

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

Fakultät Agrarwissenschaften und Landschaftsarchitektur (AuL)

Klausurfragen Bio Data Science

für Pflichtmodule

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

(Prüfungsleistung der Wahlpflichtmodule ist eine Portfolioprüfung)

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

Wintersemester 2024/25



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

Erlaubte Hilfsmittel

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

Endnote

_____ von 20 Punkten sind aus den Multiple Choice Aufgaben erreicht.
_____ von 76 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 96 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
91.5 - 96.0	1,0
87.0 - 91.0	1,3
82.0 - 86.5	1,7
77.5 - 81.5	2,0
72.5 - 77.0	2,3
67.5 - 72.0	2,7
63.0 - 67.0	3,0
58.0 - 62.5	3,3
53.5 - 57.5	3,7
48.0 - 53.0	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	9	11	12	12	12	10	10

- Es sind ____ von 76 Punkten erreicht worden.

Multiple Choice Aufgaben

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

ANOVA

1. Aufgabe

(2 Punkte)

Aus einem Feldversuch ergibt sich die Notwendigkeit der Berechnung einer einfaktoriellen ANOVA. Es ergibt sich ein $\eta^2 = 0.12$. Welche Aussage ist richtig?

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

2. Aufgabe

(2 Punkte)

Sie führen ein Feldexperiment durch um das Gewicht von Erdbeeren zu steigern. Die Pflanzen wachsen unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen. Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.2$. Welche Aussage ist richtig?

- A** ☐ Es werden 20% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- B** ☐ Es werden 80% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.
- C** ☐ Es werden 20% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen erklärt wird.
- D** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 20% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 80%.
- E** ☐ 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.

3. Aufgabe

(2 Punkte)


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

- A** ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 0 annähert kann die Nullhypothese abgelehnt werden.
- B** ☐ 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.
- C** ☐ Wenn die F-Statistik kleiner als der kritische Wert ist kann die Nullhypothese nicht abgelehnt werden. Die F-Statistik ist der Quotient der MS der Behandlung durch die MS des Fehlers.

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

4. Aufgabe

(2 Punkte)

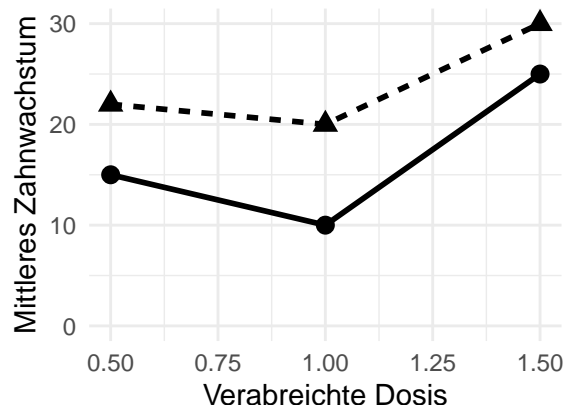
Wenn Sie mehr als zwei Gruppen als Behandlungen vorliegen haben, dann kann ein einfacher t-Test nicht für den globalen Vergleich genutzt werden. Sie entscheiden sich für eine ANOVA in . Die ANOVA analysiert dabei...

- A** ☐ ... den Unterschied zwischen der F-Statistik anhand der Varianz der Gruppen. Wenn die F-Statistik exakt 0 ist, kann die Nullhypothese abgelehnt werden.
- B** ☐ ... den Unterschied zwischen mehreren Varianzen aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.
- C** ☐ ... den Unterschied zwischen zwei paarweisen Mittelwerten aus verschiedenen Behandlungsgruppen. Wenn die signifikant ist, ist daher bekannt welcher Vergleich konkret unterschiedlich ist.
- D** ☐ ... den Unterschied zwischen der Varianz aus verschiedenen Behandlungsgruppen und der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in den Gruppen exakt zu bestimmen.
- E** ☐ ... den Unterschied zwischen der Varianz ausgelöst durch alle Behandlungsgruppen und der Varianz aus globalen Behandlungsgruppen der Kontrollen. Wenn die ANOVA nicht signifikant ist, muss ein Posthoc-Test ausgeschlossen werden.

5. Aufgabe

(2 Punkte)

Ein Versuch wurde an 52 Tieren durchgeführt, wobei jedes Tier eine von drei Vitamin-C-Dosen (0.5, 1 und 1.5 mg/Tag) über eine von zwei Verabreichungsmethoden erhielt. Die folgende Abbildung enthält die Daten aus diesem Versuch zur Bewertung der Wirkung von Vitamin E auf das Zahnwachstum bei Kanarienvögeln. Welche Aussage ist richtig, wenn Sie eine zweifaktorielle ANOVA rechnen?



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

Deskriptive Statistik & Explorative Datenanalyse

6. Aufgabe

(2 Punkte)

Berechnen Sie den Mittelwert und Standardabweichung von y mit 7, 19, 13, 26 und 11.

- A ☐ Sie erhalten 15.2 +/- 7.43
- B ☐ Sie erhalten 15.2 +/- 2.73
- C ☐ Es ergibt sich 14.2 +/- 27.6
- D ☐ Es berechnet sich 15.2 +/- 55.2
- E ☐ Sie erhalten 15.2 +/- 3.715

7. Aufgabe

(2 Punkte)

Wie lautet der Median, das 1st Quartile sowie das 3rd Quartile von y mit 11, 15, 12, 19, 24, 20 und 51.

- A ☐ Es berechnet sich 19 [12; 24]
- B ☐ Es berechnet sich 22 [13; 25]
- C ☐ Sie erhalten 19 +/- 24
- D ☐ Es berechnet sich 20 [13; 23]
- E ☐ Es ergibt sich 19 +/- 12

8. Aufgabe

(2 Punkte)

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

- A ☐ Die optimale Anzahl ist größer als hundert Beobachtungen, wobei es gerne sehr viel mehr sein können.
- B ☐ 1 Beobachtung.
- C ☐ Wir brauchen fünf oder mehr Beobachtungen.
- D ☐ 10 Beobachtungen.
- E ☐ Damit wir hier sauber eine Abbildung von einem

9. Aufgabe

(2 Punkte)

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

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

10. Aufgabe

(2 Punkte)

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

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

11. Aufgabe

(2 Punkte)

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

- A ☐ Der Mittelwert und der Median sollten gleich sein, wenn Outlier in den Daten vorliegen.
- B ☐ Da sich der Mittelwert und der Median unterscheiden, ist der Datensatz nicht zu verwenden. Mittelwert und Median müssen gleich sein.
- C ☐ Der Mittelwert und der Median sollten sich unterscheiden sein, wenn Outlier in den Daten vorliegen.
- D ☐ Wenn sich der Mittelwert und der Median nicht unterscheiden, liegen vermutlich Outlier in den Daten vor.
- E ☐ Der Mittelwert und der Median sollten gleich sein, wenn keine Outlier in den Daten vorliegen.

12. Aufgabe

(2 Punkte)

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

- A ☐ Nach der Erstellung eines Boxplots schauen wir, ob der Median in der Mitte der Box liegt. Dabei ist der Median als dicke Linie dargestellt und die Box ist das IQR.
- B ☐ Nach dem Einlesen der Daten nutzen wir einen Barplot um zu schauen, ob alle Mittelwerte über alle Behandlungen in etwa gleich groß sind. Damit ist dann auch die Varianz in allen Behandlungen in etwa gleich.
- C ☐ Wir erstellen uns für jede Behandlung einen Dotplot und schauen, ob die Dots und damit die Varianz für jede Behandlung gleich groß sind.
- D ☐ In einer explorativen Datenanalyse nutzen wir den Violinplot. Dabei sollte der Bauch am Rand liegen. Dann können wir von einer Normalverteilung ausgehen.
- E ☐ Nach dem Einlesen der Daten nutzen wir einen Boxplot um zu schauen, ob alle Boxen über alle Behandlungen in etwa gleich groß sind. Damit ist dann auch das IQR in allen Behandlungen in etwa gleich.

13. Aufgabe

(2 Punkte)

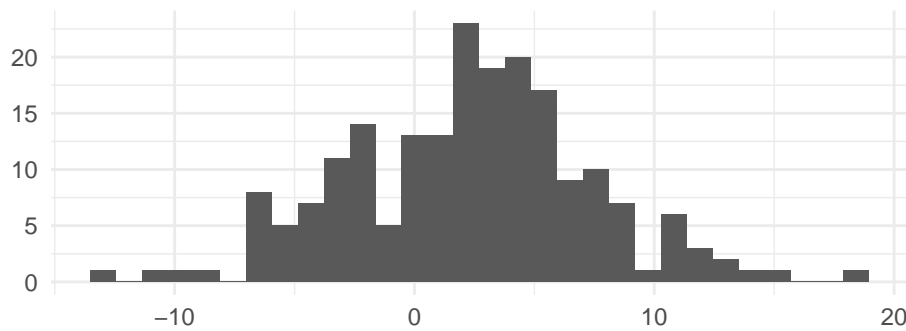
Sie wollen in Ihrer Abschlussarbeit über eine explorative Datenanalyse überprüfen, ob Ihr gemessener Endpunkt einer Normalverteilung folgt. Welche drei Abbildungen eignen sich insbesondere für die Überprüfung?

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

14. Aufgabe

(2 Punkte)

In dem folgenden Histogramm von $n = 177$ Pflanzen ist welche Verteilung abgebildet?



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

Lineare Regression & Korrelation

15. Aufgabe

(2 Punkte)

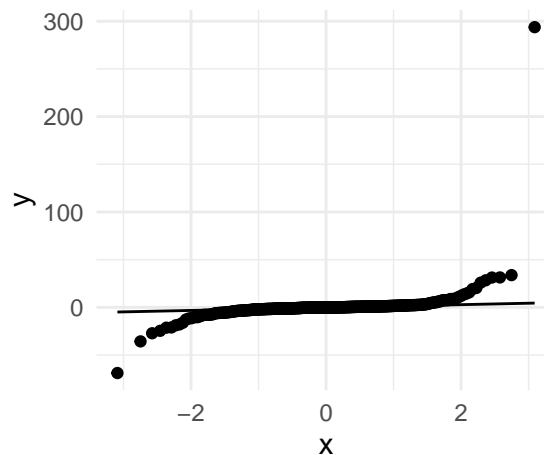
Sie haben das Modell $Y \sim X$ vorliegen und wollen nun ein kausales Modell rechnen. Welche Aussage ist richtig?

- A** ☐ Ein kausales Modell wird auf einem Trainingsdatensatz trainiert und anschliessend über eine explorative Datenanalyse validiert. Signifikanzen über β_i können hier nicht festgestellt werden.
- B** ☐ 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.
- C** ☐ Wenn ein kausales Modell gerechnet werden soll dann kann dies auf dem gesamten Datensatz geschehen. Das Ziel ist es einen Zusammenhang von X auf Y zu modellieren. Wie wirken sich die Einflussvariablen X auf den gemessenen Endpunkt Y aus?
- D** ☐ Wenn ein kausales Modell gerechnet werden soll, dann muss zum einen ein Trainingsdatensatz sowie ein Testdatensatz definiert werden. Dabei ist der Trainingsdatensatz meist 2/3 und der Testdatensatz 1/3 der Fallzahl groß. Der Testdatensatz dient zur Validierung.
- E** ☐ Wenn ein kausales Modell gerechnet werden soll dann kann dies auf dem gesamten Datensatz geschehen. Das Ziel ist es einen Zusammenhang von X auf Y zu modellieren. Wie wirken sich die Einflussvariablen Y auf die gemessenen Endpunkte $X = x_1, \dots, x_p$ aus?

16. Aufgabe

(2 Punkte)

Nach einer Regressions sollten die Residuen normalverteilt sein. Was bei einer simplen Regression noch relativ einfach visuell in einem Scatterplot zu überprüfen ist. Für komplexere Modell liefert der QQ-Plot die notwendigen Informationen über die Normalverteilung. Welche Aussage ist richtig?

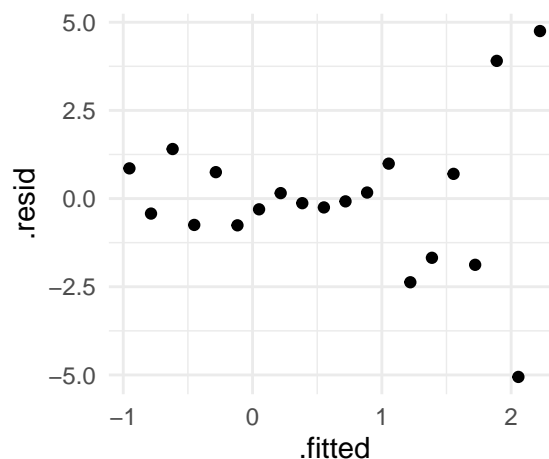


- A** ☐ Wir betrachten die Gerade. Wenn die Punkte einigermaßen gleichmäßig um die Gerade verteilt liegen, dann gehen wir von normalverteilten Residuen aus. Dies ist hier nicht der Fall. Wir haben keine normalverteilten Residuen vorliegen.
- B** ☐ Die Annahme der normalverteilten Residuen ist erfüllt.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

17. Aufgabe

(2 Punkte)

Nach einer Regressions sollten die Residuen (.resid) gleichmäßig um die Gerade verortet sein. Was bei einer simplen Regression noch relativ einfach visuell in einem Scatterplot zu überprüfen ist. Für komplexere Modell liefert der Residual Plot die notwendigen Informationen. Welche Aussage ist richtig?



- A** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil auf der Diagonalen. Damit ist das Modell erfolgreich geschätzt worden.
- B** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Es ist kein Muster zu erkennen.
- C** ☐ Wenn wir die Nulllinie betrachten so müssen die Punkte gleichmäßig unter der Nulllinie liegen. Unser Modell erfüllt somit nicht die Annahme von normalverteilten Residuen mit einem Mittelwert von > 0 und einer Streuung von s .

- D** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Kein Muster ist zu erkennen und keine Outlier zu beobachten.
- E** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Ein klares Muster ist zu erkennen und/oder einige Outlier sind zu beobachten.

18. Aufgabe

(2 Punkte)

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

19. Aufgabe


(2 Punkte)




Sie haben ein Feldexperiment mit Wasserlinsen durchgeführt und wollen nun in einer simplen linearen Regression den Einfluss der CO_2 -Konzentration in $[\mu g]$ im Wasser auf das Trockengewicht in $[kg]$ untersuchen. Sie erhalten einen β_{CO_2} Koeffizienten von 7.4×10^{-6} und einen p -Wert mit $2e - 04$. Welche Aussage zu der Signifikanz und dem Effekt ist richtig?

- A** ☐ Wenn der Effekt β_{CO_2} sehr klein ist, dann kann es an einer falsch gewählten Einheit liegen. Der Anstieg von einer Einheit in X führt ja zu einer Änderung von β_{CO_2} in y . Daher ist hier mit einer anderen Einheit in den Daten zu rechnen, so dass wir hier einen besser formatierten Effekt sehen. Der p -Wert stammt aus einer einheitslosen Teststatistik.
- B** ☐ Die Einheit der CO_2 -Konzentration ist zu klein gewählt. Dadurch sehen wir den sehr kleinen p -Wert. Der p -Wert und die Einheit von der CO_2 -Konzentration hängen antiproportional zusammen.
- C** ☐ Die Fallzahl ist zu klein angesetzt. Je kleiner die Fallzahl ist, desto höher ist die Teststatistik und damit auch der p -Wert kleiner. Wir brauchen also mehr Fallzahl um den geringen Effekt noch signifikant zu kriegen.
- D** ☐ Wenn der Effekt β_{CO_2} winzig ist, dann kann es an einer falsch gewählten Einheit liegen. Der Anstieg von einer Einheit in X führt ja zu einer Änderung von β_{CO_2} in x . Wir müssen daher die Einheit von y entsprechend anpassen.
- E** ☐ Manchmal ist die Einheit der Einflussvariable X zu groß gewählt, so dass der Anstieg von 1 Einheit in X zu einer zu großen Änderung in y führt. Daher kann der Effekt β_{CO_2} sehr klein wirken, da der p -Wert wird auf einer einheitslosen Teststatistik bestimmt wird.

20. Aufgabe

(2 Punkte)




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

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

21. Aufgabe

(2 Punkte)

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

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

Vermischte Themen

22. Aufgabe


(2 Punkte)





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

- A** ☐ Randomisierung ist die direkte Folge von Strukturgleichheit. Die Strukturgleichheit erlaubt es erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- B** ☐ Durch eine Randomisierung können wir von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.
- C** ☐ Strukturgleichheit ist durch Randomisierung gegeben. Leider hilft die Randomisierung noch nicht um von der Stichprobe auf die Grundgesamtheit zu schließen. Deshalb wurde das Falsifikationsprinzip entwickelt.
- D** ☐ Randomisierung war bis 1952 bedeutend, wurde dann aber in Folge besserer Rechnerleistung nicht mehr verwendet. Aktuelle Statistik nutzt keine Randomisierung mehr.
- E** ☐ Randomisierung bringt starke Unstrukturiertheit in das Experiment und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.

23. Aufgabe


(2 Punkte)



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

- A** ☐  Pakete sind nur in englischer Sprache verfasst. Es macht keinen Sinn  daher in Deutsch zu bedienen.
- B** ☐ Die Spracherkennung von  ist nicht in der Lage Deutsch zu verstehen.
- C** ☐ 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.
- D** ☐ Im Allgemeinen haben Programmiersprachen Probleme mit Umlauten und Sonderzeichen, die in der deutschen Sprache vorkommen. Eine Nutzung der englischen Sprache umgeht dieses Problem auf einfache Art.
- E** ☐ Alle Funktionen und auch Anwendungen sind in  in englischer Sprache. Die Nutzung von deutschen Wörtern ist nicht schick und das ist zu vermeiden.

24. Aufgabe

(2 Punkte)

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

- A** ☐ Wir lesen als erstes die Daten über `read_excel()` ein, transformieren die Spalten über `mutate()` in die richtige Form und können dann über `ggplot()` uns die Abbildungen erstellen lassen. Wichtig ist, dass wir keine Faktoren sondern nur numerische Variablen vorliegen haben.
- B** ☐ Für eine explorativen Datenanalyse (EDA) in  müssen wir als erstes die Daten über `read_excel()` einlesen. Danach müssen wir schauen, dass wir die Spalten richtig über `mutate()` transformiert haben. Insbesondere müssen Variablen mit Kategorien in einen Faktor umgewandelt werden. Am Ende nutzen wir die Funktion `ggplot()` für die eigentlich EDA.
- C** ☐ Für eine explorativen Datenanalyse (EDA) in  müssen wir als erstes die Daten über `read_excel()` einlesen. Danach müssen wir schauen, dass wir die Zeilen richtig über `mutate()` transformiert haben. Insbesondere müssen Variablen mit kontinuierlichen Werten in einen Faktor umgewandelt werden. Am Ende nutzen wir die Funktion `ggplot()` für die eigentlich EDA.
- D** ☐ 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.
- E** ☐ Die Funktionsreihenfolge ist wie folgt: `read_excel()` -> `mutate()` -> `ggplot()`. Dabei ist bei der Transformation der Daten darauf zu achten, dass keine Faktoren erstellt werden.

25. Aufgabe

(2 Punkte)

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

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

26. Aufgabe

(2 Punkte)

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

- A** ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.
- B** ☐ Die Ferkel stammen von der gleichen Sau und sind somit untereinander unabhängig.
- C** ☐ Untereinander abhängig, wenn die Mütter ebenfalls miteinander verwandt sind. Erst die Abhängigkeit 2. Grades wird in der Statistik modelliert.
- D** ☐ Die Ferkel stammen von der gleichen Sau und sind somit untereinander abhängig.
- E** ☐ Untereinander unabhängig. Sollten die Mütter verwandt sein, so ist die Varianzstruktur ähnlich und muss modelliert werden.

27. Aufgabe

(2 Punkte)

Neben der Mittelwertsdifferenz als Effektschätzer bei normalverteilten Endpunkten wird auch häufig der Effektschätzer Risk ratio bei binären Endpunkten verwendet. Welche Aussage über den Effektschätzer Risk ratio ist im folgenden Beispiel zur Behandlung von Klaueninfektionen bei Kühen richtig? Dabei sind 4 Tiere krank und 8 Tiere sind gesund.

- A** ☐ Da es sich um ein Chancenverhältnis handelt ergibt sich ein Risk ratio von 3.
- B** ☐ Es ergibt sich ein Risk ratio von 0.33, da es sich um ein Anteil handelt. Wir berechnen den Anteil der Kranken.
- C** ☐ Da es sich um ein Chancenverhältnis handelt ergibt sich ein Risk ratio von 0.5.
- D** ☐ Der Anteil der Gesunden wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Risk ratio von 0.33.
- E** ☐ Es ergibt sich ein Risk ratio von 0.5, da es sich um ein Anteil handelt.

28. Aufgabe

(2 Punkte)

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

- A** ☐ Aktuell werden immer größere Datensätze erhoben. Dadurch wird auch die Varianz immer höher was automatisch zu mehr signifikanten Ergebnissen führt.
- B** ☐ Eine große Fallzahl führt zu mehr signifikanten Ergebnissen auch bei kleinen Effekten. Daher werden fast alle Vergleich esignifikant, wenn die Fallzahl nur groß genug wird.
- C** ☐ 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.
- D** ☐ Aktuell werden zu grosse Datensätze für die gängige Statistik gemessen. Daher wendet man maschinelle Lernverfahren für kausale Modelle an. Hier ist die Relevanz gleich Signifikanz.
- E** ☐ Relevanz und Signifikanz haben nichts miteinander zu tun. Daher gibt es auch keinen Zusammenhang zwischen hoher Fahlzahl ($n > 10000$) und einem signifikanten Test. Ein Effekt ist immer relevant und somit signifikant.

Multiple Gruppenvergleiche

29. Aufgabe

(2 Punkte)

Sie haben folgende unadjustierten p-Werte gegeben: 0.02, 0.03, 0.01, 0.42 und 0.34. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.004, 0.006, 0.002, 0.084 und 0.068. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.1, 0.15, 0.05, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.
- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.1, 0.15, 0.05, 2.1 und 1.7. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.004, 0.006, 0.002, 0.084 und 0.068. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.1, 0.15, 0.05, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.

30. Aufgabe

(2 Punkte)

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

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

31. Aufgabe

(2 Punkte)

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

- A** ☐ Das R Paket {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.
- B** ☐ Das R Paket {lm}. Das Paket {lm} erstellt selbstständig Konfidenzintervalle und entsprechende p-Werte. Da wir in dem Paket nicht adjustieren müssen, ist es bei Anwendern sehr beliebt.
- C** ☐ Da Sie für Ihre Bachelorarbeit einen Barplot mit CLD brauchen nutzen Sie das R Paket {emmeans} welches Ihnen schnell die notwendigen Informationen liefert um einen Barplot zu erstellen. Die Berechnung eines CLD ist hierbei auch einfach.
- D** ☐ Das R Paket {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.
- 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.

32. Aufgabe

(2 Punkte)

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

- A** ☐ 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.
- B** ☐ Beim multiplen Testen kann es zu einer β -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 20%. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist.
- C** ☐ 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** ☐ Die Adjustierung der p-Werte nach Bonferroni erlaubt es gegen die α -Inflation vorzugehen, die häufig beim multiplen Testen auftritt. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern sehr viel höher. Das ist der Grund warum die p-Werte entsprechend adjustiert werden müssen.
- E** ☐ Beim multiplen Testen kann es zu einer α -Deflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist. Die p-Werte werden durch die Anzahl an Vergleichen geteilt

33. Aufgabe

(2 Punkte)

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

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

Statistische Testtheorie

34. Aufgabe

(2 Punkte)

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

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

35. Aufgabe

(2 Punkte)

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

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

36. Aufgabe

(2 Punkte)

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

- A** ☐ Als Kulturkonstante hat $\alpha = 5\%$ den Rang einer Naturkonstante und wurde nach langer Diskussion in der UN im Jahre 1983 festgesetzt. Damals auch schon mit der Zustimmung der UdSSR.
- B** ☐ Im Rahmen eines langen Disputs zwischen Neyman und Fischer wurde $\alpha = 5\%$ festgelegt. Leider werden die Randbedingungen und Voraussetzungen an statistische Modelle heute immer wieder ignoriert.
- C** ☐ Auf einer Statistikkonferenz in Genf im Jahre 1942 wurde dieser Cut-Off nach langen Diskussionen festgelegt. Bis heute ist der Cut Off aber umstritten, da wegen dem 2. Weltkrieg viele Wissenschaftler nicht teilnehmen konnten.
- D** ☐ Der Begründer der modernen Statistik, R. Fischer, hat die Grenze simuliert und berechnet. Dadurch ergibt sich dieser optimale Cut-Off.
- E** ☐ Da Wissenschaftler eine Schwelle für die statistische Testentscheidung benötigen wurde α historisch gewählt. Damit ist $\alpha = 5\%$ eine Kulturkonstante.

37. Aufgabe

(2 Punkte)

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

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

38. Aufgabe

(2 Punkte)

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

H_0 ablehnen obwohl die H_0 gilt

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

39. Aufgabe

(2 Punkte)

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

- A ☐ 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.
- B ☐ Das OR. Als Chancenverhältnis gibt es das Verhältnis von Relevanz und Signifikanz wieder.
- C ☐ Der p-Wert. Durch den Vergleich mit α lässt sich über die Signifikanz entscheiden und der β -Fehler erlaubt über die Power eine Einschätzung der Relevanz.
- D ☐ 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 bietet eine Entscheidung über die Signifikanz und zusätzlich kann über die Visualisierung des Konfidenzintervalls eine Relevanzschwelle definiert werden.

40. Aufgabe

(2 Punkte)

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

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

41. Aufgabe

(2 Punkte)

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

- A** ☐ In der Analogie der 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.
- B** ☐ In der Analogie des Niederschlags oder Regenmenge: ein statistischer Test gibt die Stärke eines Effektes wieder. Zum Beispiel, wie hoch ist der Mittelwertsunterschied.
- C** ☐ In der Analogie der Maximaltemperatur: Was ist der maximale Unterschied zwischen zwei Gruppen. Wir erhalten hier eine Aussage über die Spannweite und den maximalen Effekt.
- D** ☐ In der Analogie der Sonnenscheindauer: Wie lange kann mit einem entsprechenden Effekt gerechnet werden? Die Wahrscheinlichkeit für den Effekt gibt der statistische Test wieder.
- E** ☐ In der Analogie der Wahrscheinlichkeit für Regen: ein statistischer Test erlaubt die Wahrscheinlichkeit für ein Ereignis abzuschätzen. Die Stärke des Effektes können wir nicht bestimmen.

42. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit wollen Sie eine Aussage über die untersuchte Population treffen. Dazu nutzen Sie einen statistischen Test. Können Sie eine valide Aussage treffen?

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

43. Aufgabe

(2 Punkte)

Welche Aussage über die *Power* ist richtig?

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

44. Aufgabe

(2 Punkte)

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

- A** ☐ Der Forschende muss am Ende wissen, ob das Ergebnis eines Experiments relevant für seine Forschung ist. Dafür kann der Effekt eines statistischen Tests genutzt werden. Damit beschreibt der Effekt den biologischen interpretierbaren Teil einer Ausgabe eines Tests. Zum Beispiel der Unterschied zwischen zwei Anteilen.
- B** ☐ Der 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 Effekt eines statistischen Tests beschreibt die mathematisch interpretierbare Ausgabe eines Tests. Damit ist der Effekt direkt mit dem Begriff der Signifikanz verbunden. Die Entscheidung über die Signifikanz trifft der Forschende unabhängig von der Relevanz eines statistischen Tests.
- D** ☐ Der Effekt eines statistischen Tests beschreibt den Output oder die Wiedergabe eines Tests in einem Computer.
- E** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Moderne Algorithmen liefern keine Effekte mehr sondern nur noch bedingte Wahrscheinlichkeiten. Der Effekt spielt in der modernen Statistik keine Rolle mehr.

45. Aufgabe

(2 Punkte)

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

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

46. Aufgabe

(2 Punkte)

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

- A** ☐ Die Hypothesen H_0 und H_A sind rein prosaischer Natur und bilden keinen mathematischen Hintergrund ab. In der Statistik wird die wissenschaftliche Fragestellung getestet. Daher stehen auch die verständlichen Hypothesen im Mittelpunkt der biologischen Interpretation.
- B** ☐ Es gibt - bedingt durch das Falsifikationsprinzip - ein Set von k Nullhypothesen, die iterative gegen $k - 1$ Alternativhypothesen getestet werden.
- C** ☐ Es gibt ein statistisches Hypothesenpaar mit der Hypothese für und gegen die wissenschaftliche Fragestellung. Die Hypothesen werden H_{pro} und H_{contra} bezeichnet.
- D** ☐ Mit der Nullhypothese H_0 und der Alternativhypothese H_A oder H_1 gibt es zwei Hypothesen.
- E** ☐ Mit der Nullhypothese H_A und der Alternativhypothese H_0 gibt es zwei Hypothesen, die aber selten genutzt werden.

Statistische Tests für Gruppenvergleiche

47. Aufgabe

(2 Punkte)

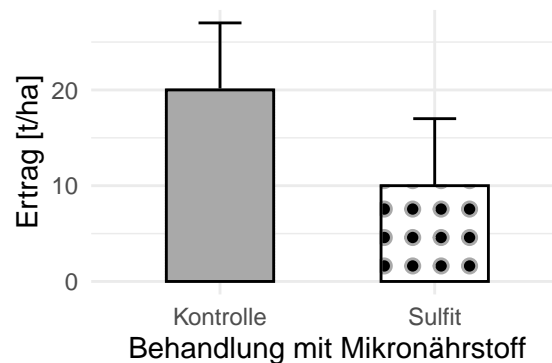
Nach einem Feldexperiment wollen Sie zwei Gruppen mit einem Welch t-Test vergleichen. Welche Aussage ist auch für den Student t-Test richtig?

- A** ☐ Der t-Test vergleicht die Mittelwerte von zwei Gruppen unter der strikten Annahme von Varianzhomogenität. Sollte keine Varianzhomogenität vorliegen, so gibt es keine Möglichkeit den t-Test in einer Variante anzuwenden.
- B** ☐ Der t-Test vergleicht zwei oder mehr Gruppen indem die Mittelwerte miteinander verglichen werden.
- C** ☐ Der t-Test vergleicht zwei Gruppen indem die Mittelwerte miteinander verglichen werden.
- D** ☐ Der t-Test vergleicht die Varianzen von mindestens zwei oder mehr Gruppen
- E** ☐ Der t-Test ist ein Vortest der ANOVA und basiert daher auf dem Vergleich von Streuungsparametern

48. Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung des Mikronährstoff Sulfit auf den Ertrag in t/ha von Papaya im Vergleich zu einer Kontrolle. Der Versuch wurde in 6 Parzellen pro Gruppe durchgeführt. Welche Aussage im Bezug auf eine statistische Auswertung ist richtig?



- A** ☐ Die Barplots deuten auf keinen signifikanten Unterschied. Der Effekt liegt vermutlich bei -10 unter einer groben Abschätzung. Wir müssen aber eine ANOVA rechnen um den Effekt wirklich bestimmen zu können.
- B** ☐ Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei -10.
- C** ☐ Der Test deutet auf keinen signifikanten Unterschied hin. Der Effekt liegt vermutlich bei -10.
- D** ☐ Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei -1.
- E** ☐ Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt kann nicht bei einem t-Test aus Barplots bestimmt werden.

49. Aufgabe

(2 Punkte)

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

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

- C** ☐ Der gepaarte t-Test wird gerechnet, wenn die Beobachtungen abhängig voneinander sind. Wir messen jede Beobachtung nur einmal und berechnen dann die Differenz zu dem Mittel der anderen Beobachtungen.
- D** ☐ Beim gepaarten t-Test kombinieren wir die Vorteile des Student t-Test für Varianzhomogenität mit den Vorteilen des Welch t-Test für Varianzheterogenität. Wir bilden dafür die Differenz der Einzelbeobachtungen.
- E** ☐ Wenn die Beobachtungen nicht unabhängig voneinander sind, rechnen wir einen gepaarten t-Test. Messen wir wiederholt an dem gleichen Tier oder Pflanze dann bilden wir die Differenz zwischen den zwei Messpunkten.

50. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit passen die Ergebnisse einer ANOVA und eines multiplen Vergleiches nicht zusammen. Nach einem Experiment mit vier Maissorten ergibt eine ANOVA ($p = 0.048$). 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** ☐ 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.
- B** ☐ Das ist kein Wunder. Die ANOVA testet auf der gesamten Fallzahl und die paarweisen t-Tests verlieren immer eine oder mehr Gruppen als Fallzahl. Mit steigender Fallzahl sind mehr signifikante Unterschiede zu erwarten. Die p-Werte unterscheiden sich numerisch auch kaum.
- C** ☐ Die adjustierten p-Werte deuten in die richtige Richtung. Zusammen mit den nicht signifikanten rohen p-Werten ist von einem Fehler in der ANOVA auszugehen.
- D** ☐ Die ANOVA testet auf der gesamten Fallzahl. Es wäre besser die ANOVA auf der gleichen Fallzahl wie die einzelnen t-Tests zu rechnen.
- E** ☐ Der Fehler liegt in den t-Tests. Wenn eine ANOVA signifikant ist, dann muss zwangsweise auch ein t-Test signifikant sein.

Teil I.

Deskriptive Statistik & Explorative Datenanalyse

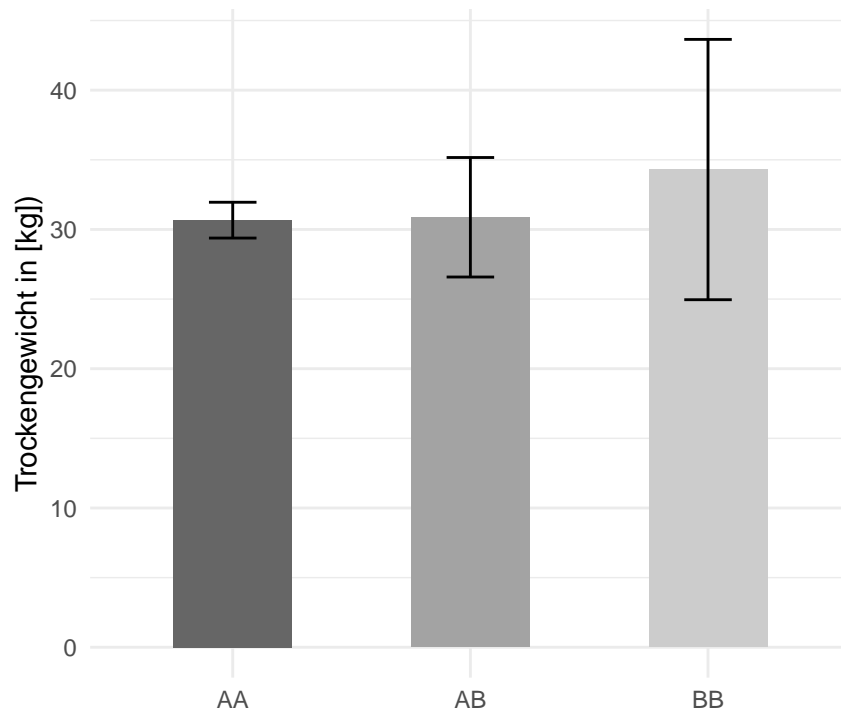
51. Aufgabe

(8 Punkte)

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



Barplots sind bedeutend in der Darstellung von wissenschaftlichen Ergebnissen. Leider hat sich Nilufar nicht gemerkt, welche statistischen Maßzahlen für einen Barplot erhoben werden müssen. Besser wäre was anderes gewesen. Nilufar liebt Blasmusik. Darin kann sie sich wirklich verlieren und immer wieder neu begeistern. Das ist in soweit doof, da nach ihrer Betreuerin erstmal ein Barplot nachgebaut werden soll, bevor es mit ihrem Projektbericht losgeht. Dann hat sie schonmal den Code vorliegen und nachher geht dann alles schneller. Na dann mal los. Nilufar schafft sich die nötige Stimmung. Wenn Deichkind ertönt, dann sucht das Huhn schleunigst Schutz unter dem Sofa. Nilufar schüttelt den Kopf. In der Behandlung für Erbsen werden verschiedene Genotypen (AA, AB und BB) sein. Erfasst wird als Endpunkt (Y) *Trockengewicht*. Nilufar soll dann *drymatter* in ihrer Exceldatei eintragen.



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

52. Aufgabe

(8 Punkte)

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



Barplots sind bedeutend in der Darstellung von wissenschaftlichen Ergebnissen. Leider hat sich Jessica nicht gemerkt, welche statistischen Maßzahlen für einen Barplot erhoben werden müssen. Besser wäre was anderes gewesen. Jessica liebt Warhammer. Darin kann sie sich wirklich verlieren und immer wieder neu begeistern. Das ist in soweit doof, da nach ihrer Betreuerin nun Barplots aus ihren Daten gebaut werden sollen, bevor es mit dem statistischen Testen weitergeht. Na dann mal los. Jessica schafft sich die nötige Stimmung. Wenn David Bowie ertönt, dann sucht der Hund schleunigst Schutz unter dem Sofa. Jessica schüttelt den Kopf. Die Behandlung für Lauch waren verschiedene Bewässerungstypen (*low*, *mid* und *high*). Erfasst wurde von Jessica als Endpunkt (Y) *Frischegewicht*. Jessica hat dann *freshmatter* in ihrer Exceldatei eintragen. Aber auch irgendwie egal. Jessica will später nochmal raus um Rad zu fahren. Druck ablassen, dass muss sie auch.

treatment	freshmatter
high	38.3
low	49.3
high	31.8
low	39.3
mid	34.0
mid	28.8
low	42.1
mid	29.1
mid	31.1
low	40.8
high	21.5
high	21.0
high	41.7
low	46.8

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

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

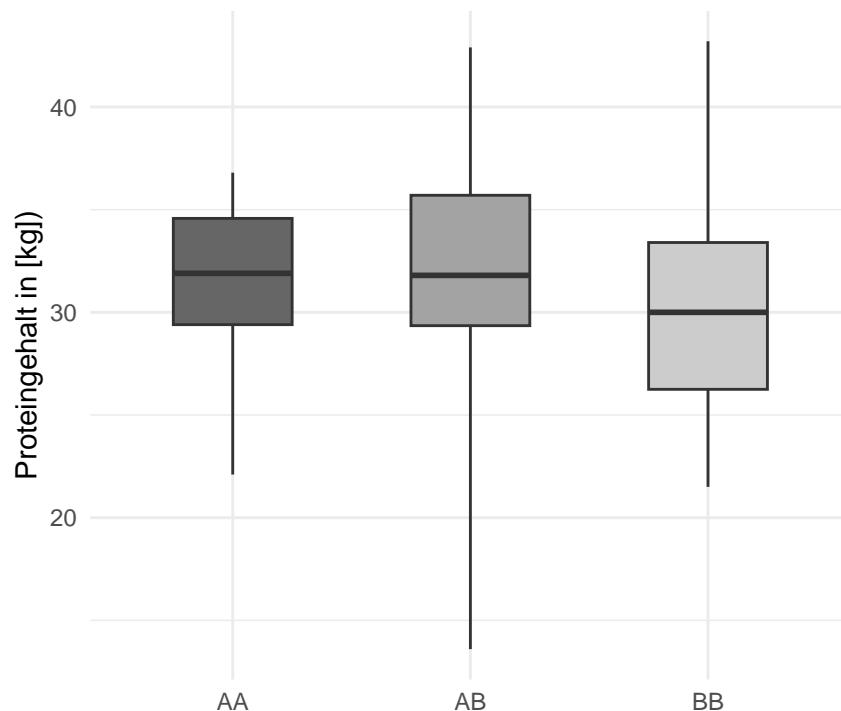
53. Aufgabe

(9 Punkte)

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



Tina steht vor einem ersten Problem, denn wenn es nach ihrem Betreuer geht, soll sie in einem einem Gewächshausexperiment Lauch auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Tina liebt Astronomie. Darin kann sie sich wirklich verlieren und immer wieder neu begeistern. Das heißt erstmal überlegen für Tina. 'Hm...', Katjes und Tocotronic. Das ist und bleibt die beste Kombination zum Nachdenken für Tina. Die Behandlung werden verschiedene Genotypen (AA, AB und BB) sein. In ihrer Exceldatei wird sie den Outcome (Y) *Proteingehalt* als *protein* aufnehmen. Vorab soll Tina aber einmal die folgenden Boxplots ihrem Betreuer nachbauen, damit sie den R Code schonmal für später vorliegen hat. Damit geht das Problem schon los. Tina und die Wut, eine unendliche Geschichte mit kniffligen Wendungen.



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

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

54. Aufgabe

(9 Punkte)

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



Tina steht vor einem ersten Problem, denn wenn es nach ihrem Betreuer geht, soll sie in einem Feldexperiment Erbsen auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Am Ende dann doch besser Astronomie. Wunderbar. Eine echte Ablenkung für Tina. Die Behandlung waren verschiedene Genotypen (AA und BB). In ihrer Exceldatei hat sie den Messwert (Y) *Trockengewicht* als *drymatter* aufgenommen. Nun soll Tina die Daten einmal als Boxplots in einer Präsentation visualisieren, damit ihrem Betreuer wieder klar wird, was sie eigentlich nochmal gemacht hat und was für ein Ergebnis in einem statistischen Test zu erwarten wäre. Anhand von Boxplots lässt sich eine Aussage über die Normalverteilung von Y treffen. Wäre da nicht noch etwas. Eine echte Herausforderung für sie war schon immer die Wut gewesen. Ein leidiges Lied. Aber egal. Einfach mal raus um zu Boxen. Ohne Ziel und Uhr. Das ist was für Tina.

treatment	drymatter
AA	22.0
AA	35.0
AA	37.8
AA	39.9
BB	35.3
BB	26.1
BB	33.5
BB	40.5
BB	31.6
AA	39.0
AA	34.9
BB	36.8
BB	32.4
AA	29.9
AA	25.2
AA	34.3
BB	33.9

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

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

55. Aufgabe

(8 Punkte)

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



In ihrer Hausarbeit möchte Jessica gerne die Daten aus einem Leistungssteigerungsversuch mit Hühnern in einem Histogramm darstellen. Das Histogramm erlaubt ihr dabei Rückschlüsse auf die Verteilung über den Messwert (Y) zu treffen. Jessica schmeißt noch eine Handvoll Schokobons in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von David Bowie. In seinem Experiment hat Jessica die auffälligen Hautflecken gezählt. Es wäre einfacher, wenn da nicht noch was wäre. Wenn der Mangel nicht wäre, ja dann wäre wohl vieles möglich für Jessica! Aber so.. Wenn David Bowie ertönt, dann sucht der Hund schleunigst Schutz unter dem Sofa. Jessica schüttelt den Kopf.

Die auffälligen Hautflecken: 3, 7, 9, 3, 2, 5, 3, 4, 4, 9, 8, 1, 2, 3, 2, 4, 4, 2, 4, 2, 4, 4, 7, 3, 5, 3, 2, 5, 8, 4

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

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

56. Aufgabe

(8 Punkte)

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



In ihrer Projektbericht möchte Jessica gerne die Daten aus einem Stallexperiment mit Hühnern in einem Histogramm darstellen. Das Histogramm erlaubt ihr dabei Rückschlüsse auf die Verteilung über den Endpunkt (Y) zu treffen. Jessica schmeißt noch eine Handvoll Schokobons in ihren Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von David Bowie. In seinem Experiment hat Jessica die mittlere Anzahl an weißen Blutkörperchen gezählt. Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für sie war schon immer der Mangel gewesen. Ein leidiges Lied. Jessica nickt im Takt von David Bowie und bemerkt dabei gar nicht was der Hund schon wieder anstellt.

Die mittlere Anzahl an weißen Blutkörperchen: 13.9, 11.3, 11.1, 13.6, 7.8, 11.5, 10.2, 7.9, 10.6, 11, 8.2, 7, 13.8, 9, 13.4, 9.8, 12.7, 11.6, 10.2, 13.5, 10.8, 10.2, 10.5, 13.9, 7.7

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

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

57. Aufgabe

(10 Punkte)

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



'Hm...', Katjes und Tocotronic. Das ist und bleibt die beste Kombination zum Nachdenken für Tina. Jetzt heißt es aber erstmal auf was anderes konzentrieren. Tina möchte gerne den Zusammenhang zwischen durchschnittlicher Tagestemperatur [C/d] und Schlachtgewicht [kg] im Kontext von Hühnern herausfinden. Hierfür hat Tina ein Stallexperiment im Emsland durchgeführt. Nach einigen unvorgesehenen Ereignissen hat sie es geschafft folgende Datentabelle zu erstellen. Wenn die Wut nicht wäre, ja dann wäre wohl vieles möglich für Tina! Aber so.. Aber das steht auch nicht im Zentrum. Nun stellt sich die Frage für sie, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Deshalb möchte Tina als erstes eine explorative Datenanalyse durchführen. Dann was anderes. Das Verrückte ist, dass die Spinne Indiana Jones wirklich liebt. Das ist Tina sehr recht, denn sie braucht Entspannung.

Durchschnittlicher Tagestemperatur [C/d]	Schlachtgewicht [kg]
18.5	24.0
32.6	32.1
15.6	15.4
24.1	28.5
20.9	20.5
16.3	24.1
27.3	27.8
23.9	25.2
17.2	23.1
25.9	26.5

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

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

58. Aufgabe

(10 Punkte)

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



Zwei kategoriale Variablen darzustellen ist nicht so einfach. Tina hatte erst über einen Mittelwert nachgedacht, dann aber die Idee verworfen. Wäre da nicht noch was anderes. Eine echte Herausforderung für sie war schon immer die Wut gewesen. Ein leidiges Lied. Dabei hatte sie sich in ein Kreuzungsexperiment im Oldenburger Land zum einen die Behandlung Ökologisch [ja/nein] und zum anderen die Messung Schlachtgewicht im Zielbereich [ja/nein] im Kontext von Hühnern angeschaut. Jetzt möchte ihre Betreuerin erstmal die langen Tabellen mit ja/nein in einer explorativen Datenanalyse zusammengefasst bekommen. Sonst geht es bei ihrer Hausarbeit nicht weiter. Was super nervig ist. Einfach mal raus um zu Boxen. Ohne Ziel und Uhr. Das ist was für Tina.

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

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

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

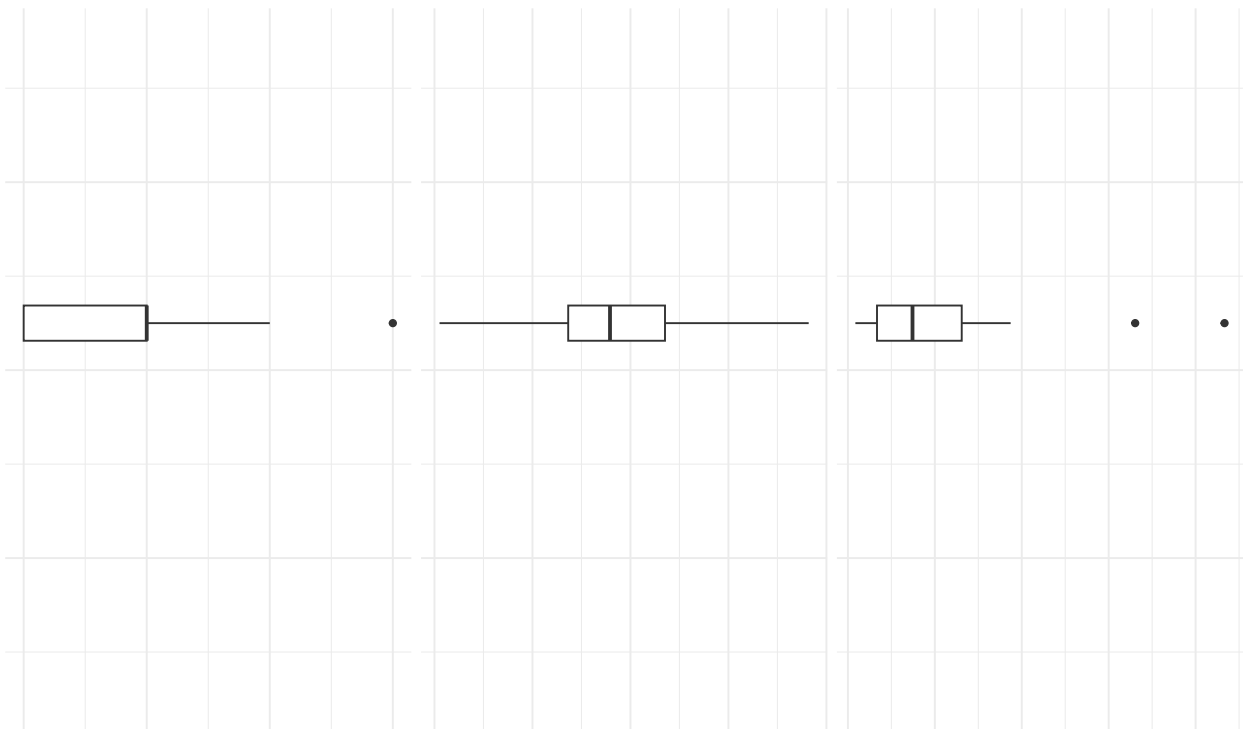
59. Aufgabe

(10 Punkte)

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



1. Zeichnen Sie über die untenstehenden Boxplots die entsprechende zugehörige Verteilung! **(3 Punkte)**
2. Zeichnen Sie unter die untenstehenden Boxplots die entsprechende zugehörige Beobachtungen als Stiche! **(3 Punkte)**
3. Wieviel Prozent der Beobachtungen fallen in das IQR? Ergänzen Sie die Abbildung entsprechend um den Bereich! **(2 Punkte)**
4. Wieviel Prozent der Beobachtungen fallen in $\pm 1s$ unter der Annahme einer Normalverteilung? Wenn möglich, ergänzen Sie die Abbildung entsprechend um den Bereich! **(2 Punkte)**



60. Aufgabe

(9 Punkte)

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



1. Skizzieren Sie 4 Normalverteilungen *in einer Abbildung* mit $\bar{y}_1 \neq \bar{y}_2 \neq \bar{y}_3 \neq \bar{y}_4$ und $s_1 = s_2 = s_3 = s_4$! **(3 Punkte)**
2. Beschriften Sie die Normalverteilungen mit den entsprechenden Parametern! **(2 Punkte)**
3. Ergänzen Sie die Bereiche in der 68% und 95% der Beobachtungen fallen! Beschriften Sie die Grenzen der Bereiche mit der statistischen Maßzahl! **(2 Punkte)**
4. Liegt Varianzhomogenität oder Varianzheterogenität vor? Begründen Sie Ihre Antwort! **(2 Punkte)**

61. Aufgabe

(9 Punkte)

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



1. Skizzieren Sie in die unten stehenden, freien Abbildungen die Verteilungen, die sich nach der Abbildungsüberschrift ergeben! **(6 Punkte)**
2. Beschriften Sie die Achsen der Abbildungen entsprechend! **(1 Punkt)**
3. Achten Sie auf die entsprechende Skalierung der beiden Verteilungen in den Abbildungen! **(2 Punkte)**

$N(1, 1)$ und $N(0, 9)$



$Pois(20)$ und $Pois(1)$



62. Aufgabe

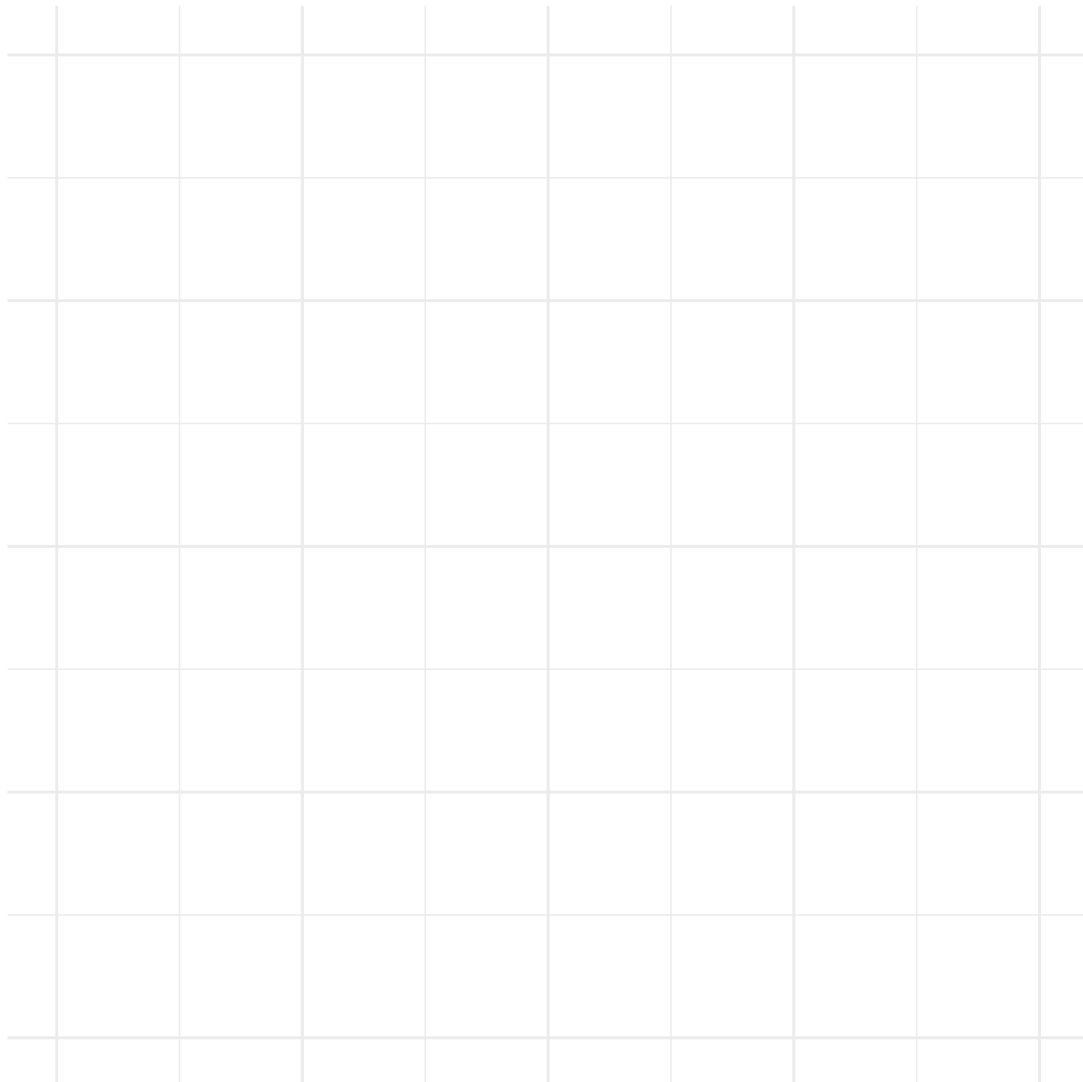
(8 Punkte)

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



Sie haben folgende Zahlenreihe y vorliegen $y = \{17, 22, 15, 19, 18, 16, 16\}$.

1. Visualisieren Sie den Mittelwert von y in der untenstehenden Abbildung! **(4 Punkte)**
2. Beschriften Sie die Y und X -Achse entsprechend! **(2 Punkte)**
3. Für die Berechnung der Varianz wird der Abstand der einzelnen Werte y_i zum Mittelwert \bar{y} quadriert. Warum muss der Abstand, $y_i - \bar{y}$, in der Varianzformel quadriert werden? Erklären Sie den Zusammenhang unter Berücksichtigung der Abbildung! **(2 Punkte)**



Teil II.

Statistisches Testen & statistische Testtheorie

63. Aufgabe

(9 Punkte)



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

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

64. Aufgabe

(9 Punkte)



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

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

5% H_0 abgelehnt H_0 beibehalten H_0 falsch

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

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

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

65. Aufgabe

(9 Punkte)

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



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

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



66. Aufgabe

(10 Punkte)



Sie rechnen einen t-Test für Gruppenvergleiche der Mittelwerte. Sie schätzen den Unterschied zwischen dem mittleren Trockengewicht nach Düngergabe zu einer unbehandelten Kontrolle.

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



67. Aufgabe

(10 Punkte)

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



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

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

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

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

Teil III.

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

68. Aufgabe

(9 Punkte)

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



Jessica ist im Wendland für einen Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) mit Hühnern. Allein diese Tatsache ist für sie eine Erzählung wert. Jessica und der Mangel, eine unendliche Geschichte mit kniffligen Wendungen. Für ihren Projektbericht musste sie einen Leistungssteigerungsversuch mit Hühnern durchführen und das sollte laut ihrem Betreuer an diesem Ort besonders gut gelingen, da man hier gut neue technische Anlagen und Behandlungen fernab der Bevölkerung testen könne. Zeugen gibt es hier jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn ihre Behandlung Bestandsdichte (*Verordnung* und *Erhht*) und der Messwert Gewichtszuwachs in der 1LW sollen mit einem t-Test ausgewertet werden. Immerhin weiß sie, dass ihr Messwert einer Normalverteilung folgt. Hm..., was entspannendes wäre gut. Einfach mal raus um Rad zu fahren. Ohne Ziel und Uhr. Das ist was für Jessica.

treatment	weight
ctrl	11.3
dose	16.1
ctrl	14.7
dose	14.4
ctrl	14.1
dose	16.2

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

1. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
2. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 2.68$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
5. 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!



'Der t-Test testet einen normalverteilten Messwert (Y).', liest Alex laut. Das hilft jetzt auch nur bedingt weiter. Wenn das Aufstehen nicht wäre, ja dann wäre wohl vieles möglich für Alex! Aber so.. Laut seinem Betreuer ist zwar ihm Messwert Fettgehalt [%/kg] normalverteilt, aber wie rechnet er jetzt einen t-Test? Für seine Abschlussarbeit musste er ein Kreuzungsexperiment mit Hühnern im Wendland durchführen. Als wäre das nicht schon anstrengend genug gewesen. Jetzt soll er auch noch testen, ob die Behandlung Lüftungssystem (*keins* und *vorhanden*) ein signifikantes Ergebnis liefert. Schon dutzende Male gesehen: Alien. Aber immer noch großartig zusammen mit Gummibärchen.

Lüftungssystem	Fettgehalt
keins	51.9
keins	62.0
keins	48.3
keins	39.8
keins	45.9
keins	50.4
vorhanden	48.9
keins	48.4
keins	47.3
vorhanden	46.8
vorhanden	40.6
keins	33.7
vorhanden	48.1
vorhanden	44.8
vorhanden	59.4
vorhanden	48.7
vorhanden	43.0
vorhanden	40.9

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.96$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *einen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann die Teststatistik T_D ? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Alex über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

70. Aufgabe

(12 Punkte)

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



Der t-Test. Tina erschauert. Tina und die Wut, eine unendliche Geschichte mit kniffligen Wendungen. Ein mächtiges Werkzeug in den Händen desjenigen, der einen normalverteilten Messwert (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Tina überhaupt aus? 'Hm...', Katjes und Tocotronic. Das ist und bleibt die beste Kombination zum Nachdenken für Tina. Tina hat ein Kreuzungsexperiment mit Hühnern durchgeführt. Dabei wurde die Behandlung Bestandsdichte (*Verordnung* und *Erhöht*) an den Hühnern getestet. Gemessen hat Tina dann als Messwert Gewichtszuwachs in der 1LW. Warum der Versuch im Wendland für ihre Abschlussarbeit stattfinden musste, ist ihr bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Gewichtszuwachs in der 1LW?

Bestandsdichte	Gewichtszuwachs
Erhöht	37.5
Erhöht	38.0
Verordnung	22.5
Erhöht	40.3
Erhöht	31.3
Verordnung	29.8
Erhöht	37.7
Verordnung	27.5
Erhöht	40.7
Erhöht	37.4
Verordnung	41.6
Erhöht	42.0
Erhöht	40.2
Verordnung	45.2
Erhöht	31.9
Verordnung	34.0
Erhöht	33.8
Verordnung	29.3

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

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

71. Aufgabe

(11 Punkte)

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



Alles voll mit Hühnern. Das haben Mark und Jonas gemeinsam in einer Hausarbeit gemacht! Worum ging es aber konkret? Beide haben als einen normalverteilten Messwert (Y) Gewichtszuwachs in der 1LW von Hühnern bestimmt. Die Daten haben beide zusammen in einem Kreuzungsexperiment erhoben. In dem Experiment ging es um eine vorher/nachher Untersuchung an den gleichen Hühnern. Als Behandlung wurde Ernährungszusatz (*ohne* und *14d*) eingesetzt. Nach der Meinung der Betreuerin muss hier ein gepaarter t-Test gerechnet werden. Mark schaut nachdenklich zu Jonas. Wenn die Unsicherheit nicht wäre, ja dann wäre wohl vieles möglich für Mark! Aber so... Steffen denkt derweil angestrengt an Snickers. Im Hintergrund wummert Andrea Berg.

ID	treatment	freshmatter
10	14d	33.4
1	ohne	29.0
8	14d	18.1
3	14d	18.2
6	ohne	32.4
5	ohne	34.4
2	14d	26.2
5	14d	28.7
9	ohne	30.6
7	ohne	37.4
6	14d	41.0
7	14d	28.7
1	14d	26.1
2	ohne	45.7
4	14d	15.0
4	ohne	40.3
3	ohne	39.4
9	14d	28.4
8	ohne	25.0

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


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

72. Aufgabe

(10 Punkte)

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



‘Wir waren im Emsland um Hühnern in einem Stallexperiment zu messen.’, Yuki legt das Dokument auf den Tisch und schaut Alex und Steffen fragend an. Beide schauen fragend zurück. Die beiden sind zu ihr gekommen, da sie sich nicht mit  auskennen und daher Hilfe bei der Interpretation des t-Tests brauchen. Kein Problem. Aber worum es in dem Versuch geht, lässt sich nur aus dem Text in seiner Hand errahnen. Vielleicht hilft da ja die Ausgabe des t-Tests in R weiter.

```
##  
## Two Sample t-test  
##  
## data: Schlachtgewicht by Ernährungszusatz  
## t = -0.6414, df = 14, p-value = 0.5316  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -10.652911 5.748149  
## sample estimates:  
## mean in group ctrl mean in group fedX  
## 41.93333 44.38571
```

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.14|$ einzeichnen! **(4 Punkte)**
5. Beschriften Sie die Abbildung! **(1 Punkt)**
6. Berechnen Sie den Effekt des t-Tests! **(1 Punkt)**

73. Aufgabe

(8 Punkte)

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



‘Mit dem R Paket {emmeans} können wir gleich die Gruppenvergleiche rechnen und uns das *compact letter displac*’ ausgeben lassen!’, verkündet Mark sichtlich stolz. ‘Nach Meinung der Betreuerin soll es aber nur erstmal ein t-Test sein. Und die Ausgabe ist schon wirr genug.’, merkt Tina an. Nilufar und Tina sind bei Mark um sich in R helfen zu lassen. Die beiden waren 1 Monate im Oldenburger Land um einen Versuch mit Hühnern in einem Leistungssteigerungsversuch durchzuführen. Ziel war es das Outcome (Y) Protein/Fettrate [%/kg] zu bestimmen.

```
##
## Two Sample t-test
##
## data: Protein/Fettrate by Ernährungszusatz
## t = 0.83679, df = 16, p-value = 0.415
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -3.856426 8.886426
## sample estimates:
## mean in group ctrl mean in group fedX
##                45.875                43.360
```

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


1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie das sich ergebende 95% Konfidenzintervall! **(2 Punkte)**
5. Beschriften Sie die Abbildung und das 95% Konfidenzintervall entsprechend! **(2 Punkte)**

74. Aufgabe

(9 Punkte)

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



‘Mit dem R Paket {emmeans} können wir gleich die Gruppenvergleiche rechnen und uns das *compact letter displac*’ ausgeben lassen!’, verkündet Steffen sichtlich stolz. ‘Nach Meinung der Betreuerin soll es aber nur erstmal ein t-Test sein. Und die Ausgabe ist schon wirr genug.’, merkt Nilufar an. Nilufar und Alex sind bei Steffen um sich in  helfen zu lassen. Die beiden waren 3 Monate im Wendland um einen Versuch mit Hühnern in einem Stallexperiment durchzuführen. Ziel war es das Outcome (Y) Schlachtgewicht [kg] zu bestimmen.

```
##  
## Two Sample t-test  
##  
## data: Schlachtgewicht by Elterlinie  
## t = -0.56731, df = 14, p-value = 0.5795  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -8.784352 5.109352  
## sample estimates:  
## mean in group Standard      mean in group Xray  
##                29.7750                31.6125
```

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

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

75. Aufgabe

(8 Punkte)

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



Es gibt ja immer die Möglichkeit sich Hilfe zu holen. Das geht natürlich auch immer in einer Abschlussarbeit. Deshalb arbeiten Alex und Tina 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. In dem Hauptversuch wurde Folgendes von den beiden gemacht. Alex und Tina haben sich Hühnern angeschaut. Dabei geht um Zusammenhang zwischen Genotypisierung (0d und 14d) und Gewichtszuwachs in der 1LW. Jetzt sollen beide einen gepaarten t-Test rechnen. Leider kennen sich beide nicht sehr gut in R aus.

```
##  
## Paired t-test  
##  
## data: Gewichtszuwachs by Genotypisierung  
## t = -5.4585, df = 8, p-value = 0.0006026  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -31.54708 -12.80848  
## sample estimates:  
## mean difference  
## -22.17778
```

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

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

Teil IV.

Die einfaktorielle & zweifaktorielle ANOVA

76. Aufgabe

(11 Punkte)

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



Jessica und Alex schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten visualisieren damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse zu erwarten sind. Die beiden waren im Emsland um einen Leistungssteigerungsversuch mit Hühnern durchzuführen. Dabei haben Jessica und Alex den Messwert Schlachtgewicht [kg] unter der Behandlung Elterlinie (*Standard*, *Yray* und *Xray*) ermittelt.

Elterlinie	Schlachtgewicht
Standard	25
Standard	25
Yray	34
Xray	41
Standard	25
Xray	41
Standard	25
Xray	38
Xray	41
Yray	35
Yray	37
Yray	31
Yray	32
Xray	42
Yray	34
Standard	25
Yray	34
Standard	25

Leider kennen sich Jessica und Alex mit Darstellung einer einfaktoriellen ANOVA überhaupt nicht aus. Deshalb brauchen beide bei der Erstellung Ihre Hilfe!

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

77. Aufgabe

(9 Punkte)

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



Mark und Yuki schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten mit einer einfaktoriellen ANOVA auswerten damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse in den multiplen Gruppenvergleichen zu erwarten sind. Die beiden waren im Emsland um einen Leistungssteigerungsversuch mit Hühnern durchzuführen. Dabei haben Mark und Yuki den Messwert Schlachtgewicht [kg] unter der Behandlung Flüssignahrung (*ctrl*, *superIn* und *fIOW*) ermittelt.

Leider kennen sich Mark 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)
Flüssignahrung	2				
error	18	546.09			
Total	20	942.67			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.55$ 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)

78. Aufgabe

(12 Punkte)

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



'Als erstes bauen wir uns aus unsere Daten die ANOVA Tabelle dann sehen wir schon, ob unser Gruppenvergleich in der ANOVA signifikant ist.', Yuki schaut Mark fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Mark tut sich auch 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 Emsland um ein Stallexperiment mit Hühnern durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Flüssignahrung (*ctrl*, *superIn* und *fIOW*) und dem Messwert Fettgehalt [%/kg] gibt.

Leider kennen sich Yuki und Mark 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)
Flüssignahrung	2	434.07			
Error	16	192.88			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.63$ 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)

Flüssignahrung	Fallzahl (n)	Mittelwert	Standardabweichung
ctrl	8	5.12	1.55
superIn	5	7.00	5.10
fIOW	6	16.00	3.79


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

79. Aufgabe

(9 Punkte)

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



'Uff... die einfaktorielle ANOVA und . Nicht so einfach... Was sagt mir jetzt die Ausgabe der ANOVA und wo sehe ich, ob da was signifikant ist?', denkt Tina und hebt die Augenbraue. Tina hatte sich einen Leistungssteigerungsversuch mit Hühnern angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) und dem Messwert Schlachtgewicht [kg] gibt. Nun möchte ihre Betreuerin ihrem Projektbericht erstmal eine ANOVA sehen und die Ergebnisse präsentiert bekommen.

```
## Analysis of Variance Table
##
## Response: Schlachtgewicht
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Lüftungssystem  3 5228.2  1742.7   43.141 2.369e-11
## Residuals      32 1292.7    40.4
```

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

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

80. Aufgabe

(12 Punkte)

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



Jessica steht im Emsland. Und das ist noch langweiliger als es sich anhört. Wäre es nur so spannend wie bei ihren Kommilitonen, die in Almería waren. Ödnis wohin man nur blickt. Oder eben Hühnern. Ja, darum geht es in ihrer Hausarbeit. Und wäre das nicht noch alles genug, ist ihr Experiment auch noch als ein Kreuzungsexperiment komplex geraten. Es wurde der Messwert Protein/Fettrate [%/kg] mit dem Behandlung Flüssignahrung (*ctrl*, *superIn* und *fLOW*) sowie der Behandlung Ernährungszusatz (*ctrl* und *getIt*) untersucht. 'Hmpf', denkt Jessica und ruft 'Und jetzt!?' in die Leere.

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

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Flüssignahrung	3	369.93			
Ernährungszusatz	1	11.53			
Flüssignahrung:Ernährungszusatz	3	227.56			
Error	18	357.4			

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

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

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

81. Aufgabe

(10 Punkte)

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



Es ist schon kurz nach fünf und Jonas wird langsam nervös. Jonas wollte heute Abend noch seine E-Sport Qualifikation schauen. Stattdessen versucht seine Betreuerin die Ausgabe der zweifaktoriellen ANOVA zu visualisieren und zu überprüfen, ob es mit der Visualisierung der Daten als Boxplots zusammenpasst. Jonas hatte im Emsland einen Leistungssteigerungsversuch mit Hühnern durchgeführt. Es gab dabei zwei Behandlungen. Einmal Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) sowie als zweite Behandlung Ernährungszusatz (*ctrl* und *getIt*). Gemessen wurde der Messwert (Y) Protein/Fettrate [%/kg]. So kompliziert kann das jetzt doch nicht sein!

```
## Analysis of Variance Table
##
## Response: Protein/Fettrate
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## Lüftungssystem	2	4.38	2.19	0.1708	0.844377
## Ernährungszusatz	1	347.74	347.74	27.1075	5.951e-05
## Lüftungssystem:Ernährungszusatz	2	167.00	83.50	6.5090	0.007463
## Residuals	18	230.91	12.83		

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

82. Aufgabe

(8 Punkte)



In der untenstehenden Tabelle ist die Formel für den F-Test aus der ANOVA und die Formel für den Student t-Test dargestellt. In der ANOVA berechnen Sie die F-Statistik F_{calc} und in dem Student t-Test die T-Statistik T_{calc} .

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

1. Erklären Sie den konzeptionellen Zusammenhang zwischen der F_{calc} Statistik und T_{calc} Statistik! **(2 Punkte)**
2. Visualisieren Sie eine nicht signifikante F_{calc} Statistik sowie eine signifikante F_{calc} Statistik anhand von $MS_{treatment}$ und MS_{error} ! Beschriften Sie die Abbildung! **(2 Punkte)**
3. Erklären Sie an der Formel des F-Tests sowie an der Abbildung warum das Minimum der F-Statistik 0 ist! **(2 Punkte)**
4. Wenn die F-Statistik 0 ist, spricht dies eher für oder gegen die Nullhypothese? Begründen Sie Ihre Antwort! **(2 Punkte)**

83. Aufgabe

(8 Punkte)



Sie rechnen eine zweifaktorielle ANOVA und erhalten einen signifikanten Interaktionseffekt zwischen den beiden Faktoren f_1 und f_2 . Der Faktor f_1 hat drei Level. Der Faktor f_2 hat dagegen nur zwei Level.

1. Visualisieren Sie in zwei getrennten Abbildungen eine schwache und keine Interaktion zwischen den Faktoren f_1 und f_2 ! **(4 Punkte)**
2. Erklären Sie den Unterschied zwischen den beiden Stärken der Interaktion! **(2 Punkte)**
3. Wenn eine signifikante Interaktion in den Daten vorliegt, wie ist dann das weitere Vorgehen bei einem Posthoc-Test? **(2 Punkte)**

84. Aufgabe

(9 Punkte)



Sie rechnen eine einfaktorielle ANOVA mit einem Faktor f_1 mit vier Leveln. Nachdem Sie die einfaktorielle ANOVA gerechnet haben, erhalten Sie einen p-Wert von 0.078 und eine F Statistik mit $F_{calc} = 1.2$. Als Sie sich die Boxplots der Behandlungen anschauen, stellen Sie fest, dass es eigentlich einen Mittelwertsunterschied zwischen dem zweiten und dritten Level geben müsste. Die *IQR*-Bereiche überlappen sich nicht und die Mediane liegen auch weit vom globalen Mittel entfernt.

1. Erklären Sie die Annahme der Normalverteilung und die Annahme der Varianzhomogenität für eine ANOVA an einer passenden Abbildung! **(3 Punkte)**
2. Visualisieren Sie die Berechnung von F_{calc} am obigen Beispiel! **(3 Punkte)**
3. Erklären Sie das Ergebnis der obigen einfaktoriellen ANOVA unter der Berücksichtigung der Annahmen an eine ANOVA! **(3 Punkte)**

Teil V.

Multiple Gruppenvergleiche

85. Aufgabe

(12 Punkte)

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



Mark und Alex untersuchen gemeinsam in ihrer Abschlussarbeit den Messwert Schlachtgewicht [kg] in Hühnern. Es ist einiges schiefgelaufen, wie es immer so passiert. Hauptsächlich waren es Katzen, auch wenn man erstmal bei dem Messwert nicht unbedingt an Katzen denken würde. Aber das ist eine andere Geschichte. Jetzt wollen Mark und Alex ihre Ergebnisse nochmal mit einer Studie von Meyer et al. (2021) vergleichen und schauen, ob was ähnliches rauskommt. Angeschaut wurde sich als Behandlung Bestandsdichte (*effizient*, *standard*, *eng*, *weit* und *kontakt*). Es ergab sich dann die folgende Tabelle der rohen p -Werte für die Vergleiche zu Meyer et al. (2021).

Rohen p -Werte	Adjustierte p -Werte	Nullhypothese ablehnen?
0.0010		
0.0012		
0.7600		
0.2300		
0.3400		

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

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

86. Aufgabe

(10 Punkte)

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



Paula sitzt schon etwas länger bei ihrer Betreuerin. So langsam macht Paula sich Gedanken, ob sie nicht doch mal anmerken sollte, dass sie von CLD noch nie was gehört hat. Aber noch kann gelauscht werden, ein Ende ist erstmal nicht in Sicht! Paula hatte in ihrer Abschlussarbeit ein Kreuzungsexperiment durchgeführt. Deshalb sitzt sie hier. Also eigentlich nein, deshalb nicht. Paula will fertig werden. Hat sie sich doch mit Ernährungszusatz (*ctrl*, *fedX*, *proteinX* und *getIt*) und Gewichtszuwachs in der 1LW schon eine Menge angeschaut. Paula beugt sich leicht nach vorne. Nein, doch keine Pause. Weiter warten auf eine Lücke im Fluss...

Behandlung	Compact letter display
ctrl	a
fedX	a
proteinX	a
getIt	a

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich anhand des *Compact letter display (CLD)* ergebenden Barplots! **(2 Punkte)**
4. Ergänzen Sie das *Compact letter display (CLD)* zu den Barplots! **(1 Punkt)**
5. Erklären Sie *einen* Vorteil und *einen* Nachteil des *Compact letter display (CLD)*! **(2 Punkte)**
6. Erstellen Sie eine Matrix mit den paarweisen *p*-Werten eines Student t-Tests, die sich näherungsweise aus dem *Compact letter display (CLD)* ergeben würde! Begründen Sie Ihre Antwort! **(3 Punkte)**

87. Aufgabe

(12 Punkte)

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



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

Ernährungszusatz	Fallzahl (n)	Mittelwert	Standardabweichung
ctrl	7	10.86	1.55
fedX	9	9.80	3.07
proteinX	9	5.14	1.40
getIt	8	7.16	3.78

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

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

88. Aufgabe

(10 Punkte)

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



'Okay, dann nochmal für mich. Ich habe jetzt alles in DataTab gemacht, aber das Wichtigste, was gemacht werden soll, nämlich das CLD, das kann ich nicht in DataTab machen?', Yuki muss sich echt beherrschen. Immerhin betreut ihr Betreuer ja erst nicht seit gestern Abschlussarbeiten und wusste ja was gemacht werden soll! Yuki hatte sich zwei Variablen mit Genotypen (00, AA, AB und BB) und Protein/Fettrate [%/kg] in ein Stallexperiment mit Hühnern angeschaut. Jetzt möchte sie eigentlich fertig werden und nicht nochmal alles neu in R und {emmeans} machen. Deshalb soll jetzt das CLD per Hand aus der Matrix der p -Wert abgeleitet werden. 'Ich glaube ich wechsel nochmal das Thema...', denkt Yuki, verwirft dann aber den Gedanken.

	00	AA	AB	BB
00	1.0000000	0.8358075	0.5733737	0.1622436
AA	0.8358075	1.0000000	0.4375440	0.1116215
AB	0.5733737	0.4375440	1.0000000	0.3566007
BB	0.1622436	0.1116215	0.3566007	1.0000000

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

1. Formulieren Sie die wissenschaftliche Fragestellung! (1 Punkt)
2. Formulieren Sie die statistischen Hypothesen! (1 Punkt)
3. Zeichnen Sie die sich anhand der Matrix der p -Werte ergebenden Barplots! (2 Punkte)
4. Ergänzen Sie das *Compact letter display (CLD)*! Begründen Sie Ihre Antwort! (4 Punkte)
5. Interpretieren Sie das *Compact letter display (CLD)* für Yuki und Alex! (2 Punkte)

Teil VI.


Der Chi-Quadrat-Test & Der diagnostische Test

89. Aufgabe

(12 Punkte)

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



Am Ende hätte Yuki dann doch einen normalverteilten Endpunkt in seiner Abschlussarbeit nehmen sollen. Vor ihm liegen jetzt die Daten von zwei Variablen als Kategorien oder wie es in  so schön heißt, als Faktoren. Aber immerhin, hofft er das was bei den Daten rausgekommen ist. Gezählt hat Yuki einiges mit $n = 110$ Beobachtungen von Hühnern. Zum einen hat er als Behandlung *Automatische Fütterung* [ja/nein] bestimmt und zum anderen die Variable *Protein/Fettrate im Zielbereich* [ja/nein] ermittelt. Nun möchte sein Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen.

	24	21	
	13	52	

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

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

90. Aufgabe

(10 Punkte)

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



'Der χ^2 -Test auf einer 2x2-Kreuztabelle berechnet.', liest Yuki in ihrer Mitschrift. So richtig helfen tut ihr das jetzt eherlichweise dann doch nicht. Yuki hatte sich in einen Leistungssteigerungsversuch $n = 158$ Beobachtungen von Hühnern angeschaut. Dabei hat er als Behandlung *Klimakontrolle* [ja/nein] bestimmt und zum anderen die Variable *Protein/Fettrate im Zielbereich* [ja/nein] ermittelt. Am Ende möchte dann ihre Betreuerin gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen.

			70
			88
	100	58	158

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

91. Aufgabe



(10 Punkte)

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



Am Ende war es für Jessica in ihrer Abschlussarbeit dann doch kein normalverteiltes Outcome. Das was jetzt etwas doof, da er sich auf eine ANOVA gefreut hatte. Prinzipiell ginge das auch irgendwie, aber nun möchte ihr Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Jessica hatte sich in ein Stallexperiment $n = 158$ Beobachtungen von Hühnern angeschaut. Dabei hat sie als Behandlung *Ökologisch* [ja/nein] bestimmt und zum anderen die Variable *Fettgehalt erreicht* [ja/nein] ermittelt. Jetzt muss Jessica mal schauen, wie sie das jetzt rechnet. Nach ihrem Experiment erhielt sie folgende 2x2 Kreuztabelle aus ihren erhobenen Daten.

```
##                Ökologisch
## Fettgehalt erreicht ja  nein
##                ja    13    5
##                nein   7    18
```

Dann rechnete Jessica den Fisher-Exakt-Test auf der 2x2-Kreuztabelle in  und erhielt folgende  Ausgabe der Funktion `fisher.test()`.

```
##
## Fisher's Exact Test for Count Data
##
## data: Fettgehalt erreicht
## p-value = 0.005898
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.462677 32.500828
## sample estimates:
## odds ratio
##  6.352594
```

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

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

92. Aufgabe

(11 Punkte)



Die Prävalenz von Klauenseuche bei Wollschweinen wird mit 4% angenommen. In 75% der Fälle ist ein Test positiv, wenn das Wollschwein erkrankt ist. In 8% der Fälle ist ein Test positiv, wenn das Wollschwein *nicht* erkrankt ist und somit gesund ist. Sie werten 4000 Wollschweine mit einem diagnostischen Test auf Klauenseuche aus.

1. Füllen und beschriften Sie den untenstehenden Doppelbaum! Beschriften Sie auch die Äste des Doppelbaumes, mit denen Ihnen bekannten Informationen! **(8 Punkte)**
2. Berechnen Sie die Wahrscheinlichkeit $Pr(K^+|T^+)$! **(2 Punkte)**
3. Was sagt Ihnen die Wahrscheinlichkeit $Pr(K^+|T^+)$ aus? **(1 Punkt)**



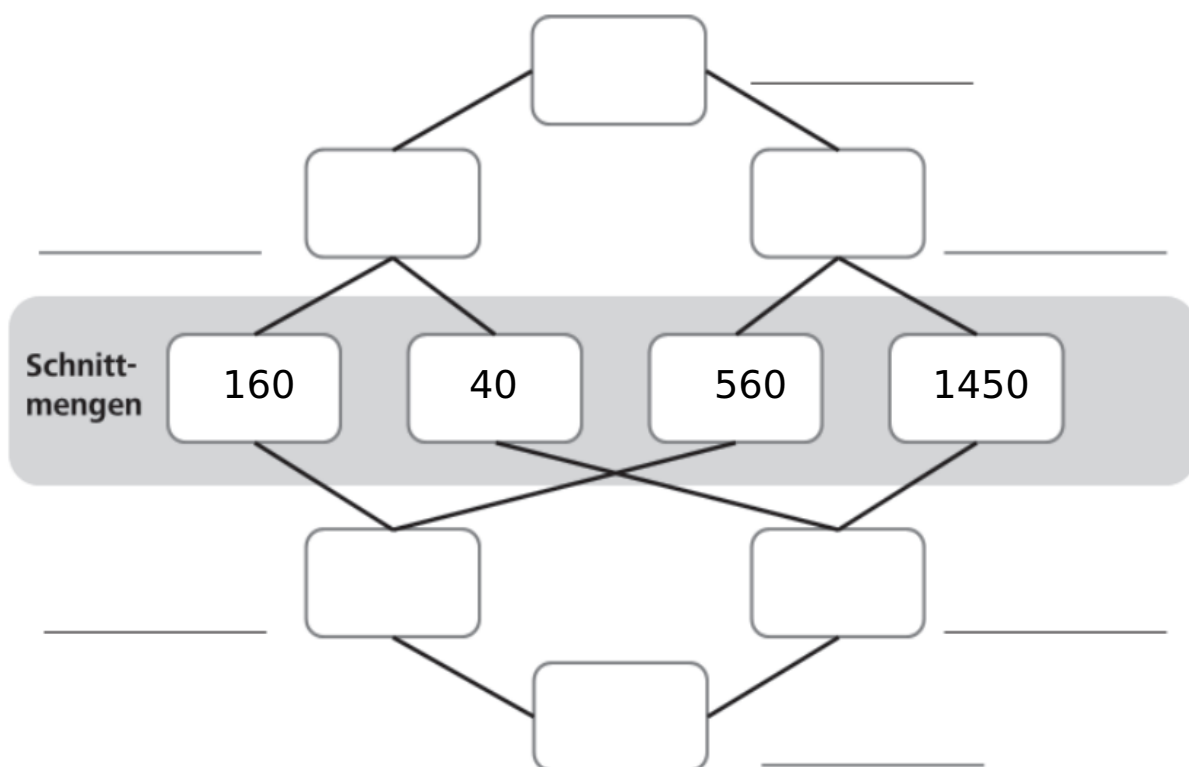
93. Aufgabe

(12 Punkte)



Folgender diagnostischer Doppelbaum nach der Testung auf Klauenseuche bei Fleckvieh ist gegeben.

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



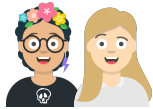
Teil VII.

Lineare Regression & Korrelation

94. Aufgabe

(10 Punkte)

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



'Ich glaube du bringst da was durcheinander. Wir nutzen zwar auch für die ANOVA die Funktion $\text{lm}()$ aber hier wollen wir, glaube ich, eine Gerade durch die Punkte zeichnen.', merkt Jessica an. 'Ich sehe keine Punkte...', antwortet Tina sichtlich übernächtigt. 'Wir müssen die Daten ja auch erst visualisieren!', spricht Jessica sehr deutlich und langsam. Die beiden hatten ein Stallexperiment in der Uckermark mit Hühnern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: mittlerer Anzahl an weißen Blutkörperchen [LEU/ml] und Fettgehalt [%/kg]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt.

Mittlerer Anzahl an weißen Blutkörperchen [LEU/ml]	Fettgehalt [%/kg]
5.4	19.9
18.5	32.7
18.6	29.3
18.5	33.0
5.4	16.0
14.8	25.1
15.6	25.0
20.3	32.4
17.8	30.8

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


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

95. Aufgabe


(12 Punkte)

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



‘Hä? Was ist denn das? Hatten wir das als Aufgabe eine lineare Regression zu rechnen? Wir bauen aus kontinuierlichen Daten eine Abbildung und interpretieren diese dann?’, fragt Jessica. Steffen schaut fragend zurück. ‘Keine Ahnung... das ist jetzt jedenfalls keine Abbildung von irgendwas sondern Zahlen in einer Tabelle...’, antwortet Steffen leicht angespannt. Die beiden hatten einen Leistungssteigerungsversuch im Oldenburger Land mit Hühnern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: mittlere Eisenkonzentration [Fe/ml] und Fettgehalt [%/kg]. Jetzt haben die beiden eigentlich alles zusammen. *Eigentlich...*, denn mit der  Ausgabe haben beide jetzt ein Problem.

term	estimate	std.error	t statistic	p-value
(Intercept)	1.09	1.84		
Mittlere Eisenkonzentration	0.56	0.18		

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



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

96. Aufgabe


(11 Punkte)

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



‘Ich glaube du bringst da was durcheinander. Wir nutzen zwar auch für die ANOVA die Funktion `lm()` aber hier wollen wir, glaube ich, eine Gerade durch die Punkte zeichnen.’, merkt Tina an. ‘Ich sehe keine Punkte... das ist doch eine Ausgabe in . Überhaupt, darum geht es doch gar nicht in unserem Versuch. Wir wollen doch keine Gerade zeichnen?’, antwortet Jonas sichtlich übernächtigt. ‘Doch wir müssen nur die Koeffizienten der linearen Regression erst richtig interpretieren und vor unserem geistigen Auge erscheint eine Gerade!’, spricht Tina sehr deutlich und langsam. Die beiden hatten ein Stallexperiment im Emsland mit Hühnern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: durchschnittliche Tagestemperatur [C/d] und Proteianteil [%/kg]. Jetzt wollen sie erstmal schauen, ob es einen Zusammenhang gibt und das soll mit der  Ausgabe möglich sein.

```
##
## Call:
## lm(formula = "Proteianteil ~ Durchschnittliche_Tagestemperatur", data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6343 -1.0380  0.2777  0.9289  1.8553
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6967     1.4501   0.48   0.634
## Durchschnittliche_Tagestemperatur  1.6185     0.1426  11.35 4.45e-14
##
## Residual standard error: 1.231 on 40 degrees of freedom
## Multiple R-squared:  0.763, Adjusted R-squared:  0.7571
## F-statistic: 128.8 on 1 and 40 DF, p-value: 4.454e-14
```

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



1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Wie groß ist der Effekt im Kontext der wissenschaftlichen Fragestellung? **(2 Punkte)**
3. Interpretieren Sie die *p*-Werte im Kontext der wissenschaftlichen Fragestellung! **(2 Punkte)**
4. Visualisieren Sie die Verteilung der Residuen! **(2 Punkte)**
5. Ist die Annahme der Normalverteilung erfüllt? Begründen Sie die Antwort! **(2 Punkte)**
6. Erklären Sie *kurz* den Begriff R-squared! Was sagt Ihnen der Wert 0.76 aus? **(2 Punkte)**

97. Aufgabe

(9 Punkte)

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



'Hä? Was ist denn das? Das wird ja immer wilder! Hatten wir als Aufgabe eine Korrelation zu berechnen? Wir bauen aus kontinuierlichen Daten eine Abbildung und interpretieren diese dann?', fragt sich Jessica laut. 'Keine Ahnung... das ist jetzt jedenfalls keine Abbildung von irgendwas sondern eine  Ausgabe mit ganz wilden Bezeichnungen...', denkt sie. Jessica hatte ein Kreuzungsexperiment im Oldenburger Land mit Hühnern durchgeführt. Dabei wurden die beiden folgenden Variablen gemessen: mittlere Eisenkonzentration [Fe/ml] und Fettgehalt [%/kg]. Jetzt hat sie eigentlich alles zusammen. *Eigentlich....*, denn mit der  Ausgabe hat Jessica jetzt ein Problem.

```
##  
## Pearson's correlation  
##  
## data: Mittlere Eisenkonzentration [Fe/ml] and Fettgehalt [%/kg]  
## t = 13.75, df = 8, p-value = 7.552e-07  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.9127987 0.9953014  
## sample estimates:  
## cor  
## 0.9794903
```

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

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

98. Aufgabe

(9 Punkte)

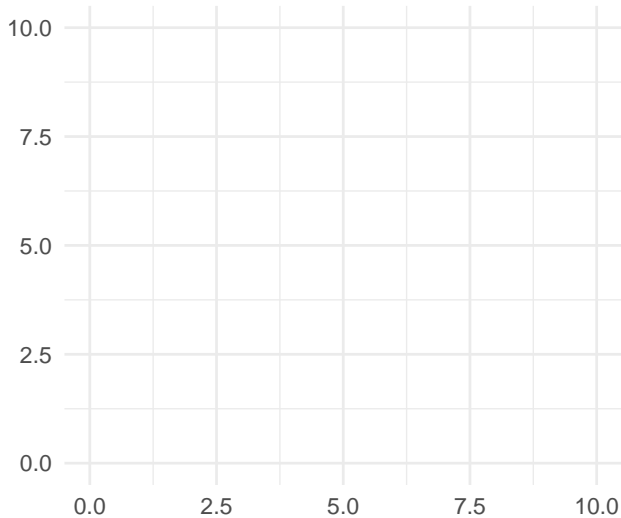


In den folgenden Abbildungen sehen Sie drei leere Scatterplots. Füllen Sie diese Scatterplots nach folgenden Anweisungen.

1. Zeichnen Sie für die angegebene ρ -Werte eine Gerade in die entsprechende Abbildung! **(3 Punkte)**
2. Zeichnen Sie für die angegebenen R^2 -Werte die entsprechende Punktwolke um die Gerade. **(3 Punkte)**
3. Sie rechnen ein statistisches Modell. Was sagen Ihnen die R^2 -Werte über das jeweilige Modell? **(3 Punkte)**

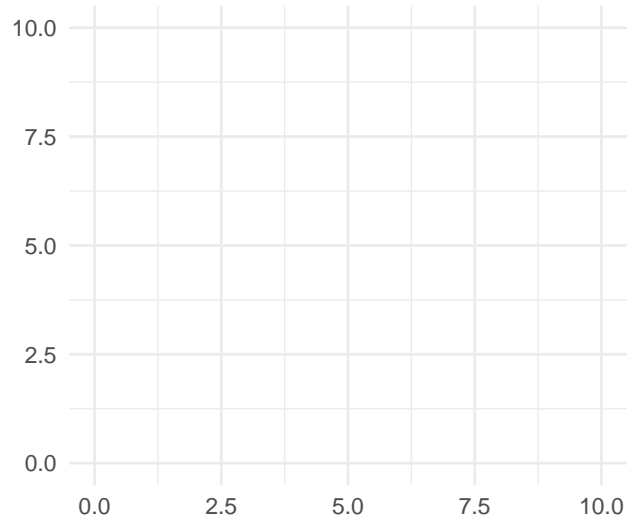
Pearsons $\rho = 0$

$R^2 = 0.75$



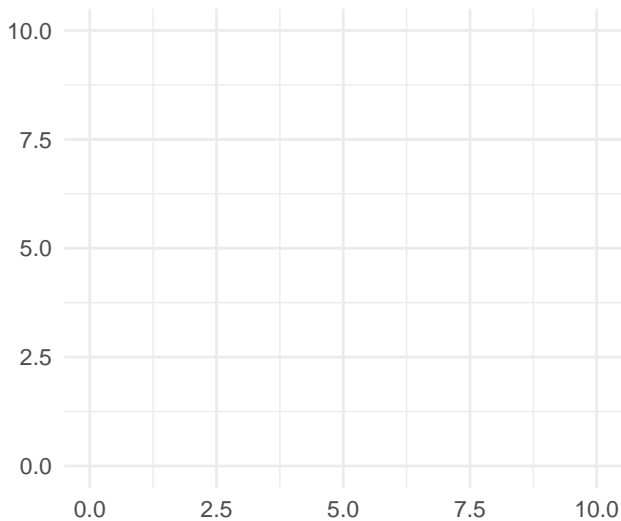
Pearsons $\rho = -0.25$

$R^2 = 0.25$



Pearsons $\rho = -0.75$

$R^2 = 0.5$



99. Aufgabe

(9 Punkte)

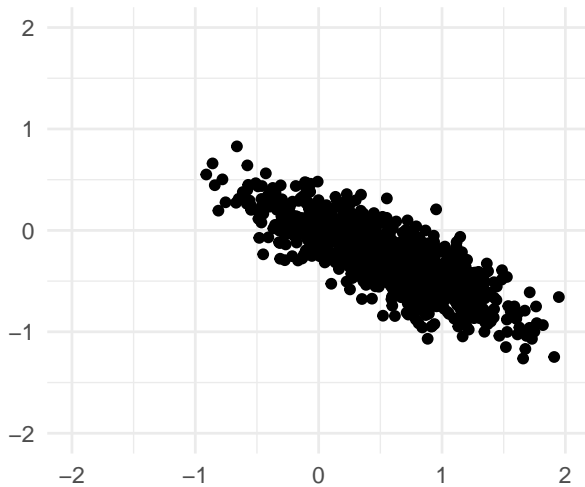


In den folgenden Abbildungen sehen Sie vier Scatterplots. Ergänzen Sie die Überschriften der jeweiligen Scatterplots.

1. Schätzen Sie die ρ -Werte in der entsprechenden Abbildung! **(4 Punkte)**
2. Schätzen Sie die R^2 -Werte in der entsprechenden Punktwolke um die Gerade! **(4 Punkte)**
3. Sie rechnen ein statistisches Modell. Was sagen Ihnen die R^2 -Werte über das jeweilige Modell? **(1 Punkt)**

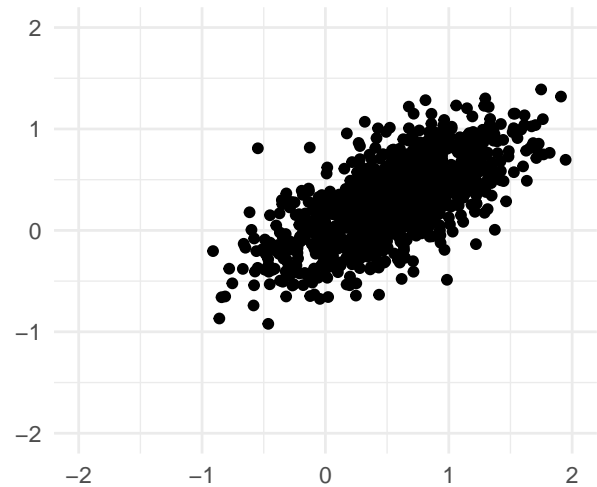
Pearsons $\rho =$

$R^2 =$



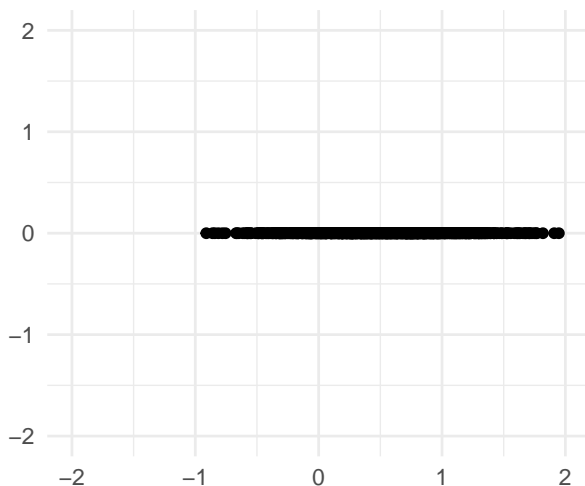
Pearsons $\rho =$

$R^2 =$



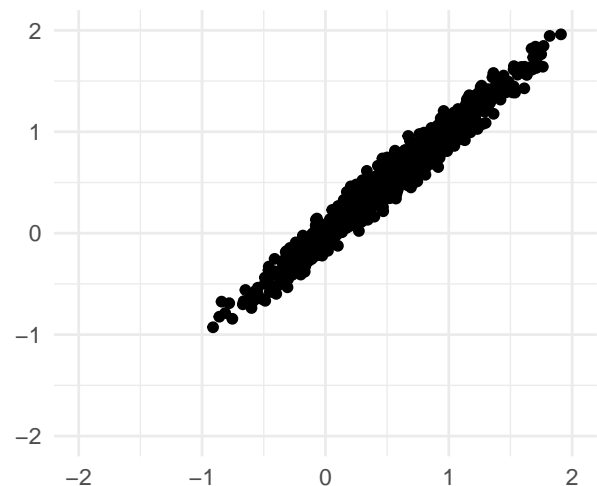
Pearsons $\rho =$

$R^2 =$



Pearsons $\rho =$

$R^2 =$



100. Aufgabe**(10 Punkte)**

Sie rechnen eine lineare Regression um nach einem Feldexperiment den Zusammenhang zwischen Trockengewicht kg/m^2 (*drymatter*) und Wassergabe l/m^2 (*water*) bei Spargel zu bestimmen. Sie erhalten folgende Datentabelle.

.id	drymatter	water	.fitted	.resid
1	26.6	11.6	27.3	
2	22.2	8.7	22.8	
3	15.4	4.5	16.2	
4	19.2	5.7	18.2	
5	31.3	12.7	29.1	
6	22.8	7.7	21.3	
7	22.9	8.5	22.5	
8	25.3	9.0	23.3	
9	28.6	12.7	29.0	
10	7.8	-0.6	8.3	
11	28.8	14.7	32.2	
12	13.6	3.6	14.8	
13	32.5	14.4	31.7	

1. Ergänzen Sie die Werte in der Spalte *.resid* in der obigen Tabelle. Geben Sie den Rechenweg und Formel mit an! **(4 Punkte)**
2. Zeichnen Sie den sich aus der obigen Tabelle ergebenden Residualplot. Beschriften Sie die Abbildung! **(4 Punkte)**
3. Gibt es auffällige Werte anhand des Residualplots? Begründen Sie Ihre Antwort! **(2 Punkte)**

101. Aufgabe

(12 Punkte)





1. Zeichnen Sie in die drei untenstehenden, leeren Abbildungen die Zeile des Regressionskreuzes der Binomialverteilung. Wählen Sie die Beschriftung der y-Achse sowie der x-Achse entsprechend aus! **(6 Punkte)**
2. Ergänzen Sie die jeweiligen statistischen Methoden zu der Abbildung! **(2 Punkte)**
3. Welchen Effektschätzer erhalten Sie aus der entsprechend linearen Regression bzw. den Gruppenvergleich? Geben Sie ein Beispiel! **(2 Punkte)**
4. Wenn Sie keinen Effekt erwarten, welchen *Zahlenraum* nimmt dann der Effektschätzer ein? Geben Sie ein Beispiel! **(2 Punkte)**

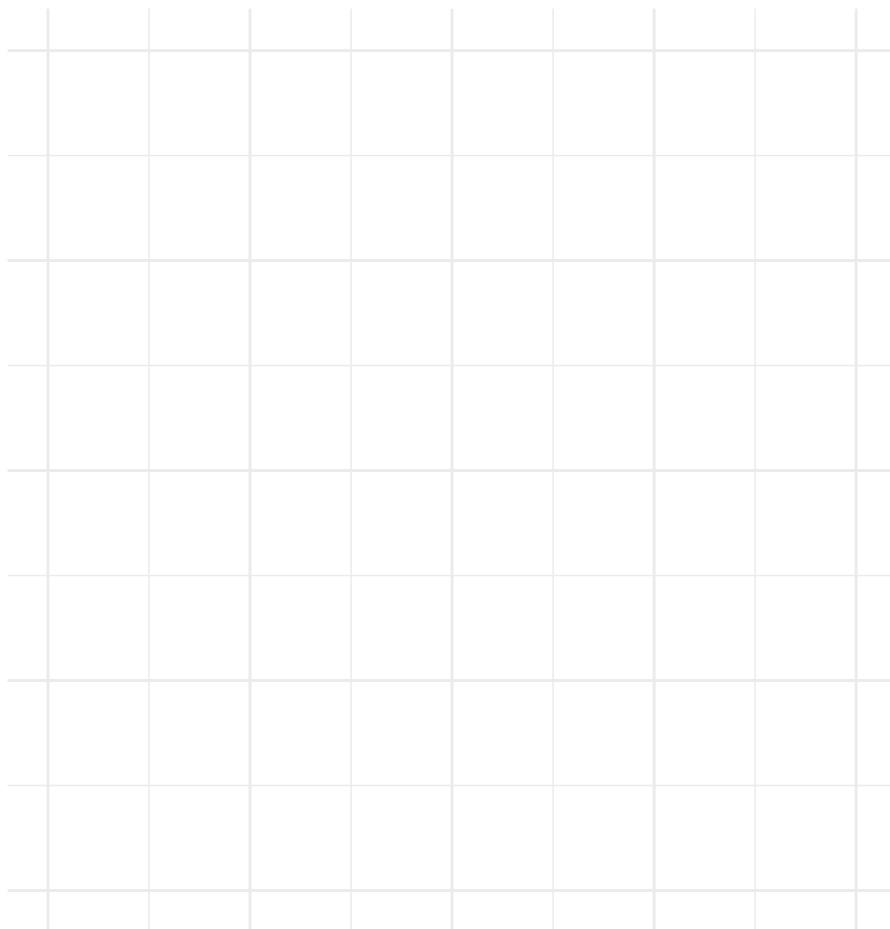
102. Aufgabe

(9 Punkte)



Ein Feldexperiment wurde mit $n = 200$ Pflanzen durchgeführt. Folgende Einflussvariablen (x) wurden erhoben: S, center und dry. Als mögliche Outcomevariablen stehen Ihnen nun folgende gemessene Endpunkte zu Verfügung: drymatter, yield, count, quality_score und dead.

1. Wählen Sie ein Outcome was zu der Verteilungsfamilie *Gaussian* gehört! **(1 Punkt)**
2. Schreiben Sie das Modell in der Form $y \sim x$ wie es in  in der Funktion `glm()` üblich ist *ohne Interaktionsterm*! **(3 Punkte)**
3. Schreiben Sie das Modell in der Form $y \sim x$ wie es in  üblich ist und ergänzen Sie *einen* Interaktionsterm nach Wahl! **(1 Punkt)**
4. Zeichnen Sie eine *starke* Interaktion in die Abbildung unten für den Endpunkt *yield*. Ergänzen Sie eine aussagekräftige Legende. Wie erkennen Sie eine Interaktion? Begründen Sie Ihre Antwort! **(4 Punkte)**




Teil VIII.

Experimentelles Design




103. Aufgabe

(16 Punkte)



Paula und Alex sind bei Steffen um sich Hilfe für eine Versuchsplanung in  zu holen. Dabei geht es um den Zusammenhang zwischen der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) sowie Elterlinie (*ctrl*, und *Xray*) und dem Messwert Protein/Fettrate [%/kg] in Hühnern. Der Versuch soll in einem Stallexperiment im Oldenburger Land durchgeführt werden. Nach dem Dozenten ist der Messwert Protein/Fettrate [%/kg] normalverteilt. Die beiden entschieden sich für ein faktorielles Versuchsdesign. Im ersten Schritt überlegt Steffen ein einfaches experimentelles Design zu probieren. Daher entscheiden sich alle drei für ein *Latin square design* mit nur einem der beiden Faktoren. 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.


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

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




104. Aufgabe

(20 Punkte)



Yuki und Tina sind bei Jonas um sich Hilfe für eine Versuchsplanung in  zu holen. Dabei geht es um den Zusammenhang zwischen der Behandlung Genotypen (*AA*, *AB* und *BB*) sowie Ernährungszusatz (*ctrl* und *getIt*) sowie zwei Blöcken und dem Messwert Protein/Fettrate [%/kg] in Hühnern. Der Versuch soll in einem Leistungssteigerungsversuch im Wendland durchgeführt werden. Nach der Dozentin ist der Messwert Protein/Fettrate [%/kg] normalverteilt. Die beiden entschieden sich für ein faktorielles Versuchsdesign. Im ersten Schritt überlegt Jonas ein komplexeres experimentelles Design zu probieren. Daher entscheiden sich alle drei für ein *Randomized complete block design* mit Berücksichtigung einer Interaktion. Das sollte für den anfang erstmal reichen. 'Und jetzt, was machen wir jetzt?', Tina schaut die anderen beiden mit großen Augen an. Die zucken mit der Schulter.

Leider kennen sich Jonas, Yuki und Tina mit dem *Randomized complete block design* überhaupt nicht aus. Deshalb brauchen die Drei bei der Erstellung Ihre Hilfe!

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




Teil IX.


Programmieren in R

105. Aufgabe

(9 Punkte)




Paula muss ihrem Projektbericht mit  arbeiten. Deshalb sitzt sie jetzt mit Ihnen zusammen und hat einige Fragen zu den Grundlagen in  an Sie! Na dann wollen Sie mal helfen. Immerhin will ihr Betreuer, dass  genutzt wird.

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


Sie antworten:

Paula: Was ist der Unterschied zwischen einem Objekt, einem Wort und einer Funktion? (1 Punkt)

Sie antworten:

Paula: Ich habe die Namen der beiden  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 ? (1 Punkt)

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

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

Sie antworten:

106. Aufgabe

(9 Punkte)



‘Unter den Blinden ist der Einäuge König!’, ruft Ihnen Tina entgegen. Das können Sie schon nicht mehr hören. Nur weil Sie einmal gesagt haben, dass Sie sich schonmal mit **R** beschäftigt haben, stehen hier alle Schlange. Aber gut, das hat Sie dann doch vorangebracht. Leider kennt sich Tina auch wieder überhaupt nicht mit **R** aus aber ihr Betreuer möchte gerne, dass die Auswertung in **R** gemacht wird. Da müssen Sie dann wohl mal nochmal ran und helfen.

Tina fragt: *Warum wurde jetzt nach dem Laden der Daten die Funktion `mutate()` genutzt? (1 Punkt)*

Sie antworten:

Tina fragt: *Nach der EDA zu urteilen liegt eine Interaktion vor, wie spezifiziere ich diese im Modell, so dass ich die Interaktion zwischen zwei Faktoren f_1 und f_2 testen kann? (1 Punkt)*

Sie antworten:

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

Sie antworten:

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

Sie antworten:

Tina fragt: *Wenn ich Daten in R mit Gruppen eingelesen habe, welche Funktion nutze ich dann meistens als erstes und warum muss ich das machen? Was muss ich da machen? (2 Punkte)*

Sie antworten:

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

Sie antworten:

Tina fragt: *Was muss ich bei der Benennung von Spalten in Excel beachten? (1 Punkt)*

Sie antworten:

Teil X.

Forschendes Lernen

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

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

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

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

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

- bar

107. Aufgabe



(20 Punkte)

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



Steffen hält die wissenschaftliche Veröffentlichung *Selle, P. H., et al. (2010). Implications of sorghum in broiler chicken nutrition* unter einem Schnaufen in die Luft. 'Worum geht es denn eigentlich in dieser Arbeit?', fragt er stirnrunzelnd. Steffen soll die Veröffentlichung nutzen um das eigene Experiment zu planen. Als eine Vorlage sozusagen. Daher möchte seine Betreuerin, dass er einmal die Veröffentlichung sinnvoll zusammenfasst. Das sollte dann doch etwas aufwendiger werden.

Leider kennt sich Steffen 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)¹ **(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)**

¹Schimel, J. (2012). Writing science: how to write papers that get cited and proposals that get funded. OUP USA.

108. Aufgabe


(20 Punkte)

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



'Uff', denkt Alex Das ist jetzt doch etwas umfangreicher. Alex soll die wissenschaftlichen Daten *data4* einmal sinnvoll analysieren und zusammenfassen. Die Daten sollen als eine Vorlage für seine eigene Arbeit und Experiment dienen. Daher möchte sein Betreuer, dass er einmal die Daten in einer PowerPoint Präsentation zusammenfasst. 'Das ist jetzt aber doch umfangreicher als gedacht.', mault Alex in sich hinein. Dann starrt er eine Weile in seinen Laptop.

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

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

Teil XI.

Mathematik

109. Aufgabe

(10 Punkte)

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



Herodot – der Schimmel aus Ivenack Während der Besetzung Mecklenburgs durch die Franzosen kamen Napoleon die Geschichten des berühmten Apfelschimmels Herodot aus Ivenack zu Gehör. Herodot lief zwar niemals Rennen, war aber eines der berühmtesten Pferde dieser Zeit. Napoleon selbst gab den Auftrag, diesen Schimmel durch die Armee nach Frankreich zu bringen. Der Legende nach sollen Arbeiter den Schimmel im hohlen Stamm einer 1000-jährigen Eiche aus Ivenack vor den Franzosen versteckt haben. Doch Herodot verrät sein Versteck durch lautes Wiehern, woraufhin die französische Armee den Schimmel beschlagnahmte und nach Frankreich führte².

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

Gehen Sie von einem radialen Wachstum der 1000-jährigen Eiche von 1.2mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 12.5m in Bruthöhe hatte.

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

Herodot hatte eine Schulterhöhe von 180cm , eine Breite von 90cm sowie eine Länge von 230cm .

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

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

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

²Die 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](#)

110. Aufgabe

(10 Punkte)

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



Von Töpfen auf Tischen In einem Experiment wollen Sie die Wuchshöhe von 180 Maispflanzen bestimmen. Bevor Sie überhaupt mit dem Experiment beginnen können, gibt es aber ein paar Abschätzungen über die Kosten und den Aufwand zu treffen. Zum einen müssen Sie die Maispflanzen einpflanzen und müssen dafür Substrat bestellen. Zum anderen müssen Sie die Maispflanzen auch bewegen und in ein Gewächshaus platzieren. Die Töpfe für die Keimung haben einen Durchmesser von 10cm und eine Höhe von 9cm. Der Kubikmeterpreis für Torf liegt bei 280 EUR.

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

111. Aufgabe

(10 Punkte)

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



Solar- & Biogasanlagen Um die Energiekosten Ihres Betriebes zu senken, wollen Sie eine Solaranlage auf den Schweinestall montieren lassen. Sie messen Ihren Stall und finden folgende Maße wieder. Die vordere Seite des Schweinestall hat eine Höhe h_v von $7m$. Die hintere Seite des Schweinestall hat eine Höhe h_b von $10m$. Der Schweinestall hat eine Tiefe t von $14m$ und eine Breite b von $40m$.

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

Ebenfalls planen Sie eine neue Biogasanlage für Ihren Betrieb. Der neue Methantank hat einen Radius r von $1.5m$. Leider gibt es ein paar bauliche Beschränkungen auf dem Grundstück. Ihr Fundament des zylindrischen Methantanks kann nur ein Gewicht von maximal $10t$ aushalten bevor der Tank wegbricht. Sie rechnen eine Sicherheitstoleranz von 25% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $220kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $270kg/m^3$. Sie betreiben Ihre Anlage bei $-92^\circ C$.

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

112. Aufgabe

(10 Punkte)

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



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

Aligatorenbirnen und Blaubeeren “Sind Sie ein Riesenfaultier oder warum kaufen Sie so viele Aligatorenbirnen?”, spricht es hinter Ihnen. Irritiert drehen Sie sich um und blicken in das puderrote Gesicht von einem Studienrat mit Stock. “Wieso?”, entfährt es Ihnen und Sie bereuen sogleich die Frage. Sofort werden Sie zu einem Whiteboard voller roter Schnüre geschliffen und müssen folgenden mathematischen untermauerten Argumenten im Aldi über sich ergehen lassen. Da kommen Sie nicht mehr raus, also können Sie auch gleich mitmachen. Das Problem liegt in Chile³. Tja, die Deutschen und Südamerika.

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

1. Wenn 5 Blaubeerschalen 9.45 Euro kosten, wie viel kosten 6 Schalen? **(2 Punkte)**
2. Wenn Sie die 6 Blaubeerschalen gekauft haben, wie viele Aligatorbirnen zu je 2.89 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 Aldi über Gemüse erfährt.

- Ein Kilo Strauchtomaten benötigt 160l Wasser. Eine Strauchtomate wiegt 90 - 115g.
- Ein Kilo Salat benötigt 130l Wasser. Ein Salatkopf wiegt 310 - 490g.
- Ein Kilo Avocado benötigt 1100l Wasser. Eine Avocado wiegt 150 - 410g.
- Ein Kilo Blaubeeren benötigt 820l Wasser. Eine Blaubeere wiegt 3 - 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 2023 blieben die Erträge von Blaubeeren mit 7.7×10^4 t in dem prognostizierten Rahmen. Die Menge steigerte sich um 6.8%. Die Exporte für Avocados stiegen in dem gleichen Zeitraum um 18.1% auf 2.3×10^5 t.

4. Wie viele Kubikmeter Wasser hat Chile in dem Exportjahr 2022 exportiert? **(2 Punkte)**

Chile ist eines der wenigen Länder der Welt, die ihr Wasser komplett privatisiert haben. Derzeit sind nur drei Prozent des Wassers des Landes für den häuslichen Verbrauch vorgesehen. In den Dörfern der Anbauregionen versorgen Tankwagen die Bevölkerung jede Woche mit Wasser, es gibt etwa 55 Liter Wasser pro Kopf für den täglichen Bedarf. In *Deutschland* liegt der Verbrauch bei 3 - 12 Liter pro Minute Händewaschen und 10 - 15 Liter pro Minute Duschen.

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 einem Studienrat mit Stock erwartet. Ganz schön viele Informationen wurden da zusammengetragen.

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

³Die 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.

113. Aufgabe

(12 Punkte)

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



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

Die Dampfnudelerde “Was für einen Unsinn!“, rufen Sie. Jetzt haben Sie kostbaren Schlaf prokrastiniert um einem Ernährungswissenschaftler auf YouTube über die Erde als Dampfnudel zu lauschen. Irgendwie passt es dann doch mit der Analogie. Übermüdet müssen Sie darüber nachdenken, warum vor 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 rechnen⁴.

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 9.65 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.289 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.21 g/cm^3 . Das Gewicht von einem heute lebenden afrikanischen Elefanten liegt bei 5t bis 7t und das Gewicht von einem Triceratops bei 6t bis 12t.

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 1.03 astronomische Einheiten (AE). Die Einheit 1 AE wird mit $1.5 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 84% Wasserstoffkernen mit einer molaren Masse von 1.05 g/mol , 9% Heliumkernen mit 3.92 g/mol sowie 7% weiteren Atomkernen mit 152.01 g/mol . Die Teilchendichte bei Eintritt in die Erdatmosphäre liegt zwischen 0.4 bis 100 Teilchen cm^{-3} pro Sekunde mit einer mittleren Teilchendichte von 5 cm^{-3} pro Sekunde.

Lösen Sie den folgenden Aufgabenteil mit einer aussagekräftigen Skizze!

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

⁴Die Quelle der Inspiration für die Aufgabe war folgender Artikel: "Skeptische Anmerkungen — Die Erde als Dampfnudel" in Der Humanistische Pressedienst

114. Aufgabe

(10 Punkte)

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



Entschuldigung, ist das Ihre Feder in meinem Auge? So hört man häufiger höfliche Gänse in Mastställen sagen. Das ist natürlich etwas ungünstig, den dann kommt es zu Picken und Kannibalismus. Denn wenn der Nachbar nervt, dann muss zu Maßnahmen gegriffen werden. Kennt jeder aus einer mittelmäßigen Wohngemeinschaft. Das wollen wir aber als vorsorgliche Gänse-Halter:innen nicht⁵. Betrachten wir also einmal das Platzangebot (eng. *space allowance*, abk. *SA*) der Gänse für vier Tätigkeiten und versuchen die notwendige Fläche zu optimieren. Wie immer gibt es dafür eine mathematische Formel:

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

mit

- SA dem benötigten Platzangebot aller aufsummierten Verhalten i .
- A_i dem benötigten Platz für ein Verhalten i .
- PB_i dem Anteil des Auftretens eines Verhaltens i .
- r_i dem Radius Gans plus dem benötigten Radius für das Verhalten i .
- R_i dem notwendigen Abstand zu den Nachbarn für das Verhalten i .
- i dem Verhalten: (1) standing, (2) foraging incl. scratching, (3) drinking/eating und (4) preening.

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

	Aldridge et al. (2021)	Baxter et al. (2022)	Jacobs et al. (2019)
standing	33cm; 26cm; 8.9%	36cm; 28cm; 8.9%	33cm; 27cm; 3.5%
foraging incl. scratching	30cm; 19cm; 0.6%	40cm; 18cm; 0.8%	31cm; 18cm; 0.2%
drinking/eating	26cm; 28cm; 8.1%	47cm; 22cm; 8.1%	35cm; 17cm; 3.6%
preening	33cm; 20cm; 16.4%	30cm; 26cm; 14.2%	37cm; 23cm; 16.4%

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

⁵Die 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.](#)

115. Aufgabe

(8 Punkte)

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



Nelken von den Molukken In der Ausstellung „Europa und das Meer“ im Deutschen Historischen Museum in Berlin gab es folgendes Zitat über die Probleme der frühen Hochseeschifffahrt.

»Ohne ausreichende Zufuhr von Vitamin C stellen sich nach 42 Tagen die ersten Symptome ein; die ersten Toten sind nach 65 Tagen zu beklagen; nach 115 Tagen rafft die Skorbut eine ganze Schiffsbesatzung dahin«

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

1. Stellen Sie den Verlauf der Anzahl an Matrosen auf einem Schiff der Flotte in der Form einer Überlebenszeitkurve dar! Beschriften Sie die Achsen entsprechend! **(2 Punkte)**
2. Was ist die Besonderheit der Überlebenszeitkurve? Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Schätzen Sie die Überlebenswahrscheinlichkeit nach 95 Tagen aus Ihrer Abbildung ab! **(1 Punkt)**

Der Chronist an Bord der Trinidad, Antonio Pigafetta, schrieb in seinem Bericht „[...] Um nicht Hungers zu sterben, aßen wir das Leder, mit dem die große Rahe zum Schutz der Taue umwunden war.“ Insbesondere die Mannschaft der Concepción erlitt große Verluste durch die Skorbut bei der Überquerung des Pazifiks, da durch Erkundungsfahrten weniger Zeit blieb, um wilden Sellerie aufzunehmen. Wilder Sellerie enthält $6000\mu\text{g}/10\text{mg}$ Vitamin C. Der Bedarf liegt bei 110mg pro Tag für Männer.

3. Berechnen Sie die notwendige Menge in t an aufzunehmenden wilden Sellerie auf die Concepción für die ununterbrochene Fahrt von drei Monate und 20 Tage über den Pazifik! **(3 Punkte)**
4. Skizzieren Sie die Überlebenszeitkurve für die Concepción im Vergleich zu der Überlebenszeitkurve der Trinidad! Beschriften Sie die Achsen! **(2 Punkte)**

116. Aufgabe

(10 Punkte)

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



Event Horizon – Am Rande des Universums Die Sonne hat eine aktuelle, angenommene Masse von $2 \times 10^{27} \text{ kg}$. Wenn die Sonne nun am Ende ihrer Lebenszeit zu einem schwarzen Loch mit dem Radius von 2500m kollabiert, wird die Sonne 35% der aktuellen Masse verloren haben. Ein Lichtteilchen mit der Masse m_f und der Fluchtgeschwindigkeit v_f will dem schwarzen Loch entkommen. Sie haben folgende Formeln für die kinetische Energie des Lichtteilchens E_{kin} und der Gravitationsenergie des schwarzen Lochs E_{grav} gegeben⁶.

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

mit

- m_f , gleich der Masse [kg] des fliehenden Objektes
- m_s , gleich der Masse [kg] des stationären Objekts
- r_s , gleich dem Radius [m] des stationären Objekts
- G , gleich der Gravitationskonstante mit $6.674 \cdot 10^{-11} \text{ m}^3 (\text{kg} \cdot \text{s}^2)^{-1}$

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

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

⁶Die Quelle der Inspiration für die Aufgabe war ein Montagnachtfilm: [Event Horizon – Am Rande des Universums](#)

117. Aufgabe

(10 Punkte)

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



Das Fermi Paradoxon Der Kernphysiker Enrico Fermi diskutierte 1950 auf dem Weg zum Mittagessen im Los Alamos National Laboratory mit seinen Kollegen angebliche UFO-Sichtungen und fragte schließlich: „Where is everybody?“. Warum seien weder Raumschiffe anderer Weltraumbewohner noch andere Spuren extraterrestrischer Technik zu beobachten? Wie lange würde eine außerirdische Zivilisation benötigen um die gesamte Milchstraße zu besuchen, wenn das maximale Reisetempo die Geschwindigkeit der Voyager 1 Sonde wäre?⁷

Wir 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 250 Jahren ist die Replikation abgeschlossen und wiederum *drei* Sonden werden ausgesendet. Gehen Sie von 5.16 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 10^{11} Sterne in der Milchstraße. Nehmen Sie eine Lichtgeschwindigkeit von $2.9 \times 10^8 \text{ m/s}$ an.

1. Skizzieren Sie in einer Abbildung die ersten vier Schritte der Vervielfältigung der Sonden in der Galaxie! Beschriften Sie die Abbildung mit der Dauer und der Anzahl an Sonden für jeden Schritt der Vervielfältigung! **(4 Punkte)**
2. Berechnen Sie die theoretische Anzahl an Vervielfältigungsschritten die benötigt werden um mit *einem einzigen Vervielfältigungsschritt* die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
3. Berechnen Sie die Dauer, die eine außerirdische Zivilisation annäherungsweise benötigt um die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
4. Bei einem vermuteten Alter der Erde von 4.3×10^9 Jahren, wie oft war dann eine Sonde einer außerirdischen Zivilisation schon zu Besuch? Korrigieren Sie Ihre Antwort mit dem Wissen, dass sich die Kontinentalplatten einmal alle 1.2×10^8 Jahre vollständig im Erdinneren umgewandelt haben! **(2 Punkte)**

⁷Die Quelle der Inspiration für die Aufgabe war folgender Wikipediaeintrag: [Fermi-Paradoxon](#)

118. Aufgabe

(10 Punkte)

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



Pyramiden bauen Es stehen die niedersächsischen Pyramidentage an und Sie sind auf abenteuerlichen Wegen für den Bau der Pyramiden zuständig. Zu allem Überfluss handelt es sich auch noch eine *Reenactment* Veranstaltung. Thema der diesjährigen Pyramidentage sind die Pyramiden von Meroe, die den Königen und Königinnen des historischen Reiches von Kusch in Nubien, dem heutigen Sudan, als Grabstätten dienten. Die Pyramiden in Meroe fallen durch ihren steilen Winkel von 71 Grad im Vergleich zu den ägyptischen Pyramiden mit 55 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 38 Königsellen. Eine Königselle misst 52.2cm.

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

1. Bei der Königspyramide von Meroe soll eine Seitenlänge der Grundfläche 38 Königsellen lang sein. Welche Höhe der Königspyramide in m ergibt sich? **(1 Punkt)**
2. Die Außenflächen der Pyramide soll begrünt werden. Für die Bepflanzung muss eine 4cm dicke Torfschicht auf die Pyramide aufgebracht werden. Berechnen Sie die ungefähre Menge an benötigten Torf in m^3 ! **(2 Punkte)**

Wie in jedem guten *Reenactment* gibt es viel Oberschicht, aber nur 2 Sklaven, die Ihnen bei dem Befüllen der Pyramide mit Schutt zu Seite stehen. Leider haben Ihre Sklaven zu allem Überfluss auch noch chronische Knieschmerzen entwickelt, als sie von der anstehenden Aufgabe erfahren haben. Gehen Sie daher von einer Effizienz der Sklaven von 85% aus. In eine Schubkarre passen 90 Liter.

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

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

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

119. Aufgabe

(12 Punkte)

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



Geocaching – Von Satelliten und Plastikdosen Es ist Wochenende und das Wetter ist *sweet*. Sie schwingen sich auf Ihr Cachermobil um mit 16km/h, geleitet von modernster Satellitentechnologie und einem Supercomputer aus dem Jahr 2000 in Ihren Händen, Plastikdosen in der Natur und an sehenswerten Orten zu finden. Sie wollen diesmal endlich die aufwärts Terrainchallenge durchführen. Die Reihenfolge der Caches nach Terrainwertung gibt daher die von Ihnen abzufahrenden Orte vor. Die Terrain- und Schwierigkeitswertungen laufen von 1 (leichteste Wertung) bis 5 (schwierigste Wertung) in 0.5 Schritten. Folgende Informationen zu den Orten und den entsprechenden Caches stehen Ihnen für Ihre Planung der Route zu Verfügung⁸.

Ort	Cache	Wertung (S T G)
A	GC3IPVO	3.5 4.5 Normal
B	GCAZ0TO	1.5 2.5 Klein
C	GCZJHAK	2.0 2.0 Normal
D	GC1JQL3	2.5 5.0 Klein
E	GCSPMGN	3.0 4.0 Mikro

Im Weiteren sind Ihnen folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \overrightarrow{AC} ist 6km. Im Weiteren ist Ihnen der Entfernungsvektor \overrightarrow{CB} mit 7.5km bekannt. Der Entfernungsvektor \overrightarrow{BE} ist das 2.1-fache des Entfernungsvektor \overrightarrow{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 20° südlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 55° östlich. Vom Ort B betrachtet, bilden die Orte C und D einen rechten Winkel am Ort B. Der Ort B liegt auf gerader Linie zwischen den Orten C und E. Somit liegt der Ort E südlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort D Ihre Cachertour.

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

$$\text{Suchzeit} = 0.15 + 0.18 \cdot \text{Schwierigkeit}$$

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

4. An der höchsten Schwierigkeit müssen Sie angeln. Ihre Angel ist ausgefahren 6m lang. Erreichen Sie einen Cache in der Höhe von 7.9m? Berechnen Sie dazu Ihre maximale mögliche Angelhöhe! Welche Annahmen mussten Sie treffen um die Aufgabe zu lösen? **(2 Punkte)**

⁸Die Quelle der Inspiration für die Aufgabe war folgende Tätigkeit: [Geocaching – Mach mit bei der weltweit größten Schatzsuche](#).

120. Aufgabe

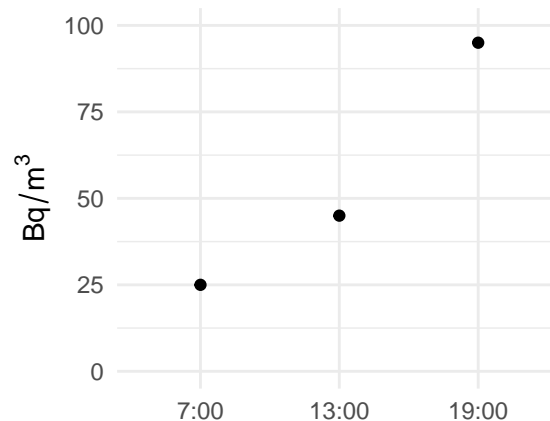
(10 Punkte)

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



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

Die atmende Wand und Brot aus Luft Als Kellerkind vom Dorf wollen das Ausmaß der Radonbelastung in ihrem Kellerzimmer bestimmen und lüften daher nicht. Spart dann auch Energie und lüften wird sowieso überschätzt. Während einer Messperiode von 7:00 Uhr bis 19:00 bestimmen Sie dreimal automatisch die Radonbelastung in Ihrem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung⁹.



1. Wie lange dauert es in Stunden bis Sie eine kritische Belastung von $280 Bq/m^3$ in Ihrem ungelüfteten Kellerraum erreicht haben? **(2 Punkte)**

Radon zerfällt mit einer Halbwertszeit von 4.1d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 140d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

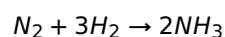
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $280 Bq/m^3$ auf unter $90 Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	78.1	28.4	
Sauerstoff	19.5	16.2	
Kohlenstoffdioxid	0.029	12.5	

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

Während Sie Ihr etwas pappiges Toastbrot mampfen kommt Ihnen die Dokumentation über Brot aus Luft in den Sinn. Sie denken darüber ein wenig nach. Für die Umwandlung von Stickstoff N_2 mit Wasserstoff H_2 zu Ammoniak NH_3 gilt folgende Reaktionsgleichung¹⁰:



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

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

⁹Die Quelle der Inspiration für die Aufgabe war folgender Artikel: [Atmende Wand](#)

¹⁰Die Quelle der Inspiration für die Aufgabe war folgender Artikel: [Haber-Bosch-Verfahren – Brot aus Luft](#)

121. Aufgabe

(10 Punkte)

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



Armee der Finsternis Ihr Studentenjob war nach Ladenschluss bei Kaufland die Regale einzuräumen. Dabei ist Ihnen in der Auslage der Sonderangebote das Necronomicon¹¹ in die Hände gefallen. Nun sind Sie ein Magier der Zeichen geworden! Also eigentlich können Sie nur Mathe und das dämliche Necronomicon hat Sie in die Vergangenheit geschleudert... aber gut, was tut man nicht alles im Jahr 91 n. Chr. für den neuen Lehnsherren Fürsten Arthur. Sie bauen natürlich einen Schrottkugelturm um sich den Horden der Finsternis mit genug Schrott erwehren zu können! Ihnen stehen zwei mächtige magische Formeln zur Unterstützung zu Verfügung.

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

mit

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

Als erstes müssen Sie die Höhe des zu bauenden Schrottkugelturmes bestimmen. Hierfür ist wichtig zu wissen, dass sich die Bleitropfen mit einem Gewicht von $40mg$ zu gleichförmigen Bleikugeln bei einer Geschwindigkeit von $14m/s$ bilden.

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

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

2. Nennen Sie die beiden geometrischen Formen aus denen sich näherungsweise ein Tropfen zusammensetzt! Erstellen Sie eine beschriftete Skizze des Tropfens! **(2 Punkte)**
3. Sie messen eine Länge des Tropfens von $4.1mm$. Die Löcher im Sieb erlauben ein Tropfendurchmesser von $1.7mm$. Welchen Durchmesser in mm haben Ihre produzierten Bleikugeln? **(3 Punkte)**

Sie haben jetzt die 2.3×10^5 Bleikugeln zusammen. Blei hat eine Dichte von $10.32g/cm^3$.

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

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

5. Wie groß in cm^2 ist Ihr quadratisches Sieb am oberen Ende des Turms, wenn Sie pro Fall ca. 700 Bleikugeln produzieren wollen und die Bleikugel im Fall $0.5cm$ Abstand haben müssen? **(1 Punkt)**

¹¹Ein wirklich gefährliches Buch ist: *Du bist genug: Vom Mut, glücklich zu sein* von Fumitake Koga und Ichiro Kishimi

122. Aufgabe

(10 Punkte)

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



Armee der Kaninchen Leider hat es mit Ihrer Surfschule in Down Under nicht geklappt. War vielleicht auch nicht so die beste Idee... aber dafür haben Sie eine Neue! Oder wie es Mike Tyson zugeschrieben wird: „Ich wurde nie niedergeschlagen, ich war immer am Aufstehen!“. Daher machen Sie jetzt einen Großhandel mit Kaninchenfleisch und damit dem teuersten Fleisch in Australien auf. Moment, hopsen hier nicht, seit Thomas Austin im Jahr 1859 ungefähr 30 Kaninchen entlassen hat, Millionen von Kaninchen rum? Wieso ist das Kaninchenfleisch dann so exklusiv? Dem wollen wir mal mathematisch nachgehen!¹²

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

$$f(t) = 9 \times 10^9 - 1.1 \times 10^9 \cdot 2^{-0.2 \cdot t + 3.1}$$

1. Skizzieren Sie die Sättigungsfunktion *annäherungsweise* in einer Abbildung! **(1 Punkt)**
2. Wie viele Kaninchen können nach der Sättigungsfunktion maximal im westlichen Australien leben? Ergänzen Sie den Wert in Ihrer Abbildung! **(2 Punkte)**
3. Wie viele Millionen Kaninchen leben nach der Sättigungsfunktion nach 18 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 12 Kaninchen ausgebracht. Da die Kaninchen keine Maßnahmen gegen die Ausbreitung vornehmen können, verläuft die Ausbreitung mit einem wöchentlichen Wachstumsfaktor von 1.6 nach folgender Formel.

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

3. Wie viele Wochen benötigen die Viren um theoretisch die gesamte Kaninchenpopulation nach 16 Jahren Wachstum zu durchseuchen? **(1 Punkt)**

Das Myxoma Virus und das RHDV töteten 98.5% der Kaninchenpopulation innerhalb weniger Wochen.

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

Thomas Austin entließ die Kaninchen im äußersten Norden von Australien. Australien hat eine West-Ost-Ausdehnung von 4400km und eine Nord-Süd-Ausdehnung von knapp 3500km. Die Kaninchen breiten sich radial mit einer Geschwindigkeit von 8.8km pro Jahr aus.

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

Eine jährliche Impfung gegen das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) kosten 11\$ pro Tier und der durchführende Arzt verlangt ca. 45\$ pro Tier.

6. In Ihrem Stall leben 900 Mastkaninchen. Mit welchen jährlichen Zusatzkosten für die Impfungen der Kaninchen müssen Sie daher kalkulieren? **(1 Punkt)**

¹²Die 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...](#)

123. Aufgabe

(12 Punkte)

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



Uckermark. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer der Kuh Fridolin und Ihnen. Grünes Gras unter Ihren Füßen und ein strammer Wind im Gesicht, egal wohin Sie schauen. Ein schmatzendes Geräusch ertönt unter Ihnen. Sie sinnieren, sollten Sie Ihre weiten Graslandschaften jetzt schon düngen? Dafür benötigen Sie die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Sie sehen nicht ein, Geld für einen Agrarmetrologen zu bezahlen. Also rechnen Sie mit folgenden Informationen zu Monatsmultiplikatoren des GLT-Wertes: Januar mit $0.4\times$, Februar mit $0.8\times$ und März mit $1.05\times$. Sie haben noch im letzten Jahr folgende Temperaturen gemessen.

Datum	C°
01. Jan 2023	0.1
01. Feb 2023	1.5
01. Mrz 2023	3.5
01. Apr 2023	4.3

1. Erstellen Sie eine Skizze aus den Informationen aus der Temperaturtabelle! **(1 Punkt)**
2. Stellen Sie die linearen Funktionen $f_1(t)$, $f_2(t)$ und $f_3(t)$ aus der obigen Temperaturtabelle auf! **(1 Punkt)**
3. Bestimmen Sie die Stammfunktionen $F_1(t)$, $F_2(t)$ und $F_3(t)$ für Ihre linearen Funktionen aus der obigen Temperaturtabelle! **(1 Punkt)**
4. Osterglocken beginnen ab einer GLT von 200°C zu blühen. An welchem Tag im 1. Quartal des Jahres 2023 war dies der Fall? *Ignorieren Sie ein eventuelles Schaltjahr in Ihrer Berechnung.* **(4 Punkte)**

Auf dem Weg zu Ihrer Boskoopplantage wurden Sie mit Ihrem Trecker von einer Gruppe elektrifizierter Rentner abgedrängt. Der Trecker muss wieder aus dem Graben! Fridolin und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Fridolin mit 120N . Die elektrifizierten Rentner bringen eine Kraft von 140N auf.

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

5. Im ersten Versuch legen Sie das Seil für Fridolin lotrecht über einen Ast oberhalb des Treckers. Die Rentner ziehen in einer geraden Linie über die Böschung hinweg am anderen Seil. Welche Kraft wird aufgebracht? **(2 Punkte)**
6. Im zweiten Versuch ziehen Fridolin und die Rentner mit einem 50° Winkel mit ihrem Seil an dem Trecker. Welche Kraft wird aufgebracht? **(2 Punkte)**
7. Mit welcher Beschleunigung ziehen Sie den 1.5t schweren Trecker *jeweils* aus dem Graben, wenn $F = m \cdot a$ gilt? **(1 Punkt)**

124. Aufgabe

(10 Punkte)

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



In der Kartonagenfabrik Wenn man sich zu spät anmeldet, dann ist die Exkursion nicht so toll. Also geht es mit Rektor Skinner und Mrs. Krabappel in die Kartonagenfabrik. Wie schon im vorherigen Semester... In der Kartonagenfabrik angekommen erfahren Sie, dass die Kartons zum Versand von Nägeln nicht hier zusammengebaut werden sondern das sich die Endfertigung in Flint, Michigan befindet. Unter anderem wird dort der berühmte *Doppelt gewellte, 5-mal-gefaltete, 0.7mm, 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 Sie wenig Zuversicht zu erkennen.

Jetzt heißt es Kartons optimieren. Der nun zu optimierende, flache Karton hat eine Länge von 30cm und eine Breite von 21cm. Die Kartonagenmaschine in Flint soll dann einen quadratischen Eckenausschnitt der Länge x falzen.

1. Erstellen Sie eine Skizze des Kartonblattrohlings! Beschriften Sie die Skizze mit den entsprechenden Längenangaben **(1 Punkt)**
2. Berechnen Sie die Falztiefe x für ein maximales Volumen des flachen Kartons! **(3 Punkte)**
3. Welches Volumen in Liter ergibt sich mit der von Ihnen berechneten Falztiefe x ? **(1 Punkt)**
4. Sie wollen noch einen bündig mit dem Boden abschließenden Deckel für den Karton stanzen lassen. Wie groß ist die Fläche des Kartondeckelblattrohlings in cm^2 ? **(2 Punkte)**

Rektor Skinner möchte sich gerne wieder in seinem Vorgarten aufhalten und nicht die ganze Zeit von Bart mit Erdnüssen beworfen werden. Deshalb möchte er einen geräumigen Teil seines Vorgartens einzäunen. Ein Teil der Umzäunung bildet seine Vorderhauswand. Wegen Lieferschwierigkeiten stehen Rektor Skinner nur 90m Zaun zu Verfügung. Sie wollen nun die maximale Fläche des abgeschirmten Vorgartens in Abhängigkeit der Seitenlängen bei der Verwendung von 90m Zaun bestimmen!

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

125. Aufgabe

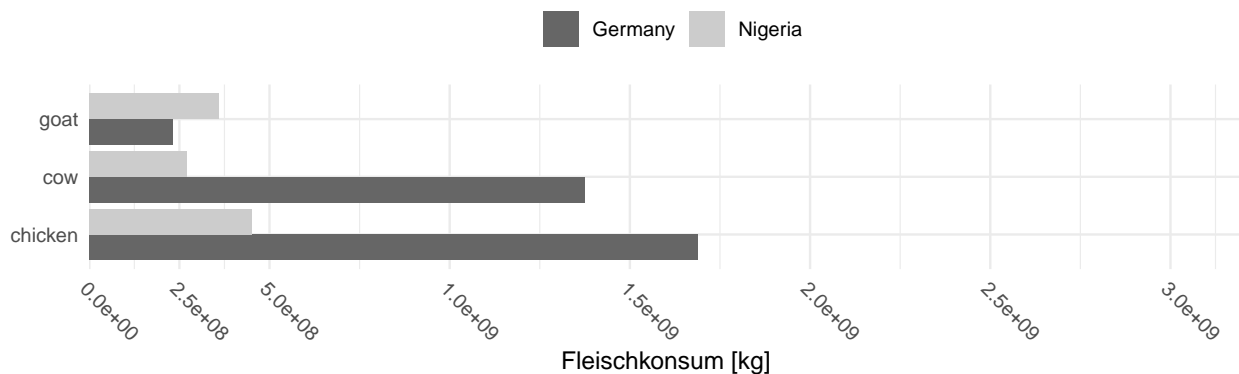
(10 Punkte)

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



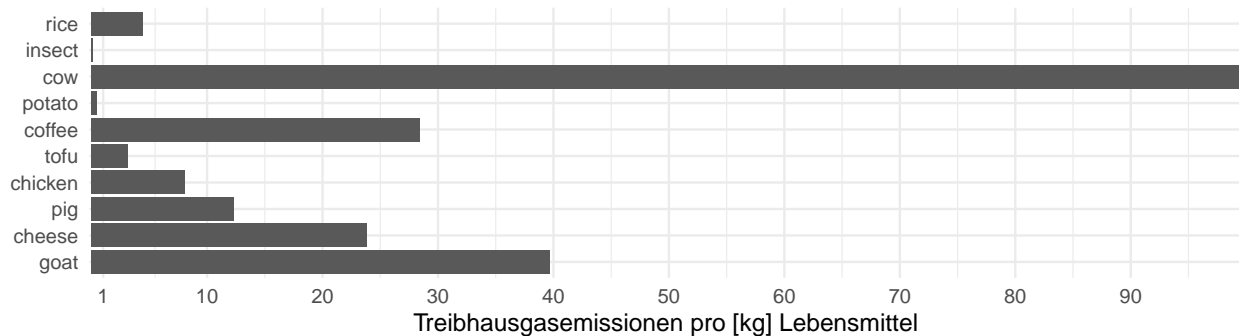
Ein Pfund Insekten, bitte! Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹³. Schauen wir uns dazu einmal den Vergleich Deutschland zu Nigeria an. Nigeria hat eine der am schnellsten wachsenden Bevölkerungen der Welt und wird vermutlich im Jahr 2100 zu den Top 5 der bevölkerungsreichsten Länder zählen. Im Jahr 2023 leben ca. 8×10^7 Menschen in Deutschland und ca. 1.8×10^8 Menschen in Nigeria. Mit den Informationen wollen wir anfangen und dann eine Prognose für den Fleischkonsum im Jahr 2050 zu treffen.

Im folgenden ist Abbildung des Fleischkonsums im Jahr 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)**

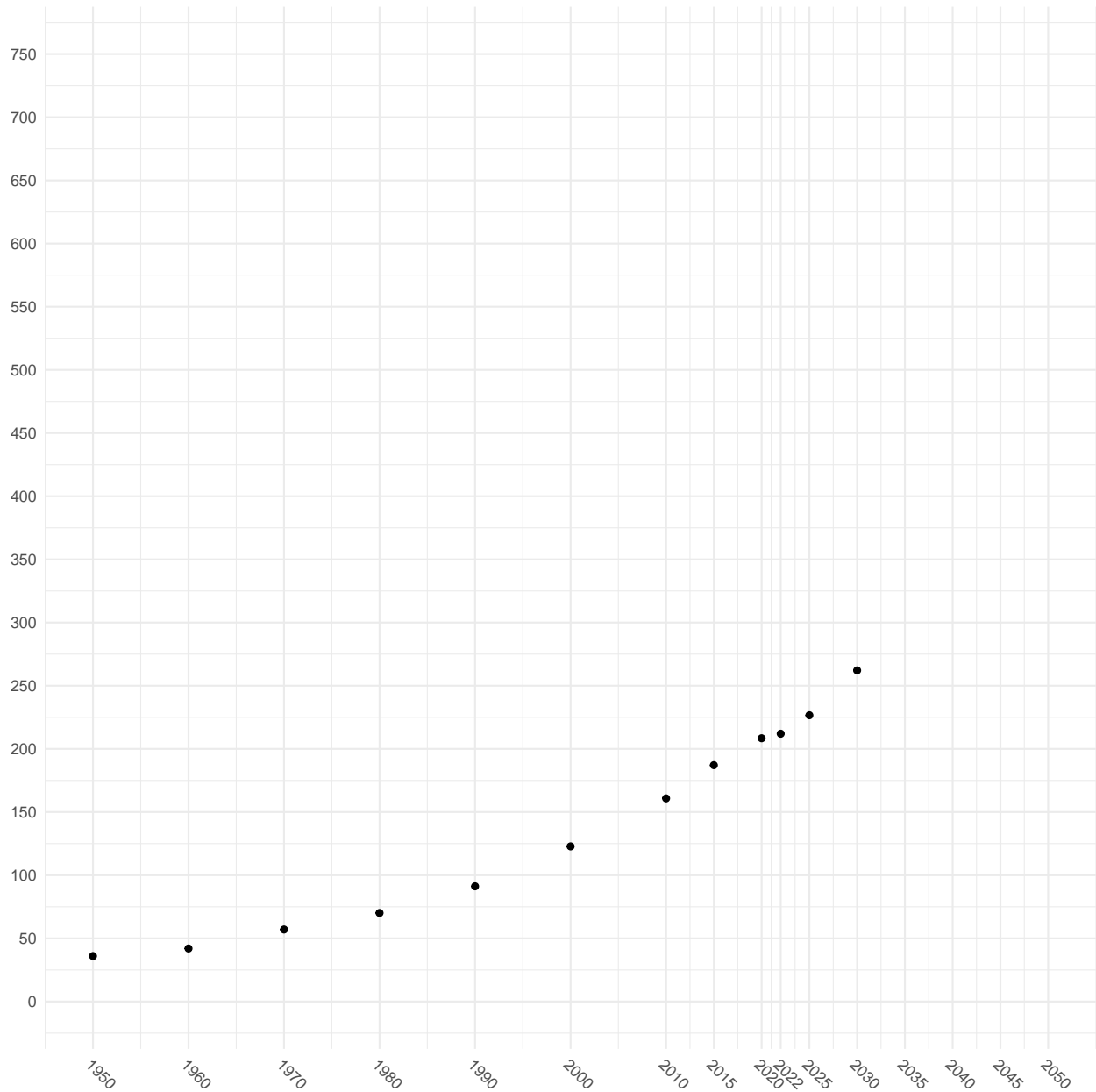
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



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)**

¹³Die Quelle der Inspiration für die Aufgabe war der folgende Artikel aus dem Spiegel: [Acht Milliarden - sind wir bald zu viele Menschen auf der Erde?](#)

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



4. Schätzen Sie graphisch die zu erwartende Bevölkerung [Millionen] in Nigeria im Jahr 2050, die sich anhand der Informationen aus der Abbildung ergibt!
 - a) Ohne Berücksichtigung der Covid-19-Pandemie! **(1 Punkt)**
 - b) Unter Berücksichtigung der Covid-19-Pandemie! **(1 Punkt)**
5. Berechnen Sie den geschätzten Fleischkonsum von Nigeria im Jahr 2050 unter der Annahme 80%-iger Angleichung der Lebensbedingungen zu Deutschland im Jahr 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)**

126. Aufgabe

(12 Punkte)

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



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

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.8% angenommen. In 95% der Fälle ist ein HIV-Test positiv, wenn der Patient erkrankt ist. In 0.5% der Fälle ist ein HIV-Test positiv, wenn der Patient *nicht* erkrankt ist und somit gesund ist. Sie stutzen. Wie wahrscheinlich ist es denn eigentlich an AIDS erkrankt zu sein (K^+), wenn Sie einen positiven AIDS-Test vorliegen haben (T^+)? Gehen Sie für die folgenden Berechnungen von $n = 3 \times 10^4$ Patienten mit einem diagnostischen Test für AIDS aus. Sie nehmen sich also einen Kuli und fangen an auf der historischen Ausgabe des Spiegels zu rechnen¹⁴.

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

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

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

¹⁴Die 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

127. Aufgabe

(12 Punkte)

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



Network-Marketing oder Schneeballschlacht! Eine Möglichkeit, leicht Geld zu verdienen, ist es anderen Menschen für Geld zu versprechen, wie man leicht reich werden kann. Am besten natürlich ohne viel Aufwand und ortsunabhängig. Dann wollen wir mal loslegen. Schnell ein YouTube-Werbevideo gedreht und auf geht es mit unserem Network-Marketing. Aber Moment, wie funktioniert Network-Marketing eigentlich und was hat das alles mit einer Schneeballschlacht zu tun? Wir wollen hier einmal in die Untiefen des „passiven Einkommens“ abtauchen¹⁵.

Das Jahr 2022 war das erfolgreichste Jahr in der Geschichte von Direct Finanzanlagen Left/Right (D-FL/R). Das Unternehmen steigerte den Umsatz um rund 27 Prozent von 260 Millionen Euro im Jahr 2021. Doch wie viel kommt bei den Partnern an? Laut D-FL/R habe das Unternehmen 3.3×10^5 aktive Partner.

1. Berechnen Sie zuerst den Umsatz der Firma D-FL/R im Jahr 2022! **(1 Punkt)**
2. Wie viel von dem Umsatz im Jahr 2022 wird im Durchschnitt von jedem aktiven Partner erwirtschaftet? **(1 Punkt)**
3. Welche *monatlicher* Umsatz ergibt sich dadurch im Durchschnitt für jeden aktiven Partner bei einer direkten Provision von 25%? **(1 Punkt)**

Ihr zu vermarkendes Produkt, hinter dem Sie voll stehen, kostet 200EUR pro Einheit im Direktverkauf. Die direkte Provision für die erste Stufe beträgt 40%. Für die zweite, dritte und vierte Stufe betragen die indirekten Provisionen jeweils 2.5%, 1.5% und 1%. Jeder Ihrer angeworbenen „Partner“ wirbt wiederum drei Partner für sich selbst an. Pro Monat werden im Schnitt drei Einheiten vom Produkt verkauft. Sie wollen nun 3200EUR im Monat *passiv* – also durch indirekte Provisionen – erwirtschaften.

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

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

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

Sie mussten zum Einstieg bei D-FL/R Einheiten des Produkts für 8000EUR kaufen. Diese Einheiten können Sie nur direkt verkaufen. Leider mussten Sie den Kauf über einen Kredit über 6% p.a. über 60 Monate finanzieren.

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

¹⁵Die 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](#)

128. Aufgabe

(10 Punkte)

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



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

In dem Spiel haben Sie nun auf einmal 6 vierseitige Würfel ($6d4$) zum würfeln in der Hand. Wenn Sie eine 4 würfeln, haben Sie einen Erfolg.

1. Berechnen Sie die Wahrscheinlichkeit *genau* 5 Erfolge zu erzielen! **(2 Punkte)**
2. Berechnen Sie die Wahrscheinlichkeit keinen Erfolg zu erzielen! **(1 Punkt)**

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

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

Jetzt wird es immer wilder, da Sie sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass Ihr Rettungswurf gegen den zaubernden Hexer funktioniert. Sie haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.65$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.7$. Sie haben mitgezählt und festgestellt, dass in 40 von 100 Fällen Ihr Rettungswurf bei einem erfolgreichen Zauber funktioniert hat.

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

129. Aufgabe

(12 Punkte)

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



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

Name	$P(\text{win})$	$P(\text{outbid})$
Günther	0.4	0.02
Frida	0.3	0.05

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 0.12 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 auf der Kirmes und spielen mit Catwoman um das große Geld. Das Glücksrad hat 24 Felder. Sie drehen das Glücksrad zweimal. Auf 10 Feldern gewinnen Sie 5000EUR sonst 1500EUR. 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 6500EUR? **(1 Punkt)**

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

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

Teil XII.

Angewandte Nutztier- und Pflanzenwissenschaften (M.Sc.)

130. Aufgabe

(6 Punkte)

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

131. Aufgabe

(8 Punkte)

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

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

mit

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

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

132. Aufgabe

(6 Punkte)

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