

Technische Universität Dortmund

Fakultät Statistik

Wintersemester 2022/2023

Fallstudien I

Projekt 3: Auswertung eines Versuchsplans

Prof. Dr. Guido Knapp

M. Sc. Yassine Talleb

Bericht von: Louisa Poggel

Mitglieder der Gruppe 1:

Caroline Baer

Daniel Sipek

Julia Keiter

Louisa Poggel

01.12.2022

Inhaltsverzeichnis

1	Einleitung	1
2	Problemstellung	1
3	Statistische Methoden	2
4	Statistische Auswertung	6
4.1	Deskriptive Analyse	6
4.2	Aufstellen des Modells und Überprüfung der Annahmen	7
4.3	Varianzanalyse und Tests	10
5	Zusammenfassung	12
	Literaturverzeichnis	13
	Anhang	14

1 Einleitung

Wichtiger Bestandteil einer erfolgreichen Datenanalyse eines Experimentes ist das richtige Erkennen des zugrunde liegenden Versuchsplanes. Daher ist es zunächst das Ziel das Versuchsdesign des landwirtschaftlichen Versuches Datensatzes *Hafer.xlsx* (Yates (1935)), der sich mit dem *Ertrag* von verschiedenen *Hafersorten* unter verschiedenen *Stickstoff-Behandlungen* beschäftigt, zu bestimmen. Es wird sich herausstellen, dass ein Split-Plot Design vorliegt. Dieses besteht aus einem *Blockfaktor*, einem whole-plot Faktor (*Hafer-sorte*) und einem split-plot Faktor (*Stickstoff-Behandlung*).

Um festzustellen ob es signifikante Unterschiede zwischen den Faktoren bestehen wird ein gemischtes Varianzanalysemodell aufgestellt. Es wird sich herausstellen, dass nur bei der *Stickstoff-Behandlung* signifikante Unterschiede zum Niveau $\alpha = 0.05$ vorliegen. Post-hoc Tests zeigen sogar, dass fast alle Faktorstufen der *Stickstoff-Behandlung* signifikant zum multiplen Niveau $\alpha = 0.05$ unterschiedlich sind.

Im folgendem werden zunächst der Versuchsaufbau und die Variablen des Datensatzes in der Problemstellung in Kapitel 2 beschrieben. Darauf folgt eine Darstellung der verwendeten statistischen Methoden in Kapitel 3. Die statistische Auswertung des Versuchsplans erfolgt in Kapitel 4, wobei zunächst eine deskriptive Analyse (4.1) erfolgt. Dann wird das Modell aufgestellt und Modellannahmen überprüft (4.2). Darauf werden Tests und eine Varianzanalyse (4.3) durchgeführt. Zum Schluss werden die zentralen Ergebnisse in Kapitel 5 zusammengefasst.

2 Problemstellung

Der in der wissenschaftlichen Arbeit von Yates (1935) vorzufindende Datensatz *Hafer.xlsx* beinhaltet Daten zu einem landwirtschaftlichen Versuchsplan im Split-Plot Design. Dort sind die nominalen Variablen *Block*, *Hafersorte*, *Stickstoff-Behandlung* (kurz: *Behandlung*) und die metrische Variable *Ertrag* mit jeweils 72 Einträgen enthalten.

Der *Block* hat die Stufen I bis VI. Die *Hafersorte* ist der whole-plot Faktor, da es sich um einen „hard-to-change“ Faktor handelt. Denn Hafer lässt sich nur innerhalb genügend großen Parzellen behandeln und anpflanzen. Pro Block erfolgt dann eine randomisierte Aufteilung durch den Faktor *Hafersorte* in die drei whole-plots. Dort wird

jeweils eine der drei *Hafersorten* namens „Victory“, „Golden.rain“ oder „Marvellous“ angepflanzt. Je whole-plot erfolgt eine erneute Randomisierung des vierstufigen sup-plot Faktors *Stickstoff-Behandlung* mit den Ausprägungen 0.0 cwt, 0.2 cwt, 0.4 cwt und 0.6 cwt. Dabei bezeichnet cwt eine alte angloamerikanische Einheit. So entstehen pro whole-plot Faktor vier $\frac{1}{80}$ acre große sup-plot Einheiten auf denen der *Ertrag* in $\frac{1}{4}$ lbs gemessen wird. Da keinerlei fehlende Werte vorliegen ist die Datenqualität sehr gut. Ziel dieses Projektes ist es, das korrekte statistische Modell zum Versuchsplan in Form eines Varianzanalysemodells unter Berücksichtigung von Haupteffekten, Fehlertermen und möglichen Wechselwirkungen aufzustellen. Mithilfe dieses Modells lässt sich untersuchen, ob signifikante Unterschiede zum Niveau $\alpha = 0.05$ zwischen den *Hafersorten* und den vier Stufen der *Stickstoff-Behandlungen* vorliegen. Sollte dies der Fall sein, wird mithilfe von Post-hoc Tests zum multiplen Niveau $\alpha = 0.05$ ergründet, welcher der Faktorausprägungen der Verursacher des signifikanten Unterschiedes ist. Dabei werden sowohl multiple Vergleiche nach Scheffe als auch nach Tukey vorgenommen.

3 Statistische Methoden

Alle folgenden statistischen Methoden werden in der Version 4.1.1 der Software R durchgeführt (R Core Team (2021)). Dabei wird bei Ergebnissen, wenn nicht anderes angegeben, auf zwei Nachkommastellen gerundet. Zur Analyse des Versuches im split-plot Design wird das in 1 definierte Varianzanalysemodell mit zwei Fehlertermen genutzt, welches auch in Dean et al. (2017) auf Seite 705 zu finden ist.

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + e_{ij}^w + e_{ijk}^s \quad (1)$$

$$i = 1, \dots, I \quad j = 1, \dots, J \quad k = 1, \dots, K \quad I, J, K \in \mathbb{N} \quad (2)$$

Dabei ist I die Anzahl der Blöcke, J die Anzahl der Stufen des whole-plot Faktors und K die Anzahl der Stufen des sub-plot Faktors. Das Modell geht von einem balancierten Design aus, bei dem alle Faktorstufe gleich viele Beobachtungen, bezeichnet mit n_1, n_2, n_3 , hat. Mit μ wird das allgemeine Mittel bezeichnet und $\alpha_i, \beta_j, \gamma_k$ sind feste Effekte der Faktoren A (Block Faktor), B (whole-plot Faktor) und C (sub-plot Faktor). Dann ist $(\beta\gamma)_{jk}$ die Wechselwirkung zwischen der Stufe j des Faktors B mit der Stufe k des Faktors C. Damit die KQ-Schätzung des Parametervektors eindeutig ist und sich

die Effekte als einen differenziellen Effekt zum allgemeinen Mittel interpretieren lassen können, müssen die Nebenbedingungen aus 3 gelten.

$$\sum_{i=1}^I \alpha_i = 0 ; \quad \sum_{j=1}^J \beta_j = 0 ; \quad \sum_{k=1}^K \gamma_k = 0 ; \quad \sum_{j=1}^J (\beta\gamma)_{jk} = 0 \quad \forall k ; \quad \sum_{k=1}^K (\beta\gamma)_{jk} = 0 \quad \forall j \quad (3)$$

Mit e_{ij}^w wird der whole-plot Fehler und mit e_{ijk}^s der sub-plot Fehler bezeichnet. Für diese gilt, dass sie normalverteilt sind mit einem Erwartungswert von null und einer Varianz von σ_w^2 bzw. σ_s^2 , d.h es gilt $e_{ij}^w \sim N(0, \sigma_w^2)$ und $e_{ijk}^s \sim N(0, \sigma_s^2)$. Somit ergibt sich für die Kovarianz von y die in 4 dargestellte Kovarianzstruktur in Blockdiagonalgestalt.

$$Cov(y_{ijk}, y_{i'j'k'}) = \begin{cases} \sigma_w^2 + \sigma_s^2 & i = i', j = j', k = k' \\ \sigma_w^2 & i = i', j = j', k \neq k' \\ 0 & sonst \end{cases} \quad (4)$$

Zur Implementierung dieses Modells in R wird, wie in Crawley (2012) auf Seite 398 beschrieben, die Funktion `aov()` mit Angabe einer besonderen Struktur der Fehler mit `Error()` genutzt.

Die Informationen für die im folgendem beschriebenen Bestandteile einer Varianzanalyse-tafel (Quadratsummen, Mittlere Quadratsummen, Freiheitsgrade, F-Statistik bzw. F-Test) stammen von Hartung et al. (2009), ergänzt mit Angaben zu split-plots von Dean et al. (2017). Um festzustellen, auf welche Faktoren sich die Gesamtvariabilität aufteilt, werden Quadratsummen bezüglich des jeweiligen Faktors genutzt (Hartung et al. (2009) S. 626 und Dean et al. (2017) S. 706). Im vorliegenden Modell müssen dabei insbesondere die zwei Fehlerterme beachtet werden. Dabei setzen sich die Quadratsummen bezüglich der Faktoren A und B, die gesamte Variabilität im whole-plot (SSW) und die Fehlerquadratsumme des whole-plots (SSE_w) wie in 5 und 6 zusammen.

$$SSA = JK \sum_{i=1}^I \hat{\alpha}_i^2 = JK \sum_{i=1}^I (\bar{y}_{i..} - \bar{y}_{...})^2 ; \quad SSB = IK \sum_{j=1}^J \hat{\beta}_j^2 = IK \sum_{j=1}^J (\bar{y}_{.j.} - \bar{y}_{...})^2 \quad (5)$$

$$SSW = K \sum_{i=1}^I \sum_{j=1}^J (\bar{y}_{ij.} - \bar{y}_{...})^2 ; \quad SSE_w = SSW - SSA - SSB \quad (6)$$

Hier bedeutet die Punktnotation beim arithmetischen Mittel, dass beim Vorliegen eines Punktes über den jeweiligen Index summiert wird. Das heißt beispielsweise $\bar{y}_{ij.} = \frac{1}{K} \sum_{k=1}^K y_{ijk}$. Die Quadratsummen bezüglich des Faktors C (SSC) und der Wechselwirkung zwischen B und C ($SSBC$) werden in 8 und 9, wie beim SSA und SSB, über

die KQ-Schätzer berechnet. Dabei wird das Minimierungsproblem aus 7 gelöst, welches mithilfe des Lagrange Ansatzes unter den Nebenbedingungen aus 3 gelöst wird (Dean et al. (2017) S. 674).

$$\min_{\mu, \alpha_i, \beta_j, \gamma_k, (\beta\gamma)_{jk}} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k - (\beta\gamma)_{jk})^2 \quad (7)$$

Die Fehlerquadratsumme des sub-plots (SSE_s) lässt sich aus der Differenz der Gesamtvariabilität (SSG) und der Variabilität des whole-plots (SSW) und des SSC und $SSBC$ wie in 10 berechnen.

$$SSC = IJ \sum_{k=1}^K \hat{\gamma}_k^2 = IJ \sum_{k=1}^K (\bar{y}_{..k} - \bar{y}_{...})^2 \quad (8)$$

$$SSBC = I \sum_{j=1}^J \sum_{k=1}^K (\widehat{\beta\gamma})_{jk}^2 = I \sum_{j=1}^J \sum_{k=1}^K (\bar{y}_{.jk} - \bar{y}_{.j.} - \bar{y}_{..k} + \bar{y}_{...})^2 \quad (9)$$

$$SSG = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (y_{ijk} - \bar{y}_{...})^2 ; \quad SSE_s = SSG - SSW - SSC - SSBC \quad (10)$$

Zur Betrachtung der Mittleren Quadratsummen werden die obig genannten Quadratsummen durch ihren Freiheitsgrad geteilt, welche in Tabelle 1 zu finden sind (Dean et al. (2017) S.706). Mithilfe der Mittleren Quadratsummen lässt sich der F-Test wie in

Tabelle 1: Freiheitsgrade

Faktor A (Block)	$I - 1$
Faktor B	$J - 1$
Fehler (whole-plot)	$(I - 1)(J - 1)$
Gesamt (whole-plot)	$IJ - 1$
Faktor C	$K - 1$
Faktor BC	$(J - 1)(K - 1)$
Fehler (split-plot)	$J(K - 1)(I - 1)$
Gesamt	$IKJ - 1$

11 definieren, der die Hypothese $H_0 = t_1 = \dots = t_p$ gegen $H_0 : \exists h \neq h'$ mit $t_h \neq t_{h'}$ und $h, h' \in 1, \dots, p$ testet (Hartung et al. (2009), S. 611 bis 612). Dabei ist t_h ein Platzhalter für die $h = 1, \dots, p, p \in \mathbb{N}$ Effekte eines beliebigen Faktors T.

$$F := \frac{MST}{MSE} \quad \varphi_F := \begin{cases} 1 & F > f_{df_1, df_2, 1-\alpha} \\ 0 & F < f_{df_1, df_2, 1-\alpha} \end{cases} \quad (11)$$

Mit MST wird die Mittlere Quadratsumme des Faktors T bezeichnet. Das kleine f ist das $1 - \alpha$ Quantil der F-Verteilung mit df_1 und df_2 Freiheitsgraden. Dabei ist df_1 der Freiheitsgrad des Faktors T und df_2 der Freiheitsgrad des Fehlers. Da im split-plot Design zwei Fehlerterme beachtet werden, muss bei dem Faktor B durch den SSE_w und bei den Faktor C und der Wechselwirkung BC durch den SSE_s geteilt werden. Falls der Test φ_F eins ist, wird H_0 abgelehnt. Das heißt es besteht ein signifikanter Unterschied zwischen den Stufen des jeweiligen Faktors. Ist der Test φ_F null wird H_0 beibehalten. Die Varianz der Fehlerterme σ_w^2 und σ_s^2 werden durch die ANOVA-Schätzer (Hartung et al. (2009) S. 630) wie in 12 geschätzt.

$$\hat{\sigma}_s^2 = MSE_s \quad \hat{\sigma}_w^2 = \frac{1}{K}(MSE_w - MSE_s) \quad (12)$$

Die Schätzer ergeben sich aus der Momentenmethode. Denn $E(MSE_s)$ ergibt sich als σ_s^2 und $E(MSE_w)$ ist $K\sigma_w^2 + \sigma_s^2$.

Falls ein signifikanter Unterschied durch den F-Test festgestellt wird, werden multiple Tests nach Scheffe und Tukey durchgeführt, um festzustellen welche Faktoren für das Ablehnen der Nullhypothese verantwortlich sind. Die Nullhypothese lautet für einen beliebigen Effekt t_h mit $h = 1 \dots p$, $p \in \mathbb{N}$ beispielsweise $H_0 : t_1 - t_2 = 0$ oder $H_0 : t_2 - t_3 = 0$. Die multiplen Tests halten dabei das multiple Niveau α ein. Das heißt die Wahrscheinlichkeit die Nullhypothese fälschlicherweise abzulehnen ist kleiner als α , egal wie viele Nullhypothesen angenommen werden.

Für die Tests werden Elementarkontraste c verwendet, für die gilt $\sum_{l=1}^L c_l = 0$, bei der $L \in \mathbb{N}$ die Länge des Vektors c ist (Dean et al. (2017), S. 708). Zudem gilt bei Elementarkontrasten $c^T t = t_h - t'_h$ mit $h \neq h'$ und $h \in \mathbb{N}$, wenn an h-ter Stelle eine 1 und an h'-ter Stelle eine -1 steht. Dann lassen sich mit $c^T t$ Testprobleme wie oben darstellen. Nun lässt sich der Scheffe-Test wie in 13 aufstellen (Hartung et al. (2009), S. 616).

$$FSD := \sqrt{(p-1) MSE \left(\frac{1}{n_1} + \frac{1}{n_2} \right) f_{df_1, df_2, 1-\alpha}} \quad (13)$$

$$\varphi_{Scheffe} := \begin{cases} 1 & |c^T \hat{t}| > FSD \\ 0 & |c^T \hat{t}| < FSD \end{cases} \quad (14)$$

Dabei ist \hat{t} der Ausschnitt der KQ-Schätzung für den Parametervektor bezüglich des Effektes t . Mit n_1 und n_2 sind die Anzahl an Messungen für eine Faktorstufe h von t gemeint, die im vorliegenden balancierten Design gleich sind. Auch der MSE muss je

nach Faktor, im vorliegenden Modell 1, als MSE_w oder MSE_s gewählt werden. Das Quantil der F-Verteilung f ist wie beim F-Test 11 definiert. Falls der Wert von $|c^T \hat{t}|$ größer ist als der, als Grenzdifferenz nach Scheffe bezeichnete FSD , bestehen signifikante Unterschiede zwischen den betrachteten Koeffizienten t_h und t'_h .

Ähnlich zum Scheffe-Test kann bei einem balancierten Design, das heißt es gilt $n_1 = n_2 = n$, auch der Tukey-Test (15) für multiple Testprobleme verwendet werden (Hartung et al. (2009), S. 616).

$$HSD := \sqrt{\frac{MSE}{n} q_{p,df_2,1-\alpha}} \quad (15)$$

$$\varphi_{Tukey} := \begin{cases} 1 & |c^T \hat{t}| > HSD \\ 0 & |c^T \hat{t}| < HSD \end{cases} \quad (16)$$

Hier gilt zusätzlich als Voraussetzung, dass die Kovarianzmatrix von $\hat{\beta}$ Blockdiagonalgestalt hat. Das heißt die Varianzen und Kovarianzen sind gleich. Zudem wird das Quantil $q_{p,df_2,1-\alpha}$ der Verteilung der studentisierten Spannweite benötigt. Interpretieren lässt sich der Tukey-Test analog zum Scheffe-Test. Konfidenzintervalle lassen sich für $c^T \hat{t}$ bei Scheffe mit $c^T \hat{t} \pm FSD$ und bei Tukey mit $c^T \hat{t} \pm HSD$ bilden (Toutenburg (2003), S. 216).

Zur Visualisierung der Wechselwirkungen zwischen zwei Faktoren (1 und 2) wird ein Interaktionsplot genutzt (Hartung et al. (2009) S. 361). Dort steht auf der y-Achse die Zielvariable und auf der x-Achse der Faktor 2. Abgetragen werden dann mehrere Linien, welche die Gruppenmittel pro Stufe des Faktors 1 kombiniert mit der Stufe des Faktors 2 verbinden. Haben die Linien alle einen gleichen, parallelen Verlauf liegt keine Wechselwirkung vor. Ist dies nicht der Fall könnten Wechselwirkungen vorliegen.

4 Statistische Auswertung

4.1 Deskriptive Analyse

In Tabelle 2 liegen Kennzahlen der zu erklärenden Variable *Ertrag* vor. Die geringe Abweichung des arithmetische Mittels (103.97 1/4 lbs) und des Medians (102.50 1/4 lbs) spricht für Symmetrie. Die Standardabweichung liegt bei etwa 26.06. Der Schiefekoeff-

fizient deutet mit 0.27 auf eine leichte Rechtsschiefe hin. Die Wölbung ist mit einem Wert von 2.61 etwas flacher als bei der Normalverteilung. Dies ist auch im Histogramm des *Ertrags* ist im Anhang auf Seite 14 in Abbildung 5 erkennbar. Die Wölbungs- und Schiefemaße wurden mithilfe des Paketes `moments` berechnet (Komsta und Novomestky (2022)). Boxplots des *Ertrages* getrennt nach *Hafersorte* (Abbildung 6) und *Stickstoff-*

Tabelle 2: Deskriptive Kennzahlen des *Ertrags* (IQR = Interquartilsabstand)

	arithm. Mittel	Median	Standardabweichung	IQR	Schiefe	Wölbung
<i>Ertrag</i>	103.97	102.50	27.06	35.25	0.27	2.61

Behandlung (Abbildung 7) sind im Anhang auf den Seiten 14 bis 15 zu finden. Es ist kein großer Unterschied zwischen den drei *Hafersorten* erkennbar, da die Mediane nah beieinander liegen. Der *Ertrag* der *Hafersorten* „Golden.rain“ und „Victory“ streut etwas mehr als bei der Sorte „Marvellous“. Hingegen lässt sich bei dem *Ertrag*, getrennt nach *Stickstoff-Behandlung*, ein Trend erkennen. Je größer die Stickstoffmenge ist, desto größer ist der *Ertrag*. Dabei liegen die Mediane des *Ertrags* recht weit auseinander. Nur die Mediane der *Stickstoff-Behandlungen* 0.0cwt und 0.6cwt weichen wenig ab.

4.2 Aufstellen des Modells und Überprüfung der Annahmen

Vor der Modellbildung wird anhand des Interaktionsplots (Abbildung 1) geprüft, ob eine Wechselwirkung zwischen *Hafersorte* und *Stickstoff-Behandlung* vorliegt.

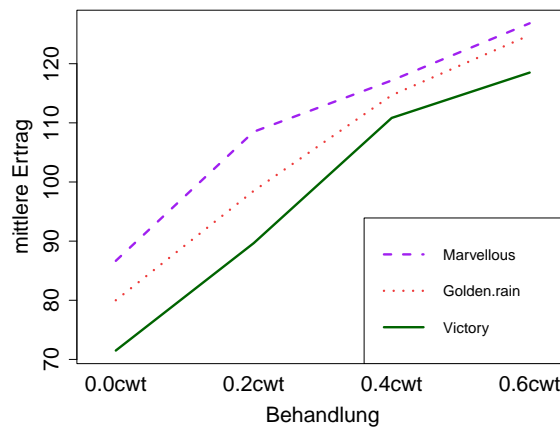


Abbildung 1: Interaktionen zwischen *Hafersorte* und *Stickstoff-Behandlung*

Dort ist zu erkennen, dass die Linien der drei *Hafersorten* einen recht ähnlichen Verlauf haben. Es sind jedoch leichte Knicke bei den *Hafersorten* „Marvellous“ und „Victory“ zu erkennen, die bei der Sorte „Golden.rain“ nicht auftreten. Daher wird die Wechselwirkung mit ins Modell aufgenommen, da ein Interaktionseffekt nicht ausgeschlossen werden kann. Somit ergibt sich für die Analyse des split-plot Plans das gemischte lineare Modell aus 17 inklusive des Interaktionseffektes zwischen *Hafersorte* und *Stickstoff-Behandlung*. Die Nebendbedingungen gelten wie im Kapitel 3 beschrieben unter Verwendung von $I = 6$, $J = 3$ und $K = 4$.

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + e_{ij}^w + e_{ijk}^s \quad (17)$$

$$i = 1, \dots, 6 \quad j = 1, \dots, 3 \quad k = 1, \dots, 4 \quad (18)$$

Nun lassen sich konkret die Parameter des Modells unter den Verteilungsannahmen der Fehlerterme folgendermaßen beschreiben:

μ : allgemeine Mittel

α_i : fester Effekt der i-ten Stufe des *Blockes*

β_j : fester Effekt der j-ten Stufe der *Hafersorte*

γ_k : fester Effekt der k-ten Stufe der *Stickstoff-Behandlung*

$(\beta\gamma)_{jk}$: feste Wechselwirkung zwischen der j-ten Stufe der *Hafersorte*
mit der k-ten Stufe der *Stickstoff-Behandlung*

e_{ij}^w : whole-plot Fehler mit $e_{ij}^w \sim N(0, \sigma_w^2)$

e_{ijk}^w : sub-plot Fehler mit $e_{ijk}^w \sim N(0, \sigma_s^2)$

Die Normalverteilungsannahmen der Fehler und die Annahme der Homoskedastizität der Varianzen der Fehler innerhalb des sup-plots bzw. whole-plots werden nun geprüft. Dazu werden zunächst Quantile-Quantile-Plots erstellt, welche die Quantile der Fehlerterme auf der y-Achse darstellen. Auf der x-Achse stehen die Quantile einer Zufallsstichprobe aus der Normalverteilung. Diese hat den Erwartungswert null und die Varianz entspricht der geschätzten Varianz des jeweiligen Fehlerterms ($\hat{\sigma}_w^2 = 106.0618$ bzw. $\hat{\sigma}_s^2 = 177.0833$). Wie diese Schätzungen zustande kommen wird in Kapitel 4.3 genauer beschrieben. In Abbildung 2 ist der QQ-Plot der whole-plot Residuen mit der Winkelhalbierenden (in rot) dargestellt. Es ist erkennbar das ein Großteil der Punkte nahe der Winkelhalbierenden liegt, was für die Erfüllung der Normalverteilungsannahme spricht.

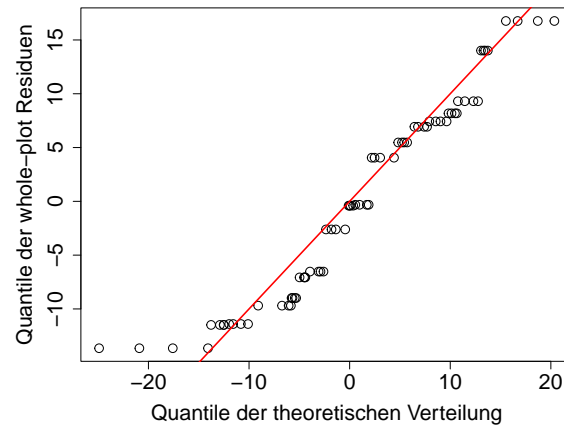


Abbildung 2: QQ-Plot - Residuen (whole-plot)

Analog dazu stellt der QQ-Plot in Abbildung 3 die sub-plot Residuen dar, welche sogar noch besser auf der roten Linie liegen. Lediglich an den Rändern gibt es stärkere Abweichungen. Insgesamt spricht aber auch hier nichts strikt gegen die Normalverteilungsannahme.

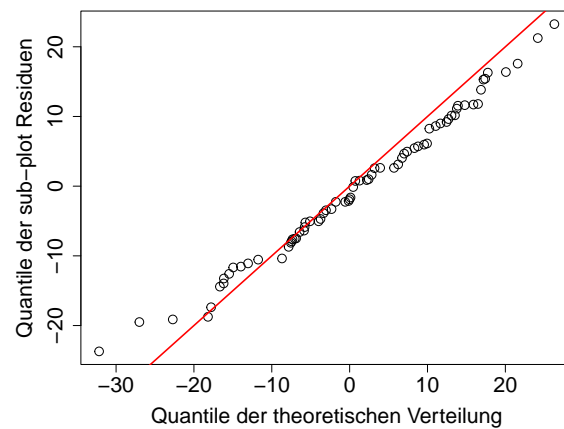


Abbildung 3: QQ-Plot - Residuen (sub-plot)

Zur Prüfung der Homskedastizität der Residuen innerhalb des sup-plot bzw. whole-plot werden in Abbildung 4 die jeweiligen Residuen gegen die angepassten Werte abgetragen. Dabei ist zu beachten, dass sich bei den gefitteten Werte im whole-plot Modell immer

vier Werte wiederholen, da es keine Schätzung für die *Behandlung* gibt. Da sich die whole-plot Residuen als die Differenz $y_{ij} - \hat{y}_{ij}$ ergeben ist dies auch dort der Fall.

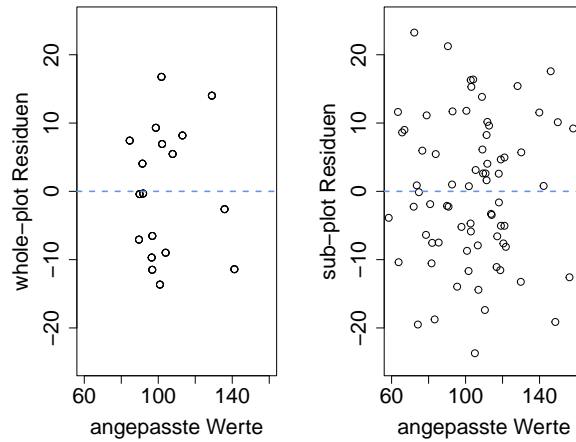


Abbildung 4: Residualplot der whole-plot Residuen (links) und der sub-plot Residuen (rechts)

Es ist aber dennoch erkennbar, dass die wenigen Punkte im Residualplot des whole-plots relativ zufällig, ohne Struktur um die null herum streuen. Auch im Residualplot bezüglich des sub-plots ist keine besondere Struktur erkennbar, sodass nichts gegen die Homoskedastizität der Varianzen spricht.

4.3 Varianzanalyse und Tests

Tabelle 3 beinhaltet die Varianzanalysetafel des Modells. Für die *Hafersorte*, *Behandlung* und die Wechselwirkung zwischen *Hafersorte* und *Behandlung* wird, neben den Quadratsummen, die F-Statistik und der p-Wert des F-Tests angegeben. Dies wäre für den Faktor *Block* auch möglich ist jedoch nicht von Interesse. Es ist erkennbar, dass nur bei der *Stickstoff-Behandlung* ein signifikanter Unterschied zwischen den Faktorstufen besteht. Denn der p-Wert liegt mit etwa $2.46 \cdot 10^{-12}$ nahezu bei Null. Die p-Werte der *Hafersorte* und der Wechselwirkung zwischen *Hafersorte* und *Behandlung* sind mit etwa 0.27 und 0.93 größer als $\alpha = 0.05$. Somit besteht dort kein signifikanter Unterschied. Führt man multiple Vergleiche nach Scheffe und Tukey durch, kommen beide Test auf das Ergebnis, dass alle *Stickstoff-Behandlungen* bis auf die *Behandlungen* 0.4 cwt und 0.6

Tabelle 3: Varianzanalysetafel (QS = Quadratsumme, MQS = Mittlere Quadratsumme, Wechselwirkung* $\hat{=}$ Wechselwirkung zwischen *Hafersorte* und *Behandlung*)

	QS	MQS	Freiheitsgrad	F-Statistik	p-Wert
Block	15 875.28	3175.06	5	-	-
Hafersorte	1786.36	893.18	2	1.49	0.27
whole-plot-error	6013.31	601.33	10	-	-
Gesamt (whole-plot)	23 674.94	1392.64	17	-	-
Behandlung	20 020.50	6673.50	3	37.69	≈ 0
Wechselwirkung*	321.75	53.63	6	0.30	0.93
sub-plot-error	7968.75	177.08	45	-	-
Gesamt	51 985.94	732.20	71	-	-

cwt signifikant zum multiplen Niveau $\alpha = 0.05$ voneinander verschieden sind. Denn die Grenzdifferenzen lassen sich als $FSD \approx 12.88$ und $HSD \approx 11.83$ berechnen. Die Tabelle 4 stellt dann die Werte der Teststatistik $|c^T \hat{t}| = |\bar{y}_{..k} - \bar{y}_{...} - (\bar{y}_{..k'} - \bar{y}_{...})| = |\bar{y}_{..k} - \bar{y}_{..k'}|$ mit $k \neq k'$ dar. Bei den Konfidenzintervallen ist erkennbar, dass nur das Konfidenzintervall

Nr.	Teststatistik	Wert der Teststatistik	KI (Scheffe)	KI (Tukey)
1	$ \bar{y}_{..1} - y_{..2} $	19.5	[6.62, 32.38]	[7.67, 31.33]
2	$ \bar{y}_{..1} - y_{..3} $	34.83	[21.95, 47.72]	[23.00, 46.67]
3	$ \bar{y}_{..1} - y_{..4} $	44.00	[31.12, 56.88]	[32.17, 55.83]
4	$ \bar{y}_{..2} - y_{..3} $	15.33	[2.45, 28.23]	[3.50, 27.17]
5	$ \bar{y}_{..2} - y_{..4} $	24.50	[11.62, 37.38]	[12.67, 36.33]
6	$ \bar{y}_{..3} - y_{..4} $	9.17	[-3.72, 22.05]	[-2.67, 20.99]

Tabelle 4: Übersicht über die Tests nach Scheffe und Tukey (KI = Konfidenzintervall)

von $|\bar{y}_{..3} - y_{..4}|$ die null beinhaltet. Eine grafische Darstellung der Konfidenzintervalle ist in Abbildung 8 auf Seite 15 im Anhang zu finden.

Eine vollständige Schätzung des Parametervektors ist in den Tabellen 5, 6 und 7 im Anhang auf den Seiten 15 bis 16 zu finden. Das allgemeine Mittel wird mit etwa 103.97 geschätzt. Bei den Schätzungen für den *Block* ist erkennbar, dass der Effekt des ersten Blockes mit etwa 31.36 besonders groß geschätzt wird. Bei den *Hafersorten* hat die Sorte „Marvellous“ mit etwa 5.82 die größte positive Schätzung des Effektes. Bei den *Stickstoff-Behandlungen* ist die Schätzung mit etwa 19.42 bei 0.6 cwt am größten. Die meisten Schätzungen der Wechselwirkungen sind betragsmäßig kleiner als zwei.

Die Schätzungen für $\hat{\sigma}_w^2$ und $\hat{\sigma}_s^2$ ergeben sich mithilfe der mittleren Quadratsummen aus Tabelle 3 als $MSE_s = \hat{\sigma}_s^2 \approx 106.06$ bzw. $\frac{1}{4}(MSE_w - MSE_s) = \hat{\sigma}_w^2 \approx 177.08$. Somit ist die geschätzte Varianz des whole-plot Fehlers etwas größer als die des sub-plot Fehlers.

5 Zusammenfassung

Bei der Analyse der Daten eines landwirtschaftlichen Versuchsplans aus dem Datensatz *Hafer.xlsx* wurde zunächst der Versuchsaufbau analysiert und ein Split-Plot Design festgestellt. Daraus resultiert aufgrund der Randomisierung innerhalb der whole-plots und sub-plots ein gemischtes lineares Modell mit zwei Fehlertermen. Außerdem beinhaltet dieses das allgemeine Mittel und feste Effekte des *Blocks*, der *Hafersorte* (whole-plot Faktor), *Stickstoff-Behandlung* (sub-plot Faktor). Zudem deuteten Interaktionsplots auf eine leichte Wechselwirkung zwischen der *Hafersorte* und *Stickstoff-Behandlung* hin, so dass diese auch ins Modell aufgenommen wurde.

Die p-Werte des F-Tests welche im Rahmen der Varianzanalyse analysiert wurden liegen bei der Wechselwirkung zwischen *Hafersorte* und *Stickstoff-Behandlung* und bei dem Effekt der *Hafersorte* mit 0.93 und 0.27 jedoch über dem vorgegebenen Niveau $\alpha = 0.05$. Somit besteht kein signifikanter Unterschied zwischen den Wechselwirkungen und den Faktorstufen der *Hafersorte*. Hingegen ist der p-Wert der *Stickstoff-Behandlung* mit 2.4610^{-12} fast null. Es besteht also ein signifikanter Unterschied zum Niveau $\alpha = 0.05$ zwischen den *Stickstoff-Behandlungen*. Mithilfe von multiplen Vergleichen nach Scheffe und Tukey ließ sich außerdem feststellen, dass alle Stufen der *Stickstoff-Behandlung* bis auf 0.4 und 0.6 cwt signifikant zum multiplen Niveau $\alpha = 0.05$ sind.

Das heißt es lässt sich bezüglich der ersten drei Faktorstufen der *Behandlung* (0.0cwt, 0.2cwt, 0.4cwt) zusammenfassen, dass je höher die Dosis gewählt wird, desto höher der Ertrag ist. Die Erhöhung der Dosis der *Stickstoff-Behandlung* von 0.4 cwt auf 0.6 cwt scheint hingegen keinen Erhöhung des Ertrages mehr hervorzurufen. Weiterführend könnte man Experimente durchführen, die den Effekt der *Stickstoff-Behandlung* anhand einer der hier genutzten *Hafersorten* untersucht. Denn es gab keine signifikante Unterschiede zwischen den *Hafersorten* oder den Wechselwirkungen zwischen *Hafersorte* und *Behandlung*. Dies würde zu einer Komplexitätsreduzierung und Kosteneinsparung des Versuches führen und man könnte die optimale Dosis der *Stickstoff-Behandlung* anhand feineren Abstufungen des Faktors (z.B. 0.0 cwt, 0.1 cwt, 0.2 cwt, 0.3cwt, 0.4 cwt, 0.5 cwt, 0.6 cwt) ergründen.

Literatur

- Crawley, M. (2012). *The R Book. Second Edition*. John Wiley & Sons, Ltd. URL: <https://books.google.de/books?id=-uBIzQEACAAJ>.
- Dean, A., D. Voss und D. Draguljić (2017). *Design and Analysis of Experiments. Second Edition*. Springer International Publishing AG: Cham.
- Hartung, J., B. Elpelt und K.-H. Klösener (2009). *Statistik Lehr- und Handbuch der angewandten Statistik. 15.Auflage*. Oldenbourg Verlag: München.
- Komsta, L. und F. Novomestky (2022). *moments: Moments, Cumulants, Skewness, Kurtosis and Related Tests*. R package version 0.14.1. URL: <https://CRAN.R-project.org/package=moments>.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Toutenburg, H. (2003). *Lineare Modelle: Theorie und Anwendung. 2. Auflage*. Springer-Verlag Berlin Heidelberg.
- Yates, F. (1935). „Complex experiments“. In: *Journal of the Royal Statistical Society Suppl.*(2), 181–247. URL: <https://www.jstor.org/stable/2983638>.

Anhang

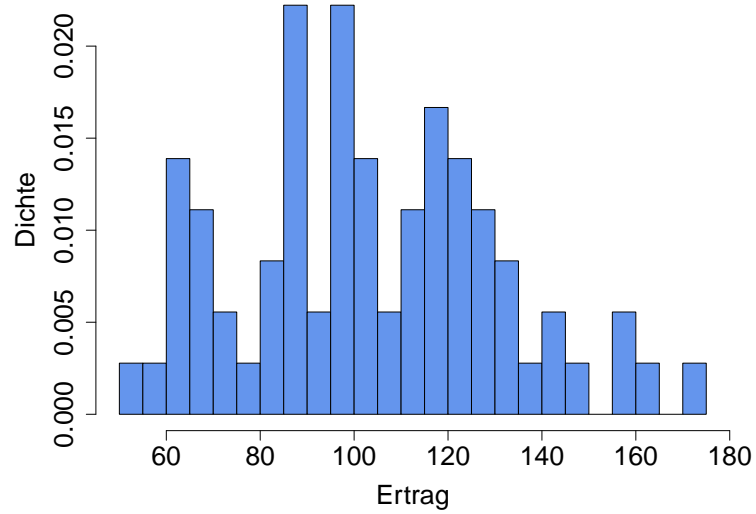


Abbildung 5: Histogramm - *Ertrag*

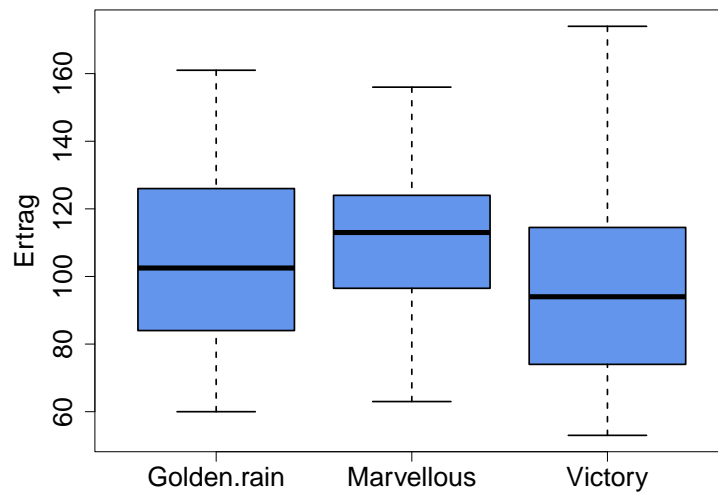


Abbildung 6: Boxplots - *Ertrag* getrennt nach *Hafersorte*

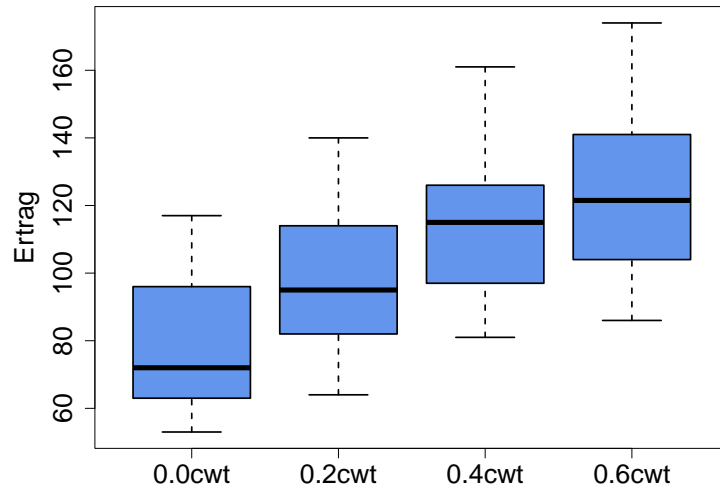


Abbildung 7: Boxplots - *Ertrag* getrennt nach *Behandlung*

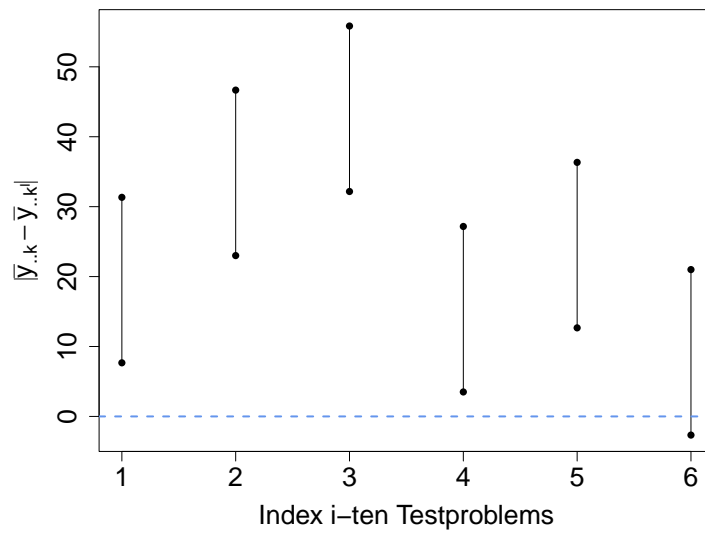


Abbildung 8: Konfidenzintervalle des Tukey-Tests

$\hat{\mu}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\alpha}_6$
103.97	31.36	3.28	-8.06	-5.81	-13.06	-7.72

Tabelle 5: KQ-Schätzungen vom allgemeinen Mittel μ und dem festen Effekt α_i

$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$
0.53	5.82	-6.35	-24.58	-5.08	10.25	19.42

Tabelle 6: KQ-Schätzungen der festen Effekte β_j und γ_k

$(\hat{\beta}\hat{\gamma})_{11}$	$(\hat{\beta}\hat{\gamma})_{12}$	$(\hat{\beta}\hat{\gamma})_{13}$	$(\hat{\beta}\hat{\gamma})_{14}$
0.08	-0.92	-0.08	0.92
$(\hat{\beta}\hat{\gamma})_{21}$	$(\hat{\beta}\hat{\gamma})_{22}$	$(\hat{\beta}\hat{\gamma})_{23}$	$(\hat{\beta}\hat{\gamma})_{24}$
1.46	3.79	-2.88	-2.38
$(\hat{\beta}\hat{\gamma})_{31}$	$(\hat{\beta}\hat{\gamma})_{32}$	$(\hat{\beta}\hat{\gamma})_{33}$	$(\hat{\beta}\hat{\gamma})_{34}$
-1.54	-2.88	2.96	1.46

Tabelle 7: KQ-Schätzungen der Wechselwirkungen $(\beta\gamma)_{jk}$