% Konfidenzintervall
 - (f) Ein 95% Konfidenzintervall mit höherer Fallzahl n in der Stichprobe als der Rest der 95% Konfidenzintervalle



66. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



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

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

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

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

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

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

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

Teil IV.

Der t-Test

67. Aufgabe

(9 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Berechnung des Student t-Test oder Welch t-Test Tina ist im Oldenburger Land für einen Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) mit Lauch. Allein diese Tatsache ist für sie eine Erzählung wert. Wenn die Wut nicht wäre, ja dann wäre wohl vieles möglich für Tina! Aber so.. Für ihrer Hausarbeit musste sie ein Gewächshausexperiment mit Lauch 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 Lüftungssystemen und Folientunneln (*ctrl* und *tornado*) und der Messwert Trockengewicht [kg/ha] sollen mit einem t-Test ausgewertet werden. Immerhin weiß sie, dass ihr Messwert einer Normalverteilung folgt. Hm..., was entspannendes wäre gut. Einfach mal raus um zu Boxen. Ohne Ziel und Uhr. Das ist was für Tina.

Lüftungssystemen	Trockengewicht
ctrl	10.5
tornado	24.9
tornado	19.1
ctrl	9.0
tornado	19.3
ctrl	16.2

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 1.96$ 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 Tina über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

68. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Berechnung des Student t-Test Der t-Test. Jessica erschauert. Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen. Ein mächtiges Werkzeug ist der t-Test in den Händen desjenigen, der ein normalverteiltes Outcome (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Jessica überhaupt aus? Jessica schmeißt noch eine Handvoll Schokobons in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von David Bowie. Jessica hat ein Feldexperiment mit Maiss durchgeführt. Dabei wurde die Behandlung Genotypen (AA und BB) an den Maiss getestet. Gemessen hat Jessica dann als Messwert Proteingehalt [g/kg]. Warum der Versuch im Emsland für ihre Abschlussarbeit stattfinden musste, ist ihr bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Proteingehalt [g/kg]?

Genotypen	Proteingehalt
AA	31.4
AA	33.8
AA	27.9
BB	37.6
BB	34.1
AA	27.7
AA	25.8
BB	15.8
BB	40.9
BB	32.7
BB	34.3
BB	33.3
BB	25.6
AA	30.8
AA	30.5
AA	30.4
BB	25.6
AA	34.3
AA	27.9
BB	27.4

Leider kennt sich Jessica 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 $T_{\alpha=5\%} = 2.04$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *einen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann der *mindeste* Effekt? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Jessica über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

69. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Berechnung des Welch t-Test Der t-Test. Steffen erschauert. Wenn die Romantik nicht wäre, ja dann wäre wohl vieles möglich für Steffen! Aber so.. Ein mächtiges Werkzeug ist der t-Test in den Händen desjenigen, der einen normalverteilten Endpunkt (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Steffen überhaupt aus? 'Hm...', Oreos und Taylor Swift. Das ist und bleibt die beste Kombination zum Nachdenken für Steffen. Steffen hat ein Kreuzungsexperiment mit Fleischrindern durchgeführt. Dabei wurde die Behandlung Flüssignahrung (*ctrl* und *fOw*) an den Fleischrindern getestet. Gemessen hat Steffen dann als Messwert Fettgehalt [%/kg]. Warum der Versuch in der Uckermark für seiner Hausarbeit stattfinden musste, ist ihm bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Fettgehalt [%/kg]?

Flüssignahrung	Fettgehalt
ctrl	39.4
ctrl	46.4
ctrl	31.6
fOw	16.3
fOw	29.3
ctrl	35.8
fOw	32.6
ctrl	36.5
ctrl	33.6
fOw	27.5
ctrl	32.1
fOw	24.9
fOw	27.3
ctrl	41.3
fOw	32.4
fOw	42.9
fOw	18.7
ctrl	42.6
fOw	14.6

Leider kennt sich Steffen 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 $T_{\alpha=10\%} = 1.96$ 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 Steffen über das Ergebnis Ihrer statistischen Analyse! Beziehen Sie dabei das 90% Konfidenzintervall mit in Ihre Antwort ein! **(2 Punkte)**

70. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Interpretation des t-Tests in R - die Teststatistik und der p-Wert Almería. Spanien. Sonne und Strand. Tina und Yuki haben ihren gemeinsamen Auslandsaufenthalt sichtlich genossen. Dann hatte sich auch noch angeboten ihre Abschlussarbeit gemeinsam in Almería durchzuführen. Es hätte sogar noch besser funktioniert, wenn Jonas nicht die Erschöpfung ein paar Mal im Weg gestanden hätte und Tina nicht das Problem gehabt hätte die Wut zu händeln. Nun müssen jetzt alle Daten in R ausgewertet werden, da R international der Standard in der Datenauswertung ist und die Betreuer in Spanien nur R können. Während beide Jonas Oliven mit Snickers füttern, hoffen Tina und Yuki mehr Informationen von Jonas über die seltsame R Ausgabe des t-Tests. Immerhin erinnern beide sich an die Behandlung Lüftungssystem (*keins* und *vorhanden*) und das es um Zandern ging. Im Hintergrund wummert Iron Maiden und Fotos zeigen Jonas mit dem Hobby Stricken.

```
##
## Two Sample t-test
##
## data: Fettgehalt by Lüftungssystem
## t = 2.5918, df = 16, p-value = 0.01967
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## 2.917484 29.132516
## sample estimates:
## mean in group keins mean in group vorhanden
## 35.350 19.325
```

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des t-Tests! **(1 Punkt)**
5. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.12|$ einzeichnen! **(4 Punkte)**
6. Beschriften Sie die Abbildung! **(1 Punkt)**

71. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Interpretation des t-Tests in R - das 95% Konfidenzintervall Almería. Spanien. Sonne und Strand. Tina und Nilufar haben ihren gemeinsamen Auslandsaufenthalt sichtlich genossen. Dann hatte sich auch noch angeboten ihre Abschlussarbeit gemeinsam in Almería durchzuführen. Es hätte sogar noch besser funktioniert, wenn Paula nicht der Perfektionismus ein paar Mal im Weg gestanden hätte und Tina nicht das Problem gehabt hätte die Wut zu händeln. Nun müssen jetzt alle Daten in R ausgewertet werden, da R international der Standard in der Datenauswertung ist und die Betreuer in Spanien nur R können. Während beide Paula Oliven mit Smarties füttern, hoffen Tina und Nilufar mehr Informationen von Paula über die seltsame R Ausgabe des t-Tests. Immerhin erinnern beide sich an die Behandlung Flüssignahrung (*ctrl* und *flOw*) und das es um Schweinen ging. Im Hintergrund wummert White Lies und Fotos zeigen Paula mit dem Hobby Harry Potter.

```
##  
## Two Sample t-test  
##  
## data: Fettgehalt by Flüssignahrung  
## t = 3.0442, df = 14, p-value = 0.00875  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## 2.876692 16.596324  
## sample estimates:  
## mean in group ctrl mean in group flOw  
## 32.52222 22.78571
```

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

72. Aufgabe

(9 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Interpretation des t-Tests in R - die Visualisierung 'Wir waren in der Uckermark um Puten in einem Kreuzungsexperiment zu messen.', Nilufar legt das Dokument auf den Tisch und schaut Steffen und Jonas fragend an. Beide schauen fragend zurück. Gäbe es die Erschöpfung nicht, dann wäre es für Jonas irgendwie einfacher hier zu helfen. Echt unangenehm. Die beiden sind zu Nilufar gekommen, da sie sich nicht mit R auskennen und daher Hilfe bei der Interpretation des t-Tests brauchen. Im Hintergrund wummert Deichkind und leere Takis Blue Heat Packungen stappeln sich auf dem Boden. 'Kein Problem', sagt Nilufar und streichelt langsam das Huhn. 'Aber worum es in dem Versuch geht, lässt sich nur aus dem Text in seiner Hand erraten.' merkt sie an. Vielleicht hilft da ja die Ausgabe des t-Tests in R weiter. Draußen geht blutrot die Sonne unter.

```
##  
## Two Sample t-test  
##  
## data: Protein/Fettrate by Lüftungssystem  
## t = 1.4874, df = 15, p-value = 0.1576  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -2.186461 12.286461  
## sample estimates:  
## mean in group keins mean in group vorhanden  
## 44.80 39.75
```

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

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

73. Aufgabe

(11 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Statistik und Versuchswesen • Biostatistik



Berechnung des gepaarten t-Test Jonas und Nilufar haben sich dazu entschieden zusammenzuarbeiten. Das sollte alles etwas einfacher machen. Jeder hat zwar ein getrenntes Themenfeld aber den Hauptversuch machen beide gemeinsam. Das hat sich schonmal als gut Idee soweit herausgestellt. In einem Projektbericht sollen beide herausfinden, ob es einen Zusammenhang zwischen Beschattung (7d und 14d) und Frischegewicht [kg/ha] gibt. Die Besonderheit ist hierbei, dass die Messungen an der gleichen Beobachtung stattfinden. Beide messen also zweimal an den gleichen Lauch. Hier muss dann wohl auf einen normalverteilten Endpunkt (Y) ein gepaarter t-Test gerechnet werden. Jonas schaut etwas flehentlich zu Nilufar. Jonas und die Erschöpfung, eine unendliche Geschichte mit kniffligen Wendungen.. Steffen denkt derweil angestrengt an Deichkind und wippt leicht mit dem Fuß.

ID	treatment	freshmatter
2	14d	33.0
4	14d	29.8
5	14d	31.5
4	7d	26.4
3	7d	24.9
1	7d	20.3
3	14d	26.8
7	14d	34.4
2	7d	21.6
8	14d	33.8
7	7d	32.4
6	7d	19.9
5	7d	27.4
6	14d	34.3
1	14d	37.5

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines gepaarten t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.96$ 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)**

74. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Statistik und Versuchswesen • Biostatistik



Interpretation des gepaarten t-Tests in R Es gibt ja immer die Möglichkeit sich Hilfe zu holen. Das geht natürlich auch immer in einer Abschlussarbeit. Deshalb arbeiten Mark und Jessica gemeinsam an einer Abschlussarbeit. Das macht dann auch die Analyse ihres Hauptversuches einfacher. Zwar hat jeder von ihnen noch ein Subthema, aber auch da kann man sich ja helfen. Das hilft dann teilweise nur bedingt. Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so.. In dem Hauptversuch wurde Folgendes von den beiden gemacht. Mark und Jessica haben sich Erbsen angeschaut. Dabei geht um Zusammenhang zwischen Beschattung (7d und 14d) und Trockengewicht [kg/ha]. Jetzt sollen beide einen gepaarten t-Test rechnen. Leider kennen sich beide nicht sehr gut in R aus. Aber wenigstens haben beide eine Menge an Marzipankugeln und in der Wohnung wummert Andrea Berg.

```
##  
## Paired t-test  
##  
## data: Trockengewicht by Beschattung  
## t = 0.68308, df = 8, p-value = 0.5138  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -6.150925 11.328702  
## sample estimates:  
## mean difference  
## 2.588889
```

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

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

Teil V.

Die einfaktorielle & zweifaktorielle ANOVA

75. Aufgabe

(11 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Visualisierung der einfaktoriellen ANOVA 'Uff... die einfaktorielle ANOVA. Und wir können jetzt anhand der Visualisierung sehen, ob da schon was signifikant ist?', Jonas hebt die Augenbraue. 'Ja, können wir. Dafür müssen wir aber erstmal in {ggplot} uns die Daten anschauen. Oder wir zeichnen es flott mit der Hand. Geht auch.', meint Mark dazu. Die beiden hatten sich auf einem Konzert von Andrea Berg kennengelernt. Jonas hatte sich in einen Leistungssteigerungsversuch verschiedene Schweinen angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Elterlinie (*Standard*, *Yray* und *Xray*) und dem Messwert Gewichtszuwachs in der 1LW gibt. Später wird noch Columbo geguckt. Mark befürwortet das!

Elterlinie	Gewichtszuwachs
Standard	30
Standard	31
Yray	41
Yray	34
Xray	34
Xray	34
Xray	35
Standard	30
Yray	41
Yray	41
Yray	38
Yray	38
Xray	35
Yray	36
Xray	34
Standard	31
Standard	31
Standard	30

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

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

76. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



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 Steffen fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Da hilft die Schlange von Steffen auch nur bedingt. Steffen tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren im Oldenburger Land um ein Gewächshausexperiment mit Erdbeeren durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Bewässerungstypen (*ctrl*, *low*, *mid* und *high*) und dem Messwert Trockengewicht [kg/ha] gibt. Nachher wollen sich beide noch mit dem Hobby Klemmbausteine von Steffen beschäftigen. Kennt Paula noch nicht, klingt aber interessant.

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

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Bewässerungstypen	3				
error	19	411.5			
Total	22	862.87			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.13$ 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

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik • Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Die einfaktorielle ANOVA in R 'Uff... die einfaktorielle ANOVA und R. Nicht so einfach... Was sagt mir jetzt die Ausgabe der ANOVA und wo sehe ich, ob da was signifikant ist?', denkt Paula und hebt die Augenbraue. Paula hatte sich ein Feldexperiment mit Lauch angeschaut. Als wäre das nicht alles schon schwer genug. Paula und der Perfektionismus, eine unendliche Geschichte mit kniffligen Wendungen. Dabei ging es beim Experiment herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lichtstufen (*none*, *200lm*, *400lm* und *600lm*) und dem Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] gibt. Nun möchte ihre Betreuerin ihrer Abschlussarbeit erstmal eine ANOVA sehen und die Ergebnisse präsentiert bekommen. Und eigentlich will sie ja was anderes... Schon dutzende Male gesehen: Jagd auf roter Oktober. Aber immer noch großartig zusammen mit Smarties.

```
## Analysis of Variance Table
##
## Response: Chlorophyllgehalt
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Lichtstufen  3 1016.53   338.84   20.963 4.097e-07
## Residuals   26   420.27    16.16
```

Leider kennen sich Paula 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! **(2 Punkte)**
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)**

78. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



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.', Jonas schaut Yuki fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Yuki schmeißt sich noch ein paar Reese's Peanut Butter Cups 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 Kreuzungsexperiment mit Fleischrindern durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Elterlinie (*ctrl*, *Standard*, *Yray* und *Xray*) und dem Messwert Schlachtgewicht [kg] gibt. Später wollen die beiden dann noch raus um zu Boldern.

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

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Elterlinie	3	77.58			
Error	29	204.3			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 2.93$ ab. Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Was bedeutet ein signifikantes Ergebnis in einer einfaktoriellen ANOVA? **(1 Punkt)**
6. Berechnen Sie *einen* Student t-Test für den *vermutlich* signifikantesten Gruppenvergleich anhand der untenstehenden Tabelle mit $T_{\alpha=5\%} = 2.03$. Begründen Sie Ihre Auswahl! **(3 Punkte)**

Elterlinie	Fallzahl (n)	Mittelwert	Standardabweichung
ctrl	5	2.00	1.41
Standard	10	0.60	2.72
Yray	10	4.40	2.46
Xray	8	3.25	3.28

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

79. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Ergebnistabelle der zweifaktoriellen ANOVA Jonas steht im Oldenburger Land. Und das ist noch langweiliger als es sich anhört. Wäre es nur so spannend wie bei seinen Kommilitonen, die in Almería waren. Ödnis wohin man nur blickt. Oder eben Schweinen. Das Meerschweinchen durchbohrt ihn mit leeren Blick. 'Woher zum Teufel!', entfährt es ihm. Aber da ist es schon weg. Ja, darum geht es in seiner Hausarbeit. Und wäre das nicht noch alles genug, ist sein Experiment auch noch als ein Kreuzungsexperiment komplex geraten. Es wurde der Messwert Fettgehalt [%/kg] mit dem Behandlung Ernährungszusatz (*ctrl*, *fedX* und *getIt*) sowie der Behandlung Genotypen (*AA* und *BB*) untersucht. 'Hmpf', denkt Jonas und ruft 'Und jetzt!?' in die Leere. Und eigentlich wollte Jonas doch noch seinem Hobby nachgehen! Am Ende dann doch besser Stricken. Wunderbar. Eine echte Ablenkung für Jonas. Leider kennen sich Jonas mit Berechnung einer zweifaktoriellen ANOVA überhaupt nicht aus. Deshalb braucht er bei der Erstellung Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung für beide Faktoren separat! (2 Punkte)
2. Formulieren Sie die statistischen Hypothesenpaare für beide Faktoren separat! (2 Punkte)
3. Füllen Sie die unterstehende zweifaktorielle ANOVA Ergebnistabelle aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Ernährungszusatz	3	588.25			
Genotypen	1	0.46			
Ernährungszusatz:Genotypen	3	237.98			
Error	18	259.96			

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

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

5. Was sagt der Term *Ernährungszusatz:Genotypen* aus? Interpretieren Sie das Ergebnis! (2 Punkte)

80. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Die zweifaktorielle ANOVA in R In ein Stallexperiment wurden Fleischrindern mit der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) sowie der Behandlung Ernährungszusatz (*ctrl* und *getIt*) untersucht. Es wurde als Messwert Schlachtgewicht [kg] bestimmt. Jetzt starrt Jonas mit auf die R Ausgabe einer zweifaktoriellen ANOVA. Leider starrt seine Betreuerin in der gleichen Art Jonas 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 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.

```
## Analysis of Variance Table
##
## Response: Schlachtgewicht
##
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Lüftungssystem      3   17.20    8.598    0.9333    0.4071
## Ernährungszusatz    1   20.51   20.513    2.2267    0.1487
## Lüftungssystem:Ernährungszusatz  3  514.07  257.033  27.9018  5.474e-07
## Residuals          24  221.09    9.212
```

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

1. Formulieren Sie die wissenschaftliche Fragestellung für beide Faktoren separat! **(2 Punkte)**
2. Formulieren Sie die statistischen Hypothesenpaare für beide Faktoren separat! **(2 Punkte)**
3. Interpretieren Sie das Ergebnis der zweifaktoriellen ANOVA! Begründen Sie Ihre Antwort! **(3 Punkte)**
4. Zeichnen Sie eine Abbildung, der dem obigen Ergebnis der zweifaktoriellen ANOVA näherungsweise entspricht! **(5 Punkte)**

81. Aufgabe


(11 Punkte)

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





Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Interaktion in der zweifaktoriellen ANOVA In ein Freilandversuch wurden Kartoffeln mit der Behandlung Bewässerungstypen (*ctrl*, *low*, *mid* und *high*) sowie der Behandlung Genotypen (*AA*, und *BB*) untersucht. Es wurde als Messwert Frischgewicht [kg/ha] bestimmt. Jetzt starrt Alex mit auf die  Ausgabe einer zweifaktoriellen ANOVA. Leider starrt seine Betreuerin in der gleichen Art Alex zurück an. Es liegt anscheinend eine signifikante Interaktion vor? 'Das wird ein langer Nachmittage.', 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. Alex hätte doch nichts mit Kartoffeln machen sollen. Kartoffeln – was soll das auch bedeutendes sein? Eigentlich wollte Alex nachher noch einen Film schauen. Irgendwie komisch, wenn er Alien anmacht, dann ist die Katze eigentlich sofort vor dem Bildschirm und starrt hinein.

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

1. Visualisieren Sie folgende mögliche Interaktionen zwischen den Behandlungen! Heben Sie die Positionen der auftretenden Interaktion (farblich) hervor! 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. Schreiben Sie das Modell mit Interaktionsterm in  für eine zweifaktorielle ANOVA! **(1 Punkt)**
5. Skizzieren Sie die Funktion in  für eine Post-hoc Analyse in {emmeans}! **(1 Punkt)**
6. Wenn eine signifikante Interaktion in den Daten vorliegt, wie ist dann das weitere Vorgehen? Berücksichtigen Sie dabei die Verwendung von {emmeans}! **(2 Punkte)**

82. Aufgabe

(11 Punkte)

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

Inhalt folgender Module: Biostatistik



Zusammenhang zwischen der ANOVA und dem Post-hoc-Test 'Mit der einfaktoriellen ANOVA lassen sich flott die Gruppen in einer Behandlungen vergleichen, wenn wir normalverteilte Daten und Varianzhomogenität vorliegen haben!', seine Betreuerin scheint die einfaktoriellen ANOVA zu verstehen. Warum jetzt er jetzt nochmal alles wiederkauen muss, wird Jonas echt nicht so klar. Wenn es doch so klar ist? 'Wir haben jetzt bei der ANOVA einen p-Wert mit 0.052 raus sowie eine F-Statistik F_D mit 1.61 berechnet. Nach den Boxplots müsste sich eigentlich ein Unterschied zwischen *tornado* und *storm* ergeben. Der Unterschied ist in {emmeans} auch signifikant mit einem p-Wert von 0.049. Wie kann das sein?', fragt Jonas etwas provokant und dreht Iron Maiden leiser. Jonas war im Emsland und hatte dort ein Stallexperiment mit Puten durchgeführt. Die Komune wo er untergekommen war, war cool gewesen. Dort gab es selbstgemachte Snickers aus Vollkorn! Nur jetzt muss eben das Experiment fertig ausgewertet werden. Jonas hatte eine Behandlungen Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) auf Puten angewendet. Gemessen wurde der Messwert (Y) Gewichtszuwachs in der 1LW. Dabei wurden die Daten D erhoben. Jetzt muss das hier zu einem Ende kommen! Jonas hat schon genug Probleme. Wenn die Erschöpfung nicht wäre, dann wäre es einfacher.

Gegebene Formeln

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

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

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

Teil VI.

Multiple Gruppenvergleiche

83. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Adjustierung multipler Vergleiche 'Moment, die haben ja das Gleiche gemacht wie wir!', ruft Tina laut aus. Jonas schaut etwas verwundert. 'Das glaube ich eher nicht. Lass uns mal unsere Daten mit den Ergebnissen von Almar et al. (2012) vergleichen.', antwortet Jonas. In ein Freilandversuch mit Spargel wurde die Behandlung Lichtstufen (*none*, 200lm, 400lm, 600lm, 700lm und 800lm) auf den Messwert Trockengewicht [kg/ha] untersucht. Jetzt müssen die beiden mal schauen, ob sie wirklich was Neues gefunden haben oder ob die Ergebnisse alle die gleichen sind wie schon bei Almar et al. (2012). Es ergab sich dann die folgende Tabelle der rohen p-Werte für die Vergleiche zu Almar et al. (2012).

Rohen p-Werte	Adjustierte p-Werte	Nullhypothese ablehnen?
0.001		
0.030		
0.070		
0.230		
0.760		
0.060		

Leider kennen sich Tina und Jonas 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. Welche Adjustierung wird im Allgemeinen vorgezogen? Die Adjustierung der p -Werte oder die Adjustierung des Signifikanzniveaus α ? Begründen Sie Ihre Antwort! **(2 Punkte)**

84. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Visualisierung des Compact Letter Displays (CLD) Yuki hatte in seiner Hausarbeit ein Gewächshausexperiment durchgeführt. Soweit so gut. Dabei hat er sich mit Maiss beschäftigt. Angeblich der neueste heiße Kram... aber das ist wiederum was anderes. So richtig mitgenommen hat Yuki das Thema dann doch nicht. Hat er sich doch mit Bewässerungstypen (*none*, *ctrl*, *low* und *high*) und Frischegewicht [kg/ha] schon eine Menge an Daten angeschaut. Nach seinem Betreuer soll er nun ein CLD bestimmen. Weder weiß er was ein CLD ist, noch war sein erster Gedanke mit Köln und die LGBTQ Community richtig...

Behandlung	Compact letter display
none	B
ctrl	C
low	AB
high	A

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

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

85. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



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

Substrattypen	Fallzahl (n)	Mittelwert	Standardabweichung
kompost	8	4.91	2.27
torf	8	5.71	3.89
40p60n	8	15.34	1.60
70p30n	7	16.26	3.66

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich ergebenden Barplots! **(1 Punkt)**
4. Berechnen Sie die Matrix der p -Werte anhand von Student t-Tests! *Nutzen Sie hierfür ein globales s_p sowie eine gemittelte Fallzahl n für die Berechnung der Teststatistik! Nehmen Sie einen kritischen Wert $T_{\alpha=5\%}$ von 1.96 an!* **(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 Yuki! **(1 Punkt)**

Teil VII.

Der Chi-Quadrat-Test

86. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Den Chi-Quadrat-Test berechnen Yuki hat sich ein Herz gefasst und war für ihrer Abschlussarbeit 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 London Grammar auf das Ohr und los gehts. Yuki ist schon eine ganze Zeit im Büro, da ihre Betreuerin möchte, dass sie jetzt auf ihren Daten mit $n = 174$ Beobachtungen von Erdbeeren einen χ^2 -Test rechnet. Das ginge, da sie als Behandlung *Herbizideinsatz* [ja/nein] bestimmt und zum anderen die Variable *Chlorophyllgehalt unter Zielwert* [ja/nein] ermittelt hat. Wie genau, das ist jetzt eine andere Frage. Eigentlich wollte Yuki nachher noch einen Film schauen. Irgendwie komisch, wenn sie Matrix anmacht, dann ist das Minischwein eigentlich sofort vor dem Bildschirm und starrt hinein.

	56	51	
	23	44	

Leider kennt sich Yuki 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 $\chi^2_{\alpha=5\%} = 6.21$! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Skizzieren Sie in einer Abbildung die χ^2 -Verteilung, wenn die H_0 wahr ist! Ergänzen Sie $\chi^2_{\alpha=5\%}$ und χ^2_D in der Abbildung! Beachten Sie folgenden Informationen zur χ^2 -Verteilung. Die χ^2 -Verteilung hat ein Maxima bei $\chi^2 = 6$ sowie ein Minima bei $\chi^2 = 11$. **(2 Punkte)**

87. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



Den Chi-Quadrat-Test in einem Fragebogen berechnen Jessica hatte sich gleich von Beginn an in ihrer Hausarbeit für eine Umfrage im Marketing interessiert. Jetzt geht es um den Haupt- und Nebenerwerb von Erlebnishöfen in Norddeutschland. Viele Höfe haben angefangen auch Großkatzen zu halten, damit mehr Kunden auf die Höfe kommen. Für den Verband der Großkatzenbesitzer e.V. möchte sie nun einen Fragebogen zur Zukunftsfähigkeit Schritt für Schritt auswerten. Dabei teilt sie zuerst die Antwortenden in die beiden Gruppen 'Höfe mit Großkatzen [ja]' und 'Höfe mit Großkatzen [nein]' ein. Daraufhin möchte sie für folgende Frage *f4verband* einmal auswerten, ob es einen Unterschied zwischen den beiden Höfen mit oder ohne Großkatzen gibt.

Sehen Sie die Haltung von Großkatzen als eine kulturelle Bereicherung?

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

<i>f4verband</i>	trifft gar nicht zu	trifft nicht zu	weder noch	trifft zu	trifft voll zu	
ja	3	8	12	21	11	
nein	5	7	13	25	9	

Leider kennt sich Jessica mit der Berechnung eines χ^2 -Test auf einer Frage in einem Fragebogen 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. Berechnen Sie die Teststatistik χ_D^2 eines Chi-Quadrat-Test! Ignorieren Sie Zellbelegungen kleiner gleich fünf in der Berechnung von χ_D^2 und Runden Sie den erwarteten Wert E ganzzahlig! **(2 Punkte)**
4. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem $\chi_{\alpha=5\%}^2 = 11.56$! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Visualisieren Sie die 2x5 Kreuztabelle *ohne* die Berücksichtigung der Antwortkategorie 'weder noch'! **(2 Punkte)**
6. Berechnen Sie den Effektschätzer *Cramers V* auf der 2x5 Kreuztabelle! **(1 Punkt)**
7. Interpretieren Sie den berechneten Effektschätzer *Cramers V* unter Berücksichtigung der Fragestellung! **(1 Punkt)**

88. Aufgabe


(11 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik



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

	24	11	
	23	52	

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Ergänzen Sie die Tabelle um die fehlenden Informationen! **(1 Punkt)**
3. Berechnen Sie die Teststatistik eines Chi-Quadrat-Test! **(3 Punkte)**
4. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem $\chi^2_{\alpha=5\%} = 6.21$! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effektschätzer *Cramers V* auf der 2x2 Kreuztabelle! **(1 Punkt)**
6. Welchen Wertebereich kann der Effektschätzer *Cramers V* annehmen? Wann liegt kein Effekt und wann ein starker Effekt vor? **(2 Punkte)**
7. Interpretieren Sie den berechneten Effektschätzer *Cramers V* unter Berücksichtigung der Fragestellung! **(1 Punkt)**

Teil VIII.

Lineare Regression & Korrelation

89. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen



Visualisierung der linearen Regression '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 Jessica an. 'Ich sehe keine Punkte...', antwortet Jonas sichtlich übernächtigt. 'Wir müssen die Daten ja auch erst visualisieren!', spricht Jessica sehr deutlich und langsam. Die beiden hatten ein Freilandversuch im Emsland mit Lauch durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche UV Einstrahlung [UV/d] und Trockengewicht [kg/ha]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt.

Durchschnittliche UV Einstrahlung [UV/d]	Trockengewicht [kg/ha]
14.9	24.9
16.3	26.2
13.4	21.3
11.3	18.3
12.1	21.8
14.9	24.0
10.8	17.3
16.2	23.8
12.8	19.5
14.7	18.1

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Erstellen Sie eine Visualisierung für die Datentabelle. Beschriften Sie die Achsen! **(2 Punkte)**
3. Schätzen Sie die Regressionsgleichung aus der obigen Abbildung ab! Begründen Sie Ihre Antwort mit einer Skizze der Methodik an ausgewählten Punkten! **(2 Punkte)**
4. Beschriften Sie die Gerade 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 Gerade aussehen? Antworten Sie mit einer Skizze der Geraden! **(1 Punkt)**

90. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen



Interpretation der Ergebnisse einer linearen Regression 'Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen!', ruft Nilufar. 'Ich sehe nur zwei Zeilen und keine Punkte. Wie soll ich da denn jetzt eine Gerade durchzeichnen?', fragt Jessica. Nilufar atmet schwer ein und starrt auf die R-Ausgabe der Funktion `lm()`. Die beiden hatten ein Stallexperiment im Wendland mit Milchvieh durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Bewegungsscore [Movement/h] und Schlachtgewicht [kg]. Das Bestimmtheitsmaß R^2 hatten die beiden mit 0.3 bestimmt. 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 R gemacht, aber wie soll das jetzt gehen?

term	estimate	std.error	t statistic	p-value
(Intercept)	4.33	2.26		
Durchschnittlicher Bewegungsscore	-0.14	0.23		

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die Regressionsgleichung! **(1 Punkt)**
3. Erstellen Sie eine Visualisierung der `lm()`-Ausgabe. Beachten Sie die Informationen zum Bestimmtheitsmaß R^2 aus dem Aufgabentext! Beschriften Sie die Achsen! **(2 Punkte)**
4. Beschriften Sie die Visualisierung mit den statistischen Maßzahlen der `lm()`-Ausgabe! **(2 Punkte)**
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)**

91. Aufgabe

(11 Punkte)

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



Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Interpretation der Ergebnisse einer linearen Regression in R 'Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen!', ruft Jessica. 'Ich sehe nur Kauderwelsch und keine Punkte. Wie soll ich da denn jetzt eine Gerade durchzeichnen? Und warum überhaupt? War das unsere Fragestellung?', fragt Alex. Jessica atmet schwer ein und starrt auf die R-Ausgabe der Funktion `lm()`. Die beiden hatten einen Leistungssteigerungsversuch im Wendland mit Hühnern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Bewegungsscore [Movement/h] und Schlachtgewicht [kg]. Jetzt will die Betreuung von den beiden die Interpretierung der Daten in Form einer linearen Regression gerechnet bekommen. Das haben beide in R gemacht, aber wie soll das jetzt gehen? Das mit der Interpretation?

```
##
## Call:
## Schlachtgewicht ~ Durchschnittlicher_Bewegungsscore
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.69614 -0.81159  0.05756  0.96038  2.08939
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.762      1.636   1.688  0.0993
## Durchschnittlicher_Bewegungsscore  2.445      0.160  15.279 <2e-16
##
## Residual standard error: 1.264 on 39 degrees of freedom
## Multiple R-squared:  0.8569, Adjusted R-squared:  0.8532
## F-statistic: 233.4 on 1 and 39 DF,  p-value: < 2.2e-16
```

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

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

92. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Interpretation der Ergebnisse einer Korrelationsanalyse in R 'Wichtig ist es, dass wir jetzt eine Gerade durch die Punkte zeichnen...', denkt Jonas. 'Ich sehe nur Kauderwelsch und keine Punkte. Ich glaube das war jetzt doch eine Korrelation, die ich rechnen sollte. Und warum überhaupt? War das unsere Fragestellung?', denkt sich Jonas. Jonas atmet schwer ein und starrt auf die R Ausgabe der Funktion `cor.test()`. Das hilft alles nur begrenzt. Das Verrückte ist, dass das Meerschweinchen Mission Impossible wirklich liebt. Das ist Jonas sehr recht, denn er braucht Entspannung. Jonas hatte einen Versuch in einer Klimakammer im Emsland mit Erdbeeren durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittlicher Niederschlag [ml/w] und Chlorophyllgehalt (SPAD-502Plus) [SPAD]. Jetzt will die Betreuung von ihm die Interpretierung der Daten in Form einer Korrelation berechnet bekommen. Das hat Jonas in R gemacht, aber wie soll das jetzt gehen? Das mit der Interpretation? Eine echte Herausforderung für ihn war schon immer die Erschöpfung gewesen. Ein leidiges Lied.

```
##
## Pearson's correlation
##
## data: Durchschnittlicher Niederschlag and Chlorophyllgehalt
## t = 6.4758, df = 8, p-value = 0.000193
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.6779493 0.9803666
## sample estimates:
##      cor
## 0.9164042
```

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

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

93. Aufgabe

(12 Punkte)

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

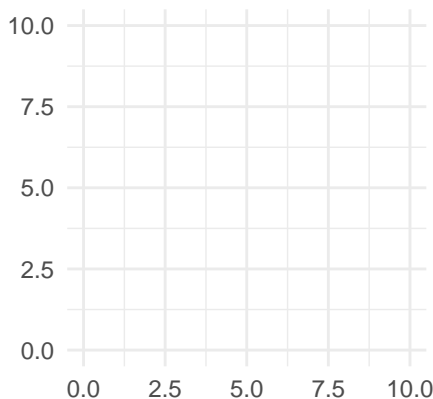
Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Visualisierung der Korrelation und des Bestimmtheitsmaßes Wenn Indiana Jones läuft, dann ist die Spinne nicht mehr da. Aber jetzt braucht sie mal Entspannung! Da hilft dann die Aufgabe auch nur bedingt. 'Hm..., drei leere Abbildungen. Was soll ich da jetzt machen?', fragt sich Tina und mampft noch ein paar Katjes in sich hinein. Tina kennt sich nur begrenzt bis gar nicht mit der linearen Regression und Korrelation aus.

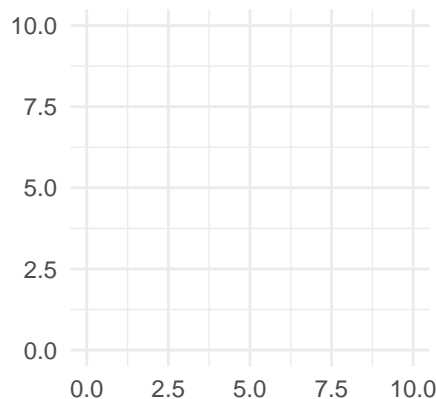
Pearsons $\rho = -1$

$R^2 = 0.25$



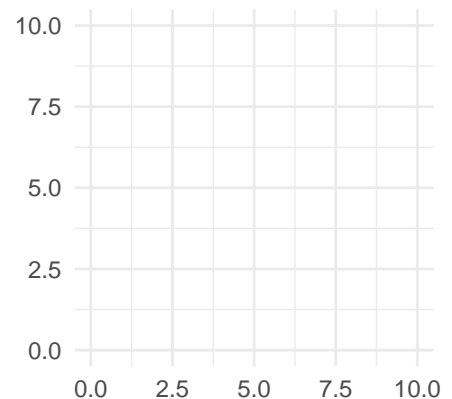
Pearsons $\rho = -0.5$

$R^2 = 0.75$



Pearsons $\rho = 0.5$

$R^2 = 0.5$



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)**

94. Aufgabe

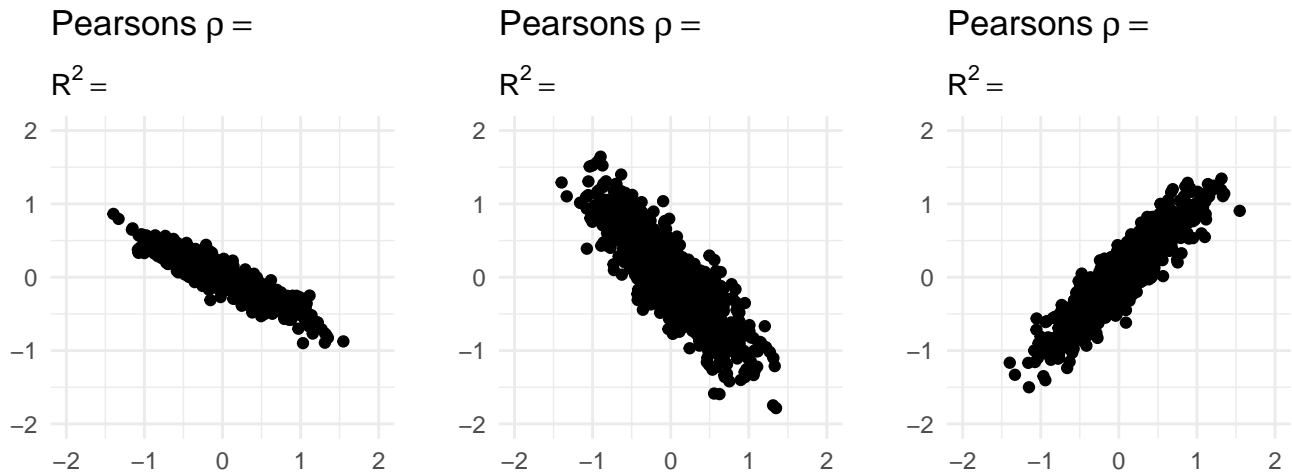
(12 Punkte)

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

Inhalt folgender Module: Statistik • Angewandte Mathematik & Statistik für Bioverfahrenstechnik • Angewandte Statistik und Versuchswesen • Biostatistik



Schätzen der Korrelation und des Bestimmtheitsmaßes Der Bildschirm strahlt blau in das Gesicht von Paula. Es ist schon spät. Und das hat einen Grund. Schon dutzende Male gesehen: Jagd auf roter Oktober. Aber immer noch großartig zusammen mit Smarties. . Paula überlegt, aber ihre Gedanken sind etwas zäh. 'Was soll das hier alles bedeuten?', fragt sich Paula. Irgendwie ist ihr nicht klar wie sie p -Werte oder R^2 -Werte abschätzen soll. Alles nicht so einfach. Paula und der Perfektionismus, eine unendliche Geschichte mit kniffligen Wendungen.



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

1. Schätzen Sie die p -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 p -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)**

95. Aufgabe

(12 Punkte)


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

Inhalt folgender Module: Biostatistik



Visualisierung des Regressionskreuzes 'Ich sehe die Sinnhaftigkeit nicht. Eine Regression in einem Kreuz darzustellen ergibt für mich keinen Sinn.', meckert Steffen so laut, dass die Cafébesucher an den anderen Tischen zu ihm rüberschauen. Steffen hatte ein Gewächshausexperiment mit Spargel durchgeführt. Dabei hatte er dann Frischengewicht [kg/ha] gemessen. Jetzt war es an ihm die Auswertung für seine Gruppe zu erledigen und dafür sollte Steffen sich an dem Regressionskreuz orientieren.

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

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

96. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Modellierung landwirtschaftlicher Daten • Biostatistik



Interpretation der Ergebnisse eines gemischten Modells in R 'Ich glaube ich bringe da was durcheinander. Ich möchte eine Gerade durch die Punkte zeichnen oder doch eine Korrelation berechnen?', merkt Alex laut an. 'Ich sehe keine Punkte... das ist doch eine Ausgabe in R. Überhaupt, darum geht es doch gar nicht in meinem Versuch. Ich wollte doch keine Gerade zeichnen?.', antwortet Alex sich sichtlich übernächtigt selber. Das Verrückte ist, dass die Katze Alien wirklich liebt. Das ist Alex sehr recht, denn er braucht Entspannung. Die Nacht war zu lang und überhaupt. Wenn die Gefälligkeit nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so.. Alex hatte einen Leistungssteigerungsversuch im Emsland mit Milchvieh durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: mittlere Anzahl an weißen Blutkörperchen [LEU/ml] und Fettgehalt [%/kg]. Jetzt will er erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der R Ausgabe möglich sein.

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: iq_score ~ body_height + (1 | cave/direction)
## Data: dragon_small_tbl
##
## REML criterion at convergence: 2233.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.40723 -0.66869 -0.01582  0.73845  2.93382
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## direction:cave (Intercept)    26.63     5.16
## cave         (Intercept)    244.70    15.64
## Residual                    203.42    14.26
## Number of obs: 270, groups: direction:cave, 18; cave, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  27.5562    26.4694   1.041
## body_height   0.0777     0.1271   0.611
##
## Correlation of Fixed Effects:
##              (Intr)
## body_height -0.969
```

Leider kennt sich Alex mit der Analyse eines gemischten Modells in R überhaupt nicht aus. Deshalb braucht er bei der Analyse Ihre Hilfe!

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar für den festen Effekt! **(1 Punkt)**
3. Skizzieren Sie das Versuchsdesign! *Nutzen Sie eine reduzierte Auswahl an Faktorkombinationen!* **(1 Punkt)**
4. Visualisieren Sie die Ausgabe des linearen gemischten Modells! *Nutzen Sie eine reduzierte Auswahl an Faktorkombinationen!* **(1 Punkt)**
5. Folgt der Messwert y einer annähernden Normalverteilung? Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Treffen Sie eine Aussage mit $T_{\alpha=5\%} = 2.04$ zu Ihrer aufgestellten Nullhypothese! **(2 Punkte)**
7. Interpretieren Sie den Effekt des festen Effekts im Kontext der wissenschaftlichen Fragestellung! **(2 Punkte)**
8. Wieviel Varianz erklärt der zufällige Haupteffekt? **(2 Punkte)**

Teil IX.


Experimentelles Design

97. Aufgabe

(16 Punkte)

Inhalt folgender Module: Angewandte Statistik und Versuchswesen • Biostatistik



Einfache experimentelle Designs Jessica und Alex sind bei Jonas um sich Hilfe für eine Versuchsplanung in  zu holen. Im Hintergrund läuft viel zu laut Iron Maiden. Dabei geht es um den Zusammenhang zwischen der Behandlung Lüftungssysteme (*ctrl*, *storm*, *thunder* und *tornado*) und dem Messwert Trockengewicht [kg/ha] in Kartoffeln. Der Versuch soll in einem Gewächshausexperiment in der Uckermark durchgeführt werden. Nach der Dozentin ist der Messwert Trockengewicht [kg/ha] normalverteilt. Die beiden entschieden sich für ein faktorielles Versuchsdesign. Im ersten Schritt überlegt Jonas ein einfaches experimentelles Design zu probieren. Daher entscheiden sich alle drei für ein *Randomized complete block design (RCBD)* mit zwei Blöcken. Das sollte für den Anfang erstmal reichen. 'Und jetzt, was machen wir jetzt?', Alex schaut die anderen beiden mit großen Augen an. Die zucken mit der Schulter. Alle mampfen Snickers.

Leider kennen sich Jonas, Jessica und Alex mit dem *Randomized complete block design (RCBD)* mit zwei Blöcken überhaupt nicht aus. Deshalb brauchen die Drei bei der Erstellung Ihre Hilfe!

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

98. Aufgabe




(20 Punkte)

Inhalt folgender Module: Biostatistik



Fortgeschrittene experimentelle Designs Die Hündin macht mal wieder Randalen in Jessicas Zimmer und rennt davon! Yuki und Paula sind bei Jessica im Wendland wo der neue, bessere Versuch stattfinden soll. Dabei soll in einem Stallexperiment im Wendland mit Puten durchgeführt werden. Auf dem Tisch stapeln sich Schokobons aus Vollkorndinkelmehl. Eine Spezialität der Komune hier. Yuki hasst Vollkorn wie Smarties geliebt werden. In dem neuen Versuch geht es um den Zusammenhang zwischen der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) sowie Bestandsdichte (*standard* und *kontakt*) sowie drei Blöcken und dem Messwert Schlachtgewicht [kg]. Immerhin ist der Messwert normalverteilt, was einiges einfacher macht. Was es nicht so einfacher macht ist, dass Paula als zusätzliche Herausforderung noch der Perfektionismus mitgebracht hat. Daher entscheiden sich alle drei für ein *Split plot design* oder auch *Spaltanlage*. 'Naja, so viel einfacher ist es dann doch nicht...', merkt Paula an und sucht die Hündin.

Leider kennen sich Jessica, Yuki und Paula mit dem *Split plot design* oder auch *Spaltanlage* überhaupt nicht aus. Deshalb brauchen die Drei bei der Erstellung Ihre Hilfe!

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

Teil X.

Forschendes Lernen

99. Aufgabe

(20 Punkte)

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

Inhalt folgender Module: Biostatistik





Die folgende Aufgabe basiert auf einer der zwei folgenden wissenschaftlichen Veröffentlichungen. Für die Prüfung wird die vertiefende Kenntnis der beiden Veröffentlichungen vorausgesetzt.

- Sánchez, M., et al. (2022). Hoverfly pollination enhances yield and fruit quality in mango under protected cultivation. *Scientia Horticulturae*, 304, 111320. [\[Link\]](#)
- Wu, G., et al. (2004). Arginine nutrition in neonatal pigs. *The Journal of Nutrition*, 134(10), 2783S-2790S. [\[Link\]](#)

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

Zerforschen einer wissenschaftlichen Veröffentlichung Vor dem Start der eigenen Arbeit möchte seine Betreuerin, dass Mark 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 Mark hätte dann schon eine Vorlage. 'Das ist jetzt aber umfangreicher als gedacht!', schnauft er und runzelt die Stirn. Im Hintergrund spielt viel zu leise Andrea Berg. Das wird dann vermutlich heute Abend nichts mehr mit seinem Hobby Geocaching. Der Hamster schaut mitleidig.

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

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

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

100. Aufgabe

(20 Punkte)

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

Inhalt folgender Module: Biostatistik



Zerforschen eines wissenschaftlichen Datensatzes Unter einem langen Schnaufen starrt Yuki auf den wissenschaftlichen Datensatz *in der Tabelle 1* in ihrem Laptop. Insgesamt wurden n Beobachtungen erhoben. 'Worum geht es denn eigentlich in diesem Datensatz?', fragt sie 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 sie nur noch ausfüllen muss. Daher möchte ihre Betreuerin, dass sie einmal die Daten sinnvoll zusammenfasst. Das sollte dann doch etwas aufwendiger werden. Das wird dann vermutlich heute Abend nichts mehr mit Matrix.

f_1		f_2		x_1		y_1		y_2	
<	>	<	>	<	>	<	>	<	>
1		1		2.3		10.1		0	
1		1		4.1		13.1		0	
1		1		5.7		16.5		1	
1		1		3.4		14.6		0	
1		2		2.8		12.1		1	
1		2		6.1		13.4		1	
⋮		⋮		⋮		⋮		⋮	
5		2		1.9		9.6		0	


Yuki füllt sich mit der Analyse der Daten in der Tabelle 1 überfordert. Deshalb braucht sie bei der Auswertung Ihre Hilfe! Glücklicherweise kennen Sie den wissenschaftlichen Datensatz aus Ihren eigenen Analysen schon im Detail und können sofort helfen.

Beantworten Sie die folgenden Fragen anhand eines selbst gewählten Beispiels!

Allgemeiner Aufgabenteil

1. Ergänzen Sie die Eigenschaften der Spalten in der Form eines tibbles! **(2 Punkte)**
2. Skizzieren Sie zwei übergeordnete Analysebereiche der Statistik! *Nutzen Sie hierfür die Variablennamen der obigen Datentabelle.* Beschriften Sie die Abbildungen! **(4 Punkte)**
3. Formulieren Sie zwei mögliche wissenschaftliche Fragestellungen in Form einer PowerPoint Folie aus der obigen Datentabelle! **(2 Punkte)**

Spezieller Aufgabenteil für die Variablen x_1 und y_1

4. In welchen der übergeordneten Analysebereiche der Statistik gehört die Auswertung Ihres Endpunktes? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Skizzieren Sie eine beispielhafte Abbildung für Ihren Endpunkt im Kontext der wissenschaftlichen Fragestellung! **(2 Punkte)**
6. Erstellen Sie das statistische Modell in der in  üblichen Schreibweise! **(1 Punkt)**
7. Skizzieren Sie die Datenanalyse für Ihren Endpunkt! **(4 Punkte)**
8. Auf welche Eigenschaften der Daten müssen Sie für Ihre statistische Analyse im Besonderen achten? **(2 Punkte)**
9. Welche statistische Maßzahl können Sie aus Ihrer Datenanalyse berichten? **(1 Punkt)**

Teil XI.

Mathematik

101. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



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

Helfen Sie der Lerngruppe *Die Pantoffeltieren* bei der Beantwortung der Forschungsfrage! Gehen Sie von einem radialen Wachstum der 1000-jährigen Eiche von 1mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 12.5m in Brusthöhe hatte.

1. Wie groß war der Durchmesser in m der Eiche im Jahr 1820 als Herodot in der Eiche versteckt werden sollte? **(2 Punkte)**
2. Skizzieren Sie in einer Abbildung einen linearen Zusammenhang und einen exponentiellen Zusammenhang für das Wachstum der 1000-jährigen Eiche. Erklären Sie die Auswirkungen der Entscheidung für linear oder exponentiell auf Ihre Berechnungen! **(2 Punkte)**

Herodot hatte eine Schulterhöhe von 195cm, eine Breite von 85cm sowie eine Länge von 220cm.

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 *sehrbequem* mit Abstand um die eigene Achse drehen konnte.

4. Berechnen Sie die Dicke der Eichenwand in cm ! Verdeutlichen Sie Ihre Berechnungen an einer aussagekräftigen Skizze für Pferd und Eiche! **(2 Punkte)**
5. Unter einer Dicke der Eichenwand von 30cm bricht die Eiche zusammen. Beantworten Sie die Forschungsfrage! Begründen Sie Ihre Antwort! **(2 Punkte)**

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

102. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



Von Töpfen auf Tischen Die Projektgruppe *D* bestehend aus Alex, Jonas, Nilufar und Mark hat sich zusammengefunden um den ersten Versuch zu planen. In einem Experiment wollen sie die Wuchshöhe von 210 Maispflanzen bestimmen. Bevor die Vier überhaupt mit dem Experiment beginnen können, gibt es aber ein paar Abschätzungen über die Kosten und den Aufwand zu treffen. Zum einen müssen sie die Maispflanzen einpflanzen und müssen dafür Substrat bestellen. Zum anderen muss die Projektgruppe die Maispflanzen auch bewegen und in ein Gewächshaus auf rechteckigen Tischen platzieren. Die schmale Tischseite fast ohne Randpflanzen 5 Pflanzen. Die Töpfe für die Keimung haben einen Durchmesser von 10cm und eine Höhe von 9cm. Der Kubikmeterpreis für Torf liegt bei 270 EUR.

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

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

103. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



Solar- & Biogasanlagen Nilufar bringt ein neues, tolles Projekt mit in die Lerngruppe *Die Pantoffeltieren* bestehend aus ihr, Alex, Jessica sowie Tina. Um die Energiekosten ihres Betriebes zu senken, will sie eine Solaranlage auf den Schweinestall montieren lassen. Dafür hat sie ihren Stall ausgemessen und findet folgende Maße wieder. Die vordere Seite des Schweinestall hat eine Höhe h_v von $6.5m$. Die hintere Seite des Schweinestall hat eine Höhe h_b von $11m$. Der Schweinestall hat eine Tiefe t von $13m$ und eine Breite b von $50m$. 'Sag mal Nilufar, ist das eine Matheaufgabe oder rechnen wir hier gerade für dich kostenlos als menschliche Computer Sachen für deinen Betrieb?', fragt Jessica mit erhobenen Augenbrauen. Tina und Alex nicken zustimmend.

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

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

Ebenfalls plant Nilufar eine neue Biogasanlage für ihren Betrieb. Der neue Methantank hat einen Radius r von $1.2m$. Leider gibt es ein paar bauliche Beschränkungen auf dem Grundstück. Das Fundament des zylindrischen Methantanks kann nur ein Gewicht von maximal $14t$ aushalten bevor der Tank wegbricht. Nilufar rechnen eine Sicherheitstoleranz von 20% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $240kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $300kg/m^3$. Nilufar betreibt ihre Anlage bei $-92^\circ C$.

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

104. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



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 Yuki. “Wieso?”, entfährt es Ihnen und Sie bereuen sogleich die Frage. Sofort werden Sie zu einem Whiteboard voller roter Schnüre geschliffen und müssen folgenden mathematischen untermauerten Argumenten im Lidl über sich ergehen lassen. Da kommen Sie nicht mehr raus, also können Sie auch gleich mitmachen. Das Problem liegt in Chile⁵. Tja, die Deutschen und Südamerika.

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

1. Wenn 3 Blaubeerschalen 4.77 Euro kosten, wie viel kosten 8 Schalen? **(2 Punkte)**
2. Wenn Sie die 8 Blaubeerschalen gekauft haben, wie viele Aligatorbirnen zu je 0.99 EUR können Sie sich dann noch für 100 EUR leisten? **(1 Punkt)**

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

- Ein Kilo Strauchtomaten benötigt 190l Wasser. Eine Strauchtomate wiegt 90 - 130g.
 - Ein Kilo Salat benötigt 100l Wasser. Ein Salatkopf wiegt 280 - 530g.
 - Ein Kilo Avocado benötigt 1050l Wasser. Eine Avocado wiegt 130 - 380g.
 - Ein Kilo Blaubeeren benötigt 830l Wasser. Eine Blaubeere wiegt 3.2 - 3.7g.
3. Berechnen Sie den Wasserverbrauch für die Produktion für jeweils eine Strauchtomate, einem Salat, einer Avocado und einer Blaubeeren. Stellen Sie das Ergebnis als Tabelle dar! **(3 Punkte)**

Chile exportiert im großem Ausmaß Blaubeeren und Avocados. In dem Exportjahr 2022 blieben die Erträge von Blaubeeren mit 7.9×10^4 t in dem prognostizierten Rahmen. Die Menge steigerte sich um 7.2%. Die Exporte für Avocados fielen in dem gleichen Zeitraum um 18.1% auf 2.3×10^5 t.

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

Chile ist eines der wenigen Länder der Welt, die ihr Wasser komplett privatisiert haben. Derzeit sind nur zwei Prozent des Wassers des Landes für den häuslichen Verbrauch vorgesehen. In den Dörfern der Anbauregionen versorgen Tankwagen die Bevölkerung jede Woche mit Wasser, es gibt etwa 48 Liter Wasser pro Kopf für den täglichen Bedarf. In *Deutschland* liegt der Verbrauch bei 35 - 115 Liter pro Waschgang einer Waschmaschine und 8 - 17 Liter pro Spülmaschinenlauf.

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

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

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

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

105. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



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

Die Dampfnudelerde “Was für einen Unsinn!“, rufen Sie. Jetzt haben Sie auf Empfehlung von von Alex 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 69 Millionen Jahren die Dinosaurier - so groß sie auch waren - nicht von der Schwerkraft zu Boden gerissen wurden. Hat der Dampfplauderer etwa recht und war die Schwerkraft vor Millionen von Jahren eine andere? Sind deshalb alle Lebewesen auf der Erde *heutzutage* so viel kleiner, weil die Schwerkraft größer ist als damals? War die Erde kleiner und hatte weniger Masse? Oder ist es nur ein Rechenfehler wie bei der Theorie der Hohlerde von Edmond Halley aus dem 17.-18. Jahrhundert? Müde reiben Sie sich die Augen. So wird es nichts mehr mit dem Schlafen, dann können Sie auch mal etwas rechnen⁶.

Betrachten wir die Schwerkraft oder Gewichtskraft, die auf Lebewesen damals und heute gewirkt haben soll. Nehmen Sie für die Fallbeschleunigung g der Erde *heutzutage* einen Wert von 9.81 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.2742 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.51 g/cm^3 . Das Gewicht von einem heute lebenden Waldelefanten mit 2.7t liegt bei 6t und das Gewicht von einem Tyrannosaurus rex (T. rex) bei 4.5t bis 8t.

1. Welchen Durchmesser müsste die Erde vor 69 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 69 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 69 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 $1.55 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 84% Wasserstoffkernen mit einer molaren Masse von 1.05 g/mol , 10% Heliumkernen mit 4.01 g/mol sowie 6% weiteren Atomkernen mit 69.18 g/mol . Die Teilchendichte bei Eintritt in die Erdatmosphäre liegt zwischen 0.4 bis 100 Teilchen cm^{-3} pro Sekunde mit einer mittleren Teilchendichte von 7 cm^{-3} pro Sekunde.

Lösen Sie den folgenden Aufgabenteil mit einer aussagekräftigen Skizze!

3. Berechnen Sie die Anzahl an massebehafteten Teilchen des Sonnenwindes, die die gesamte Erde pro Sekunde treffen! **(2 Punkte)**
4. Berechnen Sie die Masse, die die Erde pro Jahr durch die *massebehafteten* Teilchen des Sonnenwind zunimmt! **(2 Punkte)**

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

106. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



'Entschuldigung, ist das Ihre Feder in meinem Auge?' Man hört schon häufig vieles Geschnatter von höflichen Gänse in Mastställen. Enge 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 Mark, Nilufar, Jonas und Steffen aber als vorsorgliche Gänse-Halter:innen nicht⁷. Gemeinsam sind die Vier in einer Projektgruppe gelandet. Betrachten wir also gemeinsam einmal das Platzangebot (eng. *space allowance*, abk. *SA*) der Gänse für vier Tätigkeiten und versuchen die notwendige Fläche zu optimieren. Wie immer gibt es dafür eine mathematische Formel:

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

mit

- *SA* dem benötigten Platzangebot aller aufsummierten Verhalten *i*.
- *A_i* dem benötigten Platz für ein Verhalten *i*.
- *PB_i* dem Anteil des Auftretens eines Verhaltens *i*.
- *r_i* dem Radius Gans plus dem benötigten Radius für das Verhalten *i*.
- *R_i* dem notwendigen Abstand zu den Nachbarn für das Verhalten *i*.
- *i* dem Verhalten: (1) wing/leg stretching, (2) dustbathing, (3) wingflapping 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)	Jacobs et al. (2019)
wing/leg stretching	18cm; 26cm; 52.2%	27cm; 43cm; 54.1%	34cm; 21cm; 50.1%
dustbathing	37cm; 52cm; 8.1%	26cm; 34cm; 2.4%	29cm; -8cm; 0.5%
wingflapping	9cm; 16cm; 6.2%	27cm; 16cm; 5.1%	39cm; 38cm; 5.1%
sitting	5cm; 48cm; 0.8%	25cm; 37cm; 0.8%	47cm; 34cm; 0.8%

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

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

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

107. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



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

»Ohne ausreichende Zufuhr von Vitamin C stellen sich nach 40 Tagen die ersten Symptome ein; die ersten Toten sind nach 72 Tagen zu beklagen; nach 110 Tagen rafft die Skorbut eine ganze Schiffsbesatzung dahin«

Ferdinand Magellan stach im Jahre 1519 in See um eine Passage durch den südamerikanischen Kontinent zu finden. Zu seiner Flotte gehörten fünf Schiffe - das Flaggschiff Trinidad, die San Antonio, die Victoria, die Concepción und die Santiago - mit einer Besatzung von insgesamt 237 Mann.

1. Stellen Sie den Verlauf der Anzahl an Matrosen auf einem Schiff der Flotte in der Form einer Überlebenszeitkurve dar! Beschriften Sie die Achsen entsprechend! **(2 Punkte)**
2. Was ist die Besonderheit der Überlebenszeitkurve? Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Schätzen Sie die Überlebenswahrscheinlichkeit nach 100 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 Skrobut bei der Überquerung des Pazifiks, da durch Erkundungsfahrten weniger Zeit blieb, um wilden Sellerie aufzunehmen. Wilder Sellerie enthält $5000\mu\text{g}/10\text{g}$ Vitamin C. Der Bedarf liegt bei 105mg pro Tag für Männer.

3. Berechnen Sie die notwendige Menge in kg an aufzunehmenden wilden Sellerie auf die Concepción für die ununterbrochene Fahrt von drei Monate und 24 Tage über den Pazifik! **(3 Punkte)**
4. Skizzieren Sie die Überlebenszeitkurve für die Concepción im Vergleich zu der Überlebenszeitkurve der Trinidad! Beschriften Sie die Achsen! **(2 Punkte)**

108. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



Event Horizon – Am Rande des Universums Nilufar ist bei Yuki um gemeinsam *Event Horizon – Am Rande des Universums* zu streamen. Das war jetzt nicht die beste Idee. Denn Nilufar kann Horror überhaupt nicht ab. Deshalb flüchtet sie sich in Logik um ihre Emotionen zu bändigen. Yuki mampft ungerührt Reese's Peanut Butter Cups. Folgenden Gedankengang nutzt Nilufar um dem Film zu entkommen. Die Sonne hat eine aktuelle, angenommene Masse von $2 \times 10^{27} \text{ kg}$. Wenn die Sonne nun am Ende ihrer Lebenszeit zu einem schwarzen Loch mit dem Radius von 4000m kollabiert, wird die Sonne 35% der aktuellen Masse verloren haben. Ein Lichtteilchen mit der Masse m_f und der Fluchtgeschwindigkeit v_f will dem schwarzen Loch entkommen. An folgende Formeln erinnert sich Nilufar für die kinetische Energie des Lichtteilchens E_{kin} und der Gravitationsenergie des schwarzen Lochs E_{grav} ⁸.

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

mit

- m_f , gleich der Masse [kg] des fliehenden Objektes
- m_s , gleich der Masse [kg] des stationären Objekts
- r_s , gleich dem Radius [m] des stationären Objekts
- G , gleich der Gravitationskonstante mit $5.974 \cdot 10^{-11} \text{ m}^3 (\text{kg} \cdot \text{s}^2)^{-1}$

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

1. Geben Sie die Formel für die Fluchtgeschwindigkeit v_f an! **(2 Punkte)**
2. Überprüfen Sie Ihre umgestellte Formel nach v_f anhand der Einheiten! **(1 Punkt)**
3. Berechnen Sie die notwendige Fluchtgeschwindigkeit v_f des Lichtteilchens mit den angegebenen Informationen! **(2 Punkte)**
4. Gehen Sie von einer Lichtgeschwindigkeit von $2.7 \times 10^8 \text{ m/s}$ aus. Kann das Lichtteilchen der Gravitation des schwarzen Lochs entkommen? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Stellen Sie den Zusammenhang zwischen dem sich verringernden Radius r des schwarzen Lochs bei gleichbleibender Masse m_s und der notwendigen Fluchtgeschwindigkeit v_f in einer Abbildung dar! Erstellen Sie dafür eine Datentabelle mit fünf Werten für den Radius r ! **(3 Punkte)**

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

109. Aufgabe

(10 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



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

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

Die beiden treffen folgende Annahmen. Eine außerirdische Zivilisation schickt *drei* Voyager 1 ähnliche Sonden mit der Geschwindigkeit von $6.3587 \times 10^4 \text{ km/h}$ los um sich auf den erreichten Planeten selbst zu replizieren. Nach 1000 Jahren ist die Replikation abgeschlossen und wiederum *drei* Sonden werden ausgesendet. Gehen Sie von 3.57 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 2×10^{11} Sterne in der Milchstraße. Nehmen Sie eine Lichtgeschwindigkeit von $2.7 \times 10^8 \text{ m/s}$ an.

1. Skizzieren Sie in einer Abbildung die ersten vier Schritte der Vervielfältigung der Sonden in der Galaxie! Beschriften Sie die Abbildung mit der Dauer und der Anzahl an Sonden für jeden Schritt der Vervielfältigung! **(4 Punkte)**
2. Berechnen Sie die theoretische Anzahl an Vervielfältigungsschritten die benötigt werden um mit *einem einzigen Vervielfältigungsschritt* die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
3. Berechnen Sie die Dauer, die eine außerirdische Zivilisation annäherungsweise benötigt um die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
4. Bei einem vermuteten Alter der Erde von 4.5×10^9 Jahren, wie oft war dann eine Sonde einer außerirdischen Zivilisation schon zu Besuch? Korrigieren Sie Ihre Antwort mit dem Wissen, dass sich die Kontinentalplatten einmal alle 8×10^7 Jahre vollständig im Erdinneren umgewandelt haben! **(2 Punkte)**

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

110. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



Pyramiden bauen Es stehen die oldenburgischen Pyramidentage an! Sie und Jessica sind auf abenteuerlichen Wegen für den Bau der Pyramiden zuständig. Zu allem Überfluss handelt es sich auch noch eine *Reenactment* Veranstaltung. Thema der diesjährigen Pyramidentage sind die Pyramiden von Meroe, die den Königen und Königinnen des historischen Reiches von Kusch in Nubien, dem heutigen Sudan, als Grabstätten dienten. Die Pyramiden in Meroe fallen durch ihren steilen Winkel von 70 Grad im Vergleich zu den ägyptischen Pyramiden mit 51 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 32 Königsellen. Eine Königselle misst 52.4cm.

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

1. Bei der Königspyramide von Meroe soll eine Seitenlänge der Grundfläche 32 Königsellen lang sein. Einer der Skla'angestellten hat eine Höhe der Königspyramide von 21.98m berechnet. Dem müssen Sie auf jeden Fall nachgehen. Überprüfen Sie die Berechnung! **(1 Punkt)**
2. Die Außenflächen der Königspyramide soll begrünt werden. Für die Bepflanzung muss eine 3cm dicke Torfschicht auf die Königspyramide 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 und Jessica bei dem Befüllen der Königspyramide mit Schutt zu Seite stehen. Leider haben Ihre Sklaven zu allem Überfluss auch noch chronische Schulterschmerzen entwickelt, als die Sklaven von der anstehenden Aufgabe erfahren haben. Gehen Sie daher von einer Effizienz der Sklaven von 80% aus. In eine Schubkarre passen 90 Liter.

3. Wie oft müssen Ihre maladen Sklaven die Rampe mit der Schubkarre zur Spitze der Königspyramide hochfahren um die Pyramide mit Schutt zu füllen? **(1 Punkt)**
4. Berechnen Sie die Länge der Rampe zur Spitze der Königspyramide mit einem Anstellwinkel von 11° ! **(2 Punkte)**
5. Wie weit reicht Ihre Rampe vom Fuß der Königspyramide in die oldenburgische Landschaft? **(2 Punkte)**

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

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

111. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



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

Ort	Cache	Wertung (S T G)
A	GCYZR3I	5.0 4.0 Klein
B	GCN2XYL	4.0 1.0 Klein
C	GCA3EQX	3.0 1.5 Normal
D	GC6CMR0	4.5 3.5 Normal
E	GCNV7X9	2.5 2.5 Mikro

Im Weiteren sind den beiden folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \vec{AB} ist 3km. Im Weiteren ist Ihnen der Entfernungsvektor \vec{CB} mit 4.5km bekannt. Der Entfernungsvektor \vec{BE} ist das 1.3-fache des Entfernungsvektor \vec{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 25° südlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 55° östlich. Vom Ort B betrachtet, bilden die Orte C und D einen rechten Winkel am Ort B. Der Ort B liegt auf gerader Linie zwischen den Orten C und E. Somit liegt der Ort E südlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort D Ihre Cachertour.

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

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

$$\text{Suchzeit} = 0.1 + 0.15 \cdot \text{Schwierigkeit}$$

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

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

112. Aufgabe

(10 Punkte)

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

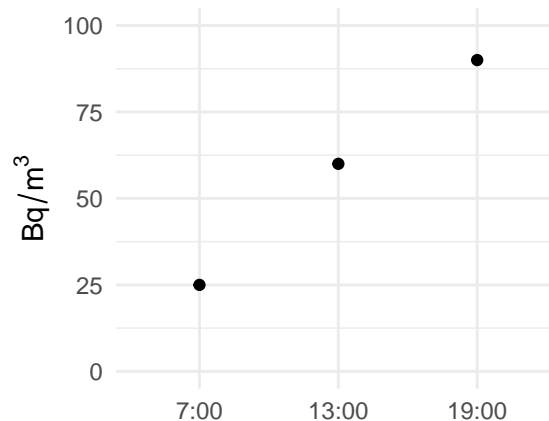


Inhalt folgender Module: Mathematik & Statistik



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

Die atmende Wand und Brot aus Luft Als Kellerkind¹¹ vom Dorf will Jonas das Ausmaß der Radonbelastung in seinem Kellerzimmer bestimmen und lüften daher nicht. Passt schon. Spart dann auch Energie und lüften wird sowieso überschätzt. Während einer Messperiode von 7:00 Uhr bis 19:00 bestimmt er dreimal automatisch die Radonbelastung in seinem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung¹². Leider helfen die Messwerte Jonas überhaupt nicht weiter. Sie müssen also helfen!



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

Radon zerfällt mit einer Halbwertszeit von 2.8d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 143d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

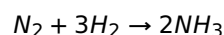
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $300 Bq/m^3$ auf unter $110 Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	77.1	28.1	
Sauerstoff	21.3	16.2	
Kohlenstoffdioxid	0.045	12.1	

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

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



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

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

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

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

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

113. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



Armee der Finsternis Der Studentenjob von Steffen war nach Ladenschluss bei IKEA die Regale einzuräumen. Dabei ist Steffen in der Auslage der Sonderangebote das Necronomicon¹⁴ in die Hände gefallen. Nun ist er eine Magierin der Zeichen geworden! Also eigentlich kann Steffen nur Mathe und das dämliche Necronomicon hat ihn in die Vergangenheit geschleudert... aber gut, was tut man nicht alles im Jahr 1054 n. Chr. für den neuen Lehnsherren Fürsten Arthur. Steffen baut natürlich einen Schrottkugelturm um sich den Horden der Finsternis mit genug Schrott erwehren zu können! Steffen stehen zwei mächtige magische Formeln zur Unterstützung zu Verfügung. Leider wird das nicht reichen, deshalb müssen Sie hier auch noch durch Zeit und Raum helfen!

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

mit

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

Als erstes müssen Sie die Höhe des zu bauenden Schrottkugelturmes bestimmen. Hierfür ist wichtig zu wissen, dass sich die *Bleitropfen* mit einem Gewicht von $10mg$ zu gleichförmigen *Bleikugeln* bei einer Geschwindigkeit von $14m/s$ bilden.

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

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

2. Nennen Sie die beiden geometrischen Formen aus denen sich näherungsweise ein Tropfen zusammensetzt! Erstellen Sie eine beschriftete Skizze des Tropfens! **(2 Punkte)**
3. Sie messen eine Länge des Tropfens von $3.2mm$. Die Löcher im Sieb erlauben ein Tropfendurchmesser von $1.8mm$. Welchen Durchmesser in mm haben Ihre produzierten *Bleikugeln*? **(3 Punkte)**

Sie haben jetzt die 6.1×10^4 *Bleikugeln* zusammen. Blei hat eine Dichte von $11.34g/cm^3$.

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

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

5. Wie groß in cm^2 ist Ihr quadratisches Sieb am oberen Ende des Turms, wenn Sie pro Fall ca. 1200 *Bleikugeln* produzieren wollen und die *Bleikugel* im Fall $0.5cm$ Abstand haben müssen? **(1 Punkt)**

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

114. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



Armee der Kaninchen Leider hat es bei Mark mit der Surfschule in Down Under nicht geklappt. War vielleicht auch nicht so die beste Idee... aber dafür hat Mark eine neue Eingebung! Oder wie es Mike Tyson zugeschrieben wird: »Ich wurde nie niedergeschlagen, ich war immer am Aufstehen!«. Daher macht Mark jetzt einen Großhandel mit Kaninchenfleisch und damit dem teuersten Fleisch in Australien auf. Moment, hopsen hier nicht, seit Thomas Austin im Jahr 1860 ungefähr 26 Kaninchen entlassen hat, Millionen von Kaninchen rum? Wieso ist das Kaninchenfleisch dann so exklusiv? Mark wird stutzig und frag Sie, dem mal mathematisch nachzugehen!¹⁵

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

$$f(t) = 8 \times 10^9 - 1.4 \times 10^9 \cdot 2^{-0.2 \cdot t + 4.1}$$

1. Skizzieren Sie die Sättigungsfunktion *annäherungsweise* in einer Abbildung! **(1 Punkt)**
2. Wie viele Kaninchen können nach der Sättigungsfunktion maximal im westlichen Australien leben? Ergänzen Sie den Wert in Ihrer Abbildung! **(2 Punkte)**
3. Wie viele Millionen Kaninchen leben nach der Sättigungsfunktion nach 15 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 1.8 nach folgender Formel.

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

3. Wie viele Wochen benötigen die Viren um theoretisch die gesamte Kaninchenpopulation nach 13 Jahren Wachstum zu durchseuchen? **(1 Punkt)**

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

4. Wie lange in Jahren dauert es bis eine Kaninchenpopulation nach einer Viruspandemie wieder auf 70% der gesättigten Kaninchenpopulation angewachsen ist? **(2 Punkte)**

Thomas Austin entließ die Kaninchen im äußersten Norden von Australien. Australien hat eine West-Ost-Ausdehnung von 4000km und eine Nord-Süd-Ausdehnung von knapp 3500km. Die Kaninchen breiten sich radial mit einer Geschwindigkeit von 7.3km pro Jahr aus.

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

Eine jährliche Impfung gegen das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) kosten 13\$ pro Tier und der durchführende Arzt verlangt ca. 45\$ pro Tier.

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

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

115. Aufgabe

(12 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



Ostfriesland. Unendliche Weiten. Wir schreiben das Jahr 2025. Dies sind die Abenteuer der Kuh Frida und Nilufar. Grünes Gras unter Nilufars Füßen und ein strammer Wind im Gesicht, egal wohin sie schaut. Ein schmatzendes Geräusch ertönt unter Nilufar. Nilufar sinniert, sollte sie ihre weiten Graslandschaften jetzt schon düngen? Dafür benötigt Nilufar die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Nilufar sieht nicht ein, Geld für einen Agrarmetrologen zu bezahlen, wenn auch Sie mitrechnen können. Also rechnen Sie beide mit folgenden Informationen zu Monatsmultiplikatoren des GLT-Wertes: Januar mit $0.4\times$, Februar mit $0.75\times$ und März mit $1.05\times$. Sie haben noch im letzten Jahr folgende Temperaturen gemessen.

Datum	C°
01. Jan 2024	0.3
01. Feb 2024	1.2
01. Mrz 2024	3.1
01. Apr 2024	6.1

1. Erstellen Sie eine Skizze aus den Informationen aus der Temperaturtabelle! **(1 Punkt)**
2. Stellen Sie die linearen Funktionen $f_1(t)$, $f_2(t)$ und $f_3(t)$ aus der obigen Temperaturtabelle auf! **(1 Punkt)**
3. Bestimmen Sie die Stammfunktionen $F_1(t)$, $F_2(t)$ und $F_3(t)$ für Ihre linearen Funktionen aus der obigen Temperaturtabelle! **(1 Punkt)**
4. Osterglocken beginnen ab einer GLT von 200°C zu blühen. An welchem Tag im 1. Quartal des Jahres 2024 war dies der Fall? *Ignorieren Sie ein eventuelles Schaltjahr in Ihrer Berechnung.* **(4 Punkte)**

Auf dem Weg zu Nilufars Jonagoldplantage werden Sie beide auf dem Trecker von einer Gruppe elektrifizierter Renter abgedrängt. Der Trecker muss wieder aus dem Graben! Frida und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Frida mit 180N . Die elektrifizierten Renter bringen eine Kraft von 210N auf.

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

5. Im ersten Versuch legen Sie das Seil für Frida lotrecht über einen Ast oberhalb des Treckers. Die Rentner ziehen in einer geraden Linie über die Böschung hinweg am anderen Seil. Welche Kraft wird aufgebracht? **(2 Punkte)**
6. Im zweiten Versuch ziehen Frida und die Rentner mit einem 40° Winkel mit ihrem Seil an dem Trecker. Welche Kraft wird aufgebracht? **(2 Punkte)**
7. Mit welcher Beschleunigung ziehen Sie den 1.5t schweren Trecker *jeweils* aus dem Graben, wenn $F = m \cdot a$ gilt? **(1 Punkt)**

116. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



In der Kartonagenfabrik Mark, Alex, Jonas und Steffen sitzen im Bus. Wenn man sich zu spät anmeldet, dann ist die Exkursion nicht so toll. Mark hatte den Anderen in der Lerngruppe zu spät Bescheid gesagt. 'Was denn, bin ich eure Nanny oder was?!', entfährt es Mark nachdem die vorwurfsvollen Blicke schon eine Weile auf ihm lasten. Also geht es eben mit Rektor Skinner und Mrs. Krabappel in die Kartonagenfabrik. Wie schon im vorherigen Semester... In der Kartonagenfabrik angekommen erfahren die Vier, dass die Kartons zum Versand von Nägeln nicht hier zusammengebaut werden sondern das sich die Endfertigung in Flint, Michigan befindet. Unter anderem wird dort der berühmte *Doppelt gewellte, 4-mal-gefaltete, 0.6mm, 30-cm-Karton* durch Falzung hergestellt. Beim letzten Mal war Rektor Skinner die Stimmung zu schlecht und deshalb geht es erst nach Hause, wenn ein paar Aufgaben gelöst sind. Martin gefällt das. An dem Vorrat an Zigaretten von Mrs. Krabappel meinen alle wenig Zuversicht zu erkennen.

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

1. Erstellen Sie eine Skizze des Karton**blatt**rohlings! 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 Kartondeckel**blatt**rohlings in cm^2 ? **(2 Punkte)**

Rektor Skinner möchte sich gerne wieder in seinem Vorgarten aufhalten und nicht die ganze Zeit von Bart mit Erdnüssen beworfen werden. Deshalb möchte er einen geräumigen Teil seines Vorgartens einzäunen. Ein Teil der Umzäunung bildet seine Vorderhauswand. Wegen Lieferschwierigkeiten stehen Rektor Skinner nur 130m Zaun zu Verfügung. Auch hier sollen Sie mal helfen, sonst fährt der Bus Sie nicht nach Hause. Sie wollen nun die maximale Fläche des abgeschirmten Vorgartens in Abhängigkeit der Seitenlängen bei der Verwendung von 130m Zaun bestimmen!

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

117. Aufgabe

(10 Punkte)

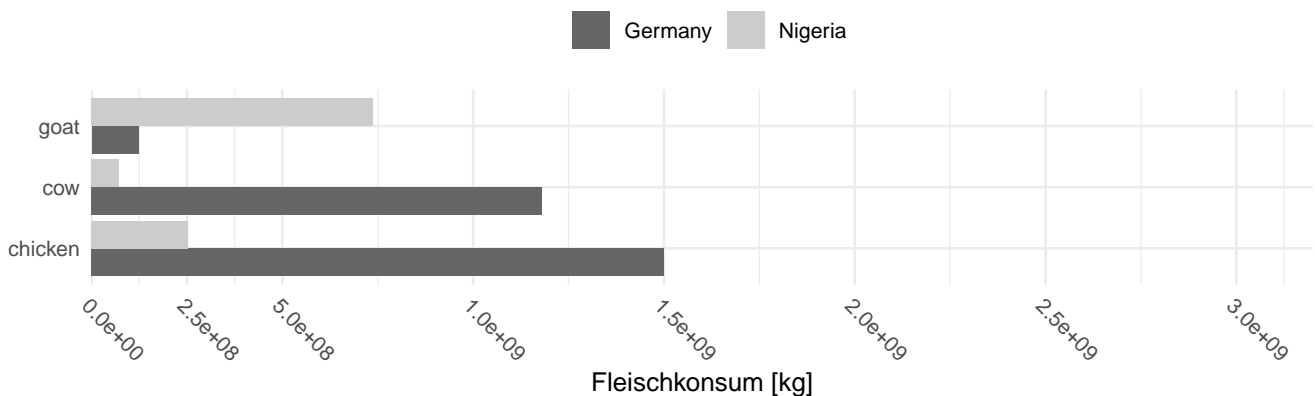
Geben Sie grundsätzlich Formeln und Rechenweg zur Lösung der Teilaufgaben mit an!

Inhalt folgender Module: Mathematik & Statistik



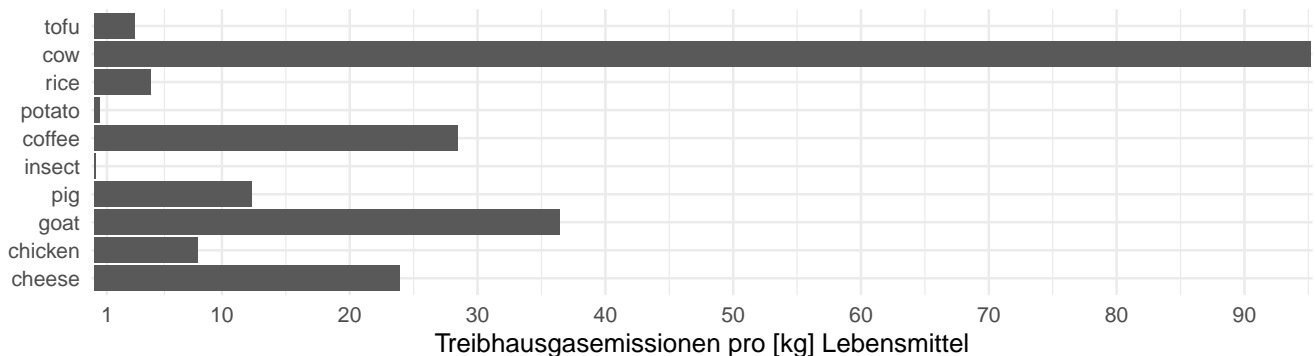
Ein Pfund Insekten, bitte! 'Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹⁶.', merkt Nilufar an. Die Lerngruppe um Paula, Alex und Yuki sind bei Nilufar um mal was außergewöhnliches zu essen. Um den Sinn der Nahrungsumstellung zu verdeutlichen, vergleicht Nilufar einmal Deutschland mit Nigeria. Nigeria hat eine der am schnellsten wachsenden Bevölkerungen der Welt und wird vermutlich im Jahr 2100 zu den Top 5 der bevölkerungsreichsten Länder zählen. Im Jahr 2023 leben ca. 8.2×10^7 Menschen in Deutschland und ca. 1.8×10^8 Menschen in Nigeria. Mit den Informationen wollen Sie und Nilufar mit der Überzeugungsarbeit anfangen und dann eine Prognose für den Fleischkonsum im Jahr 2050 zu treffen.

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



1. Stellen Sie den Fleischkonsum in Deutschland und Nigeria im Jahr 2023 *pro Kopf* in einer aussagekräftigen Tabelle dar! **(2 Punkte)**
2. Ergänzen Sie in der Tabelle eine Spalte in der Sie für den Fleischkonsum in Nigeria auf Deutschland normieren, daher ins Verhältnis Nigeria/Deutschland, setzen! **(1 Punkt)**

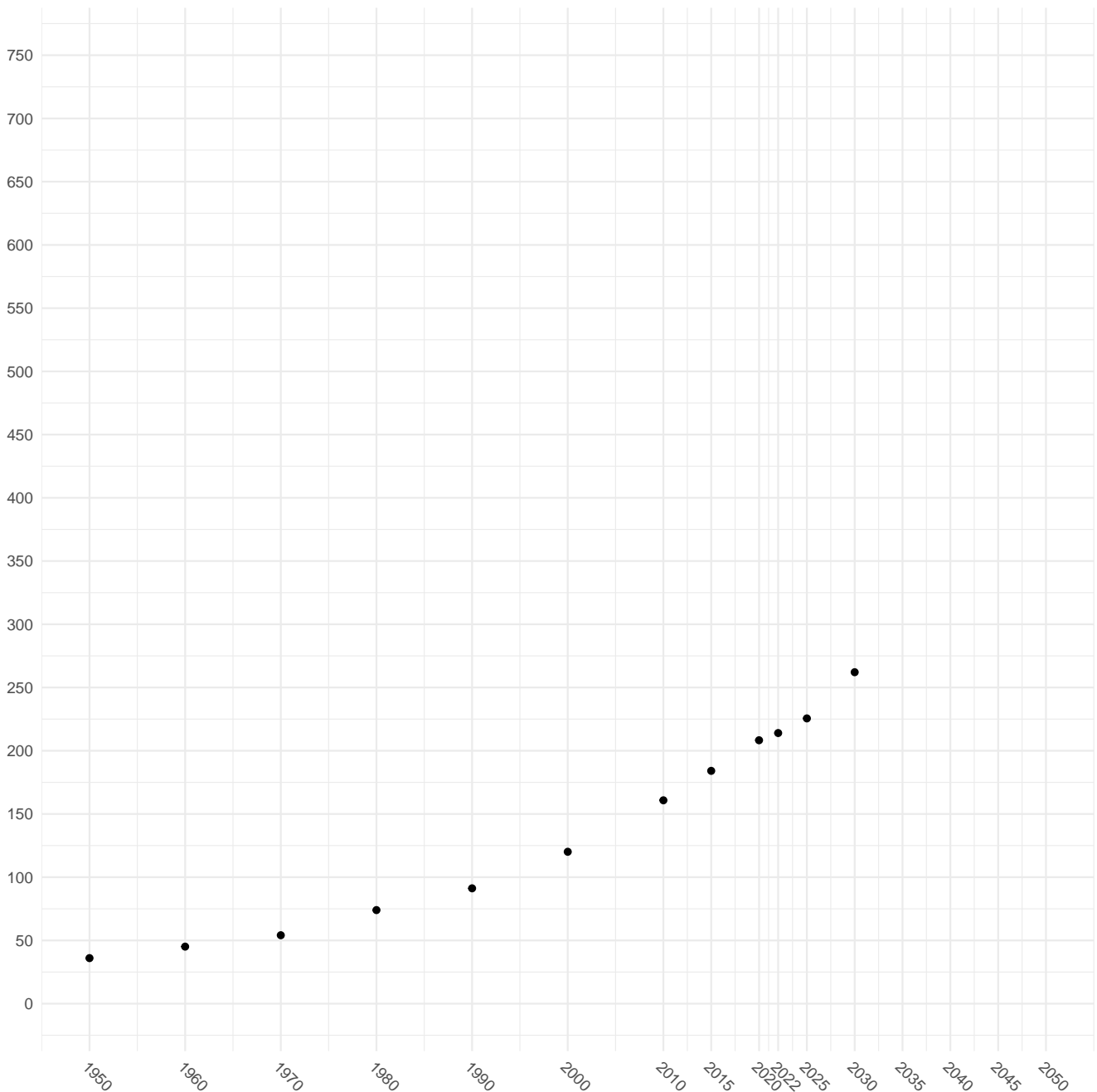
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



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

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

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



4. Schätzen Sie graphisch die zu erwartende Bevölkerung [Millionen] in Nigeria im Jahr 2050, die sich anhand der Informationen aus der Abbildung ergibt!
 - a) Ohne Berücksichtigung der Covid-19-Pandemie! **(1 Punkt)**
 - b) Unter Berücksichtigung der Covid-19-Pandemie! **(1 Punkt)**
5. Berechnen Sie den geschätzten Fleischkonsum von Nigeria im Jahr 2050 unter der Annahme 60%-iger Angleichung der Lebensbedingungen zu Deutschland im Jahr 2023! **(1 Punkt)**
6. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria im Jahr 2050 im Vergleich zum Jahr 2023, der sich durch den angeglichenen Fleischkonsum ergibt! **(1 Punkt)**
7. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria, wenn die gesamte Proteinaufnahme durch Insekten ersetzt würde! **(1 Punkt)**

118. Aufgabe

(12 Punkte)

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

Inhalt folgender Module: Mathematik & Statistik



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

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

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.8% angenommen. In 96% der Fälle ist ein HIV-Test positiv, wenn der Patient erkrankt ist. In 3% der Fälle ist ein HIV-Test positiv, wenn der Patient *nicht* erkrankt ist und somit gesund ist. Sie stutzen. Wie wahrscheinlich ist es denn eigentlich an AIDS erkrankt zu sein (K^+), wenn Sie einen positiven AIDS-Test vorliegen haben (T^+)? Gehen Sie für die folgenden Berechnungen von $n = 3 \times 10^4$ Patienten mit einem diagnostischen Test für AIDS aus. Sie nehmen sich also einen Kuli und fangen an auf der historischen Ausgabe des Spiegels zu rechnen¹⁷.

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

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

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

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

119. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



Network-Marketing oder Schneeballschlacht! Mark, Nilufar und Paula sitzen bei Alex und hören sich etwas über Network-Marketing an. Alex ist jetzt im Network-Marketing tätig. 'Jetzt reicht es. Wir sind eine Lerngruppe und du versuchst uns hier abzuziehen!', poltert Mark und fährt fort, 'Ich erklär dir mal, wie falsch du liegst!'. Eine Möglichkeit, leicht Geld zu verdienen, ist es anderen Menschen für Geld zu versprechen, wie man leicht reich werden kann. Am besten natürlich ohne viel Aufwand und ortsunabhängig. Schnell ein YouTube-Werbevideo gedreht und auf geht es mit unserem Network-Marketing. Aber Moment, wie funktioniert Network-Marketing eigentlich und was hat das alles mit einer Schneeballschlacht zu tun? Mark und Sie wollen hier einmal in die Untiefen des »passiven Einkommens« abtauchen und die Lerngruppe vor Schlimmeren bewahren¹⁸!

Das Jahr 2023 war das erfolgreichste Jahr in der Geschichte von Healthy Herbs Manufacture International (HeHeMan). Das Unternehmen steigerte den Umsatz um rund 15 Prozent von 260 Millionen Euro im Jahr 2022. Doch wie viel kommt bei den Partnern an? Laut HeHeMan habe das Unternehmen 3.8×10^5 aktive Partner.

1. Berechnen Sie zuerst den Umsatz der Firma HeHeMan im Jahr 2023! **(1 Punkt)**
2. Wie viel von dem Umsatz im Jahr 2023 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 40%? **(1 Punkt)**

Das von Alex zu vermarkende Produkt, hinter dem Alex voll steht, kostet 200EUR 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 3.25%, 2.25% und 1.75%. Jeder von Alex angeworbener »Partner« wirbt wiederum vier Partner für sich selbst an. Pro Monat werden im Schnitt zwei Einheiten vom Produkt verkauft. Alex will nun 2500EUR im Monat *passiv* – also durch indirekte Provisionen – erwirtschaften. Kann das klappen? Sie sind zusammen mit Mark skeptisch.

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

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

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

Alex musste zum Einstieg bei HeHeMan Einheiten des Produkts für 6000EUR kaufen. Diese Einheiten kann Alex nur direkt verkaufen. Das ganze Wohnzimmer ist voll davon. Leider musste Alex den Kauf über einen Kredit über 5.1% p.a. über 36 Monate finanzieren. Sie schütteln den Kopf und klären Alex über Zinsen auf.

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

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

120. Aufgabe

(10 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



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

In dem Spiel hat Alex nun auf einmal 5 sechseitige Würfel ($5d6$) zum würfeln in der Hand. Wenn Alex eine 6 würfelt, hat Alex einen Erfolg.

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

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

3. Würden Sie die Axt mit zwei zwölfseitigen Würfeln ($2d12$) als Schaden oder das Schwert mit einem achtseitigen Würfel plus 7 ($1d8+7$) als Schaden bevorzugen? Begründen Sie Ihre Antwort mathematisch! **(1 Punkt)**

Jetzt wird es immer wilder, da Jessica und Jonas sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass der Rettungswurf gegen den zaubernden Hexer funktioniert. Jessica und Jonas haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.65$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.75$. Sie haben aber mitgezählt und festgestellt, dass in 40 von 100 Fällen der Rettungswurf bei einem erfolgreichen Zauber funktioniert hat.

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

121. Aufgabe

(12 Punkte)

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



Inhalt folgender Module: Mathematik & Statistik



Retrocheck im TV Das war zu viel für Alex gestern. Die Lerngruppe mit Jessica und Tina ging viel zu lang. Während er wegdämmert, kommen in ihm seltsame Bilder hoch. 'Und hier ist sie wieder, die Show der fantastischen Preise. Seien Sie mit dabei, wenn es wieder heiß: Der Preis ist heiß!', ertönt es und Alex fragt sich, ob er nicht doch lieber bezahlter Gast bei Barbara Salesch hätten sein sollten. Aber Alex braucht das Geld und jetzt heißt es Spielschows farmen! Erstmal eine Kaffemaschine von Mitropa gewinnen. Ein Kandidat gewinnt die Kaffeemaschine von Mitropa, wenn nicht alle Kandidaten überbieten (eng. *outbid*). Mit Ihnen bilden Jessica und Tina das Team der drei Kandidaten. Alex braucht dringend Ihre Hilfe in seinen Wahnträumen. Sie gehen wie in *Interception* rein!¹⁹

Name	$P(\text{win})$	$P(\text{outbid})$
Jessica	0.2	0.043
Tina	0.4	0.08

1. Mit welcher Wahrscheinlichkeit gewinnen Sie die Kaffeemaschine von Mitropa, wenn keiner der Kandidaten überbietet? **(1 Punkt)**
2. Wenn Ihre überbietungswahrscheinlichkeit $P(\text{outbid})$ bei 0.02 liegt, mit welcher Wahrscheinlichkeit gewinnt *keiner* die Kaffeemaschine von Mitropa? **(1 Punkt)**

Glücksrad für Arme auf der Kirmes! Leider hat es für Maren Gilzer nicht gereicht. Deshalb sind Sie jetzt mit Alex auf der Kirmes und spielen mit einem Typen in einem Tentakelkostüm um das große Geld. Das Glücksrad hat 22 Felder. Sie beide drehen das Glücksrad zweimal. Auf 10 Feldern gewinnen Alex und 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)**

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

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

¹⁹South Park Inception Spoof – Wunderbare South Park Folge

Teil XII.

Angewandte Nutztier- und Pflanzenwissenschaften

(M.Sc.)

122. Aufgabe

(6 Punkte)

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

123. 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.

124. Aufgabe

(6 Punkte)

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

Teil XIII.

Platzhalter

125. Hausarbeit

(10 Punkte)

Inhalt folgender Module: Statistik • Angewandte Statistik und Versuchswesen • Modellierung landwirtschaftlicher Daten



Teil der Prüfungsleistung in dem Modul ist eine Hausarbeit zur Datenanalyse in R.

Diese Aufgabe dient in der Klausur als Platzhalter für die erreichten Punkte in der Hausarbeit.

Zum Bestehen des Moduls ist die Abgabe des Berichts nicht notwendig.

Alle Informationen zu der Hausarbeit sowie dem Stichtag der Abgabe finden Sie auf ILIAS.

Teil XIV.

Lösungen

Die folgenden Lösungen sind absolut ein Entwurf und können Fehler enthalten. Daher immer mit Vorsicht genießen. Im Weiteren ist das Runden so eine Sache. Deshalb wird es vermutlich auch Abweichungen zu deiner händischen Berechnung geben. Damit leben wir dann aber alle.

Statistik

```
library(tidyverse)
data_tbl <- read.table(file = pipe("pbpaste"), sep = " ", header = TRUE)
```

Mathematik

Herodot – der Schimmel aus Ivenack

1)	3)	4)	5)
$d_{1820} = 357.49cm$	$V_{eff} = 1.82m^3$	$s_{Eiche} = 68.74cm$	$s_{Platz} = 38.74cm$

Von Töpfen auf Tischen

2)	3a)	3b)	4)	5)
$n = 336$	$A_{Tisch} = 1.36m^2$	$A_{Topf} = 0.88m^2$	$V_{Topf} = 237.5cm^3$	64.12EUR

Solar- & Biogasanlagen

2)	3)	4)	5)
$A_{Dach} = 687.84m^2$	$\rho = 276kg/m^3$	$V = 40.58m^3$	$h = 8.97m$

Aligatorenbirnen und Blaubeere

1)	2)
12.72EUR	$n = 88$

3)	Wasserverbrauch	Mittleres Gewicht	Liter pro Stück
Strauchtomate	0.19l/g	110g	20.9l
Salat	0.1l/g	405g	40.5l
Avocado	1.05l/g	255g	267.75l
Blaubeere	0.83l/g	3.45g	2.8635l

4a)	4b)	5a)	5b)	6)
$V = 70291040t$	$V = 197788500t$	0.64x	3.84x	FAO Stat, Our World in Data

1)	2)	3)	4)	5)	6)
$h = 2.75 \cdot 8 = 21.98m$	$V_T = 11.23m^3$	$n = 694$	$r = 115.19m$	$b' = 105.08m$	15.04Ellen

Die Dampfnudelerde

Pyramiden bauen

1)	2)	3)	4)	5)	6)
$h = 2.75 \cdot 8 = 21.98m$	$V_T = 11.23m^3$	$n = 694$	$r = 115.19m$	$b' = 105.08m$	15.04Ellen