

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 72 Punkten sind aus den Rechen- und Textaufgaben erreicht.

_____ von 92 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
88.0 - 92.0	1,0
83.5 - 87.5	1,3
78.5 - 83.0	1,7
74.0 - 78.0	2,0
69.5 - 73.5	2,3
65.0 - 69.0	2,7
60.5 - 64.5	3,0
55.5 - 60.0	3,3
51.0 - 55.0	3,7
46.0 - 50.5	4,0

Es ergibt sich eine Endnote von _____.

Multiple Choice Aufgaben

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

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

- Es sind ____ von 20 Punkten erreicht worden.

Rechen- und Textaufgaben

Aufgabe	11	12	13	14	15	16	17
Punkte	9	12	12	9	10	10	10

- Es sind ____ von 72 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 Versionen. Der Text mag sich unterscheiden, die Inhalte sind aber gleich.

ANOVA

1. Aufgabe

(2 Punkte)

Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.12$. Welche Aussage ist richtig?

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

2. Aufgabe

(2 Punkte)

Sie führen ein Feldexperiment durch um das Gewicht von Brokoli zu steigern. Die Pflanzen wachsen unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen. Nach der Berechnung einer einfaktoriellen ANOVA ergibt sich ein $\eta^2 = 0.24$. Welche Aussage ist richtig?

- A** ☐ Es werden 24% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- B** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 24% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 76%.
- C** ☐ Es werden 76% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.
- D** ☐ Es werden 24% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen erklärt wird.
- E** ☐ Das η^2 beschreibt den Anteil der Varianz, der durch den Forschenden entsteht. Es gilt die Regel, dass ca. 70% der Varianz eines Versuches durch die Versuchsdurchführung entstehen sollen.

3. Aufgabe

(2 Punkte)


Die einfaktorielle ANOVA ist ein Standardverfahren in der agrawissenschaftlichen Forschung wenn es um den Vergleich von Behandlungsgruppen geht. Welche der folgenden Aussage zu der Berechnung der Teststatistik der einfaktoriellen ANOVA ist richtig?

- A** ☐ Die ANOVA berechnet die F-Statistik indem die MS des Fehlers durch die MS der Behandlung geteilt werden. Wenn die F-Statistik sich der 1 annähert kann die Nullhypothese nicht abgelehnt werden.
- 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** ☐ Die ANOVA berechnet die T-Statistik indem den Mittelwertsunterschied der Gruppen simultan durch die Standardabweichung der Gruppen teilt. Wenn die T-Statistik höher als 1.96 ist, kann die Nullhypothese abgelehnt werden.

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

4. Aufgabe

(2 Punkte)

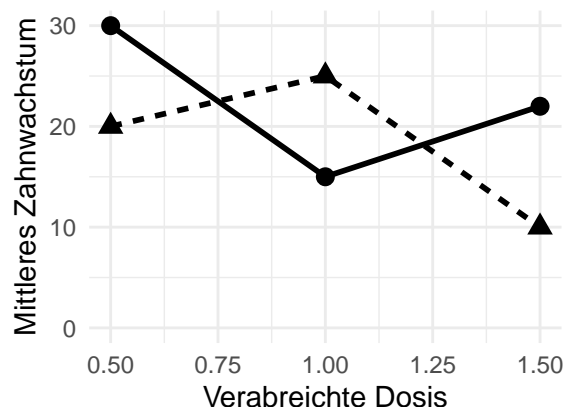
Wenn Sie mehr als zwei Gruppen als Behandlungen vorliegen haben, dann kann ein einfacher t-Test nicht für den globalen Vergleich genutzt werden. Sie entscheiden sich für eine ANOVA in . Die ANOVA analysiert dabei...

- A** ☐ ... den Unterschied zwischen mehreren Varianzen aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.
- B** ☐ ... den Unterschied zwischen der globalen Varianz und der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.
- C** ☐ ... den Unterschied zwischen der Varianz über alle Behandlungsgruppen oder der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss sich zwischen einem der beiden Varianzquellen entschieden werden.
- D** ☐ ... den Unterschied zwischen der Mittelwerte und der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist bekannt welcher Vergleich konkret unterschiedlich ist.
- E** ☐ ... den Unterschied zwischen der Varianz durch verschiedene Behandlungsgruppen unter der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, kann kein Effekt η^2 bestimmt werden.

5. Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung von Vitamin C auf das Zahnwachstum bei Schweine. Der 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. Welche Aussage ist richtig im Bezug auf eine zweifaktorielle ANOVA?



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

Deskriptive Statistik & Explorative Datenanalyse

6. Aufgabe

(2 Punkte)

Berechnen Sie den Mittelwert und Standardabweichung von y mit -1, 9, 6, 11 und 10.

- A ☐ Es ergibt sich 6 +/- 11.75
- B ☐ Sie erhalten 7 +/- 2.425
- C ☐ Es berechnet sich 8 +/- 23.5
- D ☐ Es ergibt sich 8 +/- 2.425
- E ☐ Es ergibt sich 7 +/- 4.85

7. Aufgabe

(2 Punkte)

Berechnen Sie den Median, das 1st Quartile sowie das 3rd Quartile von y mit 13, 26, 24, 11, 16, 25 und 63.

- A ☐ Sie erhalten 24 +/- 26
- B ☐ Es berechnet sich 24 [13; 26]
- C ☐ Sie erhalten 24 [11; 24]
- D ☐ Es berechnet sich 25 [14; 27]
- E ☐ Es ergibt sich 25 +/- 13

8. Aufgabe

(2 Punkte)

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

- A ☐ 1 Beobachtung.
- B ☐ 10 Beobachtungen.
- C ☐ Wir sollten eine Beobachtung mindestens pro Gruppe vorliegen haben.
- D ☐ Die untere Grenze liegt bei einer Beobachtung.
- E ☐ Wir brauchen fünf oder mehr Beobachtungen.

9. Aufgabe

(2 Punkte)

Die Varianz ist eine bedeutende deskriptive Statistik für die Analyse von Daten. Wie müssen Sie vorgehen um die Varianz zu berechnen?

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

10. Aufgabe

(2 Punkte)

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

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

11. Aufgabe

(2 Punkte)

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

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

12. Aufgabe

(2 Punkte)

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

- A** ☐ Einen Boxplot. Der Median, dargestellt als Linie, muss in der Mitte des IQR, dargestellt durch die Box, liegen.
- B** ☐ Einen Barplot. Die Mittelwerte müssen alle auf einer Höhe liegen. Die Fehlerbalken haben hier keine Informationen.
- C** ☐ Wir erstellen uns für jede Behandlung einen Boxplot und schauen, ob die Box und damit das IQR für jede Behandlung gleich groß ist.
- D** ☐ Einen Dotplot. Die Punkte müssen sich wie an einer Perlenschnur anordnen. Eine Abweichung führt zur Ablehnung der Annahme einer Varianzhomogenität.
- E** ☐ In einer explorativen Datenanalyse nutzen wir den Violinplot. Dabei sollte der Bauch am Rand liegen. Dann können wir von einer Varianzhomogenität ausgehen.

13. Aufgabe

(2 Punkte)

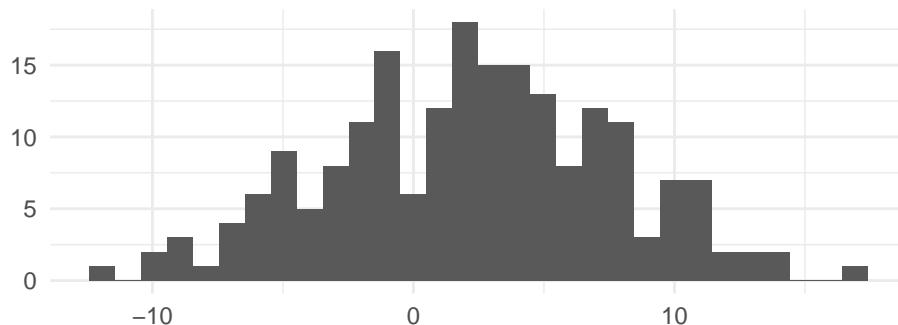
Nach der Durchführung Ihres Feldexperiments wollen Sie eine ANOVA rechnen. Dafür muss aber Ihr Messwert zumindestens approximativ einer Normalverteilung folgen. Welche der drei Abbildungen erlaubt Ihnen abzuschätzen, ob Sie eine Normalverteilung in Ihrem Endpunkt vorliegen haben?

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

14. Aufgabe

(2 Punkte)

In dem folgenden Histogramm von $n = 211$ Pflanzen ist welche Verteilung abgebildet?



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

Lineare Regression & Korrelation

15. Aufgabe

(2 Punkte)

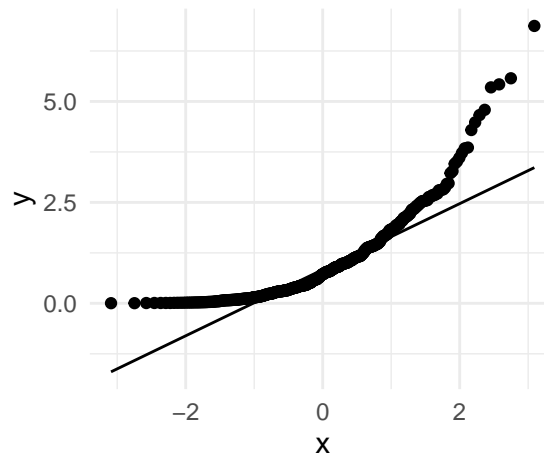
Sie haben das Modell $Y \sim X$ vorliegen und wollen nun ein prädiktives Modell rechnen. Welche Aussage ist richtig?

- A ☐ Ein prädiktives Modell möchte die Zusammenhänge von X auf Y modellieren. Hierbei geht es um die Effekte von X auf Y . Man sagt, wenn x_1 um 1 ansteigt ändert sich Y um einen Betrag β_1 .
- B ☐ Ein prädiktives Modell wird auf einem Trainingsdatensatz trainiert und anschliessend über eine explorative Datenanalyse validiert. Signifikanzen über β_i können hier nicht festgestellt werden.
- C ☐ Wenn ein prädiktives 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?
- D ☐ Wir modellieren den Zusammenhang zwischen X und Y wenn ein prädiktives Modell rerechnet wird. Dabei kann nicht der gesamte Datensatz genutzt werden. Es wird ein Trainingsdatensatz zum Trainieren des Modells benötigt.
- E ☐ Wenn ein prädiktives 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.

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?

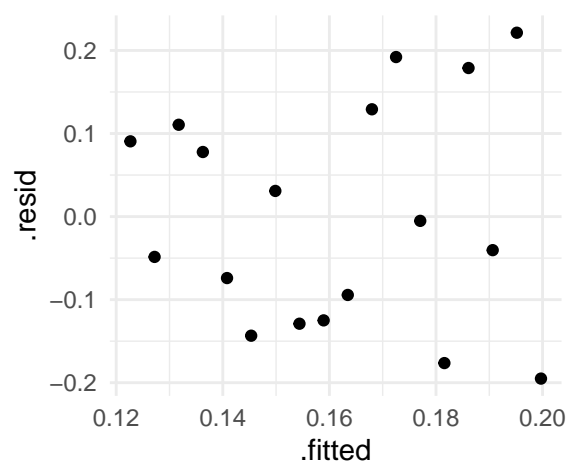


- A** ☐ Wir betrachten insbesondere die beiden Enden der Gerade. Der Rest ist mehr oder minder egal, dann ist die Annahme an die Normalverteilung der Residuen erfüllt.
- B** ☐ Wir betrachten die Gerade, die durch die einzelnen Punkte laufen sollte. Wenn die 95% der Punkte von der Geraden getroffen werden,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

17. Aufgabe

(2 Punkte)

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



- A** ☐ Wir betrachten die Nulllinie und alle Punkte sollten ohne Muster gleichmäßig um die Nulllinie liegen. Da dies der Fall ist, gehen wir von keinen Ausreißern aus.

- B** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Ein klares Muster ist zu erkennen und/oder einige Outlier sind zu beobachten.
- C** ☐ Die Punkte müssen gleichmäßig in dem positiven Bereich liegen. Dies ist hier klar nicht der Fall. Einzelne Ausreißer können beobachtet werden. Die Analyse ist gescheitert.
- D** ☐ Wenn wir die Nulllinie betrachten so müssen die Punkte gleichmäßig über der Nulllinie liegen. Unser Modell erfüllt somit nicht die Annahme von normalverteilten Residuen mit einem Mittelwert von > 0 und einer Streuung von s .
- E** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Es ist ein Muster zu erkennen und wir können damit auf die Signifikanz von x_1, \dots, x_p schließen.

18. Aufgabe

(2 Punkte)

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

- A** ☐ 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.
- B** ☐ Korrelationskoeffizienten ρ liegt zwischen 0 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als Standardisierung verstanden werden.
- C** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen -1 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos.
- D** ☐ 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.
- E** ☐ 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.

19. Aufgabe

(2 Punkte)

In einer linearen Regression kann es vorkommen, dass der Effekt repräsentiert durch den β Koeffizienten nicht so richtig von der Größenordnung zu dem p-Wert passen will. So liefert eine Untersuchung des Einflusses von der PO_2 -Konzentration in $[\mu g]$ im Wasser auf das Wachstum in $[kg]$ an Spitzkohl folgende Effekte und p-Werte: 0.00032 als p-Wert und einen β_{PO_2} Koeffizienten von 2.3×10^{-9} . Welche Aussage ist richtig?

- A** ☐ Die Fallzahl ist zu hoch angesetzt. Je höher die Fallzahl ist, desto kleiner ist die Teststatistik und damit ist dann auch der p-Wert sehr klein. Es sollte über eine Reduzierung der Fallzahl nachgedacht werden. Dann sollte der Effekt zum p-Wert passen.
- B** ☐ 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.
- C** ☐ Die Einheit der PO_2 -Konzentration ist zu klein gewählt. Die Erhöhung der PO_2 -Konzentration um 1 Einheit führt nur zu einem sehr winzigen Anstieg von β_{PO_2} im Gewicht der Wasserlinsen. Die Einheit $[\mu g]$ muss besser gewählt werden.
- D** ☐ Wenn der Effekt β_{PO_2} winzig ist, dann kann es an einer falsch gewählten Einheit liegen. Der Anstieg von einer Einheit in X führt ja zu einer Änderung von β_{PO_2} in x . Wir müssen daher die Einheit von y entsprechend anpassen.
- E** ☐ Das Gewicht und die PO_2 -Konzentration korrelieren sehr stark, deshalb wird der β_{PO_2} Koeffizient sehr klein. Mit einer ANOVA kann für die Korrelation korrigiert werden und der Effektschätzer passt dann zum p-Wert.

20. Aufgabe

(2 Punkte)

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

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

21. Aufgabe

(2 Punkte)

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

- ☐ A Das GLM ist eine Vereinfachung des LM in R. Mit dem GLM lassen sich polygonale Regressionen rechnen. Somit stehen neben der Normalverteilung noch weitere Verteilungen zu Verfügung.
- ☐ B Das GLM ist eine allgemeine Erweiterung der linearen Regression auf die Normalverteilung.
- ☐ C Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsgruppen für das X bzw. die Einflussvariablen in einer linearen Regression zu wählen.
- ☐ D In R ist mit dem *generalisierten linearen Modell (GLM)* eine Modellierung implementiert, die die Poissonverteilung für Zähldaten oder die Binomialverteilung für 0/1-Daten modellieren kann. Weitere Modellierungen sind in R auch mit zusätzlich geladenen Paketen nicht möglich.
- ☐ E Dank dem *generalisierten linearen Modell (GLM)* können auch andere Verteilungsfamilien als die Normalverteilung mit einer linearen Regression modelliert werden.

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 bringt starke Unstrukturiertheit in das Experiment und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- ☐ B Randomisierung ist die direkte Folge von Strukturgleichheit. Die Strukturgleichheit erlaubt es erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- ☐ 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 Strukturgleichheit ist durch Randomisierung gegeben. Somit kann von der Stichprobe auf die Grundgesamtheit geschlossen werden

23. Aufgabe


(2 Punkte)


Sie wollen Ihren Datensatz in  einlesen und stehen nun vor einem Problem. Sie stellen fest, dass die Hilfeseiten alle in englischer Sprache verfasst sind. Warum mag die Nutzung von Deutsch problematisch sein?

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

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** ☐ Die Funktionsreihenfolge ist wie folgt: `read_excel()` -> `mutate()` -> `ggplot()`. Dabei ist bei der Transformation der Daten darauf zu achten, dass die Faktoren richtig erstellt werden.
- C** ☐ Die Funktionsreihenfolge ist wie folgt: `read_excel()` -> `mutate()` -> `ggplot()`. Dabei ist bei der Transformation der Daten darauf zu achten, dass keine Faktoren erstellt werden.
- 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** ☐ 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.

25. Aufgabe

(2 Punkte)

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

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

26. Aufgabe

(2 Punkte)

Die Leistung von Sauen soll auf einem Zuchtbetrieb gesteigert werden. Dafür werden die Ferkel verschiedener Sauen gemessen. Die Ferkel einer Muttersau sind daher...

- A** ☐ Untereinander unabhängig. Sollten die Mütter verwandt sein, so ist die Varianzstruktur ähnlich und muss modelliert werden.
- B** ☐ Die Ferkel stammen vom gleichen Muttertier und haben vermutlich eine ähnlichere Varianzstruktur als die Ferkel von anderen Sauen. Die Ferkel sind untereinander über die Mutter abhängig.
- C** ☐ Abhängig von der Stallanlage und des Experiments können die Ferkel abhängig oder unabhängig sein. Allgemein gilt, dass Ferkel von unterschiedlichen Sauen näher miteinander verwandt sind als Ferkel von gleichen Sauen. Das Fisher-Axiom.
- D** ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.
- E** ☐ Untereinander abhängig, wenn die Mütter ebenfalls miteinander verwandt sind. Erst die Abhängigkeit 2. Grades wird in der Statistik modelliert.

27. Aufgabe

(2 Punkte)

Neben der Mittelwertsdifferenz als Effektschätzer bei normalverteilten Endpunkten wird auch häufig der Effektschätzer Odds ratio bei binären Endpunkten verwendet. Welche Aussage über den Effektschätzer Odds ratio ist im folgenden Beispiel zur Behandlung von Klaueninfektionen bei Schafe richtig? Dabei sind 3 Tiere krank und 12 Tiere sind gesund.

- A** ☐ Es ergibt sich ein Odds ratio von 4, da es sich um ein Anteil handelt.
- B** ☐ Da es sich um ein Chancenverhältnis handelt ergibt sich ein Odds ratio von 5.
- C** ☐ Der Anteil der Gesunden wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Odds ratio von 0.2.
- D** ☐ Da es sich um ein Chancenverhältnis handelt ergibt sich ein Odds ratio von 0.25.
- E** ☐ Der Anteil der Kranken wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Odds ratio von 0.2.

28. Aufgabe

(2 Punkte)

In der Bio Data Science wird häufig mit sehr großen Datensätzen gerechnet. Historisch ergibt sich nun ein Problem bei der Auswertung der Daten und deren Bewertung hinsichtlich der Signifikanz. Welche Aussage ist richtig?

- A** ☐ Aktuell werden immer größere Datensätze erhoben. Eine erhöhte Fallzahl führt automatisch auch zu mehr signifikanten Ergebnissen, selbst wenn die eigentlichen Effekte nicht relevant sind.
- B** ☐ Aktuell werden immer größere Datensätze erhoben. Dadurch wird auch die Varianz immer höher was automatisch zu mehr signifikanten Ergebnissen führt.
- C** ☐ Riesige Datensätze haben mehr Fallzahl was zur α -Inflation führt. Durch eine Adjustierung kann dem Problem entgegengewirkt werden.
- D** ☐ Relevanz und Signifikanz haben nichts miteinander zu tun. Daher gibt es auch keinen Zusammenhang zwischen hoher Fallzahl ($n > 10000$) und einem signifikanten Test. Ein Effekt ist immer relevant und somit signifikant.
- E** ☐ 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.

Multiple Gruppenvergleiche

29. Aufgabe

(2 Punkte)

Sie haben folgende unadjustierten p-Werte gegeben: 0.01, 0.03, 0.42 und 0.89. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.04, 0.12, 1.68 und 3.56. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.0025, 0.0075, 0.105 und 0.2225. Die adjustierten p-Werte werden zu einem α -Niveau von 1.25% verglichen.
- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.0025, 0.0075, 0.105 und 0.2225. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.04, 0.12, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 1.25% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.04, 0.12, 1 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.

30. Aufgabe

(2 Punkte)

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

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

31. Aufgabe

(2 Punkte)

Der multiple Vergleich als Posthoc-Test nach einer ANOVA ist in den Agrarwissenschaften heutzutage Standard. Welches R Paket wird häufig für den multiplen Vergleich genutzt? Welche Beschreibung der Eigenschaften ist korrekt?

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

32. Aufgabe

(2 Punkte)

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

- A** ☐ Beim multiplen Testen kann es zu einer α -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Welch das bekannteste Verfahren ist.
- B** ☐ 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.
- C** ☐ 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.
- D** ☐ Die Adjustierung der p-Werte nach Bonferroni erlaubt es gegen die β -Inflation vorzugehen, die häufig beim multiplen Testen auftritt. Das globale Powerniveau liegt nicht mehr bei 80% sondern sehr viel niedriger.
- E** ☐ Beim multiplen Testen kann es zu einer α -Inflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern sehr viel höher. 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.

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** ☐ Beim multiplen Testen kann es zu einer Δ -Inflation kommen. Das globale Effektniveau liegt nicht mehr bei 20%. Daher müssen die Effekte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der Effekte nach Bonferroni das bekannteste Verfahren ist.
- B** ☐ 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.
- C** ☐ Beim multiplen Testen muss der Effekt, wie der Mittelwertsunterschied Δ aus einem t-Test, nicht adjustiert werden.
- D** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ nach Bonferroni adjustiert werden. Dafür wird der Effekt mit der Anzahl an Vergleichen k multipliziert. Dies geschieht analog zu den p-Werten.
- E** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ adjustiert werden im Gegensatz zu den p-Werten.

Statistische Testtheorie

34. Aufgabe

(2 Punkte)

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

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

- D** ☐ $Pr(D|H_0)$ beschreibt die Wahrscheinlichkeit die Teststatistik T_D aus den Daten D zu beobachten, wenn die Nullhypothese wahr ist.
- E** ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit der Alternativhypothese und somit $1 - Pr(H_A)$

35. Aufgabe

(2 Punkte)

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

- A** ☐ ... dass ein schlechtes Modell durch ein weniger schlechtes Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.
- B** ☐ ... dass Annahmen an statistische Modelle meist falsch sind.
- C** ☐ ... dass Fehlerterme in statistischen Modellen nicht verifiziert werden können.
- D** ☐ ... dass in der Wissenschaft immer etwas falsch sein muss. Sonst gebe es keinen Fortschritt.
- E** ☐ ... dass ein schlechtes Modell durch ein schlechteres Modell ersetzt wird. Die Wissenschaft lehnt ab und verifiziert nicht.

36. Aufgabe

(2 Punkte)

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

- A** ☐ Da Wissenschaftler eine Schwelle für die statistische Testentscheidung benötigen wurde α in einer großen Konferenz 1945 gewählt. Damit ist $\alpha = 5\%$ eine Kulturkonstante mit einem Rank einer Naturkonstante.
- B** ☐ Der Wert ergab sich aus einer Auswertung von 1042 wissenschaftlichen Veröffentlichungen zwischen 1914 und 1948. Der Wert 5% wurde in 28% der Veröffentlichungen genutzt. Daher legte man sich auf diese Zahl fest.
- C** ☐ Die Festlegung von $\alpha = 5\%$ ist eine Kulturkonstante. Wissenschaftler benötigt eine Schwelle für eine statistische Testentscheidung, der Wert von α wurde aber historisch mehr zufällig gewählt.
- D** ☐ 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.
- E** ☐ 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.

37. Aufgabe

(2 Punkte)

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

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

38. Aufgabe

(2 Punkte)

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

H_0 ablehnen obwohl die H_0 gilt

- A ☐ Dem β -Fehler mit der Analogie eines brennenden Hauses: *Fire without alarm*.
- B ☐ *Alarm with fire*, dem α -Fehler in der Analogie von Feuer.
- C ☐ *Fire without alarm*, dem β -Fehler als Analogie eines Rauchmelders.
- D ☐ *Alarm without fire*, dem α -Fehler in der Analogie eines Rauchmelders.
- E ☐ *Fire without alarm*, dem β -Fehler als Analogie von Rauch im Haus.

39. Aufgabe

(2 Punkte)

Welche statistische Maßzahl erlaubt es Relevanz mit Signifikanz zu verbinden? 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 ☐ 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.
- C ☐ 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.
- D ☐ Das Δ . Durch die Effektstärke haben wir einen Wert für die Relevanz, die vom Anwender bewertet werden muss. Da Δ antiproportional zum p-Wert ist, bedeutet auch ein hohes Δ ein sehr kleinen p-Wert.
- E ☐ Das Konfidenzintervall. Durch die Visualisierung des Konfidenzintervalls kann eine Relevanzschwelle vom Anwender definiert werden. Zusätzlich erlaubt das Konfidenzintervall auch eine Entscheidung über die Signifikanz.

40. Aufgabe

(2 Punkte)

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

- A ☐ Wir vergleichen 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.
- 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 mit dem p-Wert und dem Signifikanzniveau α absolute Werte auf einem Zahlenstrahl und damit den Unterschied der Teststatistiken, wenn die H_0 gilt.
- D ☐ Wir vergleichen mit dem p-Wert und dem Signifikanzniveau α Wahrscheinlichkeiten und damit die Flächen unter der Kurve der Teststatistik, 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 Ergebnisse eines statistischen Tests und die damit verbundene Theorie besser zu verstehen, kann eine Analogie zur Wettervorhersage genutzt werden. Welche Analogie zu der Testtheorie trifft am meisten zu?

- A** ☐ In der Analogie der Sonnenscheindauer: Wie lange kann mit einem entsprechenden Effekt gerechnet werden? Die Wahrscheinlichkeit für den Effekt gibt der statistische Test wieder.
- B** ☐ In der Analogie der Regenwahrscheinlichkeit: ein statistischer Test gibt die Wahrscheinlichkeit für das Auftreten eines Ereignisses wieder. Die Stärke des Effektes wird nicht wiedergeben.
- C** ☐ In der Analogie der Durchschnittstemperatur: Wie oft tritt ein Effekt durchschnittlich ein? Wir erhalten eine Wahrscheinlichkeit für die Effekte. Zum Beispiel, wie hoch ist die Wahrscheinlichkeit für einen Mittelwert als Durchschnitt.
- D** ☐ In der Analogie der Maximaltemperatur: Was ist der maximale Unterschied zwischen zwei Gruppen. Wir erhalten hier eine Aussage über die Spannweite und den maximalen Effekt.
- E** ☐ In der Analogie des Niederschlags oder Regenmenge: ein statistischer Test gibt die Stärke eines Effektes wieder. Zum Beispiel, wie hoch ist der Mittelwertsunterschied.

42. Aufgabe

(2 Punkte)

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

- A** ☐ Nein, es ist nicht möglich die untersuchte Population mit einem t-Test auszuwerten. Wir erhalten dann leider keine Aussage zur Population.
- B** ☐ Ja, es ist möglich die untersuchte Population mit einem t-Test auszuwerten. Wir erhalten dann eine Aussage zur Population.
- C** ☐ Nein, wir können die untersuchte Population nicht mit einem t-Test auswerten. Wir erhalten keine Aussage zur Population. Wir können aber den Effekt als Quelle der Relevanz nutzen.
- D** ☐ Ja, wir erhalten eine Aussage. Müssen aber das Individuum im Kontext der Population adjustieren.
- E** ☐ Weder eine Aussage über die Population noch über das Individuum ist mit einem statistischen Test möglich. Wir erhalten eine Aussage über ein Experiment.

43. Aufgabe

(2 Punkte)

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

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

44. Aufgabe

(2 Punkte)

Welche Aussage über den Effekt eines statistischen Tests ist richtig?

- A** ☐ Durch den Effekt erfahren wir die statistische interpretierbare Ausgabe eines statistischen Tests. Zum Beispiel das η^2 aus einer ANOVA. Damit können wir die Signifikanz direkt mit dem Effekt verbinden. Am Ende muss der Forschende aber entscheiden, ob der Effekt entsprechend seinen Erwartungen als bedeutet zu bewerten ist.
- B** ☐ Der Effekt eines statistischen Tests beschreibt die mathematisch interpretierbare Ausgabe eines Tests. Damit ist der Effekt direkt mit dem Begriff der Signifikanz verbunden. Die Entscheidung über die Signifikanz trifft der Forschende unabhängig von der Relevanz eines statistischen Tests.
- C** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch interpretierbare Ausgabe eines Tests. Modernen Algorithmen liefern keine Effekte mehr sondern nur noch bedingte Wahrscheinlichkeiten. Der Effekt spielt in der modernen Statistik keine Rollen mehr.
- D** ☐ Der Forschende muss am Anfang wissen, ob das Ergebnis eines Experiments relevant für seine Forschung ist. Dafür kann der Effekt eines statistischen Tests genutzt werden oder auch der Prähoc-Test. Damit beschreibt der Effekt den biologischen interpretierbaren Teil eines Experiments vor der Durchführung. Zum Beispiel der Unterschied zwischen zwei Mittelwerten.
- E** ☐ Durch den Effekt erfahren wir die biologisch interpretierbare Ausgabe eines statistischen Tests. Zum Beispiel das η^2 aus einer ANOVA. Damit können wir die Relevanz direkt mit dem Effekt verbinden. Am Ende muss der Forschende aber entscheiden, ob der Effekt entsprechend seinen Erwartungen als bedeutet zu bewerten ist.

45. Aufgabe

(2 Punkte)

Welche Aussage über die Entscheidung anhand der berechneten Teststatistik 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** ☐ Ist $Pr(D|H_0)$ kleiner als das Signifikanzniveau α gleich 5% dann wird die Nullhypothese H_0 abgelehnt.
- C** ☐ Ist in dem 95%-Konfidenzintervall nicht die Null enthalten dann wird die Nullhypothese H_0 abgelehnt.
- D** ☐ Anhand der berechneten Teststatistik lässt sich wie folgt eine Entscheidung treffen. Liegt der Wert in dem Signifikanzniveauintervall α dann kann die Nullhypothese abgelehnt werden.
- E** ☐ Anhand der berechneten Teststatistik lässt sich wie folgt eine Entscheidung treffen. Liegt der Wert über oder gleich dem Signifikanzniveau α dann kann die Nullhypothese abgelehnt werden.

46. Aufgabe

(2 Punkte)

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

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

Statistische Tests für Gruppenvergleiche

47. Aufgabe

(2 Punkte)

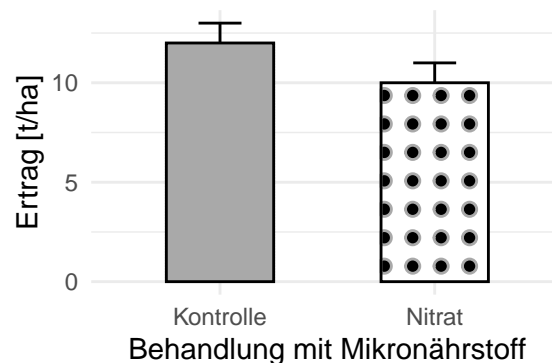
In Ihrer Abschlussarbeit rechnen Sie einen Student t-Test. Welche Aussage ist auch für den Welch t-Test richtig?

- ☐ **A** Der t-Test vergleicht die Mittelwerte von zwei Gruppen 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 ist ein Vortest der ANOVA und basiert daher auf dem Vergleich von Streuungsparametern
- ☐ **C** Der t-Test vergleicht die Varianzen von mindestens zwei oder mehr Gruppen
- ☐ **D** Der t-Test berechnet die Differenz von zwei Mittelwerten als Effekt und gibt eine Entscheidung, ob sich die beiden Mittelwerte in den Gruppen signifikant unterscheiden.
- ☐ **E** Der t-Test berechnet die Differenz von zwei Mittelwerten als Effekt und gibt eine Entscheidung, ob sich die beiden Mittelwerte *jeweils* von Null unterscheiden.

48. Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung des Mikronährstoff Nitrat auf den Ertrag in t/ha von Papaya im Vergleich zu einer Kontrolle. Der Versuch wurde in 11 Parzellen pro Gruppe durchgeführt. Welche Aussage im Bezug auf eine statistische Auswertung ist richtig?



- ☐ **A** Der Effekt und die Signifikanz lassen sich nicht aus Barplots abschätzen. Höchstens der Effekt als relativer Unterschied zwischen der Höhe der Barplots. Standard ist der mediane Unterschied aus Boxplots.
- ☐ **B** Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei -2. Wir müssen aber einen Posthoc-Test rechnen um den Effekt wirklich bestimmen zu können.
- ☐ **C** Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei -0.2.
- ☐ **D** Nach Betrachtung des Barplots liegt kein signifikanter Unterschied vor. Der Effekt liegt bei -2.
- ☐ **E** Es liegt ein signifikanter Unterschied vor. Der Effekt liegt bei -2.

49. Aufgabe

(2 Punkte)

Welche Aussage über den gepaarten t-Test für verbundene Stichproben ist richtig?

- ☐ **A** Beim gepaarten t-Test kombinieren wir die Vorteile des Student t-Test für Varianzhomogenität mit den Vorteilen des Welch t-Test für Varianzheterogenität. Wir bilden dafür die Differenz der Einzelbeobachtungen.
- ☐ **B** Der gepaarte t-Test nutzt die Varianz der Beobachtungen jeweils paarweise und bildet dafür eine verbundene Stichprobe. Dieser Datensatz d dient dann zur Differenzbildung.

- C** ☐ Der gepaarte t-Test wird gerechnet, wenn die Beobachtungen abhängig voneinander sind. Wir messen jede Beobachtung nur einmal und berechnen dann die Differenz zu dem Mittel der anderen Beobachtungen.
- D** ☐ Wenn die Beobachtungen unabhängig voneinander sind, rechnen wir einen gepaarten t-Test. Messen wir wiederholt an dem gleichen Tier oder Pflanze dann bilden wir das Produkt zwischen den zwei Messpunkten.
- E** ☐ Abhängige Beobachtungen müssen gesondert in einem gepaarten t-Test modelliert werden. Wenn wiederholt an dem gleichen Tier oder Pflanze gemessen wird, dann bilden wir die Differenz zwischen den beiden Zeitpunkten. Auf den Differenzen rechnen wir den gepaarten t-Test.

50. Aufgabe

(2 Punkte)

In Ihrer Abschlussarbeit passen die Ergebnisse einer ANOVA und eines multiplen Vergleiches nicht zusammen. Nach einem Experiment mit drei Maissorten ergibt eine ANOVA ($p = 0.049$). Sie führen anschließend die paarweisen t-Tests für alle Vergleiche durch. Nach der Adjustierung für multiples Testen ist kein p-Wert unter der α -Schwelle. Sie schauen sich auch die rohen, unadjustierten p-Werte an und finden hier als niedrigsten p-Wert $p_{3-2} = 0.053$. Welche Aussage ist richtig?

- A** ☐ Die ANOVA testet auf der gesamten Fallzahl. Die einzelnen t-Tests immer nur auf einer kleineren Subgruppe. Da mit weniger Fallzahl weniger signifikante Ergebnisse zu erwarten sind, kann eine Diskrepanz zwischen der ANOVA und den paarweisen t-Tests auftreten.
- B** ☐ 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.
- C** ☐ Das ist kein Wunder. Die ANOVA testet nicht auf der gesamten Fallzahl und die paarweisen t-Tests gewinnen immer eine oder mehr Gruppen als Fallzahl dazu. Mit steigender Fallzahl sind mehr signifikante Unterschiede zu erwarten. Die p-Werte unterscheiden sich numerisch auch kaum.
- D** ☐ Das Beispiel kann so nicht auftreten, da die ANOVA und die t-Tests algorithmisch miteinander verknüpft sind.
- E** ☐ 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.

Teil I.

Deskriptive Statistik & Explorative Datenanalyse

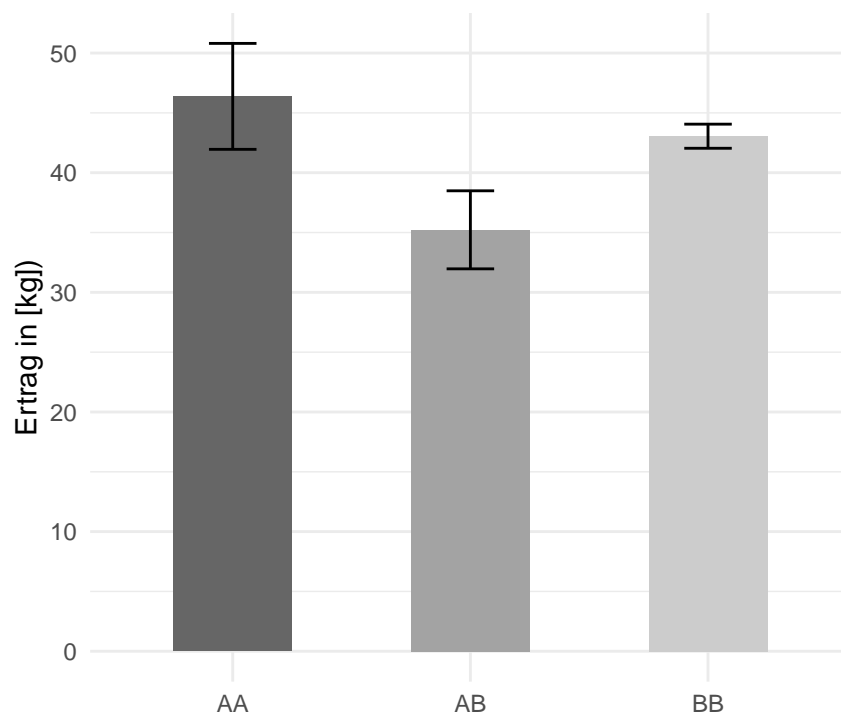
51. Aufgabe


(7 Punkte)


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



Anschaun, was andere vor einem gemacht haben, ist eine Möglichkeit schnell ans Ziel zu gelangen. Paula soll in ihrem Projektbericht Kartoffeln untersuchen. Die Behandlung in ihrem Projektbericht werden verschiedene Genotypen (AA, AB und BB) sein. Erheben wird Paula als Endpunkt (Y) *Ertrag* benannt als *yield* in ihrer Exceldatei. Von ihrer Betreuerin erhält sie nun folgende Abbildung von Barplots, die sie erstmal zur Übung nachbauen soll, bevor sie mit dem eigentlichen Versuch beginnt.



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

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

52. Aufgabe

(7 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 Steffen nicht gemerkt, welche statistischen Maßzahlen für einen Barplot erhoben werden müssen. Das ist in soweit doof, da nach seinem Betreuer nun Barplots aus seinen Daten gebaut werden sollen, bevor es mit dem statistischen Testen weitergeht. Die Behandlung für Brokoli waren verschiedene Substrattypen (*torf*, *40p60n* und *70p30n*). Erfasst wurde von Steffen als Messwert (*Y*) *Frischegewicht*. Steffen hat dann *freshmatter* in seiner Exceldatei eintragen.

treatment	freshmatter
40p60n	35.5
40p60n	37.0
70p30n	33.7
40p60n	37.8
torf	35.7
40p60n	29.3
40p60n	37.3
torf	42.5
70p30n	41.6
70p30n	35.6
torf	37.7
torf	40.0
torf	29.8

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

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

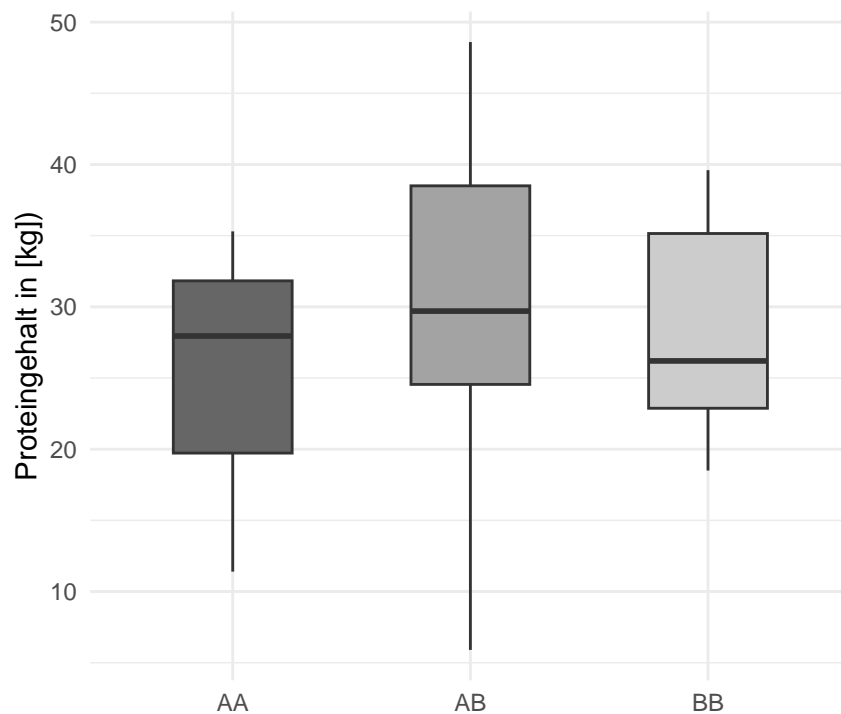
53. Aufgabe

(9 Punkte)

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



Alex steht vor einem ersten Problem, denn wenn es nach seiner Betreuerin geht, soll er in einem einem Versuch in einer Klimakammer Lauch auswerten. Soweit eigentlich alles passend. Die Behandlung werden verschiedene Genotypen (AA, AB und BB) sein. In seiner Exceldatei wird er den Outcome (Y) *Proteingehalt* als *protein* aufnehmen. Vorab soll Alex aber einmal die folgenden Boxplots seiner Betreuerin nachbauen, damit er den R Code schonmal für später vorliegen hat. Anhand von Boxplots lässt sich eine Aussage über die Normalverteilung von Y treffen. Damit geht das Problem schon los.



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

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

54. Aufgabe

(9 Punkte)

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



Alex steht vor einem ersten Problem, denn wenn es nach seiner Betreuerin geht, soll er in einem Freilandversuch Maiss auswerten. Soweit eigentlich alles passend. Die Behandlung waren verschiedene Genotypen (AA und BB). In seiner Exceldatei hat er den Messwert (Y) *Proteingehalt* als *protein* aufgenommen. Nun soll Alex die Daten einmal als Boxplots in einer Präsentation visualisieren, damit seiner Betreuerin wieder klar wird, was er 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.

treatment	drymatter
BB	39.2
AA	51.0
BB	29.7
AA	45.7
BB	37.4
AA	52.8
BB	36.3
BB	28.6
AA	39.8
AA	32.3
AA	40.6
BB	31.3
BB	34.5
AA	38.5

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

1. Zeichnen Sie in *einer* Abbildung die beiden Boxplots für die zwei Behandlungen von Maiss! 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 Maiss 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!



Jonas betrachtet die folgenden Daten nach einem Stallexperiment mit Hühnern. In dem Experiment wurden die dunklen Pigmentstörungen gezählt. Nach der Meinung seiner Betreuerin muss als erstes geschaut werden, wie diese verteilt sind. Also welcher statistischen Verteilung die dunklen Pigmentstörungen folgen. Dazu soll Jonas ein Histogramm verwenden. Dann hätte man auch einen guten Überblick über das Outcome (Y).

Die dunklen Pigmentstörungen: 3, 5, 5, 2, 6, 5, 2, 7, 3, 10, 3, 3, 5, 2, 3, 3, 3, 7, 1, 9, 2, 4, 7, 1, 2, 2, 9, 3, 3, 1

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

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

56. Aufgabe

(8 Punkte)

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



Alex betrachtet die folgenden Daten nach einem Leistungssteigerungsversuch mit Hühnern. In dem Experiment wurden die mittleren dunklen Pigmentstörungen gezählt. Nach der Meinung seinem Betreuer muss als erstes geschaut werden, wie diese verteilt sind. Also welcher statistischen Verteilung die mittleren dunklen Pigmentstörungen folgen. Dazu soll Alex ein Histogramm verwenden. Dann hätte man auch einen guten Überblick über den Messwert (Y).

Die mittleren dunklen Pigmentstörungen: 10.5, 8.7, 12.1, 9.9, 13.3, 15.2, 9, 11.5, 9.1, 11.4, 9.4, 11.4, 12.7, 12.1, 7.1, 11.2, 9, 7.4, 9.4, 9.4, 11.4, 6

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

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



Steffen liest laut: 'Wenn zwei kontinuierliche Variablen vorliegen, können diese in einer explorativen Datenanalyse...'. Steffen stoppt. Was waren noch gleich kontinuierliche Variablen? In seinem Projektbericht hatte er einen Leistungssteigerungsversuch im Teutoburgerwald durchgeführt. Dabei ging es um den Zusammenhang zwischen Gewichtszuwachs in der 1LW und durchschnittlicher Tagestemperatur [C/d] im groben Kontext von Hühnern. Nun stellt sich die Frage für ihn, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Dafür war eine explorative Datenanalyse gut!

Durchschnittlicher Tagestemperatur [C/d]	Gewichtszuwachs in der 1LW
19.7	34.0
21.4	36.5
21.8	39.4
20.4	35.7
20.8	35.3
19.1	34.7
21.1	31.8
20.4	34.1
20.4	35.6
20.2	37.0

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

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

58. Aufgabe

(10 Punkte)

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



Zwei kategoriale Variablen darzustellen ist nicht so einfach. Nilufar hatte erst über einen Mittelwert nachgedacht, dann aber die Idee verworfen. In ein Stallexperiment im Oldenburger Land hatte sie sich zum einen die Behandlung Automatische Fütterung [ja/nein] und zum anderen die Messung Fettgehalt erreicht [ja/nein] im Kontext von Hühnern angeschaut. Jetzt möchte ihr Betreuer erstmal die langen Tabellen mit ja/nein in einer explorativen Datenanalyse zusammengefasst bekommen. Sonst geht es bei ihrem Projektbericht nicht weiter.

Automatische Fütterung	Fettgehalt erreicht	er-	Automatische Fütterung	Fettgehalt erreicht	er-
ja	ja		nein	ja	
ja	ja		ja	ja	
nein	ja		nein	ja	
ja	ja		nein	ja	
nein	ja		ja	ja	
nein	nein		nein	nein	
ja	ja		ja	ja	
nein	ja		ja	nein	
ja	ja		ja	ja	
ja	ja		nein	ja	
nein	nein		ja	ja	
ja	ja		nein	ja	
nein	ja		ja	ja	
ja	ja		nein	nein	
ja	ja		nein	ja	
ja	nein		ja	ja	

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

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

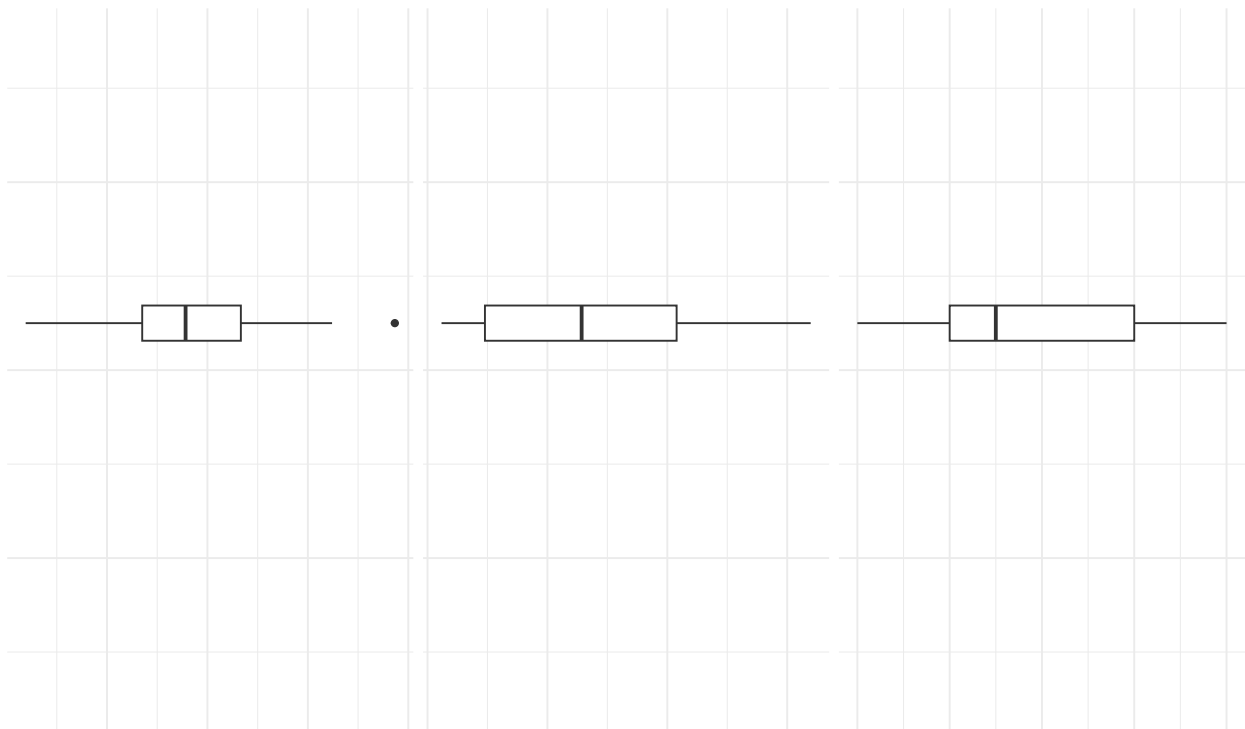
59. Aufgabe

(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 2 Normalverteilungen *in einer Abbildung* mit $\bar{y}_1 \neq \bar{y}_2$ und $s_1 \neq s_2$! **(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)**

$N(1, 0.25)$ und $N(0, 0.25)$



$Pois(2)$ und $Pois(20)$



62. Aufgabe

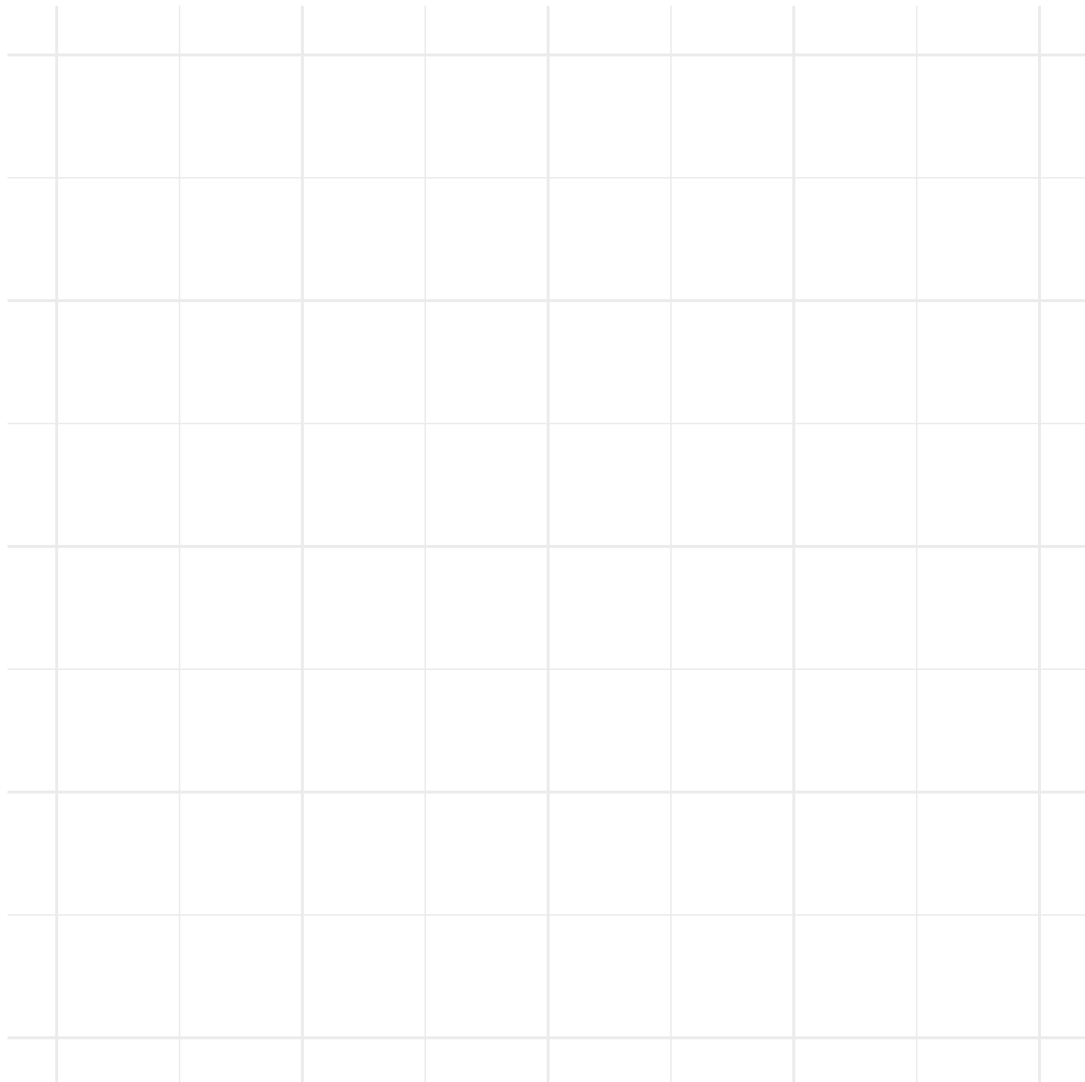
(8 Punkte)

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



Sie haben folgende Zahlenreihe y vorliegen $y = \{10, 17, 9, 9, 11, 15\}$.

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)**

Testentscheidung H_0 beibehalten β -Fehler H_0 wahr

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

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

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

65. Aufgabe

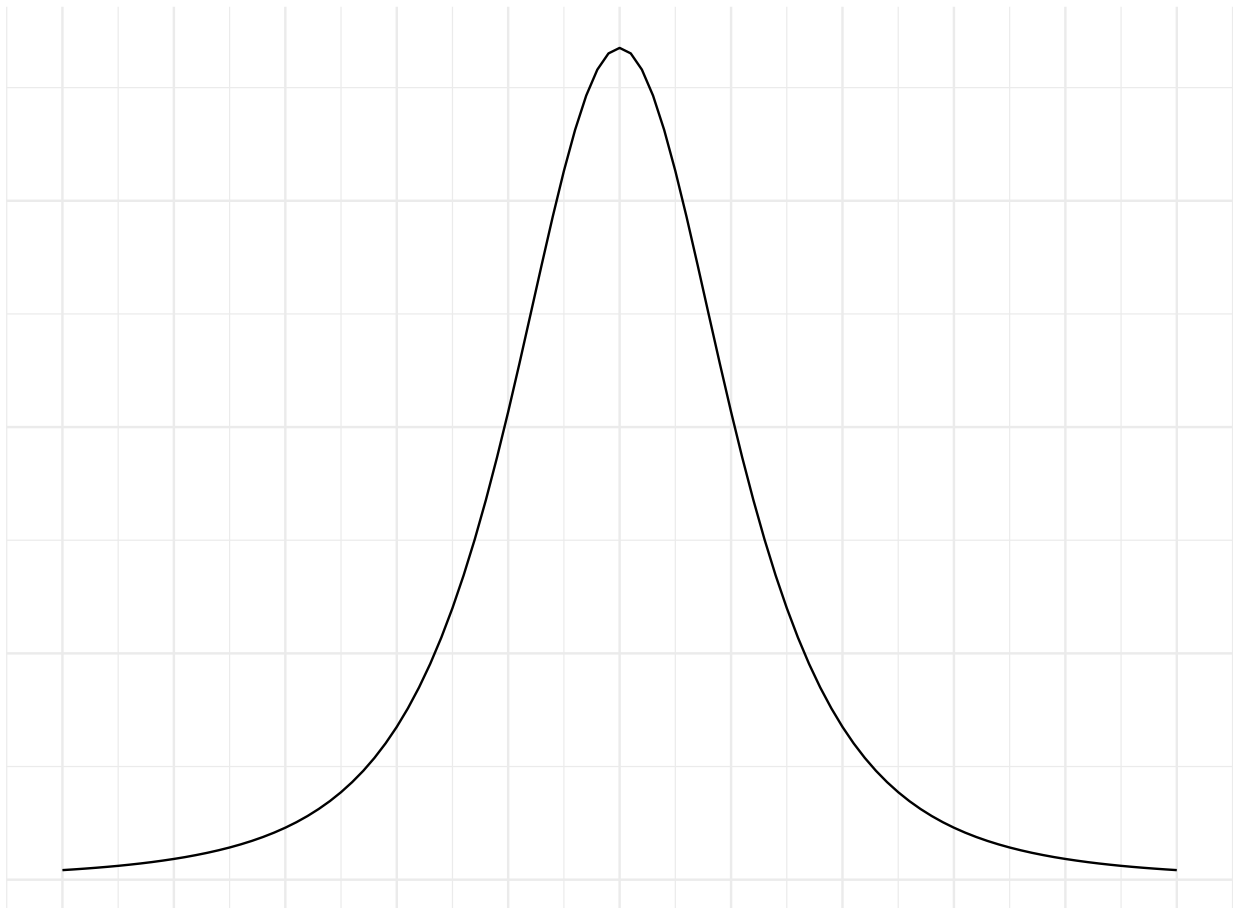
(9 Punkte)

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 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 Befall mit Parasiten zu einer unbehandelten Kontrolle.

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



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!



Mark ist in der Uckermark für einen Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) mit Hühnern. Allein diese Tatsache ist für ihn eine Erzählung wert. Für seinen Projektbericht musste er ein Stallexperiment mit Hühnern durchführen und das sollte laut seinem Betreuer an diesem Ort besonders gut gelingen, da man hier gut neue technische Anlagen und Behandlungen fernab der Bevölkerung testen könne. Zeugen gibt es hier jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn seine Behandlung Flüssignahrung (*ctrl* und *fLOW*) und der Messwert Fettgehalt [%/kg] sollen mit einem t-Test ausgewertet werden. Immerhin weiß er, dass sein Messwert einer Normalverteilung folgt.

treatment	weight
dose	14.9
ctrl	20.7
ctrl	12.2
dose	14.3
ctrl	17.1
dose	15.8

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

1. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
2. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 1.84$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des Welch t-Tests! **(1 Punkt)**
5. Formulieren Sie eine Antwort an Mark über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

69. Aufgabe

(12 Punkte)

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



Der t-Test. Steffen erschauert. Ein mächtiges Werkzeug in den Händen desjenigen, der ein normalverteiltes Outcome (Y) hat. Aber erstmal überhaupt den t-Test rechnen können. Wie sah das Experiment von Steffen überhaupt aus? Steffen hat ein Stallexperiment mit Hühnern durchgeführt. Dabei wurde die Behandlung Elterlinie (*Standard* und *Xray*) an den Hühnern getestet. Gemessen hat Steffen dann als Messwert Fettgehalt [%/kg]. Warum der Versuch in der Uckermark für seiner Hausarbeit stattfinden musste, ist ihm bis heute ein Rätsel. Egal. Gibt es jetzt einen Zusammenhang zwischen der Behandlung und Fettgehalt [%/kg]?

Elterlinie	Fettgehalt
Xray	45.6
Xray	19.8
Xray	39.2
Xray	33.9
Standard	24.3
Standard	25.0
Xray	24.1
Xray	30.7
Standard	35.0
Xray	13.4
Standard	31.7
Xray	43.0
Xray	23.5
Standard	20.8
Standard	34.5
Xray	51.1
Standard	25.4

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie den Effekt des Student t-Tests! **(1 Punkt)**
6. Wenn Sie *einen* Unterschied zwischen den Behandlungsgruppen erwarten würden, wie groß wäre dann die Teststatistik T_D ? Begründen Sie Ihre Antwort! **(2 Punkte)**
7. Formulieren Sie eine Antwort an Steffen über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

70. Aufgabe

(12 Punkte)

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



Nilufar ist im Wendland für einen Versuch mit Hühnern. Allein diese Tatsache ist für sie eine Erzählung wert. Für ihre Abschlussarbeit musste sie ein Stallexperiment mit Hühnern durchführen und das sollte laut ihrem Betreuer an diesem Nichtort besonders gut gelingen. Ablenkung gibt es jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn ihre Behandlung Ernährungszusatz (*ctrl* und *fedX*) und der Messwert Gewichtszuwachs in der 1LW sollen mit einem t-Test ausgewertet werden. Immerhin weiß sie, dass ihr Messwert einer Normalverteilung folgt.

Ernährungszusatz	Gewichtszuwachs
fedX	35.0
ctrl	20.9
fedX	46.0
fedX	49.4
fedX	59.3
fedX	43.7
ctrl	19.1
ctrl	37.8
ctrl	26.7
ctrl	26.0
ctrl	29.3
ctrl	20.8
fedX	33.1
fedX	31.4
ctrl	28.1
ctrl	30.1
fedX	39.4
ctrl	23.7
fedX	42.5

Leider kennt sich Nilufar 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 $T_{\alpha=5\%} = 1.64$ 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 Nilufar 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 Jessica und Yuki gemeinsam in einer Abschlussarbeit 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 des Betreuers muss hier ein gepaarter t-Test gerechnet werden.

ID	treatment	freshmatter
5	ohne	39.6
11	14d	62.5
6	ohne	49.6
8	14d	53.8
9	14d	34.9
11	ohne	45.9
8	ohne	33.8
4	ohne	41.3
10	14d	43.8
1	14d	31.8
4	14d	31.3
10	ohne	44.2
3	ohne	40.7
1	ohne	48.6
9	ohne	45.7
2	14d	34.9
3	14d	48.9
7	14d	39.1
5	14d	49.5
2	ohne	41.2
7	ohne	46.0
6	14d	31.3

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


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

72. Aufgabe

(10 Punkte)

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



‘Mit dem R Paket {emmeans} können wir gleich die Gruppenvergleiche rechnen und uns das *compact letter displac*’ ausgeben lassen!’, verkündet Steffen 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. Jonas und Tina sind bei Steffen um sich in  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) Fettgehalt [%/kg] zu bestimmen.

```
##  
## Two Sample t-test  
##  
## data: Fettgehalt by Elterlinie  
## t = -5.0628, df = 17, p-value = 9.61e-05  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -20.753441 -8.544286  
## sample estimates:  
## mean in group Standard      mean in group Xray  
##                28.46364                43.11250
```

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


1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.11|$ 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. Yuki und Tina sind bei Mark um sich in  helfen zu lassen. Die beiden waren 3 Monate im Emsland um einen Versuch mit Hühnern in einem Stallexperiment durchzuführen. Ziel war es das Outcome (Y) Fettgehalt [%/kg] zu bestimmen.

```
##  
## Two Sample t-test  
##  
## data: Fettgehalt by Elterlinie  
## t = 0.56469, df = 18, p-value = 0.5793  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -2.943089 5.106725  
## sample estimates:  
## mean in group Standard      mean in group Xray  
##          30.00000          28.91818
```

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



‘Programmieren ist wie eine Sprache lernen. Man muss es nur machen, dann wird man mit der Zeit immer besser!’, gibt Nilufar zwinkernd zu Protokoll. Das hilft jetzt Tina und Yuki nur bedingt, da beide jetzt die  Ausgabe interpretieren müssen und nicht vor drei Wochen, wo noch Zeit gewesen wäre. Die beiden hatten im Oldenburger Land einen Versuch mit Hühnern in einem Leistungssteigerungsversuch durchgeführt. Das war schon anstrengend genug! ‘Wir haben Gewichtszuwachs in der 1LW gemessen, vielleicht hilft das ja...’, meint Yuki leicht genervt. Alle starren auf die  Ausgabe des t-Tests.

```
##  
## Two Sample t-test  
##  
## data: Gewichtszuwachs by Flüssignahrung  
## t = 3.8715, df = 20, p-value = 0.0009499  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## 2.679169 8.939013  
## sample estimates:  
## mean in group ctrl mean in group fl0w  
## 40.55455 34.74545
```

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


1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. 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!



Paula und Steffen haben sich dazu entschieden zusammenzuarbeiten. Das sollte alles etwas einfacher machen. Jeder hat zwar ein getrenntes Themenfeld aber den Hauptversuch machen beide gemeinsam. Das hat sich schonmal als gut Idee soweit herausgestellt. In einer Abschlussarbeit sollen beide herausfinden, ob es einen Zusammenhang zwischen Bestandsdichte (*hoch* und *niedrig*) und Gewichtszuwachs in der 1LW gibt. Die Besonderheit ist hierbei, dass die Messungen an der gleichen Beobachtung stattfinden. Beide messen also zweimal an den gleichen Hühnern. Hier muss dann wohl auf einen normalverteilten Messwert (Y) ein gepaarter t-Test gerechnet werden. Leider kennen sich beide nicht sehr gut in  aus.

```
##
## Paired t-test
##
## data: Gewichtszuwachs by Bestandsdichte
## t = 0.069656, df = 8, p-value = 0.9462
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -8.204780 8.715891
## sample estimates:
## mean difference
## 0.2555556
```

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

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



‘Uff... die einfaktorielle ANOVA. Und wir können jetzt anhand der Visualisierung sehen, ob da schon was signifikant ist?’, Steffen hebt die Augenbraue. ‘Ja, können wir. Dafür müssen wir aber erstmal in {ggplot} uns die Daten anschauen. Oder wir zeichnen es flott mit der Hand. Geht auch.’, meint Jonas dazu. Steffen hatte sich in ein Stallexperiment verschiedene Hühnern angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Bestandsdichte (*eng*, *weit* und *kontakt*) und dem Messwert Protein/Fettrate [%/kg] gibt.

Bestandsdichte	Protein/Fettrate
eng	40
eng	42
weit	46
weit	45
kontakt	31
weit	45
kontakt	30
weit	45
eng	39
weit	47
kontakt	29
weit	46
kontakt	30
eng	40
eng	39
eng	40
weit	46
eng	41
kontakt	30

Leider kennen sich Steffen und Jonas 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 Achsen entsprechend! **(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!



'Als erstes bauen wir uns aus unsere Daten die ANOVA Tabelle dann sehen wir schon, ob unser Gruppenvergleich in der ANOVA signifikant ist.', Yuki schaut Paula fragend an und hofft auf eine positive Regung im Gesicht. Wird aber enttäuscht. Paula tut sich auch sehr schwer mit der einfaktoriellen ANOVA. Beide waren im Teutoburgerwald um ein Kreuzungsexperiment mit Hühnern durchzuführen. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) und dem Messwert Fettgehalt [%/kg] gibt.

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

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

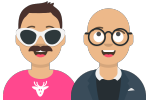
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Lüftungssystem	3	78.23			
error	21				
Total	24	485.76			

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

78. Aufgabe

(12 Punkte)

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



Alex 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. Nun möchte erstmal ihre Betreuung der Arbeit eine ANOVA-Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Alex schaut Yuki sehen erstmal gar nichts. Die beiden waren im Teutoburgerwald um ein Stall-experiment mit Hühnern durchzuführen. Dabei haben Alex und Yuki den Messwert Gewichtszuwachs in der 1LW unter der Behandlung Flüssignahrung (*ctrl*, *superIn* und *floW*) ermittelt.

Leider kennen sich Alex 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	3.23			
Error	21	604.11			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.47$ 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	10	8.40	2.27
superIn	6	8.83	6.37
floW	8	7.88	7.12


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 Paula und hebt die Augenbraue. Paula hatte sich ein Kreuzungsexperiment mit Hühnern angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Elterlinie (*ctrl*, *Standard*, *Yray* und *Xray*) 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)
## Elterlinie  3 4471.3 1490.45   53.009 6.9e-12
## Residuals 29  815.4   28.12
```

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

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



In ein Kreuzungsexperiment wurden Hühnern mit dem Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) sowie der Behandlung Bestandsdichte (*standard* und *kontakt*) untersucht. Es wurde als Messwert Protein/Fettrate [%/kg] bestimmt. Tina ahnte schon, dass es komplexer wird, als sie mit ihrer Hausarbeit angefangen hat. Das es jetzt aber so kompliziert wird, hätte sie jetzt aber auch nicht gedacht. Tina kratzt sich am Kopf. Eventuell muss sie dann doch nochmal Hilfe in der statistischen Beratung holen. Jetzt versucht sie es aber erstmal selber.

Leider kennen sich Tina 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)
Lüftungssystem	3	309.2			
Bestandsdichte	1	60.47			
Lüftungssystem:Bestandsdichte	3	132.61			
Error	18	250.53			

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

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

5. Was bedeutet ein signifikantes Ergebnis in einer zweifaktoriellen ANOVA? (2 Punkte)
6. Was sagt der Term *Lüftungssystem:Bestandsdichte* 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 Yuki wird langsam nervös. Yuki 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. Yuki hatte im Teuteburgerwald einen Leistungssteigerungsversuch mit Hühnern durchgeführt. Es gab dabei zwei Behandlungen. Einmal Genotypen (AA, AB und BB) sowie als zweite Behandlung Flüssignahrung (*ctrl* und *fIOW*). Gemessen wurde der Messwert (Y) Fettgehalt [%/kg]. So kompliziert kann das jetzt doch nicht sein!

```
## Analysis of Variance Table
##
## Response: Fettgehalt
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## Genotypen	2	8.02	4.009	0.2755	0.7623058
## Flüssignahrung	1	258.12	258.116	17.7414	0.0005240
## Genotypen:Flüssignahrung	2	413.02	206.508	14.1941	0.0001994
## Residuals	18	261.88	14.549		

Leider kennen sich Yuki 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 keine und eine schwache 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 fünf 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 ersten 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!



Nilufar und Jonas untersuchen gemeinsam in ihrer Abschlussarbeit den Messwert Protein/Fettrate [%/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 Nilufar und Jonas ihre Ergebnisse nochmal mit einer Studie von Qui et al. (2017) vergleichen und schauen, ob was ähnliches rauskommt. Angeschaut wurde sich als Behandlung Elterlinie (*ctrl*, *Standard*, *TOP*, *Yray*, *SLOW*, und *Xray*). Es ergab sich dann die folgende Tabelle der rohen p -Werte für die Vergleiche zu Qui et al. (2017).

Rohen p -Werte	Adjustierte p -Werte	Nullhypothese ablehnen?
0.080		
0.002		
0.760		
0.230		
0.060		
0.340		

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

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



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

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

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

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

87. Aufgabe

(12 Punkte)

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



'Wir müssen als erstes die Gruppen nach absteigender Effektstärke sortieren!', ruft Paula ganz aufgeregt. 'Das habe ich mir gemerkt.', blinzelt sie Jonas an. Mark hebt die Augenbraue. Als ob das jetzt so viel weiter helfen würde.

Bestandsdichte	Fallzahl (n)	Mittelwert	Standardabweichung
effizient	9	3.28	2.12
standard	8	9.66	1.36
eng	9	10.18	3.17
weit	9	5.11	3.43
kontakt	8	10.59	2.17

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 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 Paula und Jonas! **(1 Punkt)**

88. Aufgabe

(10 Punkte)

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



‘Wir müssen als erstes die Gruppen nach absteigender Effektstärke sortieren!’, ruft Yuki ganz aufgeregt. ‘Das habe ich mir gemerkt.’, blinzelt er Jonas an. Mark hebt die Augenbraue. Als ob das jetzt so viel weiter helfen würde.

	ctrl	superIn	compostIn	fIOW
ctrl	1.0000000	0.1748736	0.0010214	0.1768097
superIn	0.1748736	1.0000000	0.0000298	0.9677402
compostIn	0.0010214	0.0000298	1.0000000	0.0000409
fIOW	0.1768097	0.9677402	0.0000409	1.0000000

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie die statistischen Hypothesen! **(1 Punkt)**
3. Zeichnen Sie die sich anhand 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 Jonas! **(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 Jessica dann doch einen normalverteilten Endpunkt in ihrer Abschlussarbeit nehmen sollen. Jetzt liegen vor ihr die Daten von zwei Variablen als Kategorien oder wie es in  so schön heißt, als Faktoren. Aber immerhin, jetzt hofft er das was bei den Daten rausgekommen ist. Gezählt hat Jessica immerhin $n = 128$ Beobachtungen. Zum einen hat er als Behandlung *Automatische Fütterung* [ja/nein] bestimmt und zum anderen die Variable Schlachtgewicht im Zielbereich [ja/nein] ermittelt. Nun möchte ihr Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen.

	38	19	
	27	44	

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

1. 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!



Gegeben sind folgende Randsummen in einer 2x2 Kreuztabelle aus einem Experiment mit $n = 133$ Sauen. In dem Experiment wurde gemessen, ob eine Sau nach einer Behandlung mit einem Medikament (ja/nein) mehr als 30 Ferkel pro Jahr bekommen konnte (ja/nein).

			68
			65
	76	57	133

Leider kennt sich Paula mit der Erstellung einer explorativen Datenanalyse 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 *ein* signifikanter Effekt zu erwarten wäre! **(2 Punkte)**
4. Begründen Sie Ihr Vorgehen an der Formel des Chi-Quadrat-Tests. Erklären Sie Ihr Vorgehen an einem Beispiel! **(2 Punkte)**
5. Visualisieren Sie den Zusammenhang zwischen den beiden kategorialen Variablen! **(2 Punkte)**
6. Was ist die Mindestanzahl an Beobachtungen je Zelle? Wenn in einer der Zellen weniger Beobachtungen auftreten, welchen Test können Sie anstatt des Standard Chi-Quadrat-Tests anwenden? **(2 Punkte)**

91. Aufgabe

(10 Punkte)


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



foo

Nach einem Experiment erhalten Sie folgende 2x2 Kreuztabelle aus Ihren erhobenen Daten.

```
##               Außenkontaktkontakt
## Fettgehalt erreicht ja nein
##               ja    17    5
##               nein   4    12
```

Aus der 2x2 Kreuztabelle erhalten Sie folgende  Ausgabe der Funktion `fisher.test()`.

```
##
## Fisher's Exact Test for Count Data
##
## data: Fettgehalt erreicht
## p-value = 0.002568
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.85845 61.14631
## sample estimates:
## odds ratio
##  9.451509
```

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

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

92. Aufgabe

(11 Punkte)



Die Prävalenz von Klauenseuche bei Wollschweinen wird mit 3% angenommen. In 80% 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 1000 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)**



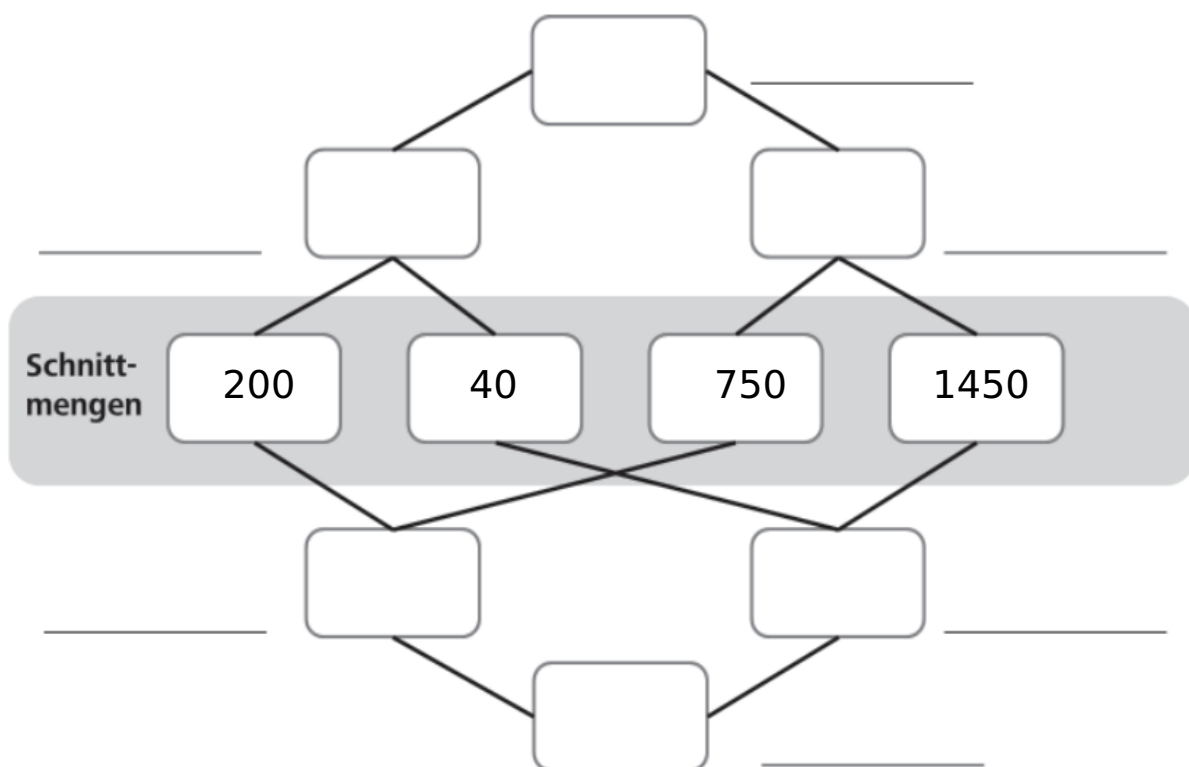
93. Aufgabe

(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.

Nicht parametrische Tests

94. Aufgabe

(12 Punkte)

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



Die Anzahl an Nematoden wurde vor und nach einer Behandlung mit einem bioaktiven Dünger gezählt. Es ergibt sich folgende Datentabelle.

Vorher	Nachher	Differenz	Vorzeichen	Rang	Positiv Rang	Negativ Rang
11	10					
13	11					
8	12					
10	11					
11	9					
8	13					
8	11					
9	13					
12	15					
10	13					
6	11					
10	10					
9	10					
7	11					
9	11					

- Ergänzen Sie die obige Tabelle mit den notwendigen Informationen, die Sie benötigen um einen Wilcoxon-Vorzeichen-Rang-Test zu rechnen! **(4 Punkte)**
- Bestimmen Sie die Teststatistik W_D mit $W_D = \min(T_-; T_+)$ und berechnen Sie den erwarteten Wert $\mu_W = \frac{n_{10} \cdot (n_{10} + 1)}{4}$! **(2 Punkte)**
- Berechnen Sie anschließend den z-Wert mit $z = \frac{W_D - \mu_W}{15.93}$! **(2 Punkte)**
- Liegt mit einer Signifikanzschwelle von $z_{\alpha=5\%} = 1.96$ ein Unterschied zwischen den beiden Zeitpunkten vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
- Berechnen Sie die Effektstärke mit $r = |\frac{z}{\sqrt{n}}|$ und interpretieren Sie die Effektstärke! **(2 Punkte)**

95. Aufgabe

(8 Punkte)

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



Nach einer Behandlung mit RootsGoneX wurde die mittlere Anzahl an Wurzeln an der invasiven Lupine (*Lupinus polyphyllus*) gezählt. Es ergab sich folgender Datensatz an mittleren Wurzelanzahl.

Treatment	Count
Kontrolle	5.6
RootsGoneX	12.5
Kontrolle	5.4
Kontrolle	7.9
Kontrolle	6.0
RootsGoneX	11.4
RootsGoneX	12.2
Kontrolle	4.6
RootsGoneX	10.8

Rechnen Sie einen Mann-Whitney-U-Test auf den obigen Daten.

1. Bestimmen Sie hierfür U_D mit $U_D = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$! (4 Punkte)

2. Geben Sie eine Aussage über die Signifikanz von U_D durch $z = \frac{U_D - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$ und dem kritischen

Wert von $z_{\alpha=5\%} = 1.96$. Begründen Sie Ihre Antwort! (2 Punkte)

3. Berechnen Sie die Effektstärke mit $r = |\frac{z}{\sqrt{n}}|$ und interpretieren Sie die Effektstärke! (2 Punkte)

96. Aufgabe

(10 Punkte)

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



Die Anzahl an Blüten der Vanillepflanze pro Box wurde nach der Gabe von zusätzlicher Phosphatlösung (Kontrolle, Dosis 20 und Dosis 40) bestimmt. Es ergeben sich folgende nach der Anzahl der Blüten geordnete Daten.

Treatment	Count	Rang Kontrolle	Rang Dosis 20	Rang Dosis 40
Dosis 40	8.8			
Dosis 40	12.4			
Kontrolle	-0.5			
Dosis 20	11.1			
Dosis 40	6.4			
Dosis 20	10.5			
Kontrolle	1.0			
Dosis 40	8.9			
Dosis 20	16.1			
Dosis 20	16.1			
Dosis 20	10.1			
Kontrolle	1.1			
Dosis 40	4.9			
Kontrolle	-0.4			
Dosis 40	9.5			
Kontrolle	2.0			

Rechnen Sie einen Kruskal-Wallis-Test auf den obigen Daten.

- Bestimmen Sie hierfür H_D mit $H_D = \frac{12}{n(n+1)} \left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} \right) - 3(n+1)$! **(6 Punkte)**
- Geben Sie eine Aussage über die Signifikanz von H_D durch den kritischen Wert von $H_{\alpha=5\%} = 5.99$! **(1 Punkt)**
- Wie lautet die statistische Nullhypothese die Sie mit dem Kruskal-Wallis-Test überprüfen? **(1 Punkt)**
- Was sagt ein signifikantes Ergebnis des Kruskal-Wallis-Test in Bezug auf die einzelnen Gruppenvergleiche aus? **(1 Punkt)**
- Nennen Sie das statistische Verfahren, welches Sie als Posthoc Test nach einem signifikanten Kruskal-Wallis-Test durchführen würden! **(1 Punkt)**

Teil VIII.

Lineare Regression & Korrelation

97. Aufgabe

(9 Punkte)

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



foo

durchschnittlicher Tagestemperatur [C/d]	Gewichtszuwachs in der 1LW
16.3	23.8
20.5	30.8
19.6	29.9
21.4	29.3
19.1	26.8
18.1	25.5
20.0	28.0
20.5	33.0
20.5	28.2
17.4	22.9

Leider kennen sich Alex und Jonas mit der Erstellung einer Datenanalyse für kontinuierliche Daten überhaupt nicht aus. Deshalb brauchen bei der Erstellung Ihre Hilfe!


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

98. Aufgabe

(10 Punkte)

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



In einem Stallexperiment mit $n = 40$ Ferkeln wurde der Gewichtszuwachs in kg unter ansteigender Lichteinstrahlung in nm gemessen. Sie erhalten den  Output einer simplen Gaussian linearen Regression sieben Wochen nach der ersten Messung.

term	estimate	std.error	t statistic	p-value
(Intercept)	-0.48	1.51		
light	0.34	0.15		

1. Zeichnen Sie die Gerade aus der obigen Tabelle in ein Koordinatenkreuz! **(1 Punkt)**
2. Beschriften Sie die Abbildung und die Gerade mit den statistischen Kenngrößen! **(2 Punkte)**
3. Formulieren Sie die Regressionsgleichung! **(2 Punkte)**
4. Berechnen Sie die t Statistik für *(Intercept)* und *light*! **(2 Punkte)**
5. Schätzen Sie den p-Wert für *(Intercept)* und *light* mit $T_{\alpha=5\%} = 1.96$ ab. Was sagt Ihnen der p-Wert aus? Begründen Sie Ihre Antwort! **(3 Punkte)**

99. Aufgabe

(10 Punkte)



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

```
##
## Call:
## lm(formula = weight ~ water, data = data_tbl)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.4444 -1.1944  0.2778  0.8889  3.5556
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  16.0000     0.7981  20.046 1.36e-10
## waterB       -3.5556     0.9955  -3.572 0.00384
##
## Residual standard error: 1.785 on 12 degrees of freedom
## Multiple R-squared:  0.5153, Adjusted R-squared:  0.4749
## F-statistic: 12.76 on 1 and 12 DF, p-value: 0.00384
```

1. Ist die Annahme der Normalverteilung an das Outcome *water* erfüllt? Begründen Sie die Antwort! **(2 Punkte)**
2. Wie groß ist der Effekt der Wassergabe? Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Liegt ein signifikanter Effekt vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Erklären Sie *kurz* den Begriff R-squared! Was sagt Ihnen der Wert 0.5153 aus? **(2 Punkte)**
5. Schreiben Sie das Ergebnis der R Ausgabe in zwei Sätzen auf, der die Information zum Effekt und der Signifikanz enthält! **(2 Punkte)**

100. Aufgabe

(9 Punkte)

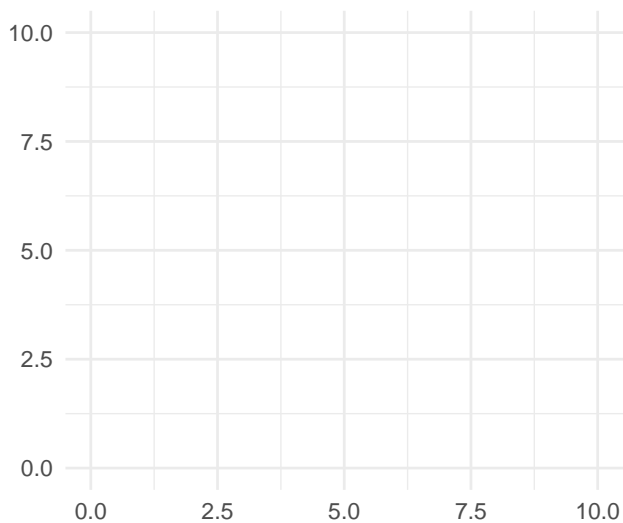


Im folgenden sehen Sie drei leere Scatterplots. Füllen Sie diese Scatterplots nach folgenden Anweisungen.

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)**

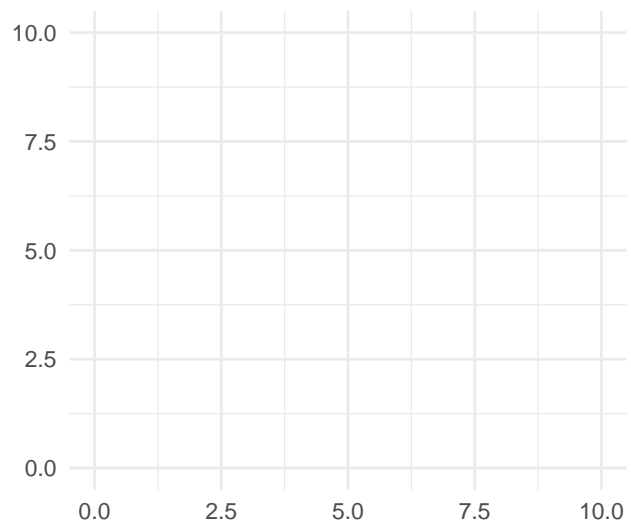
Pearsons $\rho = -0.5$

$R^2 = 0.5$



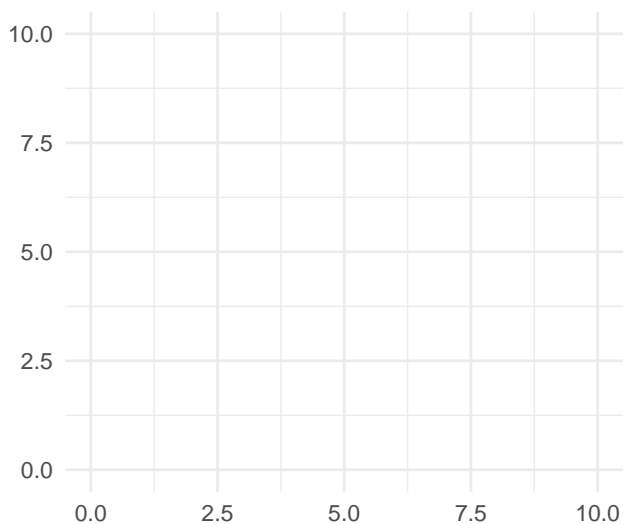
Pearsons $\rho = -0.75$

$R^2 = 0.75$



Pearsons $\rho = -0.25$

$R^2 = 1$



101. Aufgabe

(9 Punkte)

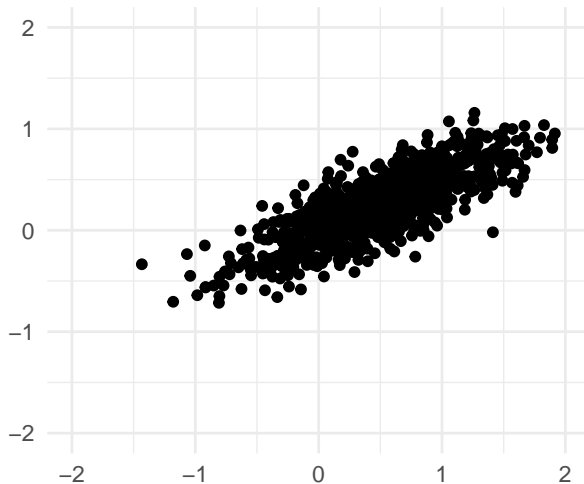


Im folgenden sehen Sie vier Scatterplots. Ergänzen Sie die Überschriften der jeweiligen Scatterplots.

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)**

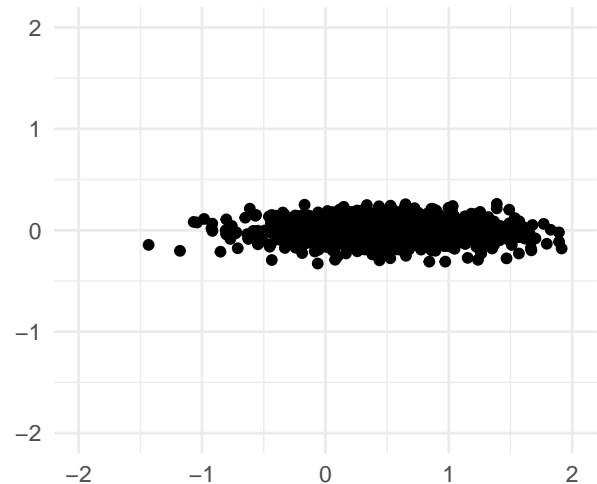
Pearsons $\rho =$

$R^2 =$



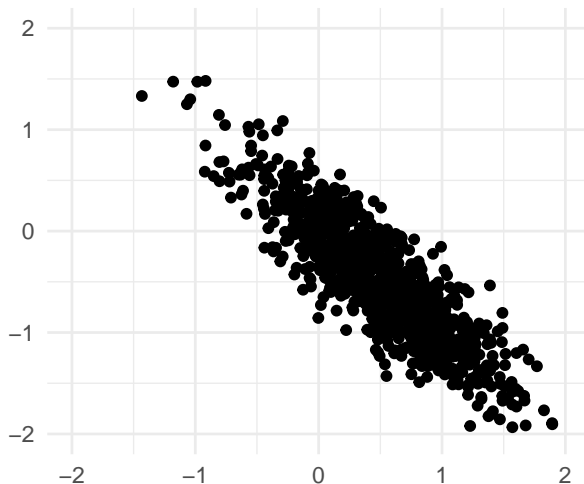
Pearsons $\rho =$

$R^2 =$



Pearsons $\rho =$

$R^2 =$



Pearsons $\rho =$


$R^2 =$



102. Aufgabe

(9 Punkte)



Sie erhalten folgende  Ausgabe der Funktion `cor.test()`.

```
##  
## Kendall's correlation  
##  
## data: height and food  
## t = -1.3898, df = 8, p-value = 0.202  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.8379602 0.2611187  
## sample estimates:  
## cor  
## -0.4410134
```

1. Formulieren Sie das statistische Hypothesenpaar! **(2 Punkt)**
2. Nennen Sie die zwei Eigenschaften des Korrelationskoeffizienten! Erklären Sie *eine* der Eigenschaften an einem Beispiel! **(2 Punkte)**
3. Interpretieren Sie den Korrelationskoeffizienten hinsichtlich des Effekts und der Signifikanz! Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Visualisieren Sie das 95% Konfidenzintervall! Beschriften Sie die Abbildung! **(2 Punkte)**
5. Sind die Variablen `height` and `food` normalverteilt? Begründen Sie Ihre Antwort! **(1 Punkt)**

103. 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	31.0	14.8	31.6	
2	16.9	3.6	13.9	
3	23.3	9.8	23.6	
4	27.6	12.2	27.4	
5	17.8	8.1	21.0	
6	31.8	14.3	30.8	
7	36.8	16.8	34.8	
8	27.5	13.4	29.3	
9	17.6	7.6	20.2	
10	29.4	13.4	29.4	
11	25.1	9.1	22.6	

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)**

104. Aufgabe

(12 Punkte)





1. Zeichnen Sie in die drei untenstehenden, leeren Abbildungen die Zeile des Regressionskreuzes der Poissonverteilung. 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)**

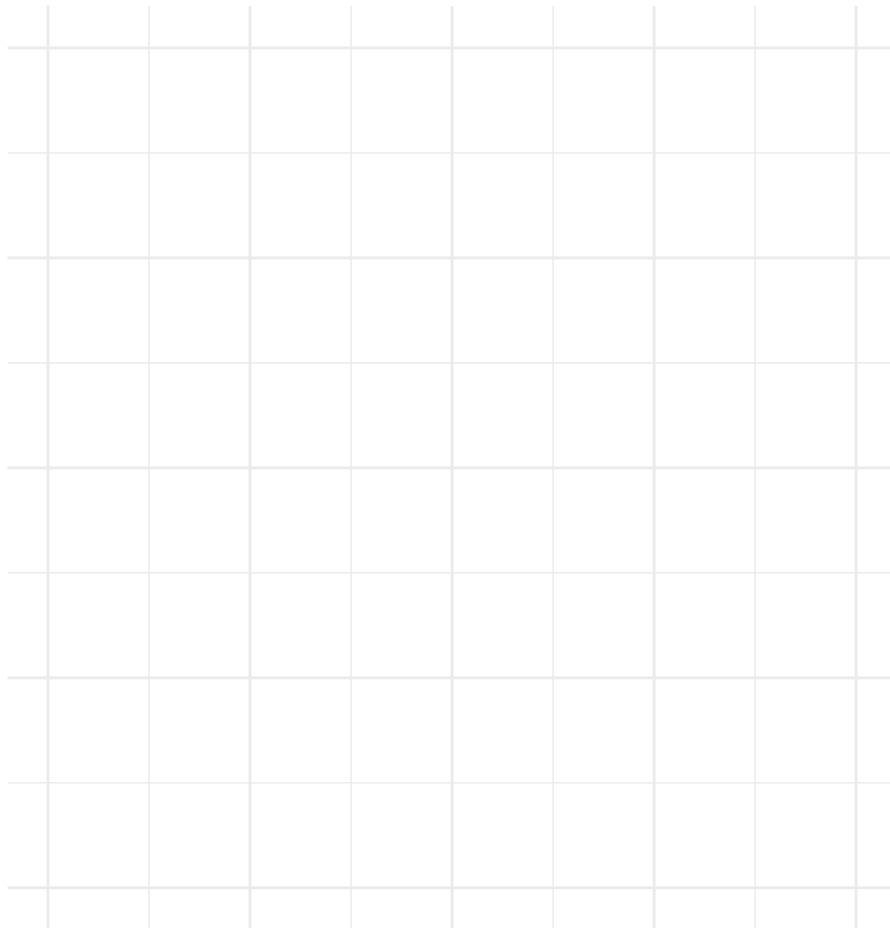
105. Aufgabe

(9 Punkte)



Ein Feldexperiment wurde mit $n = 200$ Pflanzen durchgeführt. Folgende Einflussvariablen (x) wurden erhoben: block, fertilizier und height. 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 *Poisson* 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 *schwache* 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 IX.

Experimentelles Design

106. Aufgabe

(10 Punkte)



Nach einem erfolgreichen Pilotversuch zur Wirksamkeit von Bodenbehandlungen bei Brokoli in einem Gewächshausexperiment wollen Sie nun den Versuch eine Nummer größer anlegen. Dafür entscheiden Sie sich für ein faktorielles Versuchsdesign. In Ihrem Hauptversuch stellt die Wirksamkeit von Bodenbehandlungen den ersten Faktor mit insgesamt 3 Leveln dar. Der zweite Faktor mit dem Hitzestress beinhaltet 3 Level.

Im ersten Schritt überlegen Sie ein einfaches experimentelles Design zu probieren. Sie entscheiden sich für das *Latin square design*.

1. Skizzieren Sie das *Latin square design* für Ihren Versuch! **(4 Punkte)**
2. Skizzieren Sie eine Datentabelle für den Versuch mit drei Wiederholungen! **(2 Punkte)**
3. Skizzieren Sie eine Abbildung mit Boxplots und einem angenommenen normalverteilten Outcome! **(4 Punkte)**

Teil X.

Mathematik

107. Aufgabe

(8 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 führte¹.

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

Gehen Sie von einem radialen Wachstum der 1000-jährigen Eiche von 0.9mm pro Jahr aus. Es ist bekannt, dass die Eiche im Jahr 2022 einen Umfang von 10.5m in Bruthöhe hatte.

1. Wie groß war der Durchmesser in m der Eiche im Jahr 1840 als Herodot in der Eiche versteckt werden sollte? **(2 Punkte)**
2. Skizzieren Sie in einer Abbildung einen linearen Zusammenhang und einen exponentiellen Zusammenhang für das Wachstum der 1000-jährigen Eiche. Erklären Sie die Auswirkungen der Entscheidung für linear oder exponentiell auf Ihre Berechnungen! **(2 Punkte)**

Herodot hatte eine Schulterhöhe von 195cm , eine Breite von 75cm sowie eine Länge von 220cm .

3. Berechnen Sie das effektive Volumen von Herodot in m^3 , welches Herodot in der 1000-jährigen Eiche einnehmen würde! **(1 Punkt)**

Es wurde berichtet, dass sich Herodot in der 1000-jährigen Eiche *mühsam* 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 20cm bricht die Eiche zusammen. Beantworten Sie die Forschungsfrage! Begründen Sie Ihre Antwort! **(1 Punkt)**

¹Die Quelle der Inspiration für die Aufgabe war eine Fahrt an die Ostsee und folgender Artikel: [Entdecke das erste Nationale Naturmonument Deutschlands - Ivenacker Eichen und Hutewald](#)

108. Aufgabe

(10 Punkte)

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



Von Töpfen auf Tischen In einem Experiment wollen Sie die Wuchshöhe von 120 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 8cm und eine Höhe von 10cm. 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 *Tischfläche* in m^2 gegeben der Anzahl an Pflanztöpfen inklusive Randpflanzen benötigen Sie im Gewächshaus am Anfang der Keimungsphase? **(3 Punkte)**
4. Berechnen Sie die benötigte Menge an Torf in Liter l , die Sie für das Befüllen der Pflanztöpfe benötigen! Gehen Sie von *einem Zylinder* für die Pflanztöpfe aus! **(3 Punkte)**
5. Berechnen Sie die Kosten in EUR für Ihre Torfbestellung! **(1 Punkt)**

109. Aufgabe

(10 Punkte)

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



Solar- & Biogasanlagen Um die Energiekosten Ihres Betriebes zu senken, wollen Sie eine Solaranlage auf den Rinderstall montieren lassen. Sie messen Ihren Stall und finden folgende Maße wieder. Die vordere Seite des Rinderstall hat eine Höhe h_v von $7m$. Die hintere Seite des Rinderstall hat eine Höhe h_b von $11m$. Der Rinderstall hat eine Tiefe t von $16m$ und eine Breite b von $60m$.

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

Ebenfalls planen Sie eine neue Biogasanlage für Ihren Betrieb. Der neue Methantank hat einen Radius r von $1.2m$. 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 10% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $235kg/m^3$. Bei $-100^\circ C$ hat Methan eine Dichte von $300kg/m^3$. Sie betreiben Ihre Anlage bei $-88^\circ C$.

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

110. Aufgabe

(8 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 einer ältern Dame mit aufgespannten Regenschirm. “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 Chile². Tja, die Deutschen und Südamerika.

Zuerst werden Ihre Fähigkeiten getestet, der Mathematik folgen zu können. Oder berechnen Sie gerade den Einkauf von einer ältern Dame mit aufgespannten Regenschirm?

1. Wenn 7 Blaubeerschalen 11.13 Euro kosten, wie viel kosten 8 Schalen? **(1 Punkt)**
2. Wenn Sie die 8 Blaubeerschalen gekauft haben, wie viele Aligatorbirnen zu je 2.89 EUR können Sie sich dann noch für 50 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 - 125g.
- Ein Kilo Salat benötigt 100l Wasser. Ein Salatkopf wiegt 300 - 490g.
- Ein Kilo Avocado benötigt 1050l Wasser. Eine Avocado wiegt 150 - 380g.
- Ein Kilo Blaubeeren benötigt 850l Wasser. Eine Blaubeere wiegt 3 - 3.5g.

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! **(2 Punkte)**

Chile exportiert im großem Ausmaß Blaubeeren und Avocados. In dem Exportjahr 2024 blieben die Erträge von Blaubeeren mit 7.9×10^4 t in dem prognostizierten Rahmen. Die Menge steigerte sich um 5.1%. Die Exporte für Avocados stiegen in dem gleichen Zeitraum um 23.8% auf 2.1×10^5 t.

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

Chile ist eines der wenigen Länder der Welt, die ihr Wasser komplett privatisiert haben. Derzeit sind nur ein Prozent des Wassers des Landes für den häuslichen Verbrauch vorgesehen. In den Dörfern der Anbauregionen versorgen Tankwagen die Bevölkerung jede Woche mit Wasser, es gibt etwa 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 einer ältern Dame mit aufgespannten Regenschirm 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)**

²Die Quelle der Inspiration für die Aufgabe waren folgende Reportagen: “Bis zum letzten Tropfen” in AMNESTY – Magazin der Menschenrechte vom August 2021 und “Wasserknappheit in Chile: Eine Folge der Privatisierung?” in Die Welternährung dem Fachjournal der Welthungerhilfe vom April 2022.

111. Aufgabe

(12 Punkte)

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



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

Die Dampfnudelerde “Was für einen Unsinn!“, rufen Sie. Jetzt haben Sie kostbaren Schlaf prokrastiniert um einem Ernährungswissenschaftler auf YouTube über die Erde als Dampfnudel zu lauschen. Irgendwie passt es dann doch mit der Analogie. Übermüdet müssen Sie darüber nachdenken, warum vor 66 Millionen Jahren die Dinosaurier - so groß sie auch waren - nicht von der Schwerkraft zu Boden gerissen wurden. Hat der Dampfplauderer etwa recht und war die Schwerkraft vor Millionen von Jahren eine andere? Sind deshalb alle Lebewesen auf der Erde *heutzutage* so viel kleiner, weil die Schwerkraft größer ist als damals? War die Erde kleiner und hatte weniger Masse? Oder ist es nur ein Rechenfehler wie bei der Theorie der Hohlerde von Edmond Halley aus dem 17.-18. Jahrhundert? Müde reiben Sie sich die Augen. So wird es nichts mehr mit dem Schlafen, dann können Sie auch mal etwas rechnen³.

Betrachten wir die Schwerkraft oder Gewichtskraft, die auf Lebewesen damals und heute gewirkt haben soll. Nehmen Sie für die Fallbeschleunigung g der Erde *heutzutage* einen Wert von 9.81 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.235 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.51 g/cm^3 . Das Gewicht von einem heute lebenden afrikanischen Elefanten liegt bei 5t bis 7t und das Gewicht von einem Brachiosaurus bei bis zu 30t.

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

Die Distanz zwischen Sonne und Erde entspricht 1.05 astronomische Einheiten (AE). Die Einheit 1 AE wird mit $1.55 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 85% Wasserstoffkernen mit einer molaren Masse von 1.08 g/mol , 8% Heliumkernen mit 4.11 g/mol sowie 7% weiteren Atomkernen mit 89.32 g/mol . Die Teilchendichte bei Eintritt in die Erdatmosphäre liegt zwischen 0.4 bis 100 Teilchen cm^{-3} pro Sekunde mit einer mittleren Teilchendichte von 7 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)**

³Die Quelle der Inspiration für die Aufgabe war folgender Artikel: "Skeptische Anmerkungen — Die Erde als Dampfnudel" in Der Humanistische Pressedienst

112. Aufgabe

(10 Punkte)

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



Entschuldigung, ist das Ihre Feder in meinem Auge? So hört man häufiger höfliche 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 nicht⁴. 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) foraging incl. scratching, (2) wingflapping, (3) wing/leg stretching und (4) dustbathing.

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)
foraging incl. scratching	39cm; 25cm; 0.6%	36cm; 23cm; 0.6%	30cm; 37cm; 0.9%
wingflapping	27cm; 32cm; 1.2%	46cm; 32cm; 8.1%	41cm; 25cm; 7.6%
wing/leg stretching	34cm; 26cm; 3.2%	32cm; 23cm; 3.2%	30cm; 25cm; 3.2%
dustbathing	28cm; 32cm; 0.8%	27cm; 24cm; 0.8%	29cm; 26cm; 1.2%

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)**

⁴Die Quelle der Inspiration für die Aufgabe war der folgende wissenschaftliche Artikel: [EFSA Panel on Animal Health and Welfare, et al. \(2023\) Welfare of broilers on farm. EFSA Journal 21.2.](#)

113. Aufgabe

(8 Punkte)

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



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

»Ohne ausreichende Zufuhr von Vitamin C stellen sich nach 45 Tagen die ersten Symptome ein; die ersten Toten sind nach 70 Tagen zu beklagen; nach 100 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 218 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. Schätzen Sie die Überlebenswahrscheinlichkeit nach 90 Tagen aus Ihrer Abbildung ab! **(1 Punkt)**

Der Chronist an Bord der Trinidad, Antonio Pigafetta, schrieb in seinem Bericht „[...] Um nicht Hungers zu sterben, aßen wir das Leder, mit dem die große Rahe zum Schutz der Taue umwunden war.“ Insbesondere die Mannschaft der Concepción erlitt große Verluste durch die Skorbut bei der Überquerung des Pazifiks, da durch Erkundungsfahrten weniger Zeit blieb, um wilden Sellerie aufzunehmen. Wilder Sellerie enthält $4000\mu\text{g}/10\text{mg}$ 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 18 Tage über den Pazifik! **(3 Punkte)**
4. Skizzieren Sie die Überlebenszeitkurve für die Concepción im Vergleich zu der Überlebenszeitkurve der Trinidad! Beschriften Sie die Achsen! **(2 Punkte)**

114. Aufgabe

(10 Punkte)

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



Event Horizon – Am Rande des Universums Die Sonne hat eine aktuelle, angenommene Masse von $2 \times 10^{27} \text{ kg}$. Wenn die Sonne nun am Ende ihrer Lebenszeit zu einem schwarzen Loch mit dem Radius von 2500m kollabiert, wird die Sonne 30% 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} gegeben⁵.

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

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 $2.9 \times 10^8 \text{ m/s}$ aus. Kann das Lichtteilchen der Gravitation des schwarzen Lochs entkommen? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Stellen Sie den Zusammenhang zwischen dem sich verringernden Radius r des schwarzen Lochs bei gleichbleibender Masse m_s und der notwendigen Fluchtgeschwindigkeit v_f in einer Abbildung dar! **(2 Punkte)**
6. Eine Kirchenglocke und eine Feder 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)**

⁵Die Quelle der Inspiration für die Aufgabe war ein Montagnachtfilm: [Event Horizon – Am Rande des Universums](#)

115. Aufgabe

(10 Punkte)

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



Das Fermi Paradoxon Der Kernphysiker Enrico Fermi diskutierte 1950 auf dem Weg zum Mittagessen im Los Alamos National Laboratory mit seinen Kollegen angebliche UFO-Sichtungen und fragte schließlich: „Where is everybody?“. Warum seien weder Raumschiffe anderer Weltraumbewohner noch andere Spuren extraterrestrischer Technik zu beobachten? Wie lange würde eine außerirdische Zivilisation benötigen um die gesamte Milchstraße zu besuchen, wenn das maximale Reisetempo die Geschwindigkeit der Voyager 1 Sonde wäre?⁶

Wir treffen folgende Annahmen. Eine außerirdische Zivilisation schickt vier Voyager 1 ähnliche Sonden mit der Geschwindigkeit von $5.9256 \times 10^4 \text{ km/h}$ los um sich auf den erreichten Planeten selbst zu replizieren. Nach 250 Jahren ist die Replikation abgeschlossen und wiederum vier Sonden werden ausgesendet. Gehen Sie von 4.24 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 1.5×10^{11} Sterne in der Milchstraße. Nehmen Sie eine Lichtgeschwindigkeit von $2.7 \times 10^8 \text{ m/s}$ an.

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

⁶Die Quelle der Inspiration für die Aufgabe war folgender Wikipediaeintrag: [Fermi-Paradoxon](#)

116. Aufgabe

(10 Punkte)

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



Pyramiden bauen Es stehen die mecklemburgischen Pyramidentage an und Sie sind auf abenteuerlichen Wegen für den Bau der Pyramiden zuständig. Zu allem Überfluss handelt es sich auch noch eine *Reenactment* Veranstaltung. Thema der diesjährigen Pyramidentage sind die Pyramiden von Meroe, die den Königen und Königinnen des historischen Reiches von Kusch in Nubien, dem heutigen Sudan, als Grabstätten dienten. Die Pyramiden in Meroe fallen durch ihren steilen Winkel von 74 Grad im Vergleich zu den ägyptischen Pyramiden mit 60 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 32 Königsellen. Eine Königselle misst 52.4cm.

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

1. Bei der Königspyramide von Meroe soll eine Seitenlänge der Grundfläche 32 Königsellen lang sein. Welche Höhe der Königspyramide in m ergibt sich? **(1 Punkt)**
2. Die Außenflächen der Pyramide soll begrünt werden. Für die Bepflanzung muss eine 5cm dicke Torfschicht auf die Pyramide aufgebracht werden. Berechnen Sie die ungefähre Menge an benötigten Torf in m^3 ! **(2 Punkte)**

Wie in jedem guten *Reenactment* gibt es viel Oberschicht, aber nur 3 Sklaven, die Ihnen bei dem Befüllen der Pyramide mit Schutt zu Seite stehen. Leider haben Ihre Sklaven zu allem Überfluss auch noch chronische Schulterschmerzen entwickelt, als sie von der anstehenden Aufgabe erfahren haben. Gehen Sie daher von einer Effizienz der Sklaven von 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 10° ! **(2 Punkte)**
5. Wie weit reicht Ihre Rampe vom Fuß der Pyramide in die mecklemburgische Landschaft? **(2 Punkte)**

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

6. Die Grundfläche der Pyramide ändert sich nicht. Berechnen Sie die Änderung der Höhe in Königsellen, wenn sich der Anstellwinkel der Pyramide um 7° ändert! **(2 Punkte)**

117. Aufgabe

(12 Punkte)

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



Geocaching – Von Satelliten und Plastikdosen Es ist Wochenende und das Wetter ist *sweet*. Sie schwingen sich auf Ihr Cachermobil um mit 17km/h, geleitet von modernster Satellitentechnologie und einem Supercompter aus dem Jahr 2000 in Ihren Händen, Plastikdosen in der Natur und an sehenswerten Orten zu finden. Sie wollen diesmal endlich die abwärts Schwierigkeitschallenge durchführen. Die Reihenfolge der Caches nach Schwierigkeitswertung gibt daher die von Ihnen abzufahrenden Orte vor. Die Terrain- und Schwierigkeitswertungen laufen von 1 (leichteste Wertung) bis 5 (schwierigste Wertung) in 0.5 Schritten. Folgende Informationen zu den Orten und den entsprechenden Caches stehen Ihnen für Ihre Planung der Route zu Verfügung⁷.

Ort	Cache	Wertung (S T G)
A	GCU6KZL	1.5 2.0 Normal
B	GCGCMEM	2.0 2.5 Mikro
C	GCMSD5O	4.0 1.5 Mikro
D	GCUGLQV	5.0 5.0 Klein
E	GCUZDU1	4.5 3.0 Klein

Im Weiteren sind Ihnen folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \overrightarrow{AB} ist 4km. Im Weiteren ist Ihnen der Entfernungsvektor \overrightarrow{CB} mit 6.5km 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° nördlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 35° östlich. Vom Ort B betrachtet, bilden die Orte C und D einen rechten Winkel am Ort B. Der Ort B liegt auf gerader Linie zwischen den Orten C und E. Somit liegt der Ort E nördlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort D Ihre Cachertour.

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

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

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

4. An der höchsten Schwierigkeit müssen Sie angeln. Ihre Angel ist ausgefahren 7m lang. Erreichen Sie einen Cache in der Höhe von 8.6m? Berechnen Sie dazu Ihre maximale mögliche Angelhöhe! Welche Annahmen mussten Sie treffen um die Aufgabe zu lösen? **(2 Punkte)**

⁷Die Quelle der Inspiration für die Aufgabe war folgende Tätigkeit: [Geocaching – Mach mit bei der weltweit größten Schatzsuche](#).

118. Aufgabe

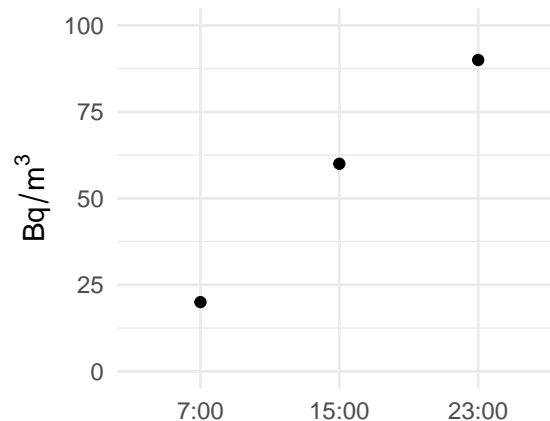
(10 Punkte)

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



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

Die atmende Wand und Brot aus Luft Als Kellerkind vom Dorf wollen das Ausmaß der Radonbelastung in ihrem Kellerzimmer bestimmen und lüften daher nicht. Spart dann auch Energie und lüften wird sowieso überschätzt. Während einer Messperiode von 7:00 Uhr bis 23:00 bestimmen Sie dreimal automatisch die Radonbelastung in Ihrem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung⁸.



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

Radon zerfällt mit einer Halbwertszeit von 1.8d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 140d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

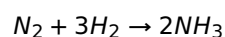
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $300Bq/m^3$ auf unter $100Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	77.1	28.1	
Sauerstoff	21.3	16.2	
Kohlenstoffdioxid	0.045	11.8	

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

Während Sie Ihr etwas pappiges Toastbrot mampfen kommt Ihnen die Dokumentation über Brot aus Luft in den Sinn. Sie denken darüber ein wenig nach. Für die Umwandlung von Stickstoff N_2 mit Wasserstoff H_2 zu Ammoniak NH_3 gilt folgende Reaktionsgleichung⁹:



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)**

⁸Die Quelle der Inspiration für die Aufgabe war folgender Artikel: [Atmende Wand](#)

⁹Die Quelle der Inspiration für die Aufgabe war folgender Artikel: [Haber-Bosch-Verfahren – Brot aus Luft](#)

119. Aufgabe

(10 Punkte)

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



Armee der Finsternis Ihr Studentenjob war nach Ladenschluss bei Aldi die Regale einzuräumen. Dabei ist Ihnen in der Auslage der Sonderangebote das Necronomicon¹⁰ 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 974 n. Chr. für den neuen Lehnsherren Henry dem Roten. Sie bauen natürlich einen Schrottkugelturm um sich den Horden der Finsternis mit genug Schrott erwehren zu können! Ihnen stehen zwei mächtige magische Formeln zur Unterstützung zu Verfügung.

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

mit

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

Als erstes müssen Sie die Höhe des zu bauenden Schrottkugelturmes bestimmen. Hierfür ist wichtig zu wissen, dass sich die Bleitropfen mit einem Gewicht von $40mg$ zu gleichförmigen Bleikugeln bei einer Geschwindigkeit von $11m/s$ bilden.

1. Wie hoch müssen Sie den Schrottkugelturm bauen lassen, damit sich runde Bleikugeln durch die Fallgeschwindigkeit von $11m/s$ bilden? **(3 Punkte)**

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

2. Nennen Sie die beiden geometrischen Formen aus denen sich näherungsweise ein Tropfen zusammensetzt! Erstellen Sie eine beschriftete Skizze des Tropfens! **(2 Punkte)**
3. Sie messen eine Länge des Tropfens von $3.5mm$. Die Löcher im Sieb erlauben ein Tropfendurchmesser von $1.5mm$. Welchen Durchmesser in mm haben Ihre produzierten Bleikugeln? **(3 Punkte)**

Sie haben jetzt die 6.1×10^4 Bleikugeln zusammen. Blei hat eine Dichte von $11.23g/cm^3$.

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

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

5. Wie groß in cm^2 ist Ihr quadratisches Sieb am oberen Ende des Turms, wenn Sie pro Fall ca. 700 Bleikugeln produzieren wollen und die Bleikugel im Fall $1.2cm$ Abstand haben müssen? **(1 Punkt)**

¹⁰Ein wirklich gefährliches Buch ist: *Du bist genug: Vom Mut, glücklich zu sein* von Fumitake Koga und Ichiro Kishimi

120. Aufgabe

(10 Punkte)

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



Armee der Kaninchen Leider hat es mit Ihrer 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 1875 ungefähr 30 Kaninchen entlassen hat, Millionen von Kaninchen rum? Wieso ist das Kaninchenfleisch dann so exklusiv? Dem wollen wir mal mathematisch nachgehen!¹¹

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

$$f(t) = 8 \times 10^9 - 9 \times 10^8 \cdot 2^{-0.2 \cdot t + 3.1}$$

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 9 Jahren auf dem australischen Kontinent? **(1 Punkt)**

Um den Kaninchen Einhalt zu gebieten wurde das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) in 20 Kaninchen ausgebracht. Da die Kaninchen keine Maßnahmen gegen die Ausbreitung vornehmen können, verläuft die Ausbreitung mit einem wöchentlichen Wachstumsfaktor von 1.6 nach folgender Formel.

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

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

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

4. Wie lange in Jahren dauert es bis eine Kaninchenpopulation nach einer Viruspandemie wieder auf 40% der gesättigten Kaninchenpopulation angewachsen ist? **(2 Punkte)**

Thomas Austin entließ die Kaninchen im äußersten Süden von Australien. Australien hat eine West-Ost-Ausdehnung von 4400km und eine Nord-Süd-Ausdehnung von knapp 3800km. 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 12\$ pro Tier und der durchführende Arzt verlangt ca. 35\$ pro Tier.

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

¹¹Die Quelle der Inspiration für die Aufgabe war der folgendes YouTube Video: [Incredible Stories – Why don't they eat wild rabbits in Australia? They have millions of them! The reason is surprising...](#)

121. Aufgabe

(12 Punkte)

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



Uckermark. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer des Hängebauschweins Frida und Ihnen. Grünes Gras unter Ihren Füßen und ein strammer Wind im Gesicht, egal wohin Sie schauen. Ein schmatzendes Geräusch ertönt unter Ihnen. Sie sinnieren, sollten Sie Ihre weiten Graslandschaften jetzt schon düngen? Dafür benötigen Sie die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Sie sehen nicht ein, Geld für einen Agrarmetrologen zu bezahlen. Also rechnen Sie mit folgenden Informationen zu Monatsmultiplikatoren des GLT-Wertes: Januar mit $0.5\times$, Februar mit $0.75\times$ und März mit $1.2\times$. Sie haben noch im letzten Jahr folgende Temperaturen gemessen.

Datum	C°
01. Jan 2023	0.3
01. Feb 2023	1.2
01. Mrz 2023	3.1
01. Apr 2023	6.1

1. Erstellen Sie eine Skizze aus den Informationen aus der Temperaturtabelle! **(1 Punkt)**
2. Stellen Sie die linearen Funktionen $f_1(t)$, $f_2(t)$ und $f_3(t)$ aus der obigen Temperaturtabelle auf! **(1 Punkt)**
3. Bestimmen Sie die Stammfunktionen $F_1(t)$, $F_2(t)$ und $F_3(t)$ für Ihre linearen Funktionen aus der obigen Temperaturtabelle! **(1 Punkt)**
4. Osterglocken beginnen ab einer GLT von 190°C zu blühen. An welchem Tag im 1. Quartal des Jahres 2023 war dies der Fall? *Ignorieren Sie ein eventuelles Schaltjahr in Ihrer Berechnung.* **(4 Punkte)**

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

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

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

122. Aufgabe

(10 Punkte)

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



In der Kartonagenfabrik Wenn man sich zu spät anmeldet, dann ist die Exkursion nicht so toll. Also geht es mit Rektor Skinner und Mrs. Krabappel in die Kartonagenfabrik. Wie schon im vorherigen Semester... In der Kartonagenfabrik angekommen erfahren Sie, dass die Kartons zum Versand von Nägeln nicht hier zusammengebaut werden sondern das sich die Endfertigung in Flint, Michigan befindet. Unter anderem wird dort der berühmte *Doppelt gewellte, 5-mal-gefaltete, 0.5mm, 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 18cm. Die Kartonagenmaschine in Flint soll dann einen quadratischen Eckenausschnitt der Länge x falzen.

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

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

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

123. Aufgabe

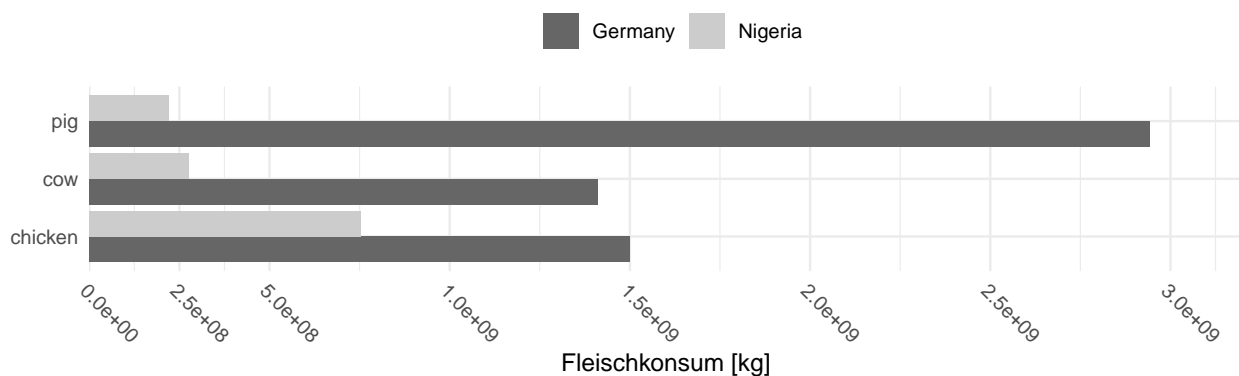
(10 Punkte)

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



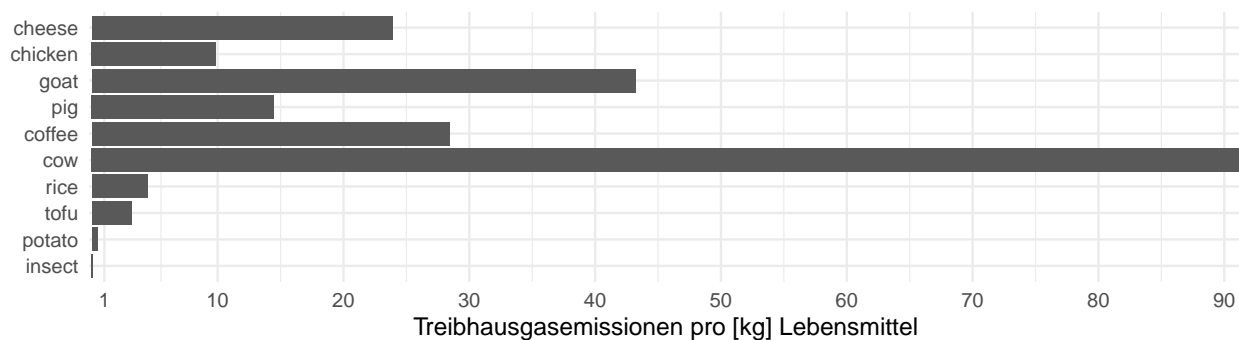
Ein Pfund Insekten, bitte! Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹². Schauen wir uns dazu einmal den Vergleich Deutschland zu Nigeria an. Nigeria hat eine der am schnellsten wachsenden Bevölkerungen der Welt und wird vermutlich im Jahr 2100 zu den Top 5 der bevölkerungsreichsten Länder zählen. Im Jahr 2023 leben ca. 8.2×10^7 Menschen in Deutschland und ca. 1.84×10^8 Menschen in Nigeria. Mit den Informationen wollen wir anfangen und dann eine Prognose für den Fleischkonsum im Jahr 2050 zu treffen.

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



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

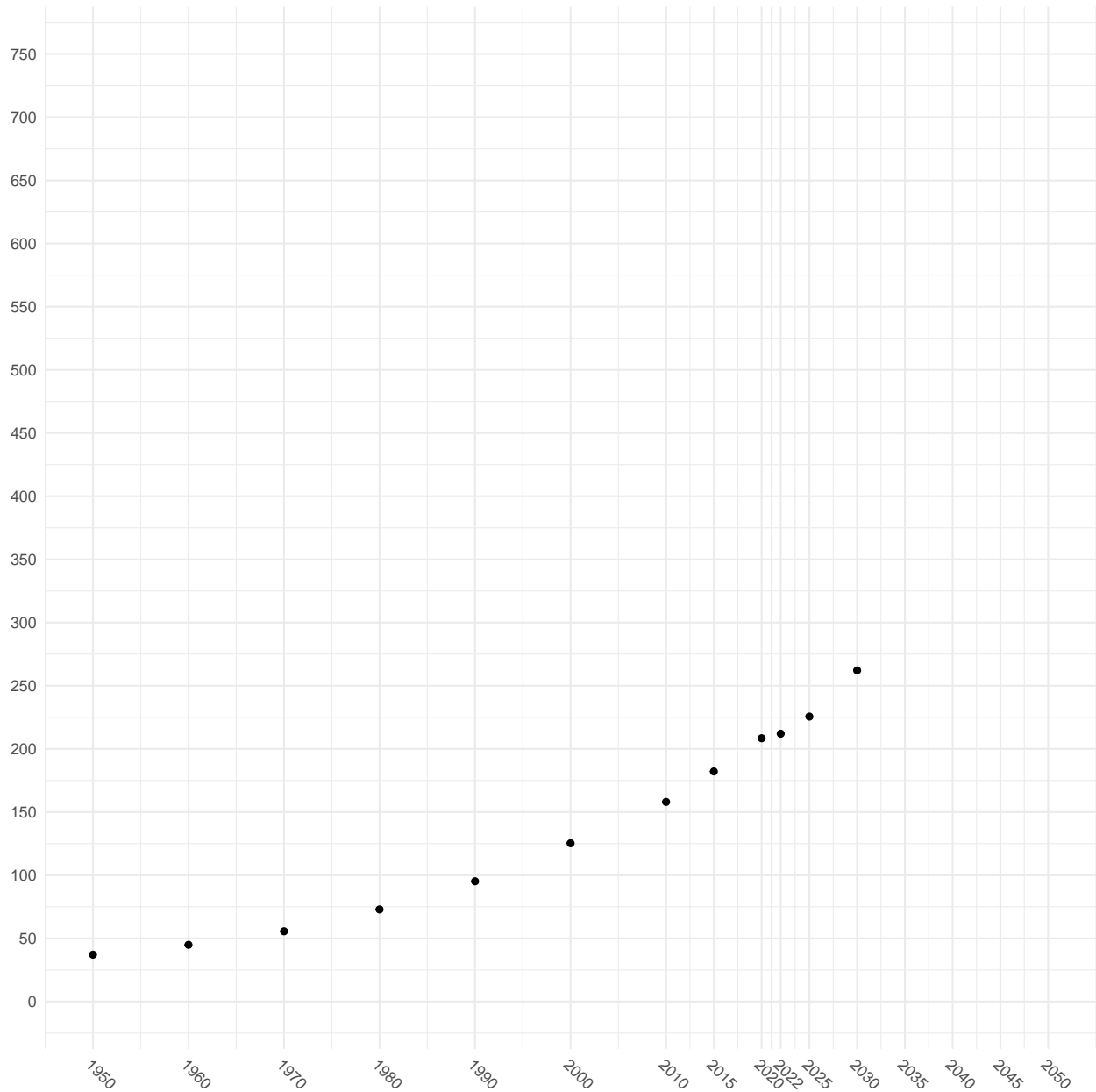
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



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

¹²Die 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)**

124. Aufgabe

(12 Punkte)

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



Tödliche Seuche AIDS – Die rätselhafte Krankheit Irritiert legen Sie die historische Ausgabe des Spiegels aus den 80zigern beiseite. Sie sind bei Ihrem Orthopäden 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 0.01% 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 $n = 2 \times 10^4$ Patienten mit einem diagnostischen Test für AIDS aus. Sie nehmen sich also einen Kuli und fangen an auf der historischen Ausgabe des Spiegels zu rechnen¹³.

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)**

¹³Die Quelle der Inspiration für die Aufgabe war der folgende wissenschaftlicher Artikel: Binder et al. (2022) Von Baumdiagrammen über Doppelbäume zu Häufigkeitsnetzen – kognitive Überlastung oder didaktische Unterstützung? Journal für Mathematik-Didaktik, 1-33

125. Aufgabe

(12 Punkte)

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



Network-Marketing oder Schneeballschlacht! Eine Möglichkeit, leicht Geld zu verdienen, ist es anderen Menschen für Geld zu versprechen, wie man leicht reich werden kann. Am besten natürlich ohne viel Aufwand und ortsunabhängig. Dann wollen wir mal loslegen. Schnell ein YouTube-Werbevideo gedreht und auf geht es mit unserem Network-Marketing. Aber Moment, wie funktioniert Network-Marketing eigentlich und was hat das alles mit einer Schneeballschlacht zu tun? Wir wollen hier einmal in die Untiefen des „passiven Einkommens“ abtauchen¹⁴.

Das Jahr 2022 war das erfolgreichste Jahr in der Geschichte von Up/Down Systems and Networking (UD-SysNet). 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 UDSysNet habe das Unternehmen 3.6×10^5 aktive Partner.

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

Ihr zu vermarkendes Produkt, hinter dem Sie voll stehen, kostet 75EUR pro Einheit im Direktverkauf. Die direkte Provision für die erste Stufe beträgt 25%. Für die zweite, dritte und vierte Stufe betragen die indirekten Provisionen jeweils 3.25%, 2.25% und 1.75%. Jeder Ihrer angeworbenen „Partner“ wirbt wiederum vier Partner für sich selbst an. Pro Monat werden im Schnitt drei Einheiten vom Produkt verkauft. Sie wollen nun 2500EUR im Monat *passiv* – also durch indirekte Provisionen – erwirtschaften.

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

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

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

Sie mussten zum Einstieg bei UDSysNet Einheiten des Produkts für 2625EUR 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)**

¹⁴Die Quellen der Inspiration für die Aufgabe waren folgendes YouTube Video: [Simplicissimus – Die meistgesuchte Betrügerin der Welt](#) und der Artikel: [Deutschlandfunk Kultur – Die Illusion, schnell reich zu werden](#)

126. Aufgabe

(10 Punkte)

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



Höhlen & Drachen Nachdem Sie sich begeistert in der Serie *Stranger Thinks* verloren haben, wollen Sie bei 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 achtseitige Würfel ($6d8$) zum würfeln in der Hand. Wenn Sie eine 8 würfeln, haben Sie einen Erfolg.

1. Berechnen Sie die Wahrscheinlichkeit *genau* 4 Erfolge zu erzielen! **(2 Punkte)**
2. Berechnen Sie die Wahrscheinlichkeit keinen Erfolg zu erzielen! **(1 Punkt)**

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

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

Jetzt wird es immer wilder, da Sie sich jetzt überlegen müssen, wie wahrscheinlich es ist, dass Ihr Rettungswurf gegen den zaubernden Hexer funktioniert. Sie haben folgende Wahrscheinlichkeiten gegeben. Die Wahrscheinlichkeit für das Ereignis A , der Rettungswurf ist erfolgreich, ist $Pr(A) = 0.6$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.7$. Sie haben mitgezählt und festgestellt, dass in 40 von 100 Fällen Ihr Rettungswurf bei einem erfolgreichen Zauber funktioniert hat.

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

127. Aufgabe

(12 Punkte)

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



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

Name	$P(\text{win})$	$P(\text{outbid})$
Annegret	0.3	0.043
Günther	0.2	0.076

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 0.05 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 20 Felder. Sie drehen das Glücksrad zweimal. Auf 6 Feldern gewinnen Sie 5000EUR sonst 1000EUR. Ganz schön viel Geld und ganz schön zwielichtig hier...

3. Skizzieren Sie das Glücksrad und ergänzen Sie die Wahrscheinlichkeiten! **(1 Punkt)**
4. Zeichnen Sie das zugehörige Baumdiagramm für das zweimalige Drehen! Ergänzen Sie die Wahrscheinlichkeiten und die entsprechenden Ereignisse **(2 Punkte)**
5. Mit welcher Wahrscheinlichkeit gewinnen Sie 6000EUR? **(1 Punkt)**

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

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

Teil XI.

Angewandte Nutztier- und Pflanzenwissenschaften (M.Sc.)

128. Aufgabe

(6 Punkte)

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

129. Aufgabe

(8 Punkte)

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

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

mit

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

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

130. Aufgabe

(6 Punkte)

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