

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

24. Juni 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 71 Punkten sind aus den Rechen- und Textaufgaben erreicht.
_____ von 91 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
87.0 - 91.0	1,0
82.5 - 86.5	1,3
78.0 - 82.0	1,7
73.5 - 77.5	2,0
69.0 - 73.0	2,3
64.5 - 68.5	2,7
60.0 - 64.0	3,0
55.5 - 59.5	3,3
51.0 - 55.0	3,7
45.5 - 50.5	4,0

Es ergibt sich eine Endnote von _____.

Multiple Choice Aufgaben

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

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

- Es sind ____ von 20 Punkten erreicht worden.

Rechen- und Textaufgaben

- Die Tabelle wird vom Dozenten ausgefüllt.

Aufgabe	11	12	13	14	15	16	17
Punkte	9	8	10	12	10	12	10

- Es sind ____ von 71 Punkten erreicht worden.

1 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 Mittelwerte und der Varianz aus verschiedenen Behandlungsgruppen. Wenn die ANOVA signifikant ist, ist bekannt welcher Vergleich konkret unterschiedlich ist.
- ☐ 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 Varianz aus verschiedenen Behandlungsgruppen und der Varianz über alle Behandlungsgruppen. Wenn die ANOVA signifikant ist, muss über einen Posthoc-Test nachgedacht werden um den signifikanten Unterschied in den Gruppen exakt zu bestimmen.
- ☐ E ... den Unterschied zwischen der Varianz 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.

2 Aufgabe

(2 Punkte)

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

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

3 Aufgabe

(2 Punkte)

Gegeben ist y mit 11, 8, 19, 16 und 17. Berechnen Sie den Mittelwert und Standardabweichung.

- ☐ A Es berechnet sich 15.2 ± 20.7
- ☐ B Es ergibt sich 13.2 ± 10.35
- ☐ C Es berechnet sich 14.2 ± 20.7
- ☐ D Sie erhalten 14.2 ± 2.13
- ☐ E Es ergibt sich 14.2 ± 4.55

4 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 des Niederschlags oder Regenmenge: ein statistischer Test gibt die Stärke eines Effektes wieder. Zum Beispiel, wie hoch ist der Mittelwertsunterschied.
- ☐ B 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.
- ☐ C In der Analogie der Maximaltemperatur: Was ist der maximale Unterschied zwischen zwei Gruppen. Wir erhalten hier eine Aussage über die Spannweite und den maximalen Effekt.

- D** ☐ In der Analogie der Regenwahrscheinlichkeit: ein statistischer Test gibt die Wahrscheinlichkeit für das Auftreten eines Ereignisses wieder. Die Stärke des Effektes wird nicht wiedergeben.
- E** ☐ In der Analogie der Regenwahrscheinlichkeit in einem bestimmten Gebiet: ein statistischer Test gibt die Wahrscheinlichkeit für ein Ereignis in einem Experiment mit den Daten D wieder und lässt sich kaum verallgemeinern.

5 Aufgabe

(2 Punkte)

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

- A** ☐ 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 $(n - 1)$.
- B** ☐ Den Median berechnen, dann die quadratischen Abstände zum Median aufsummieren, dann die Wurzel ziehen. Am Ende durch die Fallzahl $(n - 1)$ teilen
- C** ☐ Den Mittelwert berechnen und die Abstände quadrieren. Die Summe mit der Fallzahl $(n - 1)$ multiplizieren.
- D** ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend teilen wir durch die Fallzahl $(n - 1)$. Nicht zu vergessen, am Ende dann noch die Wurzel zu ziehen.
- E** ☐ Als erstes berechnen wir den Mittelwert. Dann bilden wir die Summe der quadratischen Abstände zu dem Mittelwert. Abschließend teilen wir durch die Fallzahl $(n - 1)$.

6 Aufgabe

(2 Punkte)

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

- A** ☐ Randomisierung ist die direkte Folge von Strukturgleichheit. Die Strukturgleichheit erlaubt es erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- B** ☐ Randomisierung sorgt für Strukturgleichheit und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.
- C** ☐ Randomisierung erlaubt erst die Mittelwerte zu schätzen. Ohne Randomisierung keine Mittelwerte. Ohne Mittelwerte keine Varianz und somit auch kein statistischer Test.
- D** ☐ Randomisierung erlaubt erst die Varianzen zu schätzen. Ohne eine Randomisierung ist die Berechnung von Mittelwerten und Varianzen nicht möglich. Dadurch lässt sich erst ein Experiment auswerten.
- E** ☐ Randomisierung bringt starke Unstrukturiertheit in das Experiment und erlaubt erst von der Stichprobe auf die Grundgesamtheit zurückzuschliessen.

7 Aufgabe

(2 Punkte)

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

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

8 Aufgabe

(2 Punkte)

Sie führen einen Versuch mit einer Behandlung und drei Faktorleveln durch. Danach rechnen Sie eine einfaktorielle ANOVA und es ergibt sich ein $\eta^2 = 0.12$. Welche Aussage ist richtig?

- A ☐ 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.
- B ☐ Der Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird, wird durch das η^2 beschrieben.
- C ☐ Der Anteil der Varianz, der von den Behandlungsbedingungen erklärt wird, wird durch das $1 - \eta^2$ beschrieben.
- D ☐ Das η^2 ist die Korrelation der ANOVA. Mit der Ausnahme, dass $\eta^2 = 0$ der beste Wert ist.
- E ☐ Das η^2 beschreibt den Anteil der globalen Varianz, der von den Umweltbedingungen erklärt wird.

9 Aufgabe

(2 Punkte)

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

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

10 Aufgabe

(2 Punkte)

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

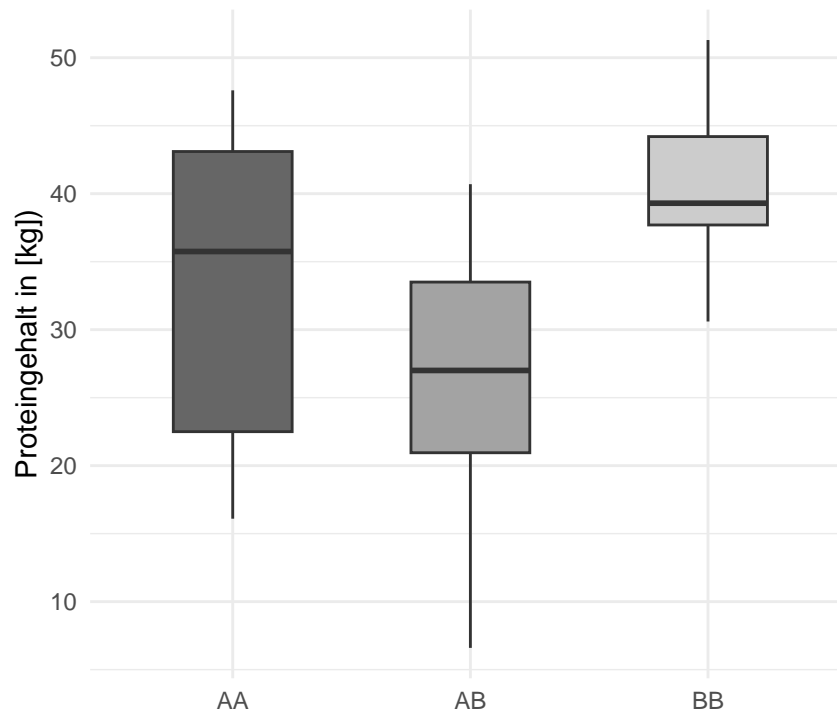
- A ☐ Es gilt $T_D = \text{signal}/\text{noise}^2$. Der Effekt *signal* wird mit der quadratischen Varianz *noise* gewichtet.
- B ☐ Es gilt $T_D = \text{signal}/\text{noise}$. Der Effekt *noise* wird mit der Varianz *signal* gewichtet.
- C ☐ Wir gewichten den Effekt *signal* mit der Varianz *noise* und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

11 Aufgabe

(9 Punkte)

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

Alex steht vor einem ersten Problem, denn wenn es nach seinem Betreuer geht, soll er in einem Gewächshausexperiment Maiss auswerten. Soweit eigentlich alles passend. Besser wäre was anderes gewesen. Alex liebt Starcraft. Darin kann er sich wirklich verlieren und immer wieder neu begeistern. Das heißt erstmal überlegen für Alex. Aus den Boxen wummert Abba und sein Mund ist verklebt von Gummibärchen. 'Herrlich', denkt Alex. Die Behandlung werden verschiedene Genotypen (AA, AB und BB) sein. In seiner Exceldatei wird er den Outcome (Y) *Proteingehalt* als *protein* aufnehmen. Vorab soll Alex aber einmal die folgenden Boxplots seinem Betreuer nachbauen, damit er den R Code schonmal für später vorliegen hat. Damit geht das Problem schon los. Eine echte Herausforderung für ihn war schon immer die Gefälligkeit gewesen. Ein leidiges Lied.



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

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

12 Aufgabe

(8 Punkte)

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

In einem Gespräch mit seinem Betreuer wird Steffen gebeten seine Daten aus einem Stallexperiment mit Fleischrindern in einem Histogramm darzustellen. Steffen schmeißt noch eine Handvoll Oreos in seinen Rachen. Im Hintergrund klirrt leise der Spiegel zum Sound von Taylor Swift. In seinem Experiment hat er die mittleren dunklen Pigmentstörungen erst fotografiert und dann ausgezählt. Laut seinem Betreuer soll das Histogramm helfen, die Verteilung der die mittleren dunklen Pigmentstörungen zu bestimmen. Es wäre einfacher, wenn da nicht noch was wäre. Eine echte Herausforderung für ihn war schon immer die Romantik gewesen. Ein leidiges Lied. Wenn Taylor Swift ertönt, dann sucht die Schlange schleunigst Schutz unter dem Sofa. Steffen schüttelt den Kopf.

Die mittleren dunklen Pigmentstörungen: 11.6, 9.3, 6.8, 7.6, 9.2, 8.7, 8.3, 8.2, 9.7, 12.7, 9.9, 8.3, 10.1, 10.3, 11.5, 11.6, 9.4, 9.4, 12.3, 8.8, 14.3, 15.3

Leider kennt sich Steffen 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)**

13 Aufgabe

(10 Punkte)

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

In ein Stallexperiment in der Uckermark hatte Mark sich zum einen die Behandlung Automatische Fütterung [ja/nein] und zum anderen die Messung Schlachtgewicht im Zielbereich [ja/nein] im Kontext von Fleischrindern angeschaut. Nun steht Mark vor dem Problem, dass er zwei kategoriale Variablen in seinem Projektbericht gemessen hat. Dazu kommt dann noch was anderes. Mark und die Unsicherheit, eine unendliche Geschichte mit kniffligen Wendungen. Da sein Betreuer erstmal die langen Tabellen mit ja/nein in einer explorativen Datenanalyse zusammengefasst und präsentiert bekommen möchte bevor es überhaupt weitergeht, muss er jetzt eine Lösung finden. Was alles auch nicht einfacher macht. Am liebsten würde er ja was anderes machen. Am Ende dann doch besser Geocaching. Wunderbar. Eine echte Ablenkung für Mark.

Automatische Fütterung	Schlachtgewicht im Zielbereich	Automatische Fütterung	Schlachtgewicht im Zielbereich
nein	ja	ja	ja
ja	nein	ja	nein
nein	ja	nein	nein
nein	ja	ja	ja
ja	nein	nein	nein
nein	ja	nein	ja
nein	nein	ja	nein
nein	ja	ja	nein
nein	nein	ja	ja
ja	nein	ja	ja
nein	ja	nein	ja
ja	nein	ja	ja
ja	ja	ja	nein
nein	ja	ja	nein
ja	nein	nein	nein

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

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

14 Aufgabe

(9 Punkte)

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

Alex ist im Oldenburger Land für einen Pilotexperiment mit sehr geringer Fallzahl ($n_1 = n_2 = 3$) mit Brokkoli. Allein diese Tatsache ist für ihn eine Erzählung wert. Alex und die Gefälligkeit, eine unendliche Geschichte mit kniffligen Wendungen. Für seine Abschlussarbeit musste er einen Versuch in einer Klimakammer mit Brokkoli durchführen und das sollte laut seiner Betreuerin an diesem Ort besonders gut gelingen, da man hier gut neue technische Anlagen und Behandlungen fernab der Bevölkerung testen könne. Zeugen gibt es hier jedenfalls keine. Gar keine. Alleine sein hilft jetzt aber nur bedingt, denn seine Behandlung Bewässerungstypen (*low* und *high*) und der Messwert Frischegewicht [kg/ha] sollen mit einem t-Test ausgewertet werden. Immerhin weiß er, dass sein Messwert einer Normalverteilung folgt. Hm..., was entspannendes wäre gut. Um zu Laufen geht Alex dann später nochmal raus. Echte Entspannung.

Bewässerungstypen	Frischegewicht
low	15.6
high	26.7
low	21.2
low	17.4
high	18.8
high	21.8


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

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

15 Aufgabe

(9 Punkte)

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

Programmieren ist nicht so einfach.  ist schon ein tolles Programm, wenn man mit dem Ding umgehen kann. Super umgehen kann damit Nilufar. Eine echte Herausforderung für Nilufar ist die Erwartung, aber das ist noch eine andere Sache. Deshalb sind aber Jonas und Alex nicht bei ihr. Sondern um sich bei einem gemeinsamen Projekt helfen zu lassen. Im Hintergrund wummert Deichkind. Beide arbeiten gemeinsam an einer Abschlussarbeit. In dem zu beschreibenden Versuch geht es im Emsland um einem Stallexperiment mit Lamas. Dabei ging darum herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Genotypen (AA und BB) und dem Messwert Gewichtszuwachs in der 1LW gibt. Da der Messwert Gewichtszuwachs in der 1LW normalverteilt ist kann ein t-Test gerechnet werden. Jonas möchte dann später noch mehr über Nilufars Hobby Hip Hop erfahren.

```
##  
## Two Sample t-test  
##  
## data: Gewichtszuwachs by Genotypen  
## t = 3.983, df = 13, p-value = 0.001561  
## alternative hypothesis: true is not equal to [condensed]  
## 95 percent confidence interval:  
## 7.661527 25.824187  
## sample estimates:  
## mean in group AA mean in group BB  
## 38.60000 21.85714
```

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

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

16 Aufgabe

(12 Punkte)

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

'Uff... die einfaktorielle ANOVA. Und wie füllen wir jetzt extitgenau die Tabelle der ANOVA aus und schauen, ob da was signifikant ist?', Tina hebt die Augenbraue. 'Das ist eine sehr gute Frage. Ich glaube man kann alles in der Tabelle relativ einfach mit wenigen Informationen berechnen.', meint Jessica dazu und schmiß sich noch ein paar Schokobons in den Rachen. Tina hatte sich in einen Leistungssteigerungsversuch verschiedene Schweinen angeschaut. Dabei ging es herauszufinden, ob es einen Zusammenhang zwischen der Behandlung Lüftungssystem (*keins*, *storm*, *tornado* und *thunder*) und dem Messwert Fettgehalt [%/kg] gibt. Nun möchte erstmal ihre Betreuerin eine ANOVA Tabelle sehen. Was immer da auch drin zu erkennen sein mag. Später wollen die beiden dann noch raus um Rad zu fahren.

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

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

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Lüftungssystem	3	146.72			
Error	24	519.14			

4. Schätzen Sie den p-Wert der Tabelle mit $F_{\alpha=5\%} = 3.01$ ab. Begründen Sie Ihre Antwort! **(2 Punkte)**
5. Was bedeutet ein signifikantes Ergebnis in einer einfaktoriellen ANOVA? **(1 Punkt)**
6. Berechnen Sie *einen* Student t-Test für den *vermutlich* signifikantesten Gruppenvergleich anhand der untenstehenden Tabelle mit $T_{\alpha=5\%} = 2.03$. Begründen Sie Ihre Auswahl! **(3 Punkte)**

Lüftungssystem	Fallzahl (n)	Mittelwert	Standardabweichung
keins	8	8.12	6.69
storm	7	8.71	1.80
tornado	7	3.00	1.63
thunder	6	7.83	5.85

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

17 Aufgabe

(10 Punkte)

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

Am Ende war es für Jessica in ihrer Abschlussarbeit dann doch kein normalverteiltes Outcome. Das was jetzt etwas doof, da er sich auf eine ANOVA gefreut hatte. Dann noch schnell Schokobons zur Stärkung und los gehts. Prinzipiell ginge das auch irgendwie, aber nun möchte ihr Betreuer gerne einen χ^2 -Test auf einer 2x2-Kreuztabelle berechnet bekommen. Jessica hatte sich in ein Freilandversuch $n = 112$ Beobachtungen von Maiss angeschaut. Dabei hat sie als Behandlung *Herbizideinsatz* [ja/nein] bestimmt und zum anderen die Variable *Chlorophyllgehalt unter Zielwert* [ja/nein] ermittelt. Jetzt muss Jessica mal schauen, wie sie das jetzt rechnet. Am Ende des Tages möchte sie dann noch ihr Hobby Warhammer genießen. Das muss auch mal sein!

			53
			59
	50	62	112

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

1. Formulieren Sie die wissenschaftliche Fragestellung! **(1 Punkt)**
2. Ergänzen Sie die Tabelle um die fehlenden Informationen! **(1 Punkt)**
3. Ergänzen Sie die Felder innerhalb der 2x2 Kreuztabelle, so dass *ein* signifikanter Effekt zu erwarten wäre! **(2 Punkte)**
4. Begründen Sie Ihr Vorgehen an der Formel des Chi-Quadrat-Tests. Erklären Sie Ihr Vorgehen an einem Beispiel! **(2 Punkte)**
5. Visualisieren Sie den Zusammenhang zwischen den beiden kategorialen Variablen! **(2 Punkte)**
6. Was ist die Mindestanzahl an Beobachtungen je Zelle? Wenn in einer der Zellen weniger Beobachtungen auftreten, welchen Test können Sie anstatt des Standard Chi-Quadrat-Tests anwenden? **(2 Punkte)**