

**Name:** \_\_\_\_\_

*Nicht bestanden:* ☐

**Vorname:** \_\_\_\_\_

**Matrikelnummer:** \_\_\_\_\_

**Endnote:** \_\_\_\_\_

**B.Eng. Wirtschaftsingenieurwesen Agrar/Lebensmittel**

## **Klausur Statistik**

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

21. Januar 2025

### Erlaubte Hilfsmittel

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

### Endnote

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

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
83.0 - 87.0	1,0
79.0 - 82.5	1,3
74.5 - 78.5	1,7
70.5 - 74.0	2,0
66.0 - 70.0	2,3
61.5 - 65.5	2,7
57.5 - 61.0	3,0
53.0 - 57.0	3,3
49.0 - 52.5	3,7
43.5 - 48.5	4,0

Es ergibt sich eine Endnote von \_\_\_\_\_.

## Multiple Choice Aufgaben

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

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

- Es sind \_\_\_\_ von 20 Punkten erreicht worden.

## Rechen- und Textaufgaben

- Die Tabelle wird vom Dozenten ausgefüllt.

<b>Aufgabe</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>
<b>Punkte</b>	10	9	9	9	10	11	9

- Es sind \_\_\_\_ von 67 Punkten erreicht worden.

## 1 Aufgabe

(2 Punkte)

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

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

## 2 Aufgabe

(2 Punkte)

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

- A ☐ ... den Unterschied zwischen der Varianz durch verschiedene Behandlungsgruppen unter der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, kann kein Effekt  $\eta^2$  bestimmt werden.
- B ☐ ... den Unterschied zwischen mehreren Varianzen aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist nicht bekannt welcher Vergleich konkret unterschiedlich ist.
- C ☐ ... den Unterschied zwischen der Varianz aus verschiedenen Behandlungsgruppen und der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in den Gruppen exakt zu bestimmen.
- D ☐ ... den Unterschied zwischen zwei paarweisen Mittelwerten aus verschiedenen Behandlungsgruppen. Wenn die signifikant ist, ist daher bekannt welcher Vergleich konkret unterschiedlich ist.
- E ☐ ... den Unterschied zwischen der F-Statistik anhand der Varianz der Gruppen. Wenn die F-Statistik exakt 0 ist, kann die Nullhypothese abgelehnt werden.

## 3 Aufgabe

(2 Punkte)

Welche Aussage über die Entscheidung anhand der berechneten Teststatistik gegen die Nullhypothese ist richtig?

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

#### 4 Aufgabe

(2 Punkte)

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

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

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

#### 6 Aufgabe

(2 Punkte)

Wie lautet der Mittelwert und Standardabweichung von  $y$  mit 13, 7, 12, 9 und 10.

- A** ☐ Es ergibt sich 10.2 +/- 2.39
- B** ☐ Sie erhalten 10.2 +/- 1.55
- C** ☐ Sie erhalten 10.2 +/- 1.195
- D** ☐ Es berechnet sich 10.2 +/- 5.7
- E** ☐ Es ergibt sich 9.2 +/- 2.85

#### 7 Aufgabe

(2 Punkte)

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

- A** ☐ Wir vergleichen die Effekte des  $p$ -Wertes mit den Effekten der Signifikanzschwelle unter der Annahme der Nullhypothese. Dabei gilt, dass wir die Nullhypothese nur ablehnen können anhand des Falsifikationsprinzips.
- B** ☐ Wir 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.

- C** ☐ Wir machen eine Aussage über die Flächen unter der 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.
- D** ☐ Wir schauen, ob der  $p$ -Wert größer ist als das Signifikanzniveau  $\alpha$  und vergleichen somit Wahrscheinlichkeiten. Die Wahrscheinlichkeiten werden als Flächen unter der Kurve der Teststatistik dargestellt, wenn die  $H_A$  gilt.
- E** ☐ Wir vergleichen mit dem  $p$ -Wert und dem Signifikanzniveau  $\alpha$  Wahrscheinlichkeiten und damit die absoluten Werte auf einem Zahlenstrahl, wenn die  $H_0$  gilt.

## 8 Aufgabe

(2 Punkte)

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

- A** ☐ Durch eine Randomisierung können wir von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.
- B** ☐ Strukturgleichheit ist durch Randomisierung gegeben. Leider hilft die Randomisierung noch nicht um von der Stichprobe auf die Grundgesamtheit zu schließen. Deshalb wurde das Falsifikationsprinzip entwickelt.
- C** ☐ Randomisierung bringt starke Unstrukturiertheit in das Experiment und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschließen.
- D** ☐ Randomisierung erlaubt erst die Mittelwerte zu schätzen. Ohne Randomisierung keine Mittelwerte. Ohne Mittelwerte keine Varianz und somit auch kein statistischer Test.
- E** ☐ Durch eine Randomisierung können wir nicht von Strukturgleichheit zwischen der Stichprobe und der Grundgesamtheit ausgehen.

## 9 Aufgabe

(2 Punkte)

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

- A** ☐ Der Effekt eines statistischen Tests beschreibt die biologisch 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.
- 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 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.
- D** ☐ Durch den Effekt erfahren wir die statistische interpretierbare Ausgabe eines statistischen Tests. Zum Beispiel das  $\eta^2$  aus einer ANOVA. Damit können wir die Signifikanz direkt mit dem Effekt verbinden. Am Ende muss der Forschende aber entscheiden, ob der Effekt entsprechend seinen Erwartungen als bedeutend zu bewerten ist.
- E** ☐ 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.

## 10 Aufgabe

(2 Punkte)

Gegeben ist  $y$  mit 27, 1, 33, 30, 23, 2 und 51. Berechnen Sie den Median, das 1<sup>st</sup> Quartile sowie das 3<sup>rd</sup> Quartile.

- A** ☐ Es ergibt sich 27 +/- 2
- B** ☐ Es berechnet sich 27 [2; 33]
- C** ☐ Sie erhalten 27 [0; 31]
- D** ☐ Sie erhalten 27 +/- 33
- E** ☐ Es berechnet sich 28 [3; 32]

## 11 Aufgabe

(10 Punkte)

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



**Visualisierung des Scatterplots** Mark liest laut: 'Wenn zwei kontinuierliche Variablen vorliegen, können diese in einer explorativen Datenanalyse...'. Mark stoppt. Mark schmeißt noch eine Handvoll Marzipankugeln in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Andrea Berg. Was waren noch gleich kontinuierliche Variablen? In seiner Abschlussarbeit hatte er ein Stallexperiment im Emsland durchgeführt. Dabei ging es um den Zusammenhang zwischen Gewichtszuwachs in der 1LW und durchschnittlichen Bewegungsscore [Movement/h] im groben Kontext von Puten. Nun stellt sich die Frage für ihn, ob es überhaupt einen Zusammenhang zwischen den gemessenen Variablen gibt. Dafür war eine explorative Datenanalyse gut! Eine echte Herausforderung für ihn war schon immer die Unsicherheit gewesen. Ein leidiges Lied. Dann was anderes. Das Verrückte ist, dass der Hamster Columbo wirklich liebt. Das ist Mark sehr recht, denn er braucht Entspannung.

Durchschnittlichen Bewegungsscore [Movement/h]	Gewichtszuwachs in der 1LW
17.9	21.7
19.1	22.7
21.0	25.0
18.6	23.6
19.4	25.9
20.9	27.0
25.2	30.0
19.6	24.3
20.1	27.3
16.8	21.7
17.9	21.9
15.2	21.7

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

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



## 12 Aufgabe

(9 Punkte)

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

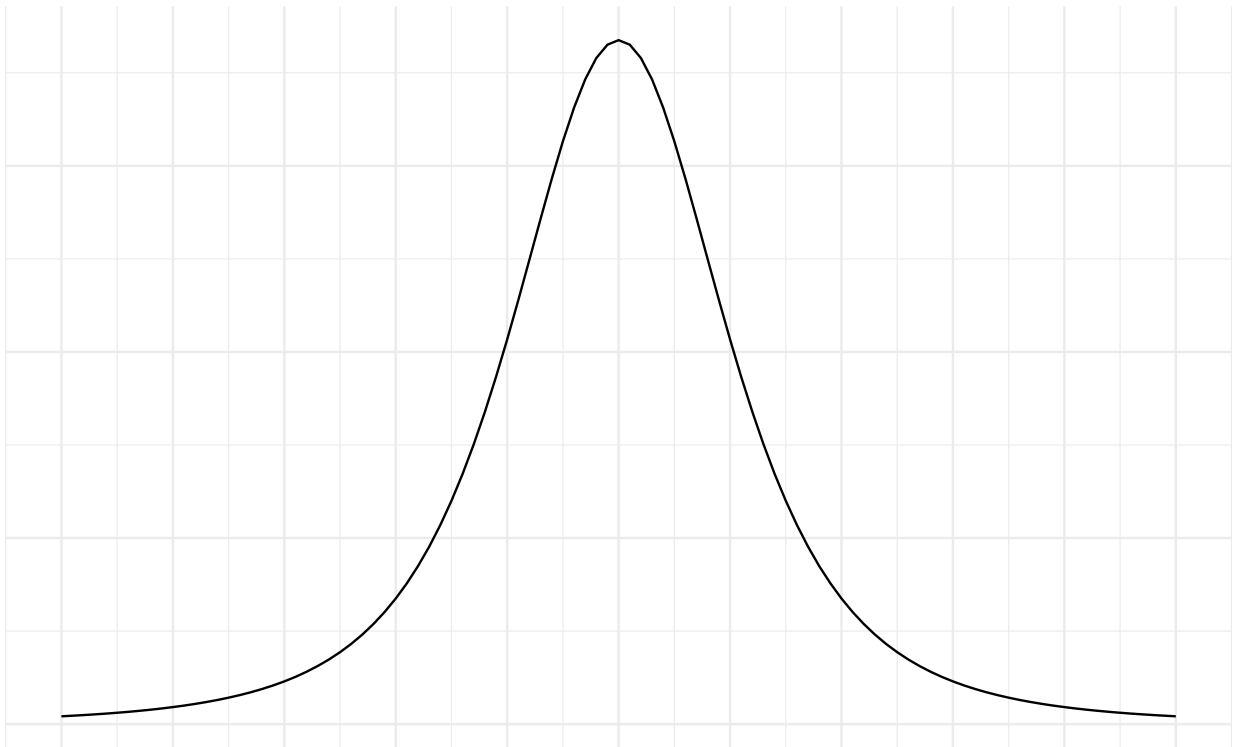


**Visualisierung der Teststatistik  $T_D$  und dem p-Wert** 'Kannst du mir nochmal an einer Visualisierung erklären, wie der Zusammenhang zwischen der Teststatistik aus den Daten  $T_D$  und dem p-Wert ist? Ich habe hier zig Fachbegriffe, kriege die aber nicht zusammen...', fragt Yuki nachdrücklich Nilufar. Das hilft aber nur bedingt, denn Nilufar hat wenig geschlafen und träumt zu den Klängen von London Grammar. Yuki hatte den ganzen Abend mit Nilufar über die Faulheit diskutiert und nun sind beide voll neben der Spur. So wird es nichts mit der Klausur. Yuki mampft noch ein paar Reese's Peanut Butter Cups und nickt ein. Jetzt brauchen die beiden gesondert Hilfe!

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

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

1. Ergänzen Sie eine beschriftete x-Achse! **(1 Punkt)**
2. Ergänzen Sie „ $\bar{y}_1 = \bar{y}_2$ “! **(1 Punkt)**
3. Ergänzen Sie „ $A = 95\%$ “! **(1 Punkt)**
4. Zeichnen Sie  $T_{\alpha=5\%}$  in die Abbildung! **(1 Punkt)**
5. Zeichnen Sie das Signifikanzniveau  $\alpha$  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)**



### 13 Aufgabe

(9 Punkte)

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



**Ergebnistabelle der einfaktoriellen ANOVA** 'Uff... die einfaktorielle ANOVA. Und wie füllen wir jetzt die Tabelle der ANOVA aus und schauen, ob da was signifikant ist?', Mark hebt die Augenbraue. 'Das ist eine sehr gute Frage. Ich glaube man kann alles in der Tabelle relativ einfach mit wenigen Informationen berechnen.', meint Nilufar dazu. Da hilft das Huhn von Nilufar auch nur bedingt. Mark hatte sich in einen Versuch in einer Klimakammer verschiedene Erbsen angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Substrattypen (*torf*, *40p60n*, *30p20n* und *70p30n*) und dem Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] gibt. Nachher wollen sich beide noch mit dem Hobby Hip Hop von Nilufar beschäftigen. Kennt Mark noch nicht, klingt aber interessant.

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

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
<b>Substrattypen</b>	3	177.73			
<b>error</b>	20				
<b>Total</b>	23	895.33			

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

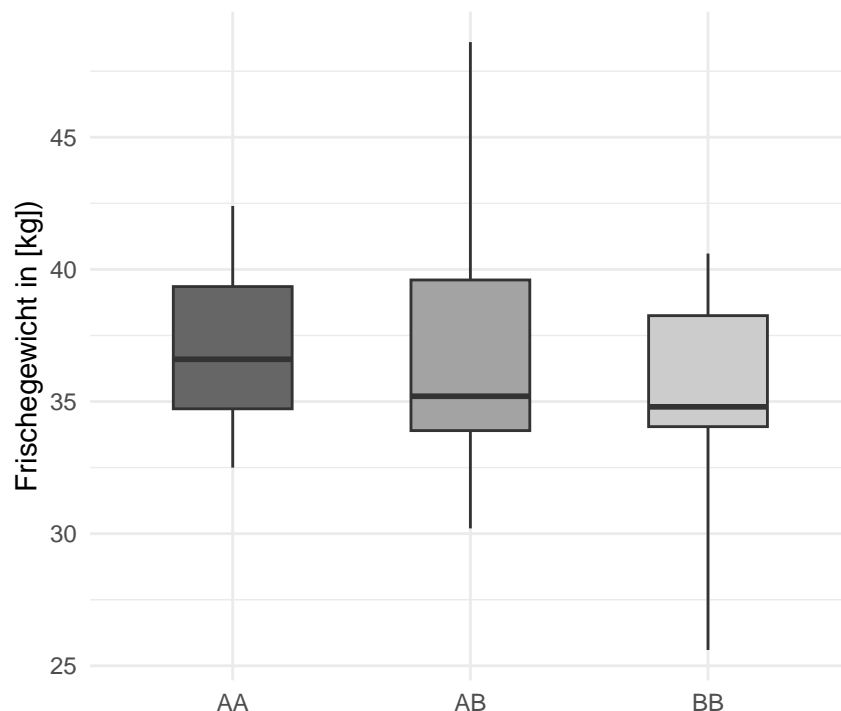
## 14 Aufgabe

(9 Punkte)

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



**Zerforschen des Boxplots** Yuki und die Faulheit, eine unendliche Geschichte mit kniffligen Wendungen. Deshalb gilt anschauen, was andere vor einem gemacht haben. Für Yuki ist es eine Möglichkeit schneller ans Ziel zu gelangen. Yuki soll in seinem Projektbericht Brokoli untersuchen. Die Behandlung in seinem Projektbericht werden verschiedene Genotypen (AA, AB und BB) sein. Erheben wird Yuki als Endpunkt (Y) *Frischegewicht* benannt als *freshmatter* in seiner Exceldatei. Von seinem Betreuer erhält er nun folgende Abbildung von Boxplots, die er erstmal zur Übung nachbauen soll, bevor er mit dem eigentlichen Versuch beginnt. Aber nur in passender Atmosphäre! Hm, lecker Reese's Peanut Butter Cups und dazu dann im Hintergrund Matrix laufen lassen.



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

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

## 15 Aufgabe

(10 Punkte)

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



**Den Chi-Quadrat-Test berechnen** Am Ende war es für Paula in ihrem Projektbericht dann doch kein normalverteiltes Outcome. Das was jetzt etwas doof, da er sich auf eine ANOVA gefreut hatte. Dann noch schnell White Lies auf das Ohr und los gehts. Prinzipiell ginge das auch irgendwie, aber nun möchte ihre Betreuerin gerne einen  $\chi^2$ -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Paula hatte sich in ein Kreuzungsexperiment  $n = 134$  Beobachtungen von Milchvieh angeschaut. Dabei hat sie als Behandlung *Klimakontrolle [ja/nein]* bestimmt und zum anderen die Variable *Fettgehalt erreicht [ja/nein]* ermittelt. Jetzt muss Paula mal schauen, wie sie das jetzt rechnet. Eigentlich wollte Paula nachher noch zum Sport. Um zu Fechten geht Paula dann später nochmal raus. Echte Entspannung.

	56	21	
	13	44	

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Ergänzen Sie die Tabelle um die fehlenden Informationen! **(1 Punkt)**
3. Visualisieren Sie den Zusammenhang zwischen den beiden kategorialen Variablen! **(2 Punkte)**
4. Berechnen Sie die Teststatistik eines Chi-Quadrat-Test! **(2 Punkte)**
5. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem  $\chi^2_{\alpha=5\%} = 3.83$ ! Begründen Sie Ihre Antwort! **(2 Punkte)**
6. Skizzieren Sie in einer Abbildung die  $\chi^2$ -Verteilung, wenn die  $H_0$  wahr ist! Ergänzen Sie  $\chi^2_{\alpha=5\%}$  und  $\chi^2_D$  in der Abbildung! Beachten Sie folgenden Informationen zur  $\chi^2$ -Verteilung. Die  $\chi^2$ -Verteilung hat ein Maxima bei  $\chi^2 = 6$  sowie ein Minima bei  $\chi^2 = 10$ . **(2 Punkte)**

## 16 Aufgabe

(11 Punkte)

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



**Visualisierung der einfaktoriellen ANOVA** Jessica und Yuki schauen sich etwas entnervt an. Gemeinsam schreiben die beiden ihre Abschlussarbeit und sollen nun als erstes einmal die Daten visualisieren damit abgeschätzt werden kann, ob überhaupt signifikante Ergebnisse zu erwarten sind. Die beiden waren im Wendland um ein Stallexperiment mit Puten durchzuführen. Dabei haben Jessica und Yuki den Messwert Schlachtgewicht [kg] unter der Behandlung Elterlinie (*Standard*, *Yray* und *Xray*) ermittelt. Kennengelernt haben sich die beiden auf einem Konzert von London Grammar. Später wird noch Matrix geguckt. Yuki befürwortet das!

Elterlinie	Schlachtgewicht
Yray	26
Standard	41
Standard	45
Yray	24
Xray	45
Xray	45
Yray	25
Standard	48
Yray	24
Yray	24
Xray	45
Standard	46
Xray	44
Standard	45
Xray	44
Standard	45
Standard	44
Xray	46
Xray	45
Yray	25

Leider kennen sich Jessica und Yuki mit Darstellung einer einfaktoriellen ANOVA überhaupt nicht aus.

1. Erstellen Sie eine Visualisierung der Datentabelle! Beschriften Sie die Abbildung! **(2 Punkte)**
2. Benennen Sie die Visualisierung mit dem korrekten, statistischen Fachbegriff! **(1 Punkt)**
3. Zeichnen Sie folgende statistischen Maßzahlen passend ein!
  - Den globalen Mittelwert  $\beta_0$  **(1 Punkt)**
  - Die Mittelwerte der einzelnen Behandlungsstufen **(1 Punkt)**
  - Die Mittelwertsdifferenz der einzelnen Behandlungsstufen mit  $\beta_{Standard}$ ,  $\beta_{Yray}$  und  $\beta_{Xray}$  **(1 Punkt)**
  - Die Residuen oder Fehler mit  $\epsilon$  **(1 Punkt)**
4. Liegt ein *vermutlicher* signifikanter Unterschied vor? Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Schätzen Sie die Effekte der Behandlungsstufen! **(2 Punkte)**

## 17 Aufgabe

(12 Punkte)

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



**Berechnung des Student t-Test** Jonas ist im Emsland für einen Versuch mit Spargel. Allein diese Tatsache ist für ihn eine Erzählung wert. Eine echte Herausforderung für ihn war schon immer die Erschöpfung gewesen. Ein leidiges Lied. Für seiner Hausarbeit musste er ein Gewächshausexperiment mit Spargel durchführen und das sollte laut seinem Betreuer an diesem Nichtort besonders gut gelingen. Ablenkung gibt es jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn seine Behandlung Genotypen (AA und BB) und der Messwert Chlorophyllgehalt (SPAD-502Plus) [SPAD] sollen mit einem t-Test ausgewertet werden. Immerhin weiß er, dass sein Messwert einer Normalverteilung folgt. Hm..., was entspannendes wäre gut. Einfach mal raus um zu Schwimmen. Ohne Ziel und Uhr. Das ist was für Jonas.

Genotypen	Chlorophyllgehalt
AA	23.2
AA	33.1
AA	16.4
BB	24.9
BB	14.9
AA	21.3
BB	22.0
BB	6.9
AA	19.7
AA	26.2
BB	31.7
AA	31.6
AA	32.4
AA	22.1
BB	30.9
BB	33.1
BB	30.8
BB	23.9

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

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