

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 75 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 95 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
91.0 - 95.0	1,0
86.0 - 90.5	1,3
81.5 - 85.5	1,7
76.5 - 81.0	2,0
72.0 - 76.0	2,3
67.5 - 71.5	2,7
62.5 - 67.0	3,0
58.0 - 62.0	3,3
53.0 - 57.5	3,7
47.5 - 52.5	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	12	12	12	9	8	12	10

- Es sind ____ von 75 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 $\eta^2 = 0.31$. Welche Aussage ist richtig?

- A** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird. Das η^2 ist damit mit dem R^2 aus der linearen Regression zu vergleichen.
- 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 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen nicht erklärt wird. Somit der Rest an nicht erklärbarer Varianz.

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 $\eta^2 = 0.25$. Welche Aussage ist richtig?

- A** ☐ Mit dem η^2 lässt sich auf die Qualität der Randomisierung und damit der Strukturgleichheit zwischen der Grundgesamtheit und der Stichprobe schließen. Es gilt dabei die Regel, dass ein η^2 -Wert von 1 zu bevorzugen ist.
- B** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird. Daher werden 25% der Varianz durch die Behandlungsgruppen erklärt.
- C** ☐ Es werden 25% der Varianz durch den Versuch erklärt. Das η^2 beschreibt den Anteil der Varianz, der durch Fehler in der Versuchsdurchführung entsteht.
- D** ☐ Das η^2 beschreibt den Anteil der Varianz, der von den Umweltbedingungen erklärt wird. Daher werden 25% der Varianz durch die Umweltbedingungen erklärt. Der Anteil der Varianz durch die Behandlungsgruppen ist dann 75%.
- E** ☐ Es werden 75% der Varianz durch die Behandlung erklärt. Das η^2 beschreibt den Anteil der Varianz, der von den unterschiedlichen Behandlungsbedingungen nicht erklärt wird.

3 Aufgabe

(2 Punkte)

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

- A** ☐ Wenn die F-Statistik kleiner als der kritische Wert ist kann die Nullhypothese nicht abgelehnt werden. Die F-Statistik ist der Quotient der MS der Behandlung durch die MS des Fehlers.
- 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 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.
- E** ☐ Die ANOVA berechnet die F-Statistik aus den SS Behandlung geteilt durch die SS Fehler.

4 Aufgabe

(2 Punkte)

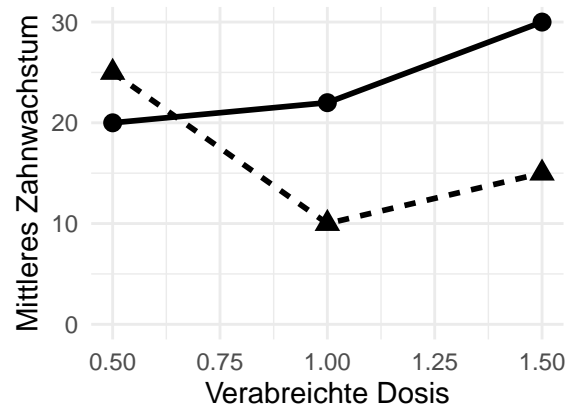
Die ANOVA ist ein statistisches Verfahren welches häufig in den Auswertungen von Experimenten in den Agrarwissenschaften angewendet wird. Dabei wird die ANOVA als ein erstes statistischen Werkzeug für die Übersicht über die Daten benutzt. Eine ANOVA testet dabei...

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

5 Aufgabe

(2 Punkte)

Die folgende Abbildung enthält die Daten aus einer Studie zur Bewertung der Wirkung von Vitamin D auf das Zahnwachstum bei Igeln. Der Versuch wurde an 67 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** ☐ Das Bestimmtheitsmaß R^2 ist klein.
- C** ☐ Die Koeffizienten sind negativ ($\beta_0 < 0$; $\beta_1 < 0$).
- D** ☐ Eine positive Interaktion liegt vor ($p \leq -0.5$)
- E** ☐ Keine Interaktion liegt vor ($p \leq 0.05$).

Deskriptive Statistik & Explorative Datenanalyse

6 Aufgabe

(2 Punkte)

Wie lautet der Mittelwert und Standardabweichung von y mit 12, 15, 9, 10 und 19.

- A ☐ Sie erhalten 13 \pm 2.01
- B ☐ Es ergibt sich 13 \pm 4.06
- C ☐ Es berechnet sich 13 \pm 16.5
- D ☐ Es ergibt sich 14 \pm 2.03
- E ☐ Es berechnet sich 14 \pm 16.5

7 Aufgabe

(2 Punkte)

Berechnen Sie den Median, das 1st Quartile sowie das 3rd Quartile von y mit 22, 32, 15, 22, 15, 25, 25, 12, 31, 18 und 51.

- A ☐ Es ergibt sich 22 [15; 31]
- B ☐ Es berechnet sich 24 [16; 32]
- C ☐ Es ergibt sich 22 \pm 15
- D ☐ Es berechnet sich 23 [16; 30]
- E ☐ Es ergibt sich 24 \pm 15

8 Aufgabe

(2 Punkte)

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

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

9 Aufgabe

(2 Punkte)

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

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

10 Aufgabe

(2 Punkte)

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

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

11 Aufgabe

(2 Punkte)

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

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

12 Aufgabe

(2 Punkte)

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

- A** ☐ Einen Dotplot. Die Punkte müssen sich wie an einer Perlenschnur aufreihen. Eine Abweichung führt zur Ablehnung der Annahme einer Normalverteilung.
- B** ☐ Wir erstellen uns für jede Behandlung einen Dotplot und schauen, ob die Dots und damit die Varianz für jede Behandlung gleich groß sind.
- C** ☐ Nach der Erstellung eines Boxplots schauen wir, ob der Median in der Mitte der Box liegt. Dabei ist der Median als dicke Linie dargestellt und die Box ist das IQR.
- D** ☐ Nach dem Einlesen der Daten nutzen wir einen Boxplot um zu schauen, ob alle Boxen über alle Behandlungen in etwa gleich groß sind. Damit ist dann auch das IQR in allen Behandlungen in etwa gleich.
- E** ☐ Einen Violinplot. Der Bauch der Violine muss hierbei einen höheren Wert annehmen als der Steg der Violine. Dann kann die Annahme einer Normalverteilung angenommen werden.

13 Aufgabe

(2 Punkte)

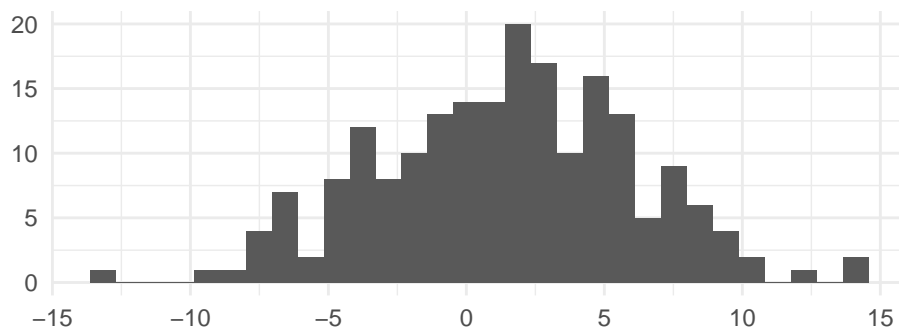
Sie wollen in Ihrer Abschlussarbeit über eine explorative Datenanalyse überprüfen, ob Ihr gemessener Endpunkt einer Normalverteilung folgt. Welche drei Abbildungen eignen sich insbesondere für die Überprüfung?

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

14 Aufgabe

(2 Punkte)

Sie haben $n = 187$ Pflanzen geerntet und wollen sich nun die Verteilung der Pflanzen einmal in einem Histogramm anschauen. Welche Verteilung ist dargestellt?



- A ☐ Wir haben eine Normalverteilung vorliegen.
- B ☐ In dem Histogramm ist eine Ordinalverteilung dargestellt.
- C ☐ Wir haben eine Gammaverteilung vorliegen.
- D ☐ Dem Histogramm entnehmen wir eine Poisson-Verteilung.
- E ☐ Eine Standardnormalverteilung.

Lineare Regression & Korrelation

15 Aufgabe

(2 Punkte)

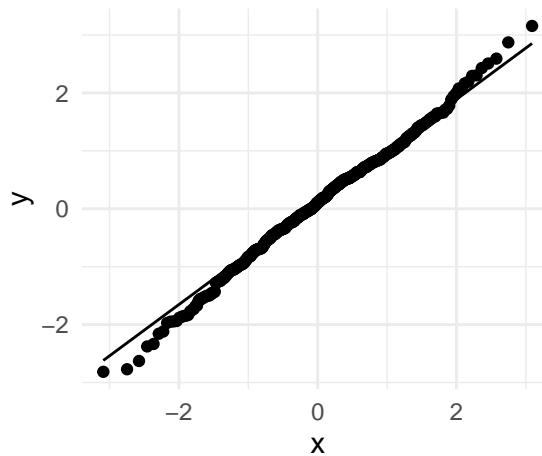
Im Allgemeinen gibt es zwei mögliche Ziele für ein Regressionsmodell. Wir können ein Vorhersagemodell oder ein kausales Modell rechnen. Welche Aussage ist für ein kausales Modell richtig?

- A ☐ Ein kausales Modell wird auf einem Trainingsdatensatz trainiert und anschliessend über eine explorative Datenanalyse validiert. Signifikanzen über β_i können hier nicht festgestellt werden.
- B ☐ Es wird ein Trainingsdatensatz zum Modellieren des Trainingsmodells benötigt. Der Testdatensatz dient rein zur Visualisierung. Dies gilt vor allem für ein kausales Modell.
- C ☐ Es wird ein Trainingsdatensatz zum Trainieren des Modells benötigt. Der Testdatensatz dient zur Validierung. Dies gilt insbesondere für ein kausales Modell.
- D ☐ Wenn ein kausales Modell gerechnet werden soll dann kann dies auf dem gesamten Datensatz geschehen. Das Ziel ist es einen Zusammenhang von X auf Y zu modellieren. Wie wirken sich die Einflussvariablen Y auf die gemessenen Endpunkte $X = x_1, \dots, x_p$ aus?
- E ☐ Ein kausales Modell möchte die Zusammenhänge von X auf Y modellieren. Hierbei geht es um die Effekte von X auf Y . Man sagt, wenn x_1 um 1 ansteigt ändert sich Y um einen Betrag β_1 .

16 Aufgabe

(2 Punkte)

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

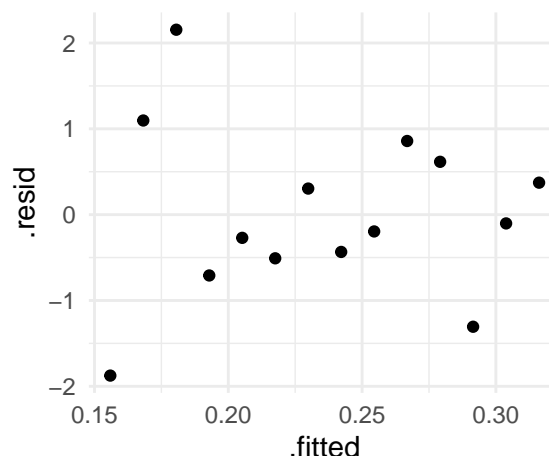


- A** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.
- B** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Die Punkte liegen zum überwiegenden Teil nicht auf der Geraden.
- C** ☐ Wir betrachten die Gerade und dabei insbesondere die beiden Enden der Gerade in dem IQR, also dem ersten und dritten Quartile. Hier sollten die Punkte auf der Geraden liegen, dann ist die Annahme an die Normalverteilung der Residuen erfüllt.
- D** ☐ Die Annahme der normalverteilten Residuen ist nicht erfüllt. Die Punkte liegen zum überwiegenden Teil nicht 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.

17 Aufgabe

(2 Punkte)

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



- A** ☐ Die Annahme der normalverteilten Residuen ist erfüllt. Kein Muster ist zu erkennen und keine Outlier zu beobachten.

- B** ☐ Wenn wir die Nulllinie betrachten so müssen die Punkte gleichmäßig unter der Nulllinie liegen. Unser Modell erfüllt somit nicht die Annahme von normalverteilten Residuen mit einem Mittelwert von > 0 und einer Streuung von s .
- C** ☐ 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.
- D** ☐ Die Punkte müssen gleichmäßig in dem positiven Bereich liegen. Dies ist hier klar nicht der Fall. Einzelne Ausreißer können beobachtet werden. Die Analyse ist gescheitert.
- E** ☐ Wenn wir die Nulllinie betrachten so liegen die Punkte nicht gleichmäßig über und unter der Nulllinie. Unser Modell erfüllt nicht die Annahme von normalverteilten Residuen mit einem Mittelwert von 0 und einer Streuung von s^2 .

18 Aufgabe

(2 Punkte)

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

- A** ☐ Der Korrelationskoeffizienten ρ wird wie das η^2 aus der ANOVA interpretiert. Der Korrelationskoeffizienten ρ beschreibt den Anteil an erklärter Varianz durch die Regression. Dabei gibt er jedoch eine Richtung an und kann auch negativ werden.
- B** ☐ Der Korrelationskoeffizienten ρ ist eine standardisierte, statistische Maßzahl, die zwischen -1 und 1 liegt. Dabei ist Korrelationskoeffizienten ρ einheitslos.
- C** ☐ Der Korrelationskoeffizienten ρ liegt zwischen -1 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ als standardisierte Steigung zu verstehen, wenn eine Standardisierung durchgeführt wurde. Diese Adjustierung nach Fischer muss am Anschluß der Berechnung der Korrelation durchgeführt werden.
- D** ☐ Der Korrelationskoeffizienten ρ zeigt keinen Zusammenhang zwischen zwei Variablen x und y bei einem Wert von 0. Einen negativen Zusammenhang Richtung -1 und somit auch einen positiven Zusammenhang Richtung 1. Je größer die Zahl allgemein, desto stärker der Effekt.
- E** ☐ Korrelationskoeffizienten ρ liegt zwischen 0 und 1. Darüber hinaus ist der Korrelationskoeffizienten ρ einheitslos und kann als Standardisierung verstanden 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 Fe_3O_4 -Konzentration in $[\mu g]$ im Wasser auf das Trockengewicht in $[kg]$ an Wasserlinsen folgende Effekte und p -Werte: $2e - 04$ als p -Wert und einen $\beta_{Fe_3O_4}$ Koeffizienten von 6.9×10^{-7} . Welche Aussage ist richtig?

- A** ☐ Manchmal ist die Einheit der Einflussvariable X zu klein gewählt, so dass der Anstieg von 1 Einheit in X zu einer zu kleinen Änderung in y führt. Daher kann der Effekt $\beta_{Fe_3O_4}$ sehr klein wirken, aber auf einer anderen Einheit sehr viel größer sein. Der p -Wert wird auf einer einheitslosen Teststatistik bestimmt.
- B** ☐ Wenn der Effekt $\beta_{Fe_3O_4}$ winzig ist, dann kann es an einer falsch gewählten Einheit liegen. Der Anstieg von einer Einheit in X führt ja zu einer Änderung von $\beta_{Fe_3O_4}$ in x . Wir müssen daher die Einheit von y entsprechend anpassen.
- C** ☐ Die Einheit der Fe_3O_4 -Konzentration ist zu klein gewählt. Dadurch sehen wir den sehr kleinen p -Wert. Der p -Wert und die Einheit von der Fe_3O_4 -Konzentration hängen antiproportional zusammen.
- D** ☐ Das Gewicht und die Fe_3O_4 -Konzentration korrelieren sehr stark, deshalb wird der $\beta_{Fe_3O_4}$ Koeffizient sehr klein. Mit einer ANOVA kann für die Korrelation korrigiert werden und der Effektschätzer passt dann zum p -Wert.
- E** ☐ 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.

20 Aufgabe

(2 Punkte)

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

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

21 Aufgabe

(2 Punkte)

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

- A** ☐ Das GLM erlaubt auch nicht normalverteilte Residuen in der Schätzung der Regressionsgrade.
- B** ☐ Das *generalisierte lineare Modell (GLM)* erlaubt auch weitere Verteilungsgruppen für das X bzw. die Einflussvariablen in einer linearen Regression zu wählen.
- C** ☐ Das GLM ist ein faktisch maschineller Lernalgorithmus, der selbstständig die Verteilungsfamilie für Y wählt.
- D** ☐ Dank dem *generalisierten linearen Modell (GLM)* können auch andere Verteilungsfamilien – außer die Normalverteilung – mit einer linearen Regression modelliert werden. Dafür werden alle Verteilungen in eine Normalverteilung überführt und anschließend standardisiert.
- 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 sorgt für Strukturgleichheit und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- B** ☐ Randomisierung erlaubt erst die Mittelwerte zu schätzen. Ohne Randomisierung keine Mittelwerte. Ohne Mittelwerte keine Varianz und somit auch kein statistischer Test.
- C** ☐ Randomisierung bringt starke Unstrukturiertheit in das Experiment und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- D** ☐ Randomisierung war bis 1952 bedeutend, wurde dann aber in Folge besserer Rechnerleistung nicht mehr verwendet. Aktuelle Statistik nutzt keine Randomisierung mehr.
- E** ☐ Durch eine Randomisierung können wir nicht von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.

23 Aufgabe

(2 Punkte)

Sie wollen Ihren Datensatz in **R** 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** ☐ Programmiersprachen haben Probleme mit Umlauten und Sonderzeichen der deutschen Sprache. Die Nutzung von englischer Sprache umgeht dieses Problem in eleganter Art.
- B** ☐ Programmiersprachen können nur englische Begriffe verarbeiten. Zusätzliche Pakete können zwar geladen werden, aber meist funktionieren diese Pakete nicht richtig. Deutsch ist International nicht bedeutend genug.
- C** ☐ Alle Funktionen und auch Anwendungen sind in **R** in englischer Sprache. Die Nutzung von deutschen Wörtern ist nicht schick und das ist zu vermeiden.
- D** ☐ Die Spracherkennung von **R** ist nicht in der Lage Deutsch zu verstehen.
- E** ☐ Es gibt keinen Grund nicht auch deutsche Wörter zu verwenden. Es ist ein Stilmittel.

24 Aufgabe

(2 Punkte)

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

- A** ☐ Wir lesen die Daten ein und mutieren die Daten. Dabei ist wichtig, dass wir nicht das Paket tidyverse nutzen, da dieses Paket veraltet ist. über die Funktion library(tidyverse) entfernen wir das Paket von der Analyse.
- B** ☐ Wir transformieren die Spalten über mutate() in ein tibble und können dann über ggplot() uns die Abbildungen erstellen lassen. Dabei beachten wir das wir keine Faktoren in den Daten haben.
- C** ☐ 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** ☐ 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.
- E** ☐ 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.

25 Aufgabe

(2 Punkte)

Es sei $s_1^2 = s_2^2$ in dem Modell $Y \sim X$. Welche Aussage ist richtig?

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

26 Aufgabe

(2 Punkte)

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

- A** ☐ Untereinander unabhängig. 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** ☐ Untereinander stark korreliert. Die Ferkel sind von einer Mutter und somit miteinander korreliert. Dies wird in der Statistik jedoch meist nicht modelliert.
- D** ☐ Untereinander abhängig, wenn die Mütter ebenfalls miteinander verwandt sind. Erst die Abhängigkeit 2. Grades wird in der Statistik modelliert.
- E** ☐ Untereinander unabhängig. Die Ferkel sind eigenständig und benötigen keine zusätzliche Behandlung.

27 Aufgabe

(2 Punkte)

In einer Studie wollen Sie den Effektschätzer Risk ratio berechnen. Sie finden in Ihrem Experiment zur Behandlung von Klaueninfektionen bei Ziegen in 5 Tieren Erkrankung der Klauen vor. 8 Tiere sind gesund. Welche Aussage ist richtig?

- A** ☐ Das Verhältnis der Chancen Risk ratio ergibt ein Chancenverhältnis von 0.38. Wir sind an der Chance krank zu sein interessiert.
- B** ☐ Es ergibt sich ein Risk ratio von 1.6, da es sich um ein Anteil handelt.
- C** ☐ Es ergibt sich ein Risk ratio von 0.62, da es sich um ein Anteil handelt.
- D** ☐ Das Verhältnis von Chancen Risk ratio ergibt ein Chancenverhältnis von 0.62.
- E** ☐ Der Anteil der Kranken wird berechnet. Da es sich um ein Anteil handelt ergibt sich ein Risk ratio von 0.38.

28 Aufgabe

(2 Punkte)

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

- A** ☐ Aktuell werden zu grosse Datensätze für die gängige Statistik gemessen. Daher wendet man maschinelle Lernverfahren für kausale Modelle an. Hier ist die Relevanz gleich Signifikanz.
- B** ☐ Riesige Datensätze haben mehr Fallzahl was zur α -Inflation führt. Durch eine Adjustierung kann dem Problem entgegengewirkt werden.
- C** ☐ Aktuell werden immer größere Datensätze erhoben. Eine erhöhte Fallzahl führt automatisch auch zu mehr signifikanten Ergebnissen, selbst wenn die eigentlichen Effekte nicht relevant sind.
- D** ☐ Relevanz und Signifikanz haben nichts miteinander zu tun. Daher gibt es auch keinen Zusammenhang zwischen hoher Fallzahl ($n > 10000$) und einem signifikanten Test. Ein Effekt ist immer relevant und somit signifikant.
- E** ☐ Big Data ist ein Problem der parametrischen Statistik. Parameter lassen sich nur auf kleinen Datensätzen berechnen, da es sich sonst nicht mehr um eine Stichprobe im engen Sinne der Statistik handelt.

Multiple Gruppenvergleiche

29 Aufgabe

(2 Punkte)

Sie haben folgende unadjustierten p-Werte gegeben: 0.01, 0.89, 0.001, 0.03 und 0.42. Sie adjustieren die p-Werte nach Bonferroni. Welche Aussage ist richtig?

- A** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.05, 1, 0.005, 0.15 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.
- B** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.05, 4.45, 0.005, 0.15 und 2.1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- C** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.002, 0.178, 2e-04, 0.006 und 0.084. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- D** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.05, 1, 0.005, 0.15 und 1. Die adjustierten p-Werte werden zu einem α -Niveau von 5% verglichen.
- E** ☐ Nach der Bonferroni-Adjustierung ergeben sich die adjustierten p-Werte von 0.002, 0.178, 2e-04, 0.006 und 0.084. Die adjustierten p-Werte werden zu einem α -Niveau von 1% verglichen.

30 Aufgabe

(2 Punkte)

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

- A** ☐ Compact line display. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt. Früher wurden keine Buchstaben sondern eine durchgezogene Linie verwendet. Bei mehr als drei Gruppen funktioniert die Linie aber graphisch nicht mehr.
- 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** ☐ Contrast letter display. Unterschiede in den Behandlungen werden durch den gleichen Buchstaben oder Symbol dargestellt. Die Interpretation des CLD führt häufig in die Irre.
- D** ☐ Compact letter detection. Gleichheit in den Behandlungen wird durch den gleichen Buchstaben oder Symbol dargestellt.
- E** ☐ Compact letter display. Das CLD ist umstritten, da es die Gleichheit der Behandlungen durch gleiche Buchstaben darstellt. Dadurch ist das CLD nicht mehr sauber auf einer Linie mit dem statistischen Testen. Wir lehnen die Nullhypothese ab und zeigen keine Gleichheit im statistischen Testen.

31 Aufgabe

(2 Punkte)

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

- A** ☐ Sie nutzen das R Paket {emmeans} für die Berechnung des multiplen Gruppenvergleichs. Die Ausgabe der Funktion emmeans() erlaubt zügig über {ggplot} einen Barplot zu erstellen und dann auch das CLD zu berechnen. Sie haben alles sofort zusammen.
- B** ☐ Das R Paket {lm}. Das Paket {lm} erstellt selbstständig Konfidenzintervalle und entsprechende p-Werte. Da wir in dem Paket nicht adjustieren müssen, ist es bei Anwendern sehr beliebt.
- C** ☐ Das R Paket {ggplot}. Wir erhalten hier sofort eine Visualisierung der Daten. Anhand der Visualisierung lässt sich eine explorative Datenanalyse durchführen, die gleichwertig zu einem Posthoc-Test ist.
- D** ☐ Das R Paket {hmisc} erlaubt die Durchführung eines multiplen Gruppenvergleichs aus verschiedenen Modellen heraus. Aus einem hmisc Objekt lässt sich recht einfach das CLD erstellen und so über Barplots eine schnelle Interpretation der statistischen Auswertung durchführen.
- E** ☐ Das R Paket {emmeans} erlaubt die Durchführung eines multiplen Gruppenvergleichs. Aus einem emmeans Objekt lässt sich leider kein CLD erstellen. Dennoch ist das Paket einfach zu bedienen und wird deshalb genutzt. Die Interpretation der statistischen Auswertung wird über einen Barplot abgebildet.

32 Aufgabe


(2 Punkte)

In den Humanwissenschaften werden multiple Vergleiche häufig anders behandelt als in den Agrarwissenschaften. In beiden Bereichen tritt jedoch das gleiche Phänomen bei multiplen Testen auf. Wie muss mit dem Phänomen umgegangen werden und wie ist es benannt?

- 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** ☐ Die Adjustierung der p-Werte nach Bonferroni erlaubt es gegen die α -Inflation vorzugehen, die häufig beim multiplen Testen auftritt. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern sehr viel höher. Das ist der Grund warum die p-Werte entsprechend adjustiert werden müssen.
- C** ☐ Beim multiplen Testen kann es zu Varianzheterogenität kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5%. Daher müssen die p-Werte entsprechend adjustiert werden. Das Verfahren nach Welch, bekannt aus dem t-Test, ist hier häufig anzuwenden.
- D** ☐ Beim multiplen Testen kann es zu einer α -Deflation kommen. Das globale Signifikanzniveau liegt nicht mehr bei 5% sondern weit darunter. Daher müssen die p-Werte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der p-Werte nach Bonferroni das bekannteste Verfahren ist. Die p-Werte werden durch die Anzahl an Vergleichen geteilt
- E** ☐ Die Adjustierung der p-Werte nach Bonferroni erlaubt es gegen die β -Inflation vorzugehen, die häufig beim multiplen Testen auftritt. Das globale Powerniveau liegt nicht mehr bei 80% sondern sehr viel niedriger.

33 Aufgabe

(2 Punkte)

In Ihrer Bachelorarbeit werten Sie einen einfaktoriellen Versuch aus. Dafür rechnen Sie in  zunächst eine ANOVA und schließen dann einen multiplen Vergleich mit t-Tests an. Welche Aussage über die Effekte in Ihrem Versuch ist richtig?

- A** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ nicht adjustiert werden. Bei einem Effekt im multiplen Testen handelt es sich um eine Wahrscheinlichkeit für das Auftreten der Nullhypothese.
- B** ☐ Beim multiplen Testen muss der Effekt, hier der Mittelwertsunterschied Δ aus den paarweisen t-Tests, nicht adjustiert werden.
- C** ☐ Wenn ein multipler Test gerechnet wird, dann muss der Effekt Δ adjustiert werden im Gegensatz zu den p-Werten.
- D** ☐ Beim multiplen Testen kann es zu einer Effektüberschätzung (Δ -Inflation) kommen. Daher müssen die Effekte angepasst werden. Dies geschieht nicht händisch sondern intern in den angewendeten Algorithmen.
- E** ☐ Beim multiplen Testen kann es zu einer Δ -Inflation kommen. Das globale Effektniveau liegt nicht mehr bei 20%. Daher müssen die Effekte entsprechend adjustiert werden. Hierfür gibt es verschiedene Verfahren, wobei das Verfahren zur Adjustierung der Effekte nach Bonferroni das bekannteste Verfahren ist.

Statistische Testtheorie

34 Aufgabe

(2 Punkte)

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

- A** ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit der Alternativhypothese und somit $1 - Pr(H_A)$
- B** ☐ Die Inverse der Wahrscheinlichkeit unter der die Nullhypothese nicht mehr die Alternativhypothese überdeckt.

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

35 Aufgabe

(2 Punkte)

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

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

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 Rang einer Naturkonstante.
- B** ☐ Als Kulturkonstante hat $\alpha = 5\%$ den Rang einer Naturkonstante und wurde nach langer Diskussion in der UN im Jahre 1983 festgesetzt. Damals auch schon mit der Zustimmung der UdSSR.
- C** ☐ Auf einer Statistikkonferenz in Genf im Jahre 1942 wurde dieser Cut-Off nach langen Diskussionen festgelegt. Bis heute ist der Cut Off aber umstritten, da wegen dem 2. Weltkrieg viele Wissenschaftler nicht teilnehmen konnten.
- D** ☐ Der Begründer der modernen Statistik, R. Fischer, hat die Grenze simuliert und berechnet. Dadurch ergibt sich dieser optimale Cut-Off.
- E** ☐ 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.

37 Aufgabe

(2 Punkte)

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

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

38 Aufgabe

(2 Punkte)

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

H_0 beibehalten obwohl die H_0 falsch ist

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

39 Aufgabe

(2 Punkte)

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

- A ☐ 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.
- B ☐ Die Teststatistik. Durch den Vergleich von T_c zu T_k ist es möglich die H_0 abzulehnen. Die Relevanz ergibt sich aus der Fläche rechts vom dem T_c -Wert.
- C ☐ Über das Konfidenzintervall. Das Konfidenzintervall beitet eine Entscheidung über die Signifikanz und zusätzlich kann über die Visualisierung des Konfidenzintervalls eine Relevanzschwelle definiert werden.
- D ☐ Über das Konfidenzintervall. Das Konfidenzintervall inkludiert eine Entscheidung über die Relevanz und zusätzlich kann über die Visualisierung des Konfidenzintervalls eine Signifikanzschwelle vom Forschenden definiert werden.
- E ☐ 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.

40 Aufgabe

(2 Punkte)

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

- A ☐ Wir vergleichen mit dem p-Wert und dem Signifikanzniveau α absolute Werte auf einem Zahlenstrahl und damit den Unterschied der Teststatistiken, wenn die H_0 gilt.
- B ☐ Wir machen ein Aussage über die Flächen und zwischen den Kurve der Teststatistiken der Hypothesen H_0 und H_A , wenn die H_0 gilt. Dabei werden Wahrscheinlichkeiten verglichen, die durch die Flächen unter der Kurve repräsentiert werden.
- C ☐ Wir machen eine Aussage über die individuelle Wahrscheinlichkeit des Eintretens der Nullhypothese H_0 . Der p-Wert wird mit dem Signifikanzniveau verglichen und bewertet.
- D ☐ Wir machen ein Aussage über die Flächen und der Kurve der Teststatistik, wenn die H_0 gilt. Dabei werden Wahrscheinlichkeiten verglichen, die durch die Flächen unter der Kurve repräsentiert werden.
- E ☐ 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.

41 Aufgabe

(2 Punkte)

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

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

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

43 Aufgabe

(2 Punkte)

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

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

44 Aufgabe

(2 Punkte)

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

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

45 Aufgabe

(2 Punkte)

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

- A** ☐ Anhand des p-Wertes lässt sich wie folgt eine Entscheidung treffen. Liegt der Wert in dem Signifikanzniveauintervall α dann kann die Nullhypothese abgelehnt werden.
- 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** ☐ Ist T_D höher als der kritische Wert $T_{\alpha=5\%}$ dann wird die Nullhypothese H_0 abgelehnt.
- E** ☐ Anhand des p-Wertes lässt sich wie folgt eine Entscheidung treffen. Liegt der Wert über oder gleich dem Signifikanzniveau α dann kann die Nullhypothese abgelehnt werden.

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** ☐ Es gibt ein statistisches Hypothesenpaar mit der Hypothese für und gegen die wissenschaftliche Fragestellung. Die Hypothesen werden H_{pro} und H_{contra} bezeichnet.
- B** ☐ Es gibt ein statistisches Hypothesenpaar mit der Nullhypothese H_0 und der Alternativhypothese H_A oder H_1 .
- 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 - bedingt durch das Falsifikationsprinzip - ein Set von k Nullhypothesen, die iterative gegen $k - 1$ Alternativhypothesen getestet werden.
- E** ☐ Mit der Nullhypothese H_A und der Alternativhypothese H_0 gibt es zwei Hypothesen, die aber selten genutzt werden.

Statistische Tests für Gruppenvergleiche

47 Aufgabe

(2 Punkte)

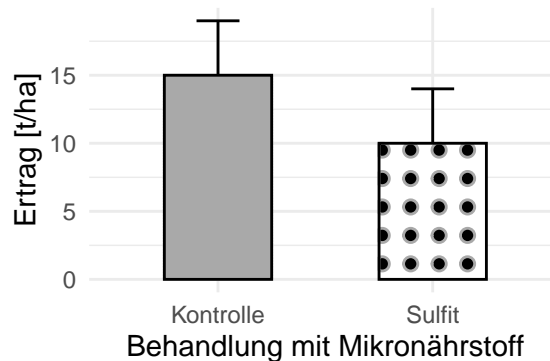
Nach einem Feldexperiment wollen Sie zwei Gruppen mit einem Welch t-Test vergleichen. Welche Aussage ist auch für den Student t-Test richtig?

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

48 Aufgabe

(2 Punkte)

In einer Studie zur Bewertung der Wirkung des Mikronährstoff Sulfite auf den Ertrag in t/ha von Papaya im Vergleich zu einer Kontrolle entstand folgende Abbildung. Der Versuch wurde in 14 Parzellen pro Gruppe durchgeführt. Welche Aussage ist im Bezug auf einen t-Test richtig?



- A** ☐ Der Test deutet auf einen signifikanten Unterschied hin. Der Effekt liegt vermutlich bei -5.
- B** ☐ Die Barplots deuten auf keinen signifikanten Unterschied. Der Effekt liegt vermutlich bei -5. Wir müssen aber einen Posthoc-Test rechnen um den Effekt wirklich bestimmen zu können.
- C** ☐ Der Test deutet auf kein signifikanten Unterschied hin. Der Effekt liegt vermutlich bei -5.
- D** ☐ Die Barplots deuten auf einen signifikanten Unterschied. Der Effekt liegt vermutlich bei -5 unter einer groben Abschätzung. Wir müssen aber eine ANOVA rechnen um den Effekt wirklich bestimmen zu können.
- E** ☐ 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.

49 Aufgabe

(2 Punkte)

Sie rechnen einen gepaarten t-Test, da Ihre Beobachtungen verbunden sind. Welche der folgenden Aussagen ist richtig?

- A** ☐ 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.
- B** ☐ Der gepaarte t-Test wird genutzt, wenn die Differenzen der Beobachtungen verbunden sind und wir dadurch die Unabhängigkeit nicht mehr vorliegen haben.

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

50 Aufgabe

(2 Punkte)

Nach einem Experiment mit fünf Weizensorten ergibt eine ANOVA ($p = 0.045$) einen signifikanten Unterschied für den Ertrag. Sie führen anschließend die paarweisen t-Tests für alle Vergleiche der verschiedenen Weizensorten durch. Nach der Adjustierung für multiples Testen ist kein p-Wert unter der α -Schwelle. Sie schauen sich auch die rohen, unadjustierten p-Werte an und finden hier als niedrigsten p-Wert $p_{3-2} = 0.053$. Welche Aussage ist richtig?

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


Deskriptive Statistik & Explorative Datenanalyse

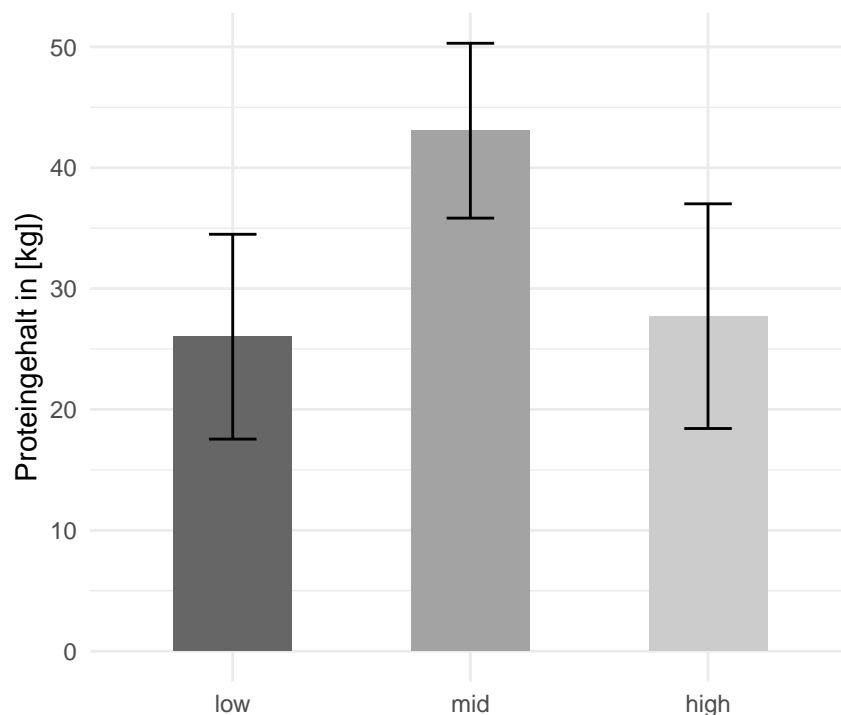
51 Aufgabe

(7 Punkte)


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



Eric steht vor einem ersten Problem, denn wenn es nach seiner Betreuerin geht, soll er in einem Feldexperiment Kartoffeln auswerten. Soweit eigentlich alles passend. Die Behandlung werden verschiedene Bewässerungstypen (*low*, *mid* und *high*) sein. In seiner Exceldatei wird er den Endpunkt (Y) *Proteingehalt* als *protein* aufnehmen. Vorab soll Eric aber einmal die folgenden Barplots seiner Betreuerin nachbauen, damit er den  Code schonmal für später vorliegen hat. Damit geht das Problem schon los.



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

1. Erstellen Sie eine Tabelle mit den statistischen Maßzahlen aus der obigen Abbildung der drei 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!



Anschauen, was andere vor einem gemacht haben, ist eine Möglichkeit schnell ans Ziel zu gelangen. Deshalb hat sich Tina viele Poster in der Fakultät angeschaut und ist zum Schluß gekommen, dass Barplots eine häufig genutzte Abbildung sind. Tina soll nun in ihrer Hausarbeit Lauch untersuchen. Die Behandlung in ihrer Hausarbeit sind verschiedene Düngestufen (*ctrl*, *low* und *high*). Erhoben wurden von Tina als Outcome (*Y*) *Frischegewicht* benannt als *freshmatter* in ihrer Exceldatei. Erwartungsgemäß erhält sie von ihrem Betreuer den Auftrag die erhobenen Daten als Barplots darzustellen. Dann kann Tina auch schonmal abschätzen, was bei einem statistischen Test rauskommen könnte.

treatment	freshmatter
high	22.1
ctrl	40.7
low	37.3
ctrl	35.5
low	27.3
high	27.9
low	28.3
high	10.9
high	31.9
low	22.7
low	21.5
ctrl	40.3

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

1. Zeichnen Sie in *einer* Abbildung die Barplots für die Behandlung von Lauch! 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 Lauch erwarten würden, wie sehen dann die Barplots aus? Antworten Sie mit einer Skizze der Barplots! **(1 Punkt)**

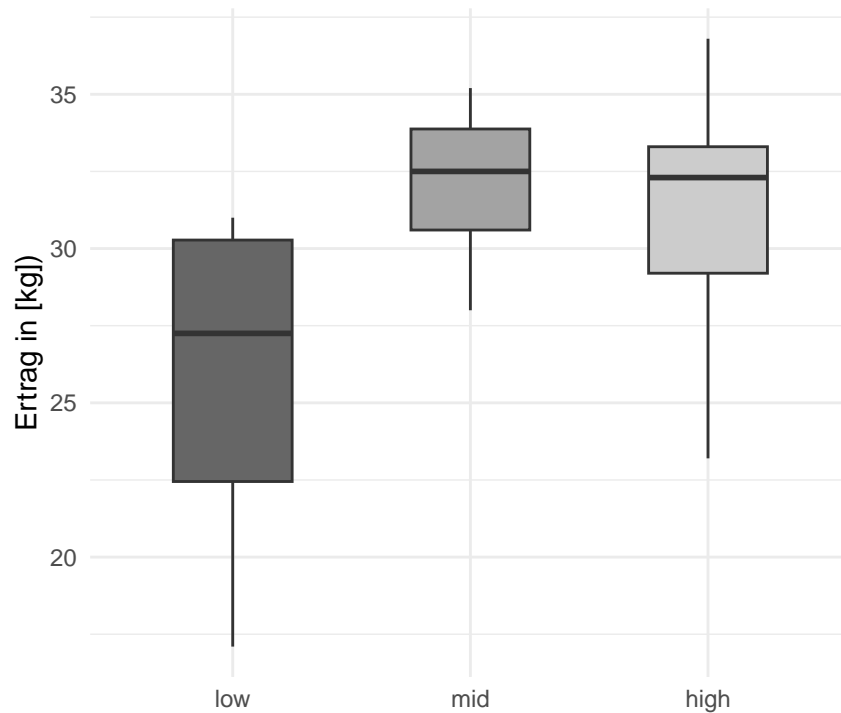
53 Aufgabe

(9 Punkte)

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



Jessica steht vor einem ersten Problem, denn wenn es nach ihrer Betreuerin geht, soll sie in einem Freilandversuch Kartoffeln auswerten. Soweit eigentlich alles passend. Die Behandlung werden verschiedene Bewässerungstypen (*low*, *mid* und *high*) sein. In ihrer Exceldatei wird sie den Endpunkt (*Y*) *Ertrag* als *yield* aufnehmen. Vorab soll Jessica aber einmal die folgenden Boxplots ihrer Betreuerin nachbauen, damit sie den R Code schonmal für später vorliegen hat. Anhand von Boxplots lässt sich eine Aussage über die Varianzhomogenität über die Behandlungsgruppen treffen. Damit geht das Problem schon los.



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

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



Nikolas steht vor einem ersten Problem, denn wenn es nach seiner Betreuerin geht, soll er in einem Versuch in einer Klimakammer Brokoli ausgewertet. Soweit eigentlich alles passend. Die Behandlung waren verschiedene Genotypen (AA und BB). In seiner Exceldatei hat er den Endpunkt (Y) *Trockengewicht* als *drymatter* aufgenommen. Nun soll Nikolas 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	26.5
BB	37.7
BB	29.9
AA	25.3
AA	19.2
BB	26.3
AA	29.0
AA	29.5
BB	24.5
AA	33.5
AA	18.4
AA	22.0
AA	31.8
BB	23.7
BB	24.5

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

55 Aufgabe

(8 Punkte)

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



Adele betrachtet die folgenden Daten nach einem Feldexperiment mit Maiss. In dem Experiment wurden die Mehltauspots gezählt. Nach der Meinung ihrer Betreuerin muss als erstes geschaut werden, wie diese verteilt sind. Also welcher statistischen Verteilung die Mehltauspots folgen. Dazu soll Adele ein Histogramm verwenden. Dann hätte man auch einen guten Überblick über den Messwert (Y).

Die Mehltauspots: 4, 6, 1, 2, 2, 2, 5, 2, 2, 4, 2, 3, 0, 3, 4, 7, 1, 5, 2, 2, 2, 6, 7, 3, 6, 3, 6, 7, 2, 4

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

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

56 Aufgabe

(8 Punkte)

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



In einem Gespräch mit seiner Betreuerin wird Manuel gebeten seine Daten aus einem Leistungssteigerungsversuch mit Fleischrinder in einem Histogramm darzustellen. In seinem Experiment hat er die mittleren auffälligen Hautflecken erst fotografiert und dann ausgezählt. Laut seiner Betreuerin soll das Histogramm helfen, die Verteilung der die mittleren auffälligen Hautflecken zu bestimmen.

Die mittleren auffälligen Hautflecken: 13, 10.1, 9.9, 12.4, 9.1, 5.2, 7, 9.5, 10.3, 8.6, 9.4, 9, 7.4, 10.2, 9.6, 10.7, 6, 10.5, 5.6, 10.3, 8.6, 10.2, 12.4, 8.3, 8.7

Leider kennt sich Manuel 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!



Sabine möchte gerne den Zusammenhang zwischen durchschnittlichen Niederschlag [ml/w] und Trockengewicht [kg/ha] im Kontext von Kartoffeln herausfinden. Hierfür hat Sabine ein Gewächshausexperiment im Wendland durchgeführt. Nach einigen unvorgesehenen Ereignissen hat sie es geschafft folgende Datentabelle zu erstellen. Nun stellt sich die Frage für sie, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Deshalb möchte Sabine als erstes eine explorative Datenanalyse durchführen.

Trockengewicht [kg/ha]	Durchschnittlichen Niederschlag [ml/w]
21.9	21.3
23.8	23.5
17.6	20.7
15.1	19.6
21.0	17.3
16.4	13.7
13.5	13.1
21.4	19.4
23.1	22.7
15.3	17.9
19.6	16.2
15.2	16.2

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

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



In ein Gewächshausexperiment im Oldenburger Land hatte Sabine sich zum einen die Behandlung Pestizideinsatz [ja/nein] und zum anderen die Messung Chlorophyllgehalt unter Zielwert [ja/nein] im Kontext von Lauch angeschaut. Nun steht Sabine vor dem Problem, dass sie zwei kategoriale Variablen in ihrer Abschlussarbeit gemessen hat. Da ihr Betreuer erstmal die langen Tabellen mit ja/nein in einer explorativen Datenanalyse zusammengefasst und präsentiert bekommen möchte bevor es überhaupt weitergeht, muss sie jetzt eine Lösung finden.

Pestizideinsatz	Chlorophyllgehalt unter Zielwert	Pestizideinsatz	Chlorophyllgehalt unter Zielwert
nein	nein	ja	ja
nein	ja	nein	ja
nein	nein	nein	nein
nein	nein	ja	ja
nein	nein	ja	ja
ja	ja	ja	ja
ja	ja	nein	ja
nein	ja	nein	nein
nein	nein	ja	ja
ja	nein	ja	ja
nein	ja	ja	ja
nein	nein	ja	ja
ja	nein	nein	nein
ja	ja	ja	ja
ja	ja	ja	ja
nein	nein	nein	nein
ja	nein	nein	ja

Leider kennt sich Sabine 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)**

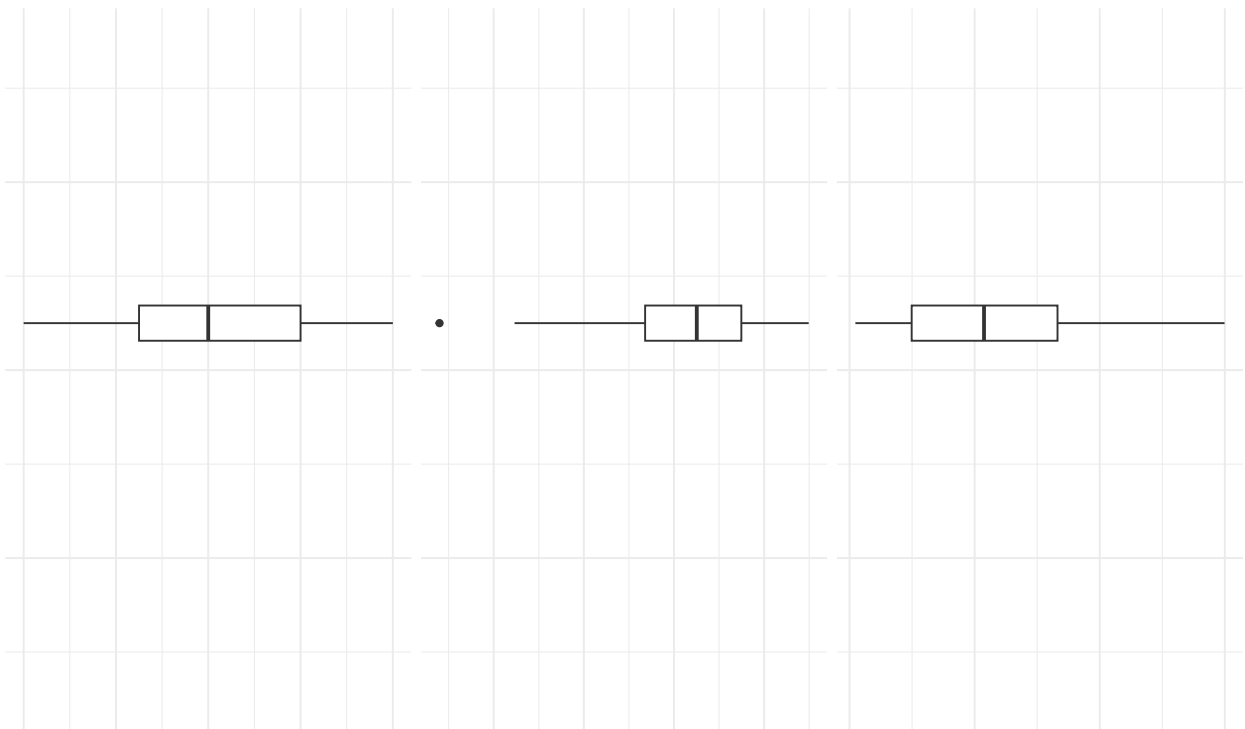
59 Aufgabe

(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 $\pm 2s$ 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 $s_1 = s_2$! **(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)**

$N(0, 1)$ und $N(4, 1)$



$Pois(15)$ und $Pois(1)$



62 Aufgabe

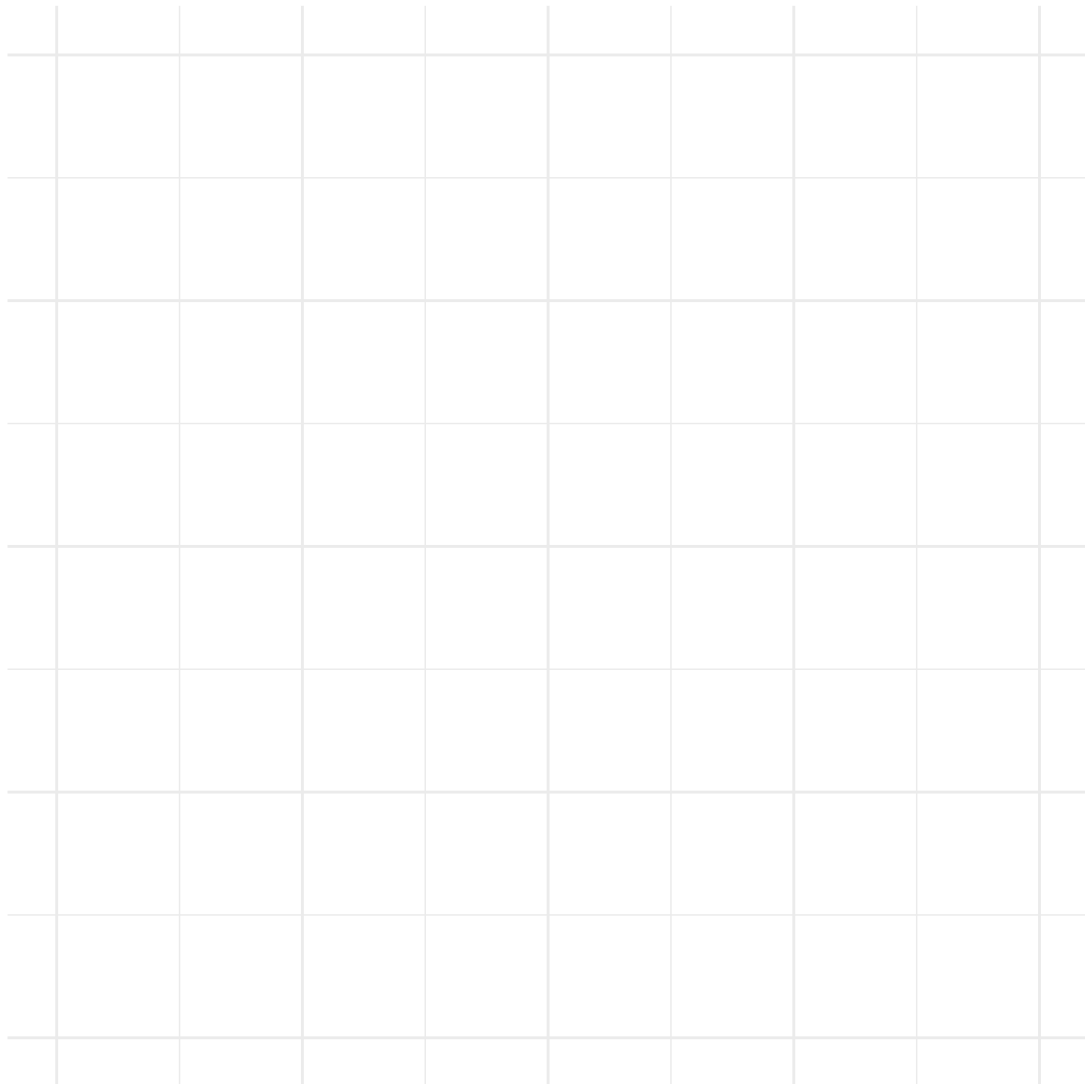
(8 Punkte)

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



Sie haben folgende Zahlenreihe y vorliegen $y = \{17, 16, 14, 21, 19, 19, 19\}$.

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)**



Statistisches Testen

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)**

H_0 falsch H_0 beibehalten α -Fehler 20%

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

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

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

65 Aufgabe

(9 Punkte)

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



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

1. Ergänzen Sie eine beschriftete x-Achse! **(1 Punkt)**
2. Ergänzen Sie „ $\bar{y}_1 = \bar{y}_2$ “! **(1 Punkt)**
3. Ergänzen Sie „ $A = 0.95$ “! **(1 Punkt)**
4. Zeichnen Sie $T_{\alpha=5\%}$ in die Abbildung! **(1 Punkt)**
5. Zeichnen Sie das Signifikanzniveau α in die Abbildung! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Zeichnen Sie $-T_D$ in die Abbildung! **(1 Punkt)**
7. Zeichnen Sie einen signifikant p-Wert in die Abbildung! Begründen Sie Ihre Antwort! **(2 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 signifikantes, relevantes 95% Konfidenzintervall
 - (b) Ein signifikantes, relevantes 90% Konfidenzintervall.
 - (c) Ein nicht signifikantes, nicht relevantes 95% Konfidenzintervall
 - (d) Ein signifikantes, nicht relevantes 95% Konfidenzintervall
 - (e) Ein 95% Konfidenzintervall mit höherer Fallzahl n in der Stichprobe als der Rest der 95% Konfidenzintervalle
 - (f) Ein 95% Konfidenzintervall mit niedriger Fallzahl n in der Stichprobe als der Rest 95% der 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)

Der t-Test

68 Aufgabe

(9 Punkte)

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



'Der t-Test testet einen normalverteilten Endpunkt (Y).', liest Mark laut. Das hilft jetzt auch nur bedingt weiter. Laut seiner Betreuerin ist zwar ihm Messwert Fettgehalt [%/kg] normalverteilt, aber wie rechnet er jetzt einen t-Test? Für seine Abschlussarbeit zum Testen einer neuen technischen Anlage musste er ein Stallexperiment mit Lamas im Wendland durchführen. Als wäre das nicht schon anstrengend genug gewesen bei dem anspruchsvollen Pilotprojekt mit sehr geringer Fallzahl ($n_1 = n_2 = 3$). Jetzt soll er auch noch testen, ob die Behandlung Ernährungszusatz (*ctrl* und *fedX*) ein signifikantes Ergebnis liefert.

treatment	weight
dose	14.0
dose	15.6
ctrl	17.9
dose	13.9
ctrl	21.2
ctrl	26.4

Leider kennt sich Mark 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. Bestimmen Sie die Teststatistik T_D eines Student t-Tests! **(3 Punkte)**
3. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Berechnen Sie den Effekt des Student 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!



Tina ist im Oldenburger Land für einen Versuch mit Zandern. Allein diese Tatsache ist für sie eine Erzählung wert. Für ihrer Hausarbeit musste sie einen Leistungssteigerungsversuch mit Zandern 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 Fettgehalt [%/kg] sollen mit einem t-Test ausgewertet werden. Immerhin weiß sie, dass ihr Messwert einer Normalverteilung folgt.

Ernährungszusatz	Fettgehalt
ctrl	13.9
fedX	44.8
fedX	27.3
ctrl	15.3
ctrl	25.8
fedX	42.7
fedX	40.4
ctrl	26.2
ctrl	23.9
fedX	50.1
fedX	36.0
fedX	38.2
ctrl	15.8
fedX	45.5
ctrl	16.3
fedX	38.5
ctrl	28.0
fedX	50.2

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

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

70 Aufgabe

(12 Punkte)

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



'Der t-Test testet einen normalverteilten Messwert (Y).', liest Sabine laut. Das hilft jetzt auch nur bedingt weiter. Laut ihrer Betreuerin ist zwar ihr Messwert Schlachtgewicht [kg] normalverteilt, aber wie rechnet sie jetzt einen t-Test? Für ihre Abschlussarbeit musste sie einen Leistungssteigerungsversuch mit Hühnern im Oldenburger Land durchführen. Als wäre das nicht schon anstrengend genug gewesen. Jetzt soll sie auch noch testen, ob die Behandlung Ernährungszusatz (*ctrl* und *fedX*) ein signifikantes Ergebnis liefert.

Ernährungszusatz	Schlachtgewicht
ctrl	46.0
ctrl	42.5
ctrl	46.1
fedX	37.9
fedX	37.9
ctrl	48.0
fedX	39.8
ctrl	42.8
ctrl	41.7
fedX	35.6
ctrl	39.9
ctrl	43.3
fedX	43.3
fedX	37.9
ctrl	52.3
fedX	41.9
fedX	42.3
ctrl	39.8
fedX	42.3

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Formulieren Sie das statistische Hypothesenpaar! **(1 Punkt)**
3. Bestimmen Sie die Teststatistik T_D eines Welch t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.96$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Berechnen Sie das 95% 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 Sabine über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

71 Aufgabe

(11 Punkte)

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



[1] "Als Jonas seiner Betreuerin zugehört hatte, klang das mit der Messung an der gleichen Beobachtung"

Alles voll mit Brokkoli. Das hat Jonas in seinen Projektbericht gemacht. Brokkoli. Worum ging es aber konkret? Jonas hat als einen normalverteilten Endpunkt (Y) Proteingehalt [g/kg] von Brokkoli bestimmt. Die Daten hat er in einem ein Freilandversuch erhoben. In dem Experiment ging es um eine vorher/nachher Untersuchung an den gleichen Brokkoli. Als Behandlung wurde Düngung (*vorher* und *nachher*) eingesetzt. Nach der Meinung seiner Betreuerin muss hier ein gepaarter t-Test gerechnet werden.

Das Schlachtgewicht von wölfen hängt maßgeblich von der Gewichtszunahme in den ersten sieben Tagen ab. Um einen Fütterungsversuch in verschiedenen Ställen zu planen würde ein Pilotversuch durchgeführt. Das Gewicht von wölfen wurde *vor* der Behandlung mit xxx und 1 Woche *nach* der Behandlung gemessen. Es ergibt sich die folgende Datentabelle.

ID	treatment	freshmatter
4	vorher	12.8
3	vorher	21.6
5	vorher	42.3
9	nachher	33.1
10	nachher	27.7
2	vorher	26.2
7	vorher	21.6
11	nachher	19.0
8	nachher	22.5
6	vorher	15.4
7	nachher	30.4
1	nachher	28.0
9	vorher	22.8
3	nachher	11.9
1	vorher	23.3
6	nachher	18.1
4	nachher	23.8
2	nachher	34.2
5	nachher	17.9
8	vorher	17.0


Leider kennt sich Jonas mit der Berechnung eines gepaarten 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 gepaarten t-Tests! **(3 Punkte)**
4. Treffen Sie mit $T_{\alpha=5\%} = 1.64$ eine Aussage zur Nullhypothese! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Schätzen Sie den p -Wert des gepaarten t-Tests ab! Begründen Sie Ihre Antwort mit einer Skizze! **(2 Punkte)**
6. Formulieren Sie eine Antwort an Jonas über das Ergebnis Ihrer statistischen Analyse! **(2 Punkte)**

72 Aufgabe

(10 Punkte)



Sie erhalten folgende  Ausgabe der Funktion `t.test()` zu einem Versuch mit Kartoffeln.


```
##  
## Two Sample t-test  
##  
## data: waterintake [ml/d] by infusion  
## t = -0.8546, df = 12, p-value = 0.4095  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## -5.679204 2.479204  
## sample estimates:  
## mean in group high mean in group trt1  
## 15.4 17.0
```

1. Formulieren Sie das statistische Hypothesenpaar! **(2 Punkte)**
2. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Skizzieren Sie eine Abbildung in der Sie T_D , $Pr(D|H_0)$, $A = 0.95$, sowie $T_{\alpha=5\%} = |2.18|$ einzeichnen! **(4 Punkte)**
4. Beschriften Sie die Abbildung entsprechend! **(1 Punkt)**
5. Berechnen Sie den Effekt der Behandlungen! **(1 Punkt)**

73 Aufgabe

(8 Punkte)



Sie erhalten folgende  Ausgabe der Funktion `t.test()` zu einem Versuch mit Erdbeeren.

```
##
## Two Sample t-test
##
## data: waterintake [ml/d] by infusion
## t = 1.828, df = 10, p-value = 0.09749
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -0.5628711 5.7057282
## sample estimates:
## mean in group mid mean in group low
##          16.57143          14.00000
```

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

74 Aufgabe

(9 Punkte)



Sie erhalten folgende  Ausgabe der Funktion `t.test()` zu einem Versuch mit Maiss.


```
##
##  Welch t-Test
##
## data:  waterintake [ml/d] by infusion
## t = -0.51903, df = 14, p-value = 0.6118
## alternative hypothesis: true  is not equal to [condensed]
## 95 percent confidence interval:
##  -6.028367  3.679161
## sample estimates:
## mean in group ctrl mean in group trt1
##           16.11111           17.28571
```

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

75 Aufgabe

(8 Punkte)



Sie erhalten folgende  Ausgabe der Funktion `t.test()` zu einem Versuch mit Erbsen.

```
##
## Paired t-test
##
## data: low and high
## t = -0.23625, df = 6, p-value = 0.8211
## alternative hypothesis: true is not equal to [condensed]
## 95 percent confidence interval:
## -3.244945 2.673517
## sample estimates:
## mean difference
## -0.2857143
```

1. Formulieren Sie das statistische Hypothesenpaar! **(2 Punkte)**
2. Liegt ein signifikanter Unterschied zwischen den Gruppen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Skizzieren Sie den sich ergebenden Datensatz mit $n = 4$ Beobachtungen! Die Daten müssen *nicht* die Mittelwertsdifferenz d erfüllen! **(2 Punkte)**
4. Skizzieren Sie den sich ergebenden Boxplot der Differenzen! Welche Annahmen an die Daten haben Sie getroffen? Begründen Sie Ihre Antwort! **(2 Punkte)**

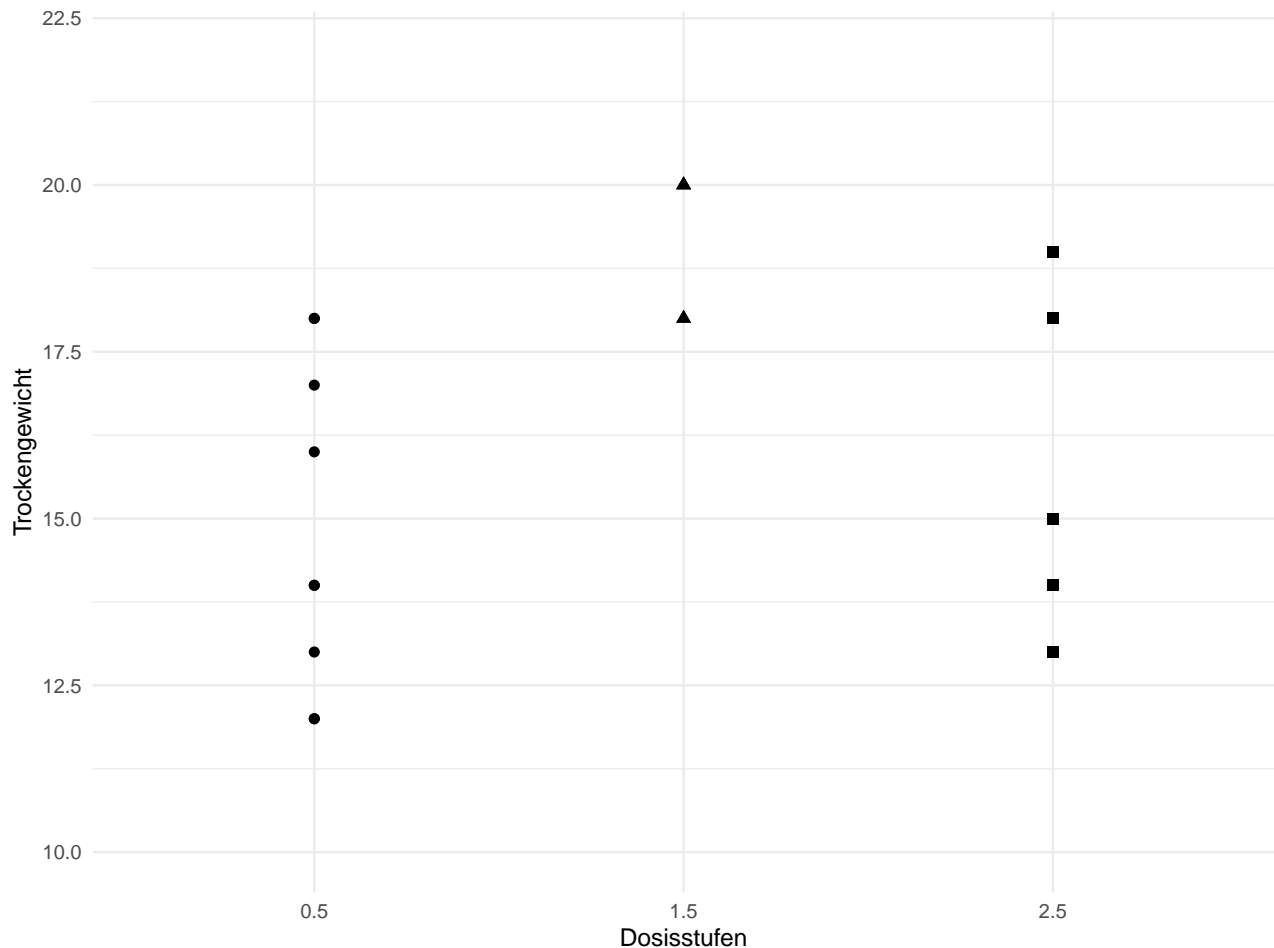
Die ANOVA

76 Aufgabe

(8 Punkte)



In einem Freilandversuch wurde der Ertrag von Kartoffeln unter drei verschiedenen Pestizid-Dosen 0.5g/l, 1.5g/l und 2.5g/l gemessen. Im Folgenden ist der Datensatz einmal visualisiert.



- Zeichnen Sie folgende statistischen Maßzahlen in die Abbildung ein! Beschriften Sie die statistischen Maßzahlen! **(6 Punkte)**
 - Globale Mittelwert: β_0
 - Mittelwerte der einzelnen Dosen: $\bar{y}_{0.5}$, $\bar{y}_{1.5}$ und $\bar{y}_{2.5}$
 - Effekt der einzelnen Dosen: $\beta_{0.5}$, $\beta_{1.5}$, und $\beta_{2.5}$
 - Residuen oder Fehler: ϵ
- Liegt ein *vermutlicher* signifikanter Unterschied zwischen den Dosisstufen vor? Begründen Sie Ihre Antwort! **(2 Punkte)**

77 Aufgabe

(9 Punkte)

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



Der Datensatz `week7_growth_tbl` enthält den Fettgehalt von Fleischrindern, die unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen erzielt wurden. Als Behandlung wurden verschiedene Nahrungszusätze in unterschiedlichen Dosen verfüttert. Als Behandlung haben Sie somit den Faktor *group* mit den Faktorstufen *pos*, *neg* und *low* vorliegen.

1. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle mit den gegebenen Informationen von **Df** und **Sum Sq** aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2				
error	20	67.46			
total	22	241.22			

2. Schätzen Sie den p-Wert der Tabelle mit der Information von $F_{\alpha=5\%} = 3.49$ ab. Begründen Sie Ihre Antwort! (2 Punkte)
3. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? (2 Punkte)
4. Skizzieren Sie Körpergröße von drei Tieren pro Behandlung für eine signifikante, einfaktorielle ANOVA! (2 Punkte)

78 Aufgabe

(12 Punkte)

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



Der Datensatz `fertilizer_growth_tbl` enthält das Trockengewicht der Kartoffeln, die unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen erzielt wurden. Dabei wurden die Kartoffeln unter verschiedenen Konzentrationen von einem alternativen Dünger angebaut. Als Behandlung haben Sie daher den Faktor *group* mit den Faktorstufen *high*, *ctrl* und *pos* vorliegen.

1. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle mit den gegebenen Informationen von **Df** und **Sum Sq** aus! (3 Punkte)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	75.7			
error	22	68.86			

2. Schätzen Sie den p-Wert der Tabelle mit der Information von $F_{\alpha=5\%} = 3.44$ ab. Begründen Sie Ihre Antwort! (2 Punkte)
3. Was bedeutet ein signifikantes Ergebnis in einer einfaktoriellen ANOVA im Bezug auf die möglichen Unterschiede zwischen den Gruppen? Beziehen Sie sich auf den obigen Fragetext bei Ihrer Antwort! (2 Punkte)
4. Berechnen Sie *einen* Student t-Test mit für den *vermutlich* signifikantesten Gruppenvergleich anhand der untenstehenden Tabelle mit $T_{\alpha=5\%} = 2.03$. Begründen Sie Ihre Auswahl! (3 Punkte)


group	n	mean	sd
high	7	15.14	0.90
ctrl	9	18.00	2.06
pos	9	14.00	1.94

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

79 Aufgabe

(9 Punkte)



Der Datensatz *plant_tbl* enthält das Outcome *height* für ein Experiment in einer Klimakammer mit Lauchstangen, welches unter drei verschiedenen Düngerbedingungen erzielt wurden. Die Düngerbedingungen sind in dem Faktor *trt* mit den Faktorstufen *A*, *B* und *C* codiert. Sie erhalten folgenden Output in .

```
## Analysis of Variance Table
##
## Response: height
##           Df Sum Sq Mean Sq F value    Pr(>F)
## trt         2  45.582   22.7911    6.9855 0.004479
## Residuals  22  71.778    3.2626
```

1. Stellen Sie die statistische H_0 und H_A Hypothese für die obige einfaktorielle ANOVA auf! **(2 Punkte)**
2. Interpretieren Sie das Ergebnis der einfaktoriellen ANOVA! **(2 Punkt)**
3. Berechnen Sie den Effektschätzer η^2 . Was sagt Ihnen der Wert von η^2 aus? **(2 Punkte)**
4. 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!



Der Datensatz *tooth_tbl* enthält Daten aus einer Studie zur Bewertung der Wirkung von Vitamin C auf das Zahnwachstum bei Meerschweinchen. Der Versuch wurde an verschiedenen Schweinen durchgeführt, wobei jedes Tier eine von 4 Vitamin-C-Dosen *dose* über eine von 2 Verabreichungsmethoden *supp* erhielt. Die Zahnlänge wurde als normalverteiltes Outcome gemessen.

1. Füllen Sie die unterstehende zweifaktorielle ANOVA Ergebnistabelle mit den gegebenen Informationen von **Df** und **Sum Sq** aus! **(4 Punkte)**
2. Schätzen Sie den p-Wert der Tabelle mit der Information von den $F_{\alpha=5\%}$ -Werten mit $F_{supp} = 4.26$ und $F_{dose} = 3.40$ sowie $F_{supp:dose} = 5.23$ ab. Begründen Sie Ihre Antwort! **(4 Punkte)**


	Df	Sum Sq	Mean Sq	F value	Pr(>F)
supp	1	28.27			
dose	3	88.25			
supp:dose	3	310.97			
error	32	258.01			

3. Was bedeutet ein signifikantes Ergebnis in einer zweifaktoriellen ANOVA im Bezug auf die möglichen Unterschiede zwischen den Gruppen? Beziehen Sie sich dabei einmal auf den Faktor *supp* und einmal auf den Faktor *dose*! **(2 Punkte)**
4. Was sagt der Term *supp:dose* aus? Interpretieren Sie das Ergebnis des abgeschätzten p-Wertes! **(2 Punkte)**


81 Aufgabe

(10 Punkte)



Der Datensatz *gain_weight_tbl* enthält Daten aus einer Studie zur Bewertung der Wirkung vom Vitamin Selen auf das Wachstum von Rindern. Der Versuch wurde an 62 Rindern durchgeführt, wobei jedes Tier eine von drei Selen-Dosen *dose* (0.2 ng/Tag, 2 ng/Tag und 15 ng/Tag) über eine von zwei Verabreichungsmethoden *form* erhielt (Wasser oder Festnahrung). Sie erhalten folgende Ausgabe in .

```
## Analysis of Variance Table
##
## Response: weight
##      Df Sum Sq Mean Sq F value    Pr(>F)
## dose      2 291.550  145.775  21.3889 1.753e-05
## form      1  18.676   18.676   2.7403  0.1152
## dose:form  2 230.537  115.269  16.9128 7.354e-05
## Residuals 18 122.678    6.815
```

1. Stellen Sie die statistische H_0 und H_A Hypothese für die obige zweifaktorielle ANOVA für den Faktor *dose* auf! **(2 Punkte)**
2. Interpretieren Sie das Ergebnis der zweifaktoriellen ANOVA. Gehen Sie im besonderen auf den Term *dose : form* ein! **(2 Punkte)**
3. Zeichnen Sie eine Abbildung, der dem obigen Ergebnis der zweifaktoriellen ANOVA näherungsweise entspricht! **(4 Punkte)**
4. Beschriften Sie die Abbildung entsprechend der  Ausgabe! **(2 Punkte)**

82 Aufgabe

(8 Punkte)



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

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

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

83 Aufgabe

(8 Punkte)



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

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

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

Multiple Gruppenvergleiche

85 Aufgabe

(8 Punkte)

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



In einem Experiment zur Dosiswirkung wurden verschiedene Dosisstufen mit einer Kontrollgruppe verglichen. Es wurden verschiedene t-Test für den Mittelwertsvergleich gerechnet und es ergab sich folgende Tabelle mit den rohen p-Werten (*Raw p-value*).

Vergleich	Raw p-value	Adjusted p-value	Reject Null
dose 20 - ctrl	0.012		
dose 35 - ctrl	0.760		
dose 50 - ctrl	0.080		
dose 10 - ctrl	0.001		


1. Füllen Sie die Spalte „Adjusted p-value“ mit den adjustierten p-Werten nach Bonferoni aus! **(4 Punkte)**
2. Entscheiden Sie, ob nach der Adjustierung die Nullhypothese weiter abgelehnt werden kann. Tragen Sie Ihre Entscheidung in die obige Tabelle ein. Begründen Sie Ihre Antwort! **(2 Punkte)**
3. Erklären Sie warum die p-Werte bei multiplen Vergleichen adjustiert werden müssen! **(2 Punkte)**

86 Aufgabe

(8 Punkte)

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



In einem Experiment für den Proteingehalt von Wasserlinsen in g/l mit fünf Dosisstufen (ctrl, low, mid, high und pos) erhalten Sie folgendes *Compact letter display* (CLD) als  Ausgabe aus den rohen, unadjustierten *p*-Werten.

```
## ctrl high low mid pos  
## "a" "b" "c" "d" "d"
```


1. Zeichnen Sie eine Abbildung, der sich ergebenden Barplots! **(2 Punkte)**
2. Ergänzen Sie das *Compact letter display* (CLD) zu der Abbildung! **(1 Punkt)**
3. Erklären Sie *einen* Vorteil und *einen* Nachteil des *Compact letter display* (CLD)! **(2 Punkte)**
4. Erstellen Sie eine Matrix mit den paarweisen *p*-Werten, die sich näherungsweise aus dem *Compact letter display* (CLD) ergeben würde! Begründen Sie Ihre Antwort! **(3 Punkte)**

87 Aufgabe

(10 Punkte)

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



In einem Experiment für den Zuckergehalt von Erdbeeren in g/kg mit vier Dosisstufen (ctrl, low, mid und high) erhalten Sie folgende Matrix als  Ausgabe mit den rohen, unadjustierten p -Werten.

```
##          ctrl      high      low      mid
## ctrl 1.0000000 0.0000025 1.18e-05 0.1279992
## high 0.0000025 1.0000000 0.00e+00 0.0002339
## low  0.0000118 0.0000000 1.00e+00 0.0000001
## mid  0.1279992 0.0002339 1.00e-07 1.0000000
```

Im Weiteren erhalten Sie folgende Informationen über die Fallzahl n , den Mittelwert $mean$ und die Standardabweichung sd in den jeweiligen Dosisstufen.

trt	n	mean	sd
ctrl	9	10.29	0.98
high	9	3.89	3.24
low	9	16.10	2.72
mid	9	8.54	1.94

1. Zeichnen Sie in eine Abbildung, die sich ergebenden Barplots! *Sortieren Sie dabei die Gruppen nach absteigender Effektstärke!* **(3 Punkte)**
2. Adjustieren Sie die rohen p -Werte nach Bonferroni. Begründen Sie Ihre Antwort! **(3 Punkte)**
3. Ergänzen Sie das *Compact letter display (CLD)* zu der Abbildung. Nutzen Sie dazu die rohen p -Werte! **(2 Punkte)**
4. Interpretieren Sie das *Compact letter display (CLD)*! **(2 Punkte)**

Der χ^2 -Test & Der diagnostische Test

88 Aufgabe

(11 Punkte)

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



Nach einem Feldexperiment ergibt sich die folgende 2x2 Datentabelle mit einem Pestizid (ja/nein) der Marke RoundUp, dargestellt in den Zeilen, und dem infizierten Pflanzenstatus (ja/nein) von Mais, dargestellt in den Spalten. Insgesamt wurden $n = 152$ Pflanzen untersucht.

	Erkrankt (ja)	Erkrankt (nein)	
Pestizid (ja)	56	21	
Pestizid (nein)	23	52	

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Ergänzen Sie die Tabelle um die Randsummen! **(1 Punkt)**
3. Berechnen Sie die Teststatistik eines Chi-Quadrat-Test auf der 2x2 Tafel! **(3 Punkte)**
4. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem $\chi^2_{\alpha=5\%} = 3.841$! Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Skizzieren Sie die χ^2 -Verteilung, wenn die H_0 wahr ist! Ergänzen Sie $\chi^2_{\alpha=5\%}$ und χ^2_{calc} in der Abbildung! **(2 Punkte)**
6. Berechnen Sie den Effektschätzer *Cramers V*! Interpretieren Sie den Effektschätzer! **(2 Punkte)**

89 Aufgabe

(7 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 $n = 161$ 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).

	>30 Ferkel (ja)	≤30 Ferkel (nein)	
Medikament (ja)			69
Medikament (nein)			92
	96	65	161

1. Ergänzen Sie die Felder innerhalb der 2x2 Kreuztabelle in dem Sinne, dass *ein* signifikanter Effekt zu erwarten wäre! **(2 Punkte)**
2. Erklären und Begründen Sie Ihr Vorgehen an der Formel des Chi-Quadrat-Tests mit

$$\chi^2 = \sum \frac{(O-E)^2}{E}.$$

Sie können dies an einem Beispiel erklären! **(2 Punkte)**

3. Was ist die Mindestanzahl an Beobachtungen je Zelle? Wenn in einer der Zellen weniger Beobachtungen auftreten, welchen Test können Sie anstatt des „normalen“ Chi-Quadrat-Tests anwenden? **(2 Punkte)**
4. Warum hat die obige Vierfeldertafel einen Freiheitsgrad von $df = 1$? **(1 Punkt)**


90 Aufgabe

(10 Punkte)



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

```
##           Infiziert
## Gruppe      yes no
## Kontrolle      4 12
## UV-bestrahlt  10  3
```

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

```
##
## Fisher's Exact Test for Count Data
##
## data:  mat
## p-value = 0.009221
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.01244701 0.70309883
## sample estimates:
## odds ratio
##  0.1102808
```

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

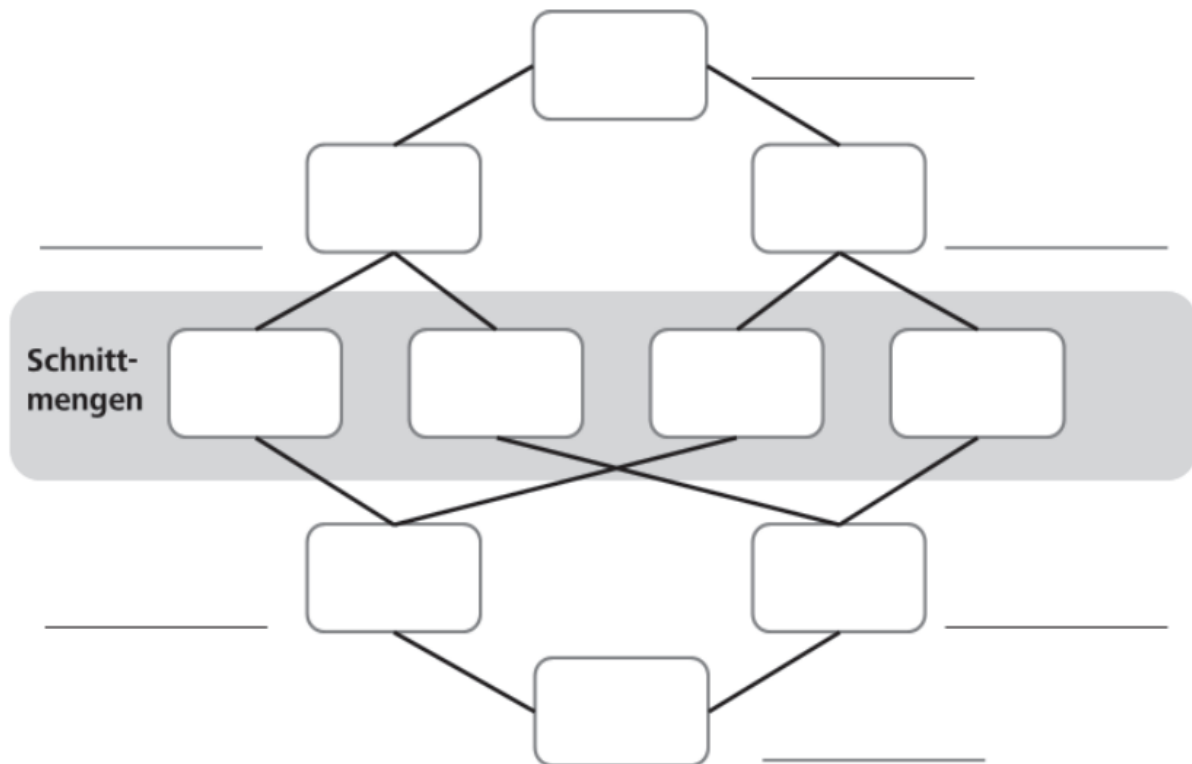
91 Aufgabe

(11 Punkte)



Die Prävalenz von Klauenseuche bei Wollschweinen wird mit 3% angenommen. In 85% der Fälle ist ein Test positiv, wenn das Wollschwein erkrankt ist. In 8.5% 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.

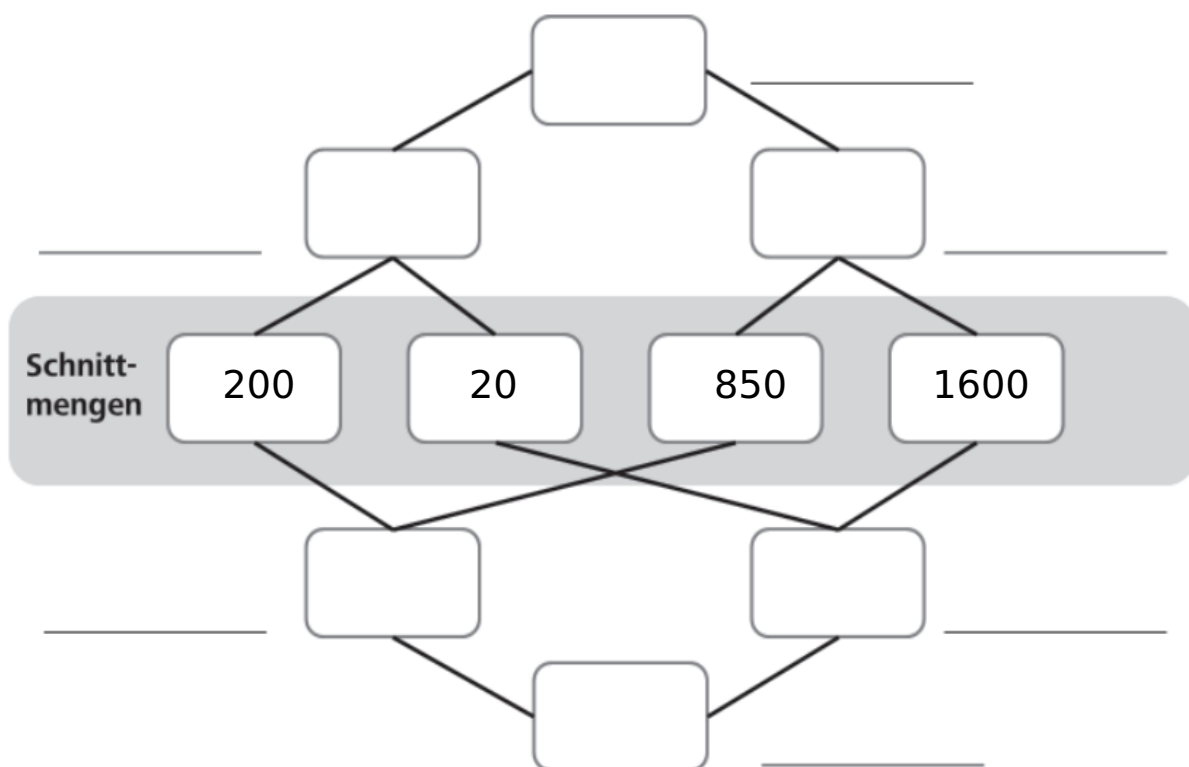
1. Füllen und beschriften Sie den untenstehenden Doppelbaum! Beschriften Sie auch die Äste des Doppelbaumes, mit denen Ihnen bekannten Informationen! **(8 Punkte)**
2. Berechnen Sie die Wahrscheinlichkeit $Pr(K^+|T^+)$! **(2 Punkte)**
3. Was sagt Ihnen die Wahrscheinlichkeit $Pr(K^+|T^+)$ aus? **(1 Punkt)**





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)**



Nicht parametrische Tests

93 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
9	10					
9	12					
12	13					
14	11					
10	14					
8	14					
10	15					
11	13					
13	13					
7	13					
7	11					
9	13					
9	13					
8	9					
5	12					

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

94 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
RootsGoneX	14.4
Kontrolle	12.0
RootsGoneX	13.6
RootsGoneX	12.6
Kontrolle	11.1
Kontrolle	7.9
Kontrolle	10.4
Kontrolle	10.7
Kontrolle	9.7
RootsGoneX	14.6
RootsGoneX	13.7
Kontrolle	11.5
RootsGoneX	14.2

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)

95 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 zusätzlichen Phosphorlösung (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	10.9			
Kontrolle	3.7			
Dosis 40	12.9			
Dosis 20	10.4			
Dosis 20	10.8			
Kontrolle	2.5			
Dosis 40	8.2			
Kontrolle	2.7			
Kontrolle	4.0			
Dosis 20	9.6			
Kontrolle	1.6			
Dosis 20	10.1			
Dosis 40	10.0			
Kontrolle	3.1			
Kontrolle	1.8			
Dosis 40	13.6			

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)**

Lineare Regression & Korrelation

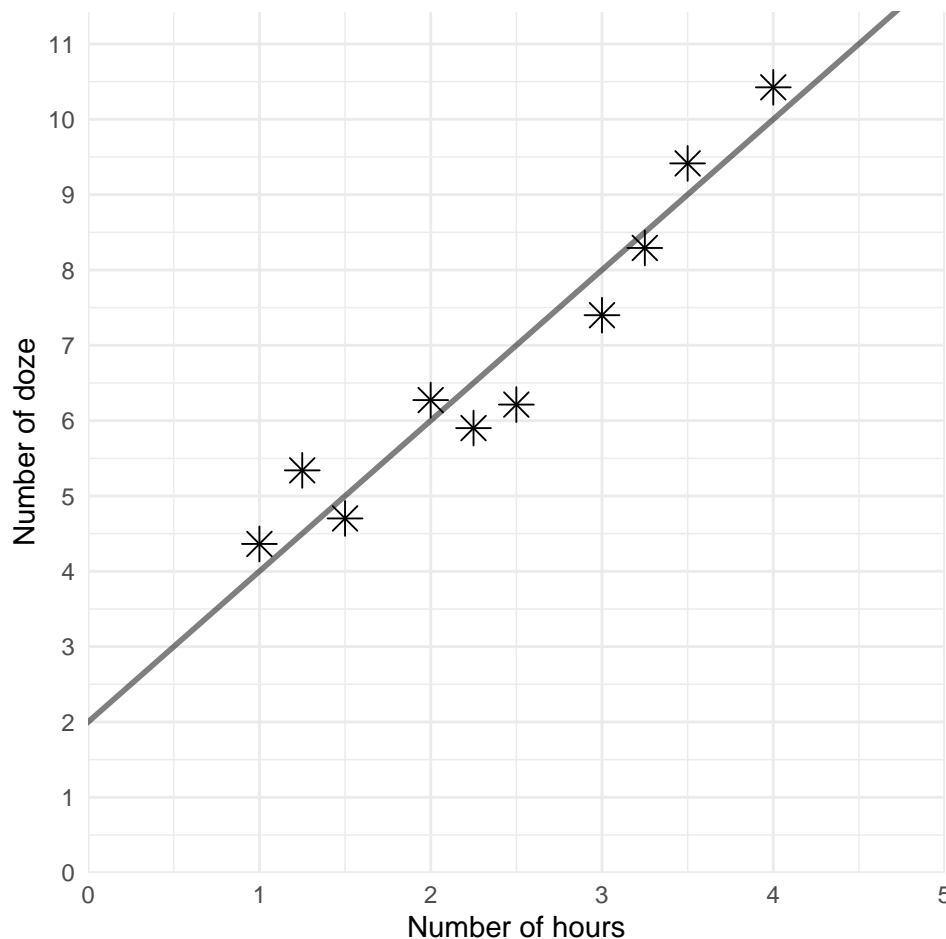
96 Aufgabe

(7 Punkte)

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



In einer Studie zur „Arbeitssicherheit auf dem Feld“ wurde gemessen wie viele Stunden auf einem Feld gefahren wurden und wie oft der Fahrer dabei drohte einzunicken. Es ergab sich folgende Abbildung.




1. Erstellen Sie die Regressionsgleichung aus der obigen Abbildung in der Form $y \sim \beta_0 + \beta_1 \cdot x$! **(2 Punkte)**
2. Beschriften Sie die Grade mit den Parametern der linearen Regressionsgleichung! **(2 Punkte)**
3. Liegt ein Zusammenhang zwischen der Anzahl an gefahrenen Runden und der Müdigkeit vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
4. Wenn kein Zusammenhang zu beobachten wäre, wie würde die Grade aussehen? **(1 Punkt)**

97 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)	1.91	2.17		
light	1.52	0.21		

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)**

98 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
## -4.7143 -1.1429  0.0714  1.7143  3.2857
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  15.1429     0.9897   15.300 3.10e-09
## waterB       8.5714     1.3997    6.124 5.15e-05
##
## Residual standard error: 2.619 on 12 degrees of freedom
## Multiple R-squared:  0.7576, Adjusted R-squared:  0.7374
## F-statistic: 37.5 on 1 and 12 DF, p-value: 5.147e-05
```

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 0.7576 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)**

99 Aufgabe

(9 Punkte)

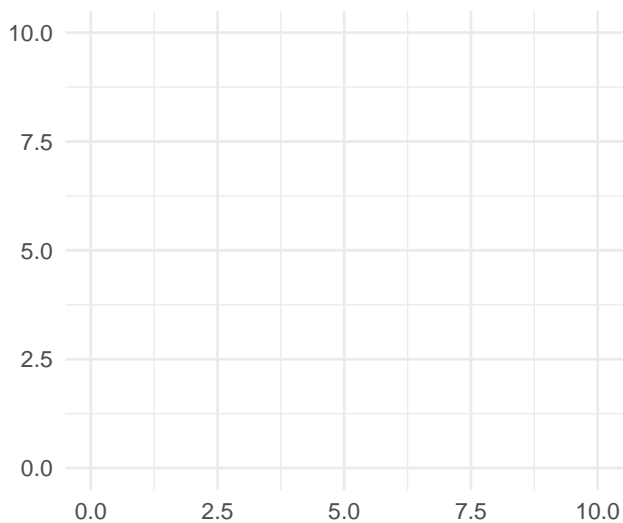


Im folgenden sehen Sie drei leere Scatterplots. Füllen Sie diese Scatterplots nach folgenden Anweisungen.

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)**

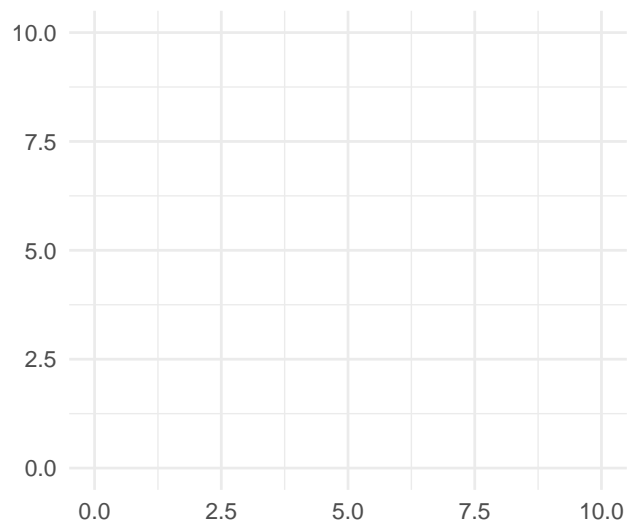
Pearsons $\rho = 0.25$

$R^2 = 0.5$



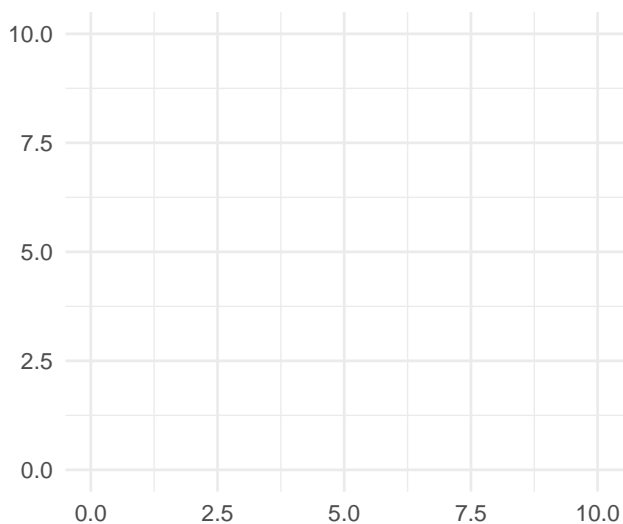
Pearsons $\rho = -0.25$

$R^2 = 0.25$



Pearsons $\rho = -0.75$

$R^2 = 1$



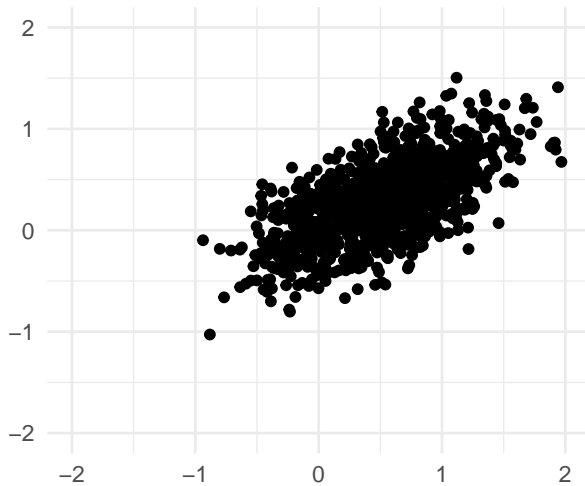


Im folgenden sehen Sie vier Scatterplots. Ergänzen Sie die Überschriften der jeweiligen Scatterplots.

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)**

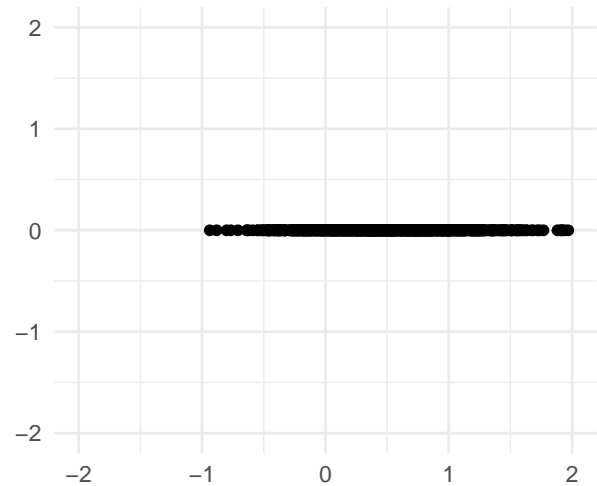
Pearsons $\rho =$

$R^2 =$



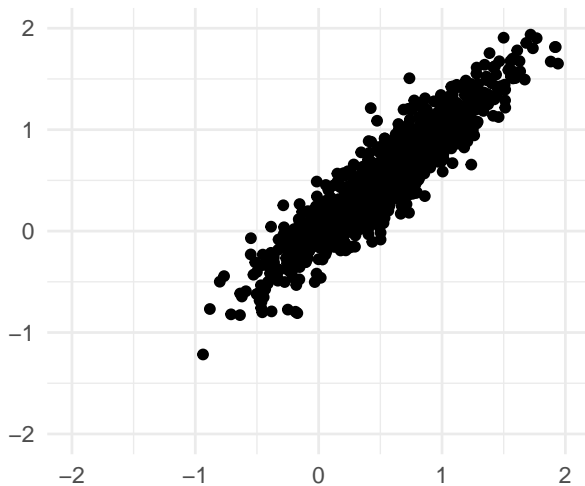
Pearsons $\rho =$

$R^2 =$



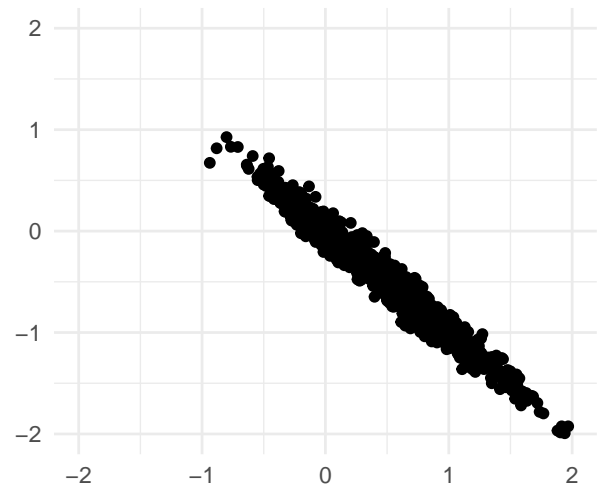
Pearsons $\rho =$

$R^2 =$



Pearsons $\rho =$


$R^2 =$



101 Aufgabe

(9 Punkte)



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

```
##  
## Pearson's correlation  
##  
## data: weight and fertilizer  
## t = -3.5146, df = 8, p-value = 0.007908  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.9450978 -0.2932913  
## sample estimates:  
## cor  
## -0.7790544
```

1. Formulieren Sie die wissenschaftliche Fragestellung! **(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 die Teststatistik und den p-Wert! Beschriften Sie die Abbildung! **(2 Punkte)**
5. Sind die Variablen `weight` and `fertilizer` normalverteilt? Begründen Sie Ihre Antwort! **(1 Punkt)**



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	28.2	14.9	31.7	
2	15.4	4.5	14.6	
3	34.8	15.7	33.1	
4	32.4	14.4	30.9	
5	25.8	11.6	26.2	
6	30.4	13.9	30.1	
7	23.1	9.9	23.5	
8	32.5	15.5	32.7	
9	25.0	10.1	23.8	
10	13.3	4.9	15.1	
11	12.8	3.1	12.3	

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)**

103 Aufgabe



(12 Punkte)

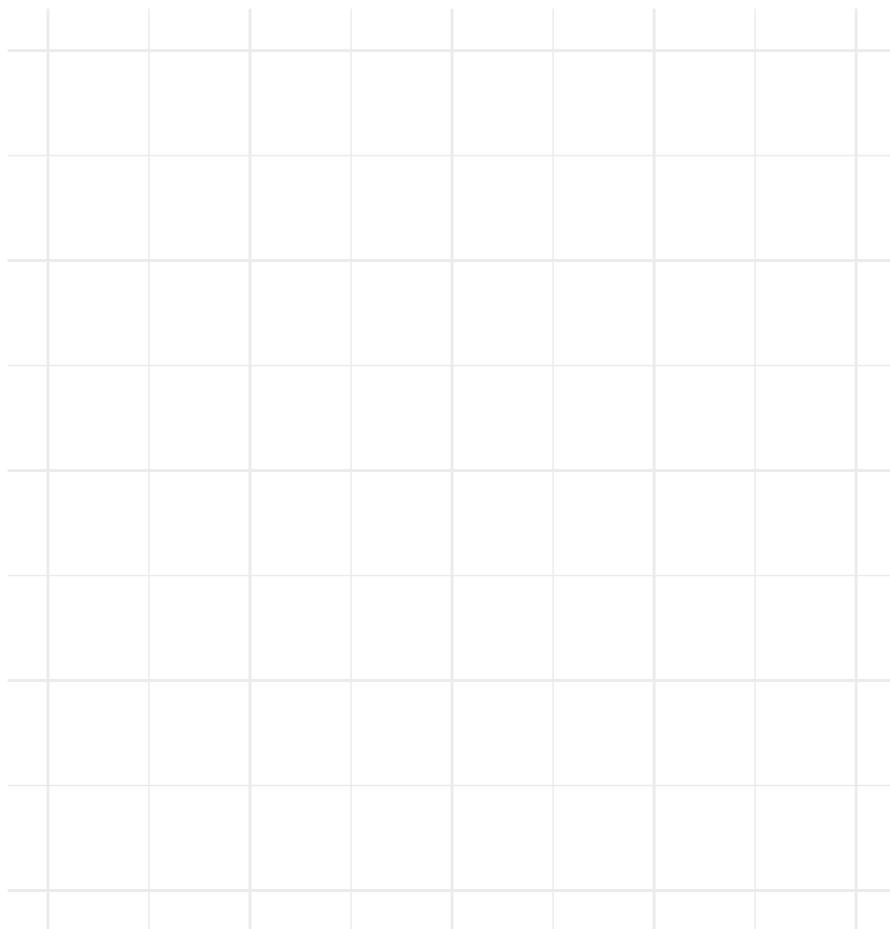


1. Zeichnen Sie in die drei untenstehenden, leeren Abbildungen die Zeile des Regressionskreuzes der Binomialverteilung. Wählen Sie die Beschriftung der y-Achse sowie der x-Achse entsprechend aus! **(6 Punkte)**
2. Ergänzen Sie die jeweiligen statistischen Methoden zu der Abbildung! **(2 Punkte)**
3.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



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

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



Experimentelles Design

105 Aufgabe

(10 Punkte)



Nach einem erfolgreichen Pilotversuch zur Wirksamkeit von Pestiziden bei Brokoli in einem Freilandversuch wollen Sie nun den Versuch eine Nummer größer anlegen. Dafür entscheiden Sie sich für ein faktorielles Versuchsdesign. In Ihrem Hauptversuch stellt die Wirksamkeit von Pestiziden den ersten Faktor mit insgesamt 4 Leveln dar. Der zweite Faktor mit dem Hitzestress beinhaltet 4 Level.

Im ersten Schritt überlegen Sie ein einfaches experimentelles Design zu probieren. Sie entscheiden sich für das *Randomized complete block design (RCBD)*.

1. Skizzieren Sie das *Randomized complete block design (RCBD)* für Ihren Versuch! **(4 Punkte)**
2. Skizzieren Sie eine Datentabelle für den Versuch mit drei Wiederholungen! **(2 Punkte)**
3. Skizzieren Sie eine Abbildung mit Barplots und einem angenommenen normalverteilten Outcome! **(4 Punkte)**

Mathematik

106 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 14m in Brusthöhe 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 180cm , eine Breite von 80cm sowie eine Länge von 230cm .

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

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

4. Berechnen Sie die Dicke der Eichenwand in cm ! Verdeutlichen Sie Ihre Berechnungen an einer aussagekräftigen Skizze für Pferd und Eiche! **(2 Punkte)**
5. Unter einer Dicke der Eichenwand von 25cm 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](#)

107 Aufgabe

(10 Punkte)

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



Von Töpfen auf Tischen In einem Experiment wollen Sie die Wuchshöhe von 180 Maispflanzen bestimmen. Bevor Sie überhaupt mit dem Experiment beginnen können, gibt es aber ein paar Abschätzungen über die Kosten und den Aufwand zu treffen. Zum einen müssen Sie die Maispflanzen einpflanzen und müssen dafür Substrat bestellen. Zum anderen müssen Sie die Maispflanzen auch bewegen und in ein Gewächshaus platzieren. Die Töpfe für die Keimung haben einen Durchmesser von 8cm und eine Höhe von 7cm. Der Kubikmeterpreis für Torf liegt bei 290 EUR.

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

108 Aufgabe

(10 Punkte)

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



Solar- & Biogasanlagen Um die Energiekosten Ihres Betriebes zu senken, wollen Sie eine Solaranlage auf den Schweinestall montieren lassen. Sie messen Ihren Stall und finden folgende Maße wieder. Die vordere Seite des Schweinestall hat eine Höhe h_v von $7m$. Die hintere Seite des Schweinestall hat eine Höhe h_b von $11m$. Der Schweinestall hat eine Tiefe t von $13m$ und eine Breite b von $70m$.

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

Ebenfalls planen Sie eine neue Biogasanlage für Ihren Betrieb. Der neue Methantank hat einen Radius r von $1m$. Leider gibt es ein paar bauliche Beschränkungen auf dem Grundstück. Ihr Fundament des zylindrischen Methantanks kann nur ein Gewicht von maximal $10t$ aushalten bevor der Tank wegbricht. Sie rechnen eine Sicherheitstoleranz von 25% ein beinhaltend das Gewicht des Methantanks. In flüssiger Form hat Methan bei $-80^\circ C$ eine Dichte von $235kg/m^3$. 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)**

109 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 einem Studenten im Karohemd. “Wieso?“, entfährt es Ihnen und Sie bereuen sogleich die Frage. Sofort werden Sie zu einem Whiteboard voller roter Schnüre geschliffen und müssen folgenden mathematischen untermauerten Argumenten im Lidl über sich ergehen lassen. Da kommen Sie nicht mehr raus, also können Sie auch gleich mitmachen. Das Problem liegt in Chile². Tja, die Deutschen und Südamerika.

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

1. Wenn 4 Blaubeerschalen 7.96 Euro kosten, wie viel kosten 8 Schalen? **(1 Punkt)**
2. Wenn Sie die 8 Blaubeerschalen gekauft haben, wie viele Aligatorbirnen zu je 1.79 EUR können Sie sich dann noch für 100 EUR leisten? **(1 Punkt)**

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

- Ein Kilo Strauchtomaten benötigt 160l Wasser. Eine Strauchtomate wiegt 100 - 130g.
 - Ein Kilo Salat benötigt 100l Wasser. Ein Salatkopf wiegt 280 - 490g.
 - Ein Kilo Avocado benötigt 980l Wasser. Eine Avocado wiegt 130 - 410g.
 - Ein Kilo Blaubeeren benötigt 830l Wasser. Eine Blaubeere wiegt 3.1 - 3.8g.
3. Berechnen Sie den Wasserverbrauch für die Produktion für jeweils eine Strauchtomate, einem Salat, einer Avocado und einer Blaubeeren. Stellen Sie das Ergebnis als Tabelle dar! **(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 reduzierte sich um 6.1%. Die Exporte für Avocados stiegen in dem gleichen Zeitraum um 17.3% auf 2.3×10^5 t.

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

Chile ist eines der wenigen Länder der Welt, die ihr Wasser komplett privatisiert haben. Derzeit sind nur drei Prozent des Wassers des Landes für den häuslichen Verbrauch vorgesehen. In den Dörfern der Anbauregionen versorgen Tankwagen die Bevölkerung jede Woche mit Wasser, es gibt etwa 50 Liter Wasser pro Kopf für den täglichen Bedarf. In *Deutschland* liegt der Verbrauch bei 10 - 15 Liter pro Minute Duschen und 9 - 14 Liter pro Spülgang.

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

Das alles hätten Sie nicht von einem Studenten im Karohemd 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.

110 Aufgabe

(12 Punkte)

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



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

Die Dampfnudelerde “Was für einen Unsinn!“, rufen Sie. Jetzt haben Sie kostbaren Schlaf prokrastiniert um einem Ernährungswissenschaftler auf YouTube über die Erde als Dampfnudel zu lauschen. Irgendwie passt es dann doch mit der Analogie. Übermüdet müssen Sie darüber nachdenken, warum vor 69 Millionen Jahren die Dinosaurier - so groß sie auch waren - nicht von der Schwerkraft zu Boden gerissen wurden. Hat der Dampfplauderer etwa recht und war die Schwerkraft vor Millionen von Jahren eine andere? Sind deshalb alle Lebewesen auf der Erde *heutzutage* so viel kleiner, weil die Schwerkraft größer ist als damals? War die Erde kleiner und hatte weniger Masse? Oder ist es nur ein Rechenfehler wie bei der Theorie der Hohlerde von Edmond Halley aus dem 17.-18. Jahrhundert? Müde reiben Sie sich die Augen. So wird es nichts mehr mit dem Schlafen, dann können Sie auch mal etwas 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 9.65 m/s^2 an. Im Weiteren hat die Erde einen ungefähren Durchmesser von $1.289 \times 10^4 \text{ km}$ und eine mittlere Dichte ρ von 5.86 g/cm^3 . Das Gewicht von einem heute lebenden afrikanischen Elefanten liegt bei 5t bis 7t und das Gewicht von einem Triceratops bei 6t bis 12t.

1. Welchen Durchmesser müsste die Erde vor 69 Millionen Jahren gehabt haben, wenn Dinosaurier und Elefanten die gleiche Gewichtskraft \vec{F}_G damals und heute erfahren hätten? *Beantworten Sie die Frage anhand der folgenden Teilaufgaben!*
 - a) Berechnen Sie die Fallbeschleunigung von vor 69 Millionen Jahren unter der obigen Annahme gleich wirkender Gewichtskraft \vec{F}_G auf Elefant und Dinosaurier! **(1 Punkt)**
 - b) Berechnen Sie Masse der heutigen Erde! **(2 Punkte)**
 - c) Schließen Sie über die Masse auf den Durchmesser der Erde vor 69 Millionen Jahren! **(2 Punkte)**
2. Beantworten Sie die Eingangsfrage mit 1-2 Antwortsätzen! **(1 Punkt)**

Die Distanz zwischen Sonne und Erde entspricht 1.05 astronomische Einheiten (AE). Die Einheit 1 AE wird mit $1.52 \times 10^8 \text{ km}$ angegeben. Der *massebehaftete* Sonnenwind besteht aus 79% Wasserstoffkernen mit einer molaren Masse von 1.08 g/mol , 9% Heliumkernen mit 3.92 g/mol sowie 12% weiteren Atomkernen mit 152.01 g/mol . Die Teilchendichte bei Eintritt in die Erdatmosphäre liegt zwischen 0.4 bis 100 Teilchen cm^{-3} pro Sekunde mit einer mittleren Teilchendichte von 8 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

111 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) sitting, (2) wingflapping, (3) foraging incl. scratching 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)
sitting	30cm; 16cm; 25.4%	45cm; 19cm; 54.1%	35cm; 27cm; 50.1%
wingflapping	31cm; 34cm; 0.6%	30cm; 28cm; 0.6%	35cm; 24cm; 0.9%
foraging incl. scratching	35cm; 32cm; 8.1%	31cm; 31cm; 8.1%	34cm; 30cm; 7.6%
dustbathing	36cm; 19cm; 7.3%	31cm; 20cm; 1%	28cm; 19cm; 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.](#)

112 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 50 Tagen die ersten Symptome ein; die ersten Toten sind nach 60 Tagen zu beklagen; nach 110 Tagen rafft die Skorbut eine ganze Schiffsbesatzung dahin«

Ferdinand Magellan stach im Jahre 1519 in See um eine Passage durch den südamerikanischen Kontinent zu finden. Zu seiner Flotte gehörten fünf Schiffe - das Flaggschiff Trinidad, die San Antonio, die Victoria, die Concepción und die Santiago - mit einer Besatzung von insgesamt 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 $5000\mu\text{g}/100\text{mg}$ Vitamin C. Der Bedarf liegt bei 105mg 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)**

113 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 $2 \times 10^{30} \text{ kg}$. 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. Ein Flugzeug und ein Lolli 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](#)

114 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 750 Jahren ist die Replikation abgeschlossen und wiederum *vier* Sonden werden ausgesendet. Gehen Sie von 7.81 Lichtjahren als mittlerer Abstand der Sterne in der Milchstraße aus. Es gibt 10^{11} Sterne in der Milchstraße. Nehmen Sie eine Lichtgeschwindigkeit von $2.7 \times 10^8 \text{ m/s}$ an.

1. Skizzieren Sie in einer Abbildung die ersten drei Schritte der Vervielfältigung der Sonden in der Galaxie! Beschriften Sie die Abbildung mit der Dauer und der Anzahl an Sonden für jeden Schritt der Vervielfältigung! **(4 Punkte)**
2. Berechnen Sie die theoretische Anzahl an Vervielfältigungsschritten die benötigt werden um mit *einem einzigen Vervielfältigungsschritt* die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
3. Berechnen Sie die Dauer, die eine außerirdische Zivilisation annäherungsweise benötigt um die gesamten Sterne der Milchstraße mit Sonden zu besuchen! **(2 Punkte)**
4. Bei einem vermuteten Alter der Erde von 4.1×10^9 Jahren, wie oft war dann eine Sonde einer außerirdischen Zivilisation schon zu Besuch? Korrigieren Sie Ihre Antwort mit dem Wissen, dass sich die Kontinentalplatten einmal alle 1.2×10^8 Jahre vollständig im Erdinneren umgewandelt haben! **(2 Punkte)**

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

115 Aufgabe

(10 Punkte)

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



Pyramiden bauen Es stehen die niedersächsischen Pyramidentage an und Sie sind auf abenteuerlichen Wegen für den Bau der Pyramiden zuständig. Zu allem Überfluss handelt es sich auch noch eine *Reenactment* Veranstaltung. Thema der diesjährigen Pyramidentage sind die Pyramiden von Meroe, die den Königen und Königinnen des historischen Reiches von Kusch in Nubien, dem heutigen Sudan, als Grabstätten dienten. Die Pyramiden in Meroe fallen durch ihren steilen Winkel von 72 Grad im Vergleich zu den ägyptischen Pyramiden mit 53 Grad auf. Die durchschnittliche Seitenlänge der Grundfläche einer Pyramide beträgt 44 Königsellen. Eine Königselle misst 52.2cm.

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

1. Bei der Königspyramide von Meroe soll eine Seitenlänge der Grundfläche 44 Königsellen lang sein. Welche Höhe der Königspyramide in m ergibt sich? **(1 Punkt)**
2. Die Außenflächen der Pyramide soll begrünt werden. Für die Bepflanzung muss eine 7cm dicke Torfschicht auf die Pyramide aufgebracht werden. Berechnen Sie die ungefähre Menge an benötigten Torf in m^3 ! **(2 Punkte)**

Wie in jedem guten *Reenactment* gibt es viel Oberschicht, aber nur 5 Sklaven, die Ihnen bei dem Befüllen der Pyramide mit Schutt zu Seite stehen. Leider haben Ihre Sklaven zu allem Überfluss auch noch chronische Schulterschmerzen entwickelt, als sie von der anstehenden Aufgabe erfahren haben. Gehen Sie daher von einer Effizienz der Sklaven von 90% aus. In eine Schubkarre passen 100 Liter.

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

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

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

116 Aufgabe

(12 Punkte)

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



Geocaching – Von Satelliten und Plastikdosen Es ist Wochenende und das Wetter ist *sweet*. Sie schwingen sich auf Ihr Cachermobil um mit 16km/h, geleitet von modernster Satellitentechnologie und einem Supercomputer aus dem Jahr 2000 in Ihren Händen, Plastikdosen in der Natur und an sehenswerten Orten zu finden. Sie wollen diesmal endlich die 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	GCGK00B	2.5 1.0 Normal
B	GCC1EHF	1.5 2.0 Klein
C	GCISAIC	3.0 3.5 Klein
D	GCOF1WS	2.0 3.0 Normal
E	GCIMSMS	4.5 4.5 Mikro

Im Weiteren sind Ihnen folgende Informationen zu den Entfernungen der Orte zugänglich. Der Entfernungsvektor \overrightarrow{AB} ist 5km. Im Weiteren ist Ihnen der Entfernungsvektor \overrightarrow{CB} mit 4.5km bekannt. Der Entfernungsvektor \overrightarrow{BE} ist das 1.3-fache des Entfernungsvektor \overrightarrow{CB} . Wenn Sie von dem Ort A den Ort C anpeilen, so liegt der Ort B ungefähr 25° nördlich. Wenn Sie von dem Ort C den Ort B anpeilen, so liegt der Ort D ungefähr 55° östlich. Vom Ort B betrachtet, bilden die Orte C und D einen rechten Winkel am Ort B. Der Ort B liegt auf gerader Linie zwischen den Orten C und E. Somit liegt der Ort E nördlich von B. Die Strecke zwischen A und E ist nicht passierbar. Sie starten an dem Ort C Ihre Cachertour.

1. Lösen Sie diese Aufgabe mit Hilfe einer aussagekräftigen Skizze der Orte und Caches. Bezeichnen Sie die Strecken und die Winkel Ihrer Skizze entsprechend! **(2 Punkte)**
2. Welche Strecke in km legen Sie bei der Bewältigung der 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

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

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 5m lang. Erreichen Sie einen Cache in der Höhe von 6.7m? 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](#).

117 Aufgabe

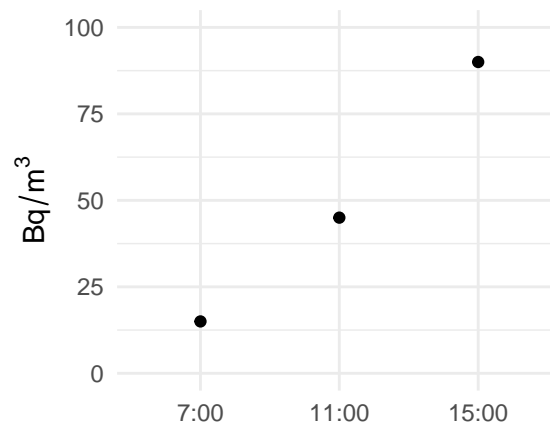
(10 Punkte)

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



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

Die atmende Wand und Brot aus Luft Als Kellerkind vom Dorf wollen das Ausmaß der Radonbelastung in ihrem Kellerzimmer bestimmen und lüften daher nicht. Spart dann auch Energie und lüften wird sowieso überschätzt. Während einer Messperiode von 7:00 Uhr bis 15:00 bestimmen Sie dreimal automatisch die Radonbelastung in Ihrem Kellerraum in Bq/m^3 . Es ergibt sich folgende Abbildung⁸.



1. Wie lange dauert es in Stunden bis Sie eine kritische Belastung von $380 Bq/m^3$ in Ihrem ungelüfteten Kellerraum erreicht haben? **(2 Punkte)**

Radon zerfällt mit einer Halbwertszeit von 3.5d zu Polonium. Polonium wiederum zerfällt mit einer Halbwertszeit von 180d zu Blei. Nur Radon und Polonium tragen zur radioaktiven Strahlenbelastung bei.

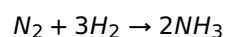
2. Wie lange dauert es in Stunden bis Ihre kritische Radonbelastung von $380 Bq/m^3$ auf unter $110 Bq/m^3$ gefallen ist? **(4 Punkte)**

Folgende Tabelle enthält die Informationen zur Zusammensetzung der normalen Umgebungsluft.

	Vol-%	M [g/mol]	ppm
Stickstoff	79.7	28.4	
Sauerstoff	20.45	15.8	
Kohlenstoffdioxid	0.035	12.1	

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

Während 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](#)

118 Aufgabe

(10 Punkte)

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



Armee der Finsternis Ihr Studentenjob war nach Ladenschluss bei Penny die Regale einzuräumen. Dabei ist Ihnen in der Auslage der Sonderangebote das 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 1040 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 $12m/s$ bilden.

1. Wie hoch müssen Sie den Schrottkugelturm bauen lassen, damit sich runde Bleikugeln durch die Fallgeschwindigkeit von $12m/s$ bilden? **(3 Punkte)**

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

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

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

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

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

5. Wie groß in cm^2 ist Ihr quadratisches Sieb am oberen Ende des Turms, wenn Sie pro Fall ca. 700 Bleikugeln produzieren wollen und die Bleikugel im Fall $0.8cm$ Abstand haben müssen? **(1 Punkt)**

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

119 Aufgabe

(10 Punkte)

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



Armee der Kaninchen Leider hat es mit Ihrem Krokodilreservat 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 1860 ungefähr 32 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.

$$f(t) = 10^{10} - 1.2 \times 10^9 \cdot 2^{-0.3 \cdot t + 2.7}$$

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

Um den Kaninchen Einhalt zu gebieten wurde das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) in 20 Kaninchen ausgebracht. Da die Kaninchen keine Maßnahmen gegen die Ausbreitung vornehmen können, verläuft die Ausbreitung mit einem wöchentlichen Wachstumsfaktor von 2 nach folgender Formel.

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

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

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

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

Thomas Austin entließ die Kaninchen im äußersten Westen von Australien. Australien hat eine West-Ost-Ausdehnung von 4000km und eine Nord-Süd-Ausdehnung von knapp 3600km. Die Kaninchen breiten sich radial mit einer Geschwindigkeit von 7.3km pro Jahr aus.

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

Eine jährliche Impfung gegen das Myxoma Virus und das Rabbit Haemorrhagic Disease Virus (RHDV) kosten 10\$ pro Tier und der durchführende Arzt verlangt ca. 35\$ pro Tier.

6. In Ihrem Stall leben 900 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...](#)

120 Aufgabe

(12 Punkte)

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



Lüneburger Heide. Unendliche Weiten. Wir schreiben das Jahr 2024. Dies sind die Abenteuer des Schafs Frida und Ihnen. Grünes Gras unter Ihren Füßen und ein strammer Wind im Gesicht, egal wohin Sie schauen. Ein schmatzendes Geräusch ertönt unter Ihnen. Sie sinnieren, sollten Sie Ihre weiten Graslandschaften jetzt schon düngen? Dafür benötigen Sie die *Grünlandtemperatur*! Die Grünlandtemperatur (GLT) ist die Summe aller positiven Tagesmitteltemperaturen seit Jahresbeginn. Ab einer GLT von 200° kann mit der Stickstoffdüngung begonnen werden. Sie sehen nicht ein, Geld für einen Agrarmetrologen zu bezahlen. Also rechnen Sie mit folgenden Informationen zu Monatsmultiplikatoren des GLT-Wertes: Januar mit $0.6\times$, Februar mit $0.7\times$ und März mit $1.1\times$. Sie haben noch im letzten Jahr folgende Temperaturen gemessen.

Datum	C°
01. Jan 2023	0.3
01. Feb 2023	1.1
01. Mrz 2023	3.1
01. Apr 2023	5.8

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

Auf dem Weg zu Ihrer Boskoopplantage wurden Sie mit Ihrem Trecker von einer Gruppe elektrifizierter Rentner abgedrängt. Der Trecker muss wieder aus dem Graben! Frida und die elektrifizierten Rentner ziehen an zwei, separaten Seilen. Dabei zieht Frida mit 180N . Die elektrifizierten Rentner bringen eine Kraft von 140N auf.

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

5. Im ersten Versuch legen Sie das Seil für 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 1.5t schweren Trecker *jeweils* aus dem Graben, wenn $F = m \cdot a$ gilt? **(1 Punkt)**

121 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, 8-mal-gefaltete, 0.7mm, 50-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 50cm und eine Breite von 23cm. Die Kartonagenmaschine in Flint soll dann einen quadratischen Eckenausschnitt der Länge x falzen.

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

Rektor Skinner möchte sich gerne wieder in seinem Vorgarten aufhalten und nicht die ganze Zeit von Bart mit Erdnüssen beworfen werden. Deshalb möchte er einen geräumigen Teil seines Vorgartens einzäunen. Ein Teil der Umzäunung bildet seine Vorderhauswand. Wegen Lieferschwierigkeiten stehen Rektor Skinner nur 110m Zaun zu Verfügung. Sie wollen nun die maximale Fläche des abgeschirmten Vorgartens in Abhängigkeit der Seitenlängen bei der Verwendung von 110m 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)**

122 Aufgabe

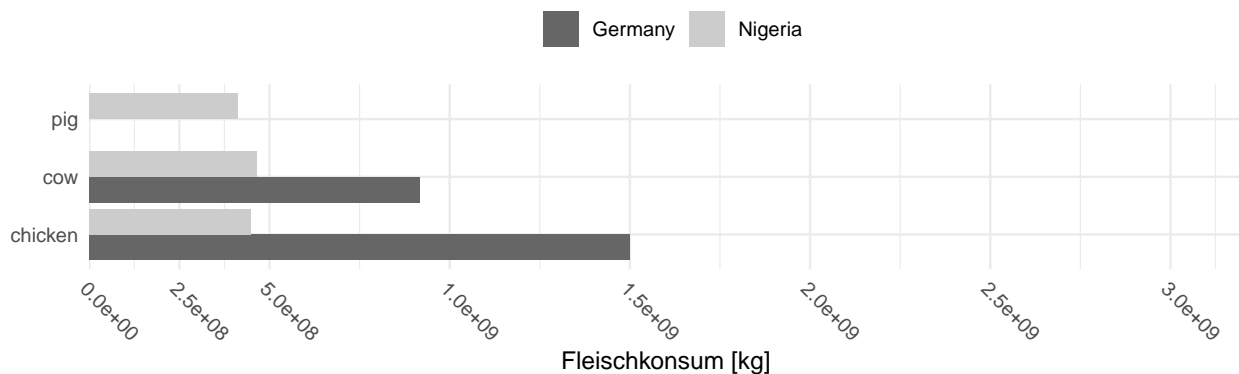
(10 Punkte)

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



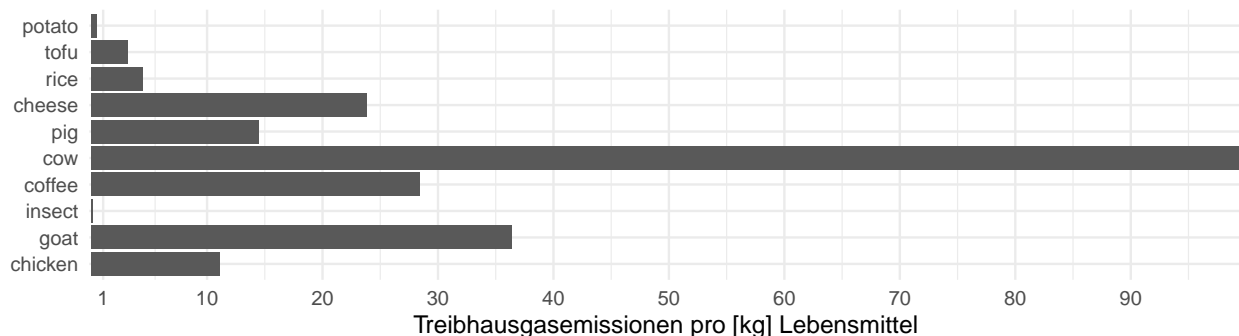
Ein Pfund Insekten, bitte! Das wird wohl häufiger gehört werden, wenn wir die Menschheit mit Proteinen ausreichend ernähren wollen¹². Schauen wir uns dazu einmal den Vergleich Deutschland zu Nigeria an. Nigeria hat eine der am schnellsten wachsenden Bevölkerungen der Welt und wird vermutlich im Jahr 2100 zu den Top 5 der bevölkerungsreichsten Länder zählen. Im Jahr 2020 leben ca. 8.2×10^7 Menschen in Deutschland und ca. 1.79×10^8 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 2020 in Deutschland und Nigeria in [kg] einmal dargestellt.



1. Stellen Sie den Fleischkonsum in Deutschland und Nigeria im Jahr 2020 *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)**

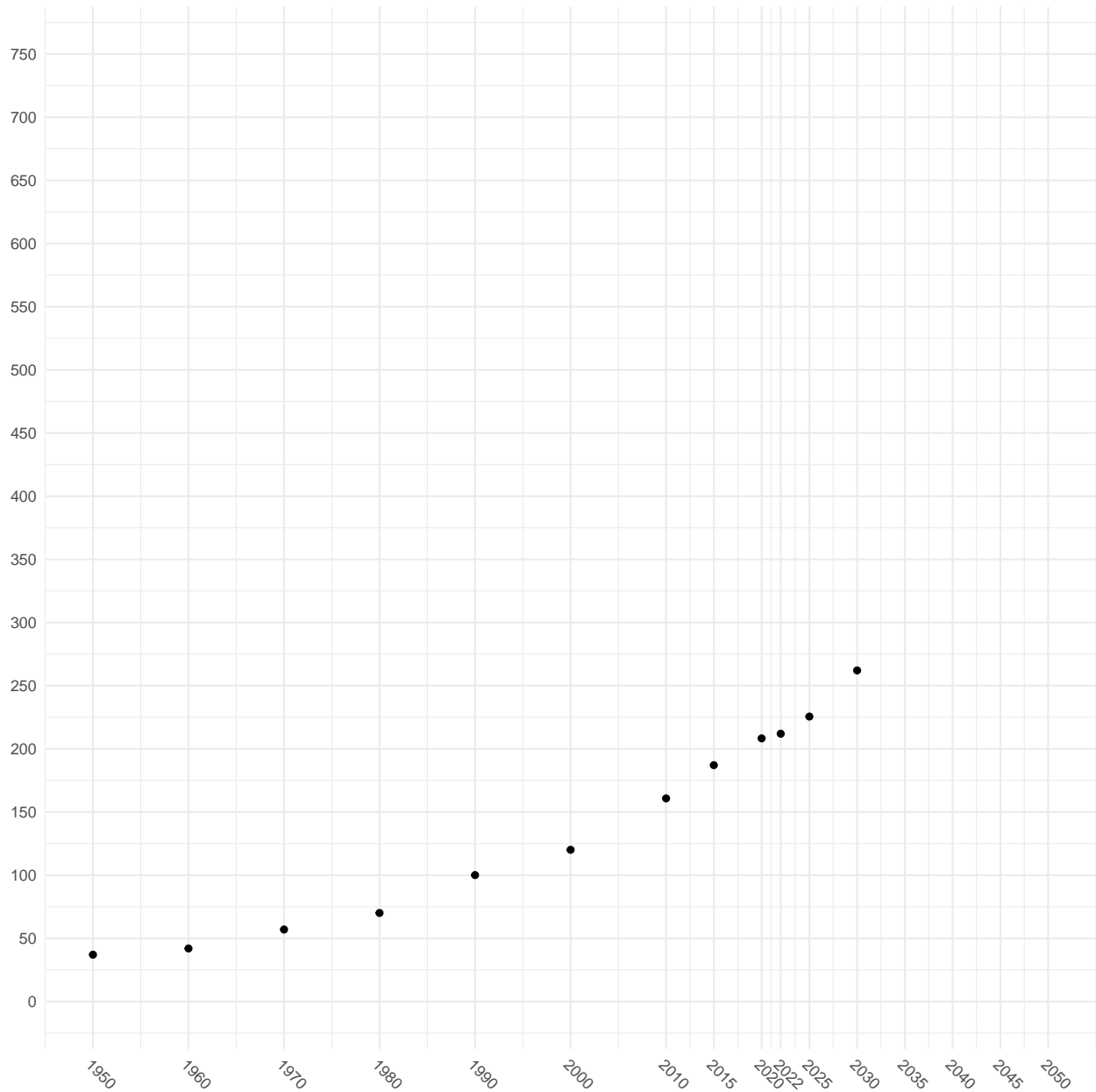
In der nächsten Abbildung finden Sie die CO₂-Emission in [kg] nach Lebensmittel, die durch die Produktion entsteht, abgebildet.



3. Stellen Sie in einer Tabelle die Treibhausgasemissionen an CO₂ pro Kopf, die durch den Fleischkonsum in Deutschland und Nigeria im Jahr 2020 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 70%-iger Angleichung der Lebensbedingungen zu Deutschland im Jahr 2020! **(1 Punkt)**
6. Berechnen Sie die prozentuale Steigerung der Treibhausgasemissionen an CO_2 in Nigeria im Jahr 2050 im Vergleich zum Jahr 2020, 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)**

123 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 HNO-Arzt und wollen einen AIDS-Test machen lassen. Woanders leider keinen Termin gekriegt... Immerhin denken Sie und Ihre Partnerin über Nachwuchs nach und da geht es eben nur durch ungeschützten Sex. Was wissen Sie nun aber über AIDS und dem diagnostischen AIDS-Test, den Sie nun machen werden?

Die Prävalenz von AIDS bei einem Menschen in Europa wird mit 0.75% angenommen. In 90% 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 $n = 4 \times 10^4$ Patienten mit einem diagnostischen Test für AIDS aus. Sie nehmen sich also einen Kuli und fangen an auf der historischen Ausgabe des Spiegels zu 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

124 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 Healthy Herbs Manufacture International (HeHeMan). Das Unternehmen steigerte den Umsatz um rund 27 Prozent von 310 Millionen Euro im Jahr 2021. Doch wie viel kommt bei den Partnern an? Laut HeHeMan habe das Unternehmen 2.8×10^5 aktive Partner.

1. Berechnen Sie zuerst den Umsatz der Firma HeHeMan 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 20%. 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 drei 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 HeHeMan Einheiten des Produkts für 2625EUR kaufen. Diese Einheiten können Sie nur direkt verkaufen. Leider mussten Sie den Kauf über einen Kredit über 5.5% 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](#)

125 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 Things* 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 7 vierseitige Würfel ($7d4$) zum würfeln in der Hand. Wenn Sie eine 4 würfeln, haben Sie einen Erfolg.

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

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

3. Würden Sie die Axt mit zwei sechseitigen Würfeln ($2d6$) als Schaden oder das Schwert mit einem achtseitigen Würfel plus 3 ($1d8+3$) 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 $Pr(A) = 0.7$, die Wahrscheinlichkeit für das Ereignis B , der Zauberwurf des Hexers ist erfolgreich, ist $Pr(B) = 0.75$. Sie haben mitgezählt und festgestellt, dass in 40 von 100 Fällen Ihr Rettungswurf bei einem erfolgreichen Zauber funktioniert hat.

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

126 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 Frida und Elke das Team der drei Kandidaten.

Name	$P(\text{win})$	$P(\text{outbid})$
Frida	0.4	0.076
Elke	0.1	0.12

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 0.043 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 einem einäugigen Piraten um das große Geld. Das Glücksrad hat 22 Felder. Sie drehen das Glücksrad zweimal. Auf 6 Feldern gewinnen Sie 4000EUR sonst 1500EUR. Ganz schön viel Geld und ganz schön zwielichtig hier...

3. Skizzieren Sie das Glücksrad und ergänzen Sie die Wahrscheinlichkeiten! **(1 Punkt)**
4. Zeichnen Sie das zugehörige Baumdiagramm für das zweimalige Drehen! Ergänzen Sie die Wahrscheinlichkeiten und die entsprechenden Ereignisse **(2 Punkte)**
5. Mit welcher Wahrscheinlichkeit gewinnen Sie 5500EUR? **(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)**

Angewandte Nutztier- und Pflanzenwissenschaften (M.Sc.)

Im Rahmen der Klausur zu dem Modul Biostatistik werden auch Fragen nur für die Studierenden des Schwerpunktes Nutztierwissenschaften gestellt. Im Folgenden daher eine lose Sammlung von möglichen Fragen zu diesem Themenkomplex.

127 Aufgabe

(6 Punkte)

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

128 Aufgabe

(8 Punkte)

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

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

mit

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

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

129 Aufgabe

(6 Punkte)

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