



HOCHSCHULE OSNABRÜCK
UNIVERSITY OF APPLIED SCIENCES

Fakultät Agrarwissenschaften und Landschaftsarchitektur

- ☐ Bachelorprüfung
☐ Masterprüfung

Klausur im Modul: _____
Prüfer*in: _____
Name, Vorname: _____
Matrikelnr.: _____ Fachsemester.: _____
Tisch Nr.: _____ Zugel. Hilfsmittel: _____

Hiermit bestätige ich, dass ich prüfungsfähig bin. Weiter bestätige ich die Kenntnisnahme der umseitigen Klausurbelehrung.

Die für den Prüfungszeitraum ggf. geltenden Hygienebestimmungen habe ich zur Kenntnis genommen.

Ort, Datum, Unterschrift der*des Studierenden

Bewertung - von den Prüfenden auszufüllen -

1. Prüfer*in: _____ Unterschrift	2. Prüfer*in: _____ Unterschrift
Punkte: _____ von _____	Punkte: _____ von _____
Note: _____	Note: _____



Klausurbelehrung

Lesen Sie die nachstehende Belehrung bitte sorgfältig durch. Durch Ihre Unterschrift auf dem Klausurendeckblatt bestätigen Sie, die Bestimmungen des Allgemeinen Teils der Prüfungsordnung der Hochschule Osnabrück (ATPO) in der aktuell gültigen Fassung über Rücktritt, Täuschung und Ordnungsverstöße zur Kenntnis genommen zu haben.

Die nachstehenden Hinweise dienen lediglich der Information und verschaffen einen Überblick. Die für Ihre Prüfungen maßgeblichen Regelungen ergeben sich aus den einschlägigen Ordnungen.

Prüfungsfähigkeit aufgrund einer eigenen Krankheit (vgl. § 15 Abs. 2 ATPO)

- Prüflinge, die eine Klausur empfangen, nehmen an der Klausur teil und erklären sich mit Unterschrift auf dem Deckblatt für prüfungsfähig. Prüflinge, die sich gesundheitlich nicht in der Lage fühlen, die Klausur zu bewältigen, müssen vor Beginn der Klausur den Prüfungsraum verlassen. Ein Prüfungsrücktritt nach Beginn der Klausur ist in der Regel ausgeschlossen.

Täuschung und Ordnungsverstöße (vgl. § 15 Abs. 3, 4 ATPO)

- Das Mitsichführen von Unterlagen mit fachlichem Bezug ist untersagt, außer sie sind ausdrücklich zugelassen worden.
- Informations- und kommunikationsfähige Geräte sind abzuschalten, es sei denn, ein Gebrauch ist ausdrücklich zugelassen worden.
- Der Prüfling trägt die Verantwortung dafür, dass sowohl er als Person als auch sein Prüfungsplatz von zur Täuschung geeigneten Materialien frei ist.
- Es gilt ein Sprechverbot für alle Prüflinge untereinander.
- Das Verlassen des Klausurraums ist nur zu dringenden WC-Gängen erlaubt. Es darf jeweils nur ein Prüfling zur gleichen Zeit den Klausurraum verlassen. Das Sprechen mit anderen Studierenden oder dritten Personen im Rahmen eines WC-Gangs ist untersagt.
- Es dürfen nur die Schreibblätter verwendet werden, die von den Prüfungsaufsichten ausgegeben werden, es sei denn, andere Schreibblätter wurden ausdrücklich zugelassen.
- Das Weiterschreiben nach Bekanntgabe des Bearbeitungszeitendes ist untersagt.
- Der Prüfling trägt die Verantwortung für die Abgabe seiner vollständigen Klausur.
- Verstöße gegen diese Verbote können als Täuschungsversuche gewertet werden. Die Prüfungsleistung wird mit „nicht bestanden“ bzw. „nicht ausreichend“ bewertet.
- Auch hier nicht genannte, weitere Formen von Täuschungshandlungen können vergleichbare Konsequenzen nach sich ziehen.
- Prüflinge, die den ordnungsgemäßen Ablauf der Prüfungen stören, können von der aufsichtführenden Person von der Klausur ausgeschlossen werden. Bei festgestellten Ordnungsverstößen wird die Prüfungsleistung mit „nicht bestanden“ bzw. „nicht ausreichend“ bewertet.

Mitwirkungspflichten (allgemeiner Grundsatz)

- Für die aufsichtführende Person nicht ohne Weiteres erkennbare Störungen (störende Geräusche, wackeliger Stuhl/Tisch etc.) sind durch den Prüfling anzuzeigen.

Sonstiges

- Abgelegte Prüfungsleistungen ohne Anmeldung und Zulassung werden nicht bewertet (§ 12 Abs. 3 S. 5 ATPO).

Name: _____

Nicht bestanden: ☐

Vorname: _____

Matrikelnummer: _____

Endnote: _____

B.Sc. Bioverfahrenstechnik in Agrar- und Lebensmittelwirtschaft

Klausur Angewandte Statistik für Bioverfahrenstechnik

Hochschule Osnabrück

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

Klausur vom 11. Juli 2022

Ergebnis der Klausur

_____ von 20 Punkten sind aus dem Multiple Choice Teil erreicht.

_____ von 61 Punkten sind aus dem Rechen- und Textteil erreicht.

_____ von 81 Punkten in Summe.

Es wird folgender Notenschlüssel angewendet.

Punkte	Note
78 - 81	1,0
75 - 77	1,3
70 - 74	1,7
65 - 69	2,0
59 - 64	2,3
54 - 58	2,7
49 - 53	3,0
44 - 48	3,3
41 - 43	3,7
40	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.**
- Es werden nur Antworten berücksichtigt, die in dieser Tabelle angekreuzt sind!

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

- 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							

- Es sind ____ von 61 Punkten erreicht worden.

1 Aufgabe

(2 Punkte)

In einer linearen Regression werden die ϵ oder Residuen geschätzt. Welcher Verteilung folgen die Residuen bei einer optimalen Modellierung?

- A ☐ Die Residuen sind normalverteilt mit $\mathcal{N}(0, 1)$.
- B ☐ Die Residuen sind binomialverteilt.
- C ☐ Die Residuen sind normalverteilt mit $\mathcal{N}(0, s^2)$.
- D ☐ Die Residuen folgen einer Poissonverteilung mit $\text{Pois}(0)$.
- E ☐ Die Residuen sind normalverteilt mit $\mathcal{N}(\bar{y}, s^2)$.

2 Aufgabe

(2 Punkte)

Sie haben das abstrakte Modell $Y \sim X$ mit X als Faktor mit zwei Levels vorliegen. Welche Aussage über $s_1^2 \neq s_2^2$ ist richtig?

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

3 Aufgabe

(2 Punkte)

Das Falsifikationsprinzip besagt...

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

4 Aufgabe


(2 Punkte)



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

- A ☐ Die Wahrscheinlichkeit der Daten unter der Nullhypothese in der Grundgesamtheit.
- B ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit der Alternativhypothese und somit $1 - Pr(H_A)$
- C ☐ Die Inverse der Wahrscheinlichkeit unter der die Nullhypothese nicht mehr die Alternativhypothese überdeckt.
- D ☐ $Pr(D|H_0)$ ist die Wahrscheinlichkeit die Daten D zu beobachten wenn die Nullhypothese wahr ist.
- E ☐ Die Wahrscheinlichkeit für die Nullhypothese, wenn die Daten wahr sind.

5 Aufgabe

(2 Punkte)

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

- A ☐ Die Spracherkennung von  ist nicht in der Lage Deutsch zu verstehen.
- 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 ☐ Im Allgemeinen haben Programmiersprachen Probleme mit Umlauten und Sonderzeichen, die in der deutschen Sprache vorkommen. Eine Nutzung der englischen Sprache umgeht dieses Problem auf einfache Art.
- D ☐ Alle Funktionen und auch Anwendungen sind in  in englischer Sprache. Die Nutzung von deutschen Wörtern ist nicht schick und das ist zu vermeiden.
- E ☐ Es gibt keinen Grund nicht auch deutsche Wörter zu verwenden. Es ist ein Stilmittel.

6 Aufgabe

(2 Punkte)

Berechnen Sie den Mittelwert und Standardabweichung von y mit 11, 13, 9, 10 und 15.

- A ☐ Es ergibt sich 11.6 ± 2.41
- B ☐ Es ergibt sich 11.6 ± 1.205
- C ☐ Es ergibt sich 10.6 ± 2.9
- D ☐ Es ergibt sich 12.6 ± 1.205
- E ☐ Es ergibt sich 11.6 ± 5.8

7 Aufgabe

(2 Punkte)

Die Randomisierung von Beobachtungen bzw. Samples 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 Varianzen zu schätzen. Ohne eine Randomisierung ist die Berechnung von Mittelwerten und Varianzen nicht möglich.
- 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.
- E ☐ Randomisierung war bis 1952 bedeutend, wurde dann aber in Folge besserer Rechnerleistung nicht mehr verwendet. Aktuelle Statistik nutzt keine Randomisierung mehr.

8 Aufgabe

(2 Punkte)

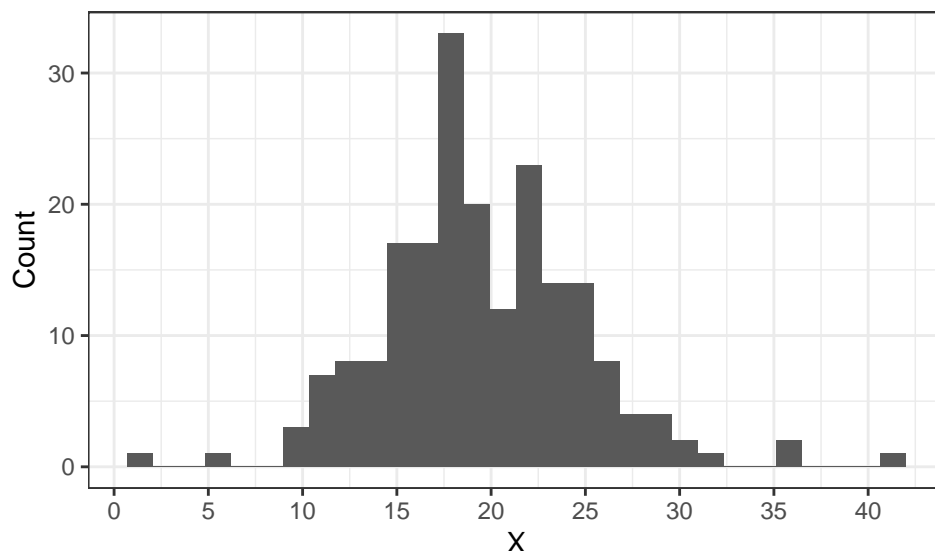
Welche Aussage über den t-Test ist richtig?

- ☐ A Der t-Test vergleicht die Mittelwerte von zwei Gruppen.
- ☐ B Der t-Test ist ein Vortest der ANOVA und basiert daher auf dem Vergleich von Streuungsparametern
- ☐ C 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.
- ☐ D Der t-Test vergleicht die Varianzen von mindestens zwei oder mehr Gruppen
- ☐ E Der t-Test testet generell zu einem erhöhten α -Niveau von 20%.

9 Aufgabe

(2 Punkte)

In dem folgenden Histogramm von $n = 200$ Pflanzen ist welche Verteilung mit welchen korrekten Verteilungsparametern dargestellt?



- ☐ A Eine rechtsschiefe, multivariate Normalverteilung.
- ☐ B Eine Standardnormalverteilung mit $N(0,1)$.
- ☐ C Es handelt sich um eine Poisson-Verteilung mit $\text{Pois}(20)$.
- ☐ D Es handelt sich um eine Normalverteilung mit $N(20, 5)$.
- ☐ E Es handelt sich um eine Binomial-Verteilung mit $\text{Binom}(10)$.

10 Aufgabe

(2 Punkte)

In der Theorie zur statistischen Testentscheidung kann „ H_0 beibehalten obwohl die H_0 falsch ist“ in welche richtige Analogie gesetzt werden?

- ☐ A In die Analogie eines Rauchmelders: *Alarm with fire*.
- ☐ B In die Analogie eines brennenden Hauses ohne Rauchmelder: *House without noise*.

- C** ☐ In die Analogie eines Rauchmelders: *Alarm without fire*, dem α -Fehler.
- D** ☐ In die Analogie eines Feuerwehrautos: *Car without noise*.
- E** ☐ In die Analogie eines Rauchmelders: *Fire without alarm*, dem β -Fehler.

11 Aufgabe

(9 Punkte)

Nach einem Feldexperiment mit zwei Düngestufen (A und B) ergibt sich die folgende Datentabelle mit dem gemessenen Trockengewicht (*drymatter*).

trt	drymatter
A	20
A	18
B	20
B	16
A	16
B	25
A	11
A	11
B	24
B	21
B	22
A	23
B	18
A	23

1. Zeichnen Sie in *einer* Abbildung die beiden Boxplots für die zwei Düngestufen A und B! Beschriften Sie die Achsen entsprechend! **(6 Punkte)**
2. Beschriften Sie *einen* der beiden Boxplots mit den gängigen statistischen Maßzahlen. Beschriften Sie auch die Achsen! **(2 Punkte)**
3. Wenn Sie *keinen Effekt* zwischen de Düngestufen erwarten würden, wie sehen dann die beiden Boxplots aus? **(1 Punkt)**

12 Aufgabe

(10 Punkte)

In einem Feldexperiment für die Bodendurchlässigkeit wurde der Niederschlag pro Parzelle sowie der durchschnittliche Ertrag gemessen. Es ergibt sich folgende Datentabelle.

water	drymatter
14	12
15	19
18	21
13	21
14	21

1. Erstellen Sie den Scatter-Plot für die Datentabelle. Beschriften Sie die Achsen entsprechend! **(4 Punkte)**
2. Zeichnen Sie eine Gerade durch die Punkte! **(1 Punkt)**
3. Beschriften Sie die Gerade mit den gängigen statistischen Maßzahlen! **(4 Punkte)**
4. Wenn kein Effekt von dem Niederschlag auf das Trockengewicht vorhanden wäre, wie würde die Gerade verlaufen und welche Werte würden die statistischen Maßzahlen annehmen? **(1 Punkt)**

13 Aufgabe

(10 Punkte)

Nach einem Experiment ergibt sich die folgende 2x2 Datentabelle mit einem Pestizid (ja/nein), dargestellt in den Zeilen. Im Weiteren mit dem infizierten Pflanzenstatus (ja/nein) in den Spalten. Insgesamt wurden $n = 142$ Pflanzen untersucht.

	Erkrankt (ja)	Erkrankt (nein)	
Pestizid (ja)	56	21	
Pestizid (nein)	27	38	

1. Ergänzen Sie die Tabelle um die Randsummen! **(1 Punkt)**
2. Formulieren Sie die Fragestellung! **(1 Punkt)**
3. Formulieren Sie das Hypothesenpaar! **(2 Punkte)**
4. Berechnen Sie die Teststatistik eines Chi-Quadrat-Test auf der 2x2 Tafel. Geben Sie Formeln und Rechenweg mit an! **(4 Punkte)**
5. Treffen Sie eine Entscheidung im Bezug zu der Nullhypothese gegeben einem $T_k = 3.841$! **(1 Punkt)**
6. Skizzieren Sie eine 2x2 Tabelle mit $n = 26$ Pflanzen in dem *vermutlich* die Nullhypothese nicht abgelehnt werden kann! **(1 Punkt)**

14 Aufgabe

(4 Punkte)

1. Skizzieren Sie in die unten stehende, freie Abbildung die Abbildung, die sich nach der Überschrift ergibt! **(2 Punkte)**
2. Beschriften Sie die Achsen entsprechend! **(2 Punkte)**

Residual plot with 3 outlier fullfiling the normality assumption.



15 Aufgabe

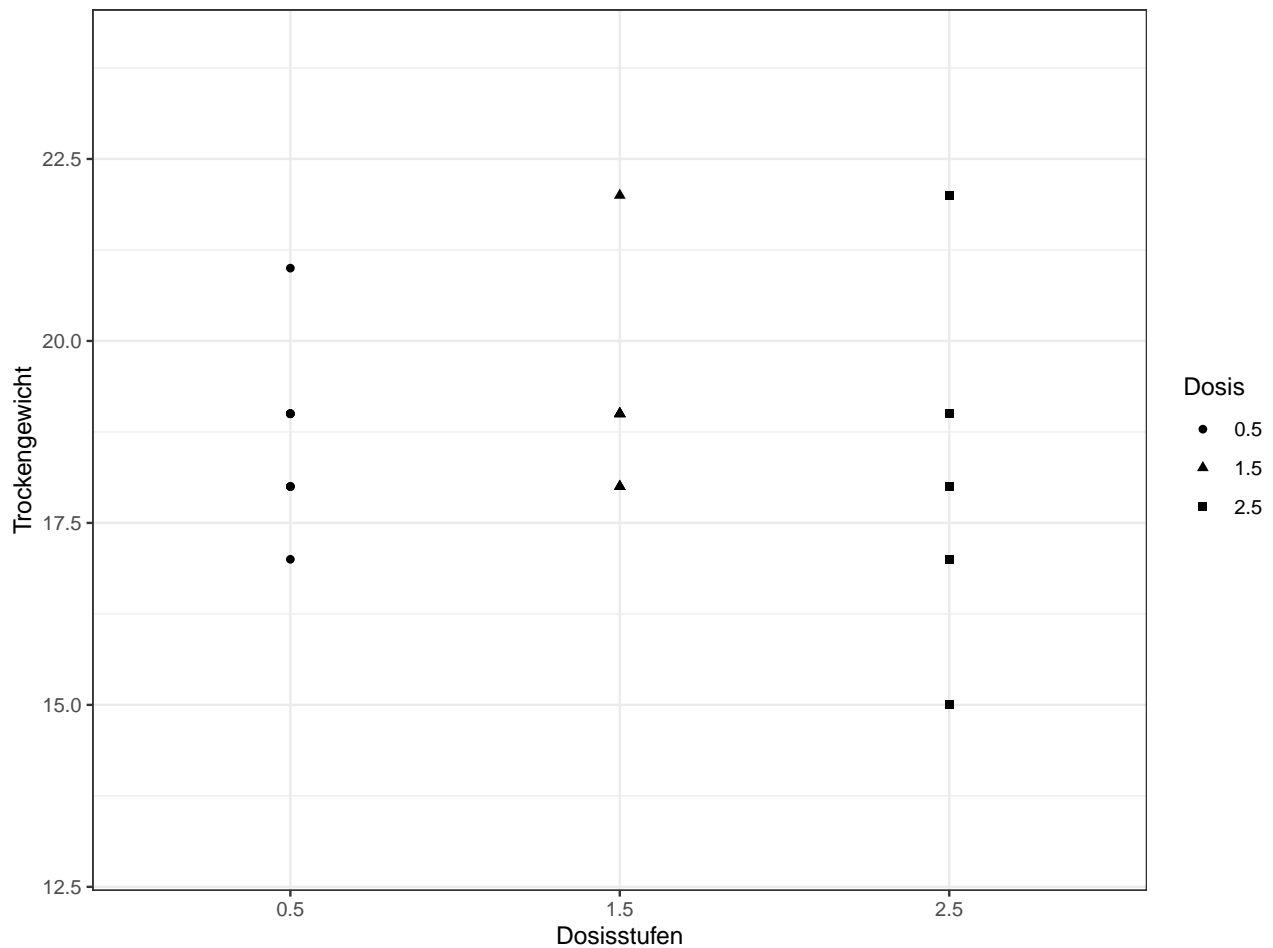
(6 Punkte)

1. Skizzieren Sie 3 Normalverteilungen *in einer Abbildung* mit $\mu_1 \neq \mu_2 \neq \mu_3$ und $s_1 = s_2 = s_3$! **(2 Punkte)**
2. Beschriften Sie die Normalverteilungen mit den entsprechenden Parametern! **(2 Punkte)**
3. Liegt Varianzhomogenität oder Varianzheterogenität vor? Begründen Sie Ihre Antwort! **(2 Punkte)**

16 Aufgabe

(8 Punkte)

In einem Experiment wurde der Ertrag von Erbsen unter drei verschiedenen Pestizid-Dosen 0.5 g/l, 1.5 g/l und 2.5 g/l gemessen. Unten stehenden sehen Sie die Visualisierung des Datensatzes.



1. Zeichnen Sie folgende statistischen Masszahlen in die Abbildung ein! (6 Punkte)

- Total (grand) mean: β_0
- Mittelwerte der Dosen: $\bar{x}_{0.5}$, $\bar{x}_{1.5}$ und $\bar{x}_{2.5}$
- Effekt der einzelnen Level der Dosen: $\beta_{0.5}$, $\beta_{1.5}$, und $\beta_{2.5}$
- Residuen oder Fehler: ϵ

2. Schätzen Sie den p-Wert einer einfaktoriellen ANOVA ab. Liegt ein *vermutlicher* signifikanter Unterschied zwischen den Dosisstufen vor? Begründen Sie Ihre Antwort! (2 Punkte)

17 Aufgabe

(14 Punkte)

Der Datensatz PlantGrowth enthält das Gewicht der Pflanzen (*weight*), die unter einer Kontrolle und zwei verschiedenen Behandlungsbedingungen erzielt wurden – dem Faktor *group* mit den Faktorstufen *ctrl*, *trt1*, *trt2*.

1. Füllen Sie die unterstehende einfaktorielle ANOVA Ergebnistabelle aus mit den gegebenen Informationen von Df und Sum Sq! **(4 Punkte)**
2. Schätzen Sie den p-Wert der Tabelle mit der Information von $F_k = 3.35$ ab. Begründen Sie Ihre Antwort! **(2 Punkte)**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	4			
Residuals	27	11.25			

3. Was bedeutet ein signifikantes Ergebnis in einer einfaktoriellen ANOVA im Bezug auf die möglichen Unterschiede zwischen den Gruppen? **(2 Punkte)**
4. Berechnen Sie *einen* Student t-Test mit $T = \frac{\bar{x}_1 - \bar{x}_2}{s_p \cdot \sqrt{2/n_g}}$ für den *vermutlich* signifikantesten Gruppenvergleich anhand der untenstehenden Tabelle mit $T_k = 2.03$. Begründen Sie Ihre Auswahl! **(4 Punkte)**

group	n	mean	sd
ctrl	10	5.05	0.63
trt1	10	4.63	0.81
trt2	10	5.53	0.44

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

