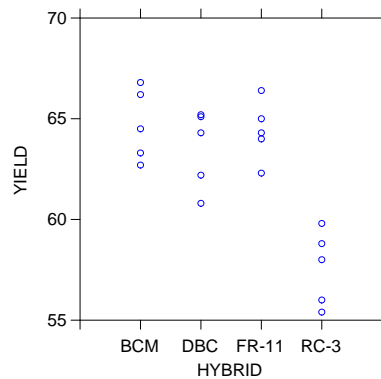


## Musterlösung zur Übung 2

1. a) Aus der graphischen Darstellung geht hervor, dass der Ertrag der Sorte RC-3 geringer ist als die Erträge der übrigen Sorten. Auch zwischen den anderen drei Sorten gibt es Unterschiede, aufgrund des Plots bleibt jedoch unklar, ob diese Unterschiede „signifikant“ sind. Extreme Beobachtungen liegen keine vor. Die Modellvoraussetzung konstanter Fehlervarianz scheint nicht verletzt zu sein: die Streuung der Daten ist in den vier Gruppen etwa gleich gross. (Es wurden keine verschiedenen Symbole für die Orte gewählt.)



- b) Wir haben hier zwei gekreuzte Faktoren mit einer Beobachtung pro Zelle. Das Modell lautet

$$\text{Modell: } y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \quad i = 1, \dots, 4, j = 1, \dots, 5$$

$\alpha_i$  = fester Effekt der Sorte  
 $\beta_j$  = fester Effekt des Ortes

Da nur eine Beobachtung pro Zelle vorliegt, können wir keine Wechselwirkungen  $(\alpha\beta)_{ij}$  schätzen.

- c) DEP VAR: YIELD    N: 20    MULTIPLE R: 0.960    SQUARED MULTIPLE R: 0.921

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
HYBRID\$	167.442	3	55.814	38.972	0.000
LOCATION\$	33.322	4	8.331	5.817	0.008
ERROR	17.186	12	1.432		

Die F-Tests testen die beiden Nullhypothesen

- $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$  (die Sorte hat keinen Einfluss auf den Ertrag)
- $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$  (der Ort hat keinen Einfluss auf den Ertrag)

Der F-Wert für den Faktor HYBRID\$ ist gegeben durch

$$F = \text{MS}(\text{HYBRID\$})/\text{MS}(\text{Fehler}) = 55.814/1.432 = 38.972$$

für den Faktor LOCATION\$ durch

$$F = \text{MS}(\text{LOCATION\$})/\text{MS}(\text{Fehler}) = 8.331/1.432 = 5.817$$

In beiden Fällen sind die entsprechenden P-Werte viel kleiner als 0.05. Beide Nullhypothesen werden deshalb auf dem 5%-Niveau verworfen.

**Bemerkung zum Blockfaktor:** In diesem Versuch interessiert man sich primär dafür, ob die verschiedenen Sorten unterschiedliche Erträge ergeben, nicht aber, ob an verschiedenen Orten unterschiedliche Erträge erzielt werden. Der Ort spielt deshalb hier die Rolle eines Blockfaktors. Ohne Berücksichtigung des Blockfaktors (Modell:  $y_{ij} = \mu + \alpha_i + e_{ij}$ ) ergäbe sich folgende Varianzanalysetabelle:

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
HYBRID\$	167.442	3	55.814	17.681	0.000
ERROR	50.508	16	3.157		

Der F-Wert beträgt 17.681. Er ist damit deutlich kleiner als der F-Wert mit Berücksichtigung des Blockfaktors (F-Wert = 38.972). Die Berücksichtigung des Blockfaktors steigert die Effizienz des Experimentes.

d) COL/ROW HYBRID\$ 1 BCM 2 DBC 3 FR-11 4 RC-3

MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3	4
1	0.000			
2	-1.180	0.000		
3	-0.300	0.880	0.000	
4	-7.100	-5.920	-6.800	0.000

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.

MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.145	1.000		
3	0.699	0.268	1.000	
4	0.000	0.000	0.000	1.000

BONFERRONI ADJUSTMENT.

MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.870	1.000		
3	1.000	1.000	1.000	
4	0.000	0.000	0.000	1.000

TUKEY HSD MULTIPLE COMPARISONS.

MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.436	1.000		
3	0.978	0.660	1.000	
4	0.000	0.000	0.000	1.000

Alle drei Tests zeigen signifikante Unterschiede zwischen der Sorte RC-3 und den Sorten BCM, DBC und FR-11. Die P-Werte werden mit 0.000 angegeben. Zwischen den anderen Sorten gibt es paarweise keine signifikanten Unterschiede.

**Bemerkung zur Theorie:**

Obwohl jeder Entscheid, eine bestimmte Nullhypothese abzulehnen, in diesem Beispiel bei den drei Testverfahren gleich ausfällt, gibt es doch Unterschiede bei den P-Werten. Es lässt sich leicht vorstellen, dass bei einer anderen Datenlage einmal der eine Test zu einem signifikanten Ergebnis führen kann, während ein anderer die gleiche Nullhypothese nicht zu verwerfen vermag.

Dies wird klar, wenn man Z.B. die Berechnungsart der Bonferroni P-Werte betrachtet:

$$P_{\text{Bonf}} = \min(1, m \cdot P_{\text{Fisher}})$$

mit m gleich Anzahl Paarvergleiche (im Bsp. m=6).

D.h. die P-Werte der Bonferroni-Tests basieren auf dem Fisher-Test: der P-Wert beim Fisher-Test wird mit der Anzahl möglicher Paarvergleiche (hier 6) multipliziert und falls grösser als 1 auf 1 abgerundet (P-Werte liegen immer zwischen 0 und 1).

Falls wir die Nullhypothese  $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$  als richtig annehmen und alle paarweisen Vergleiche auf dem Niveau  $\alpha$  durchführen, so führt die Bonferroni-Korrektur dazu, dass die globale Irrtumswahrscheinlichkeit

$$P_{H_0}[\text{mindestens einer der Tests fällt signifikant aus, obwohl } H_0 \text{ gilt}]$$

kleiner gleich  $\alpha$  ist. Ohne diese Korrektur ist die globale Irrtumswahrscheinlichkeit wesentlich grösser.

Ein multipler Test heisst konservativ, falls er das multiple Niveau einhält. Als multiples Niveau bezeichnet man die Wahrscheinlichkeit, mindestens eine wahre Hypothese zu verwerfen, unabhängig davon, wieviele und welche Nullhypothesen richtig sind. Der Bonferroni-Test ist konservativ. Tests, die das multiple Niveau nicht einhalten, heissen liberal. Konservative Tests verwerfen im Vergleich zu den liberalen Tests weniger Nullhypothesen.

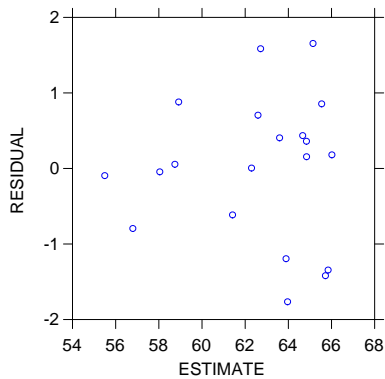
e) Tabelle zur Zusammenfassung der Resultate:

Methode	RC-3 $\mu_4$	DBC $\mu_2$	FR-11 $\mu_3$	BCM $\mu_1$
Fisher	b	a	a	a
Bonferroni	b	a	a	a
Tukey	b	a	a	a

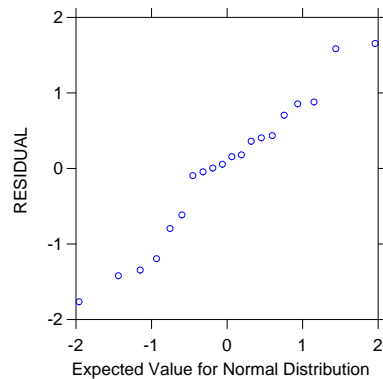
(Beachte: die Stufe der Sorte sind nach der Grösse des Mittelwertes geordnet).

Wie bereits in der Teilaufgabe d) festgestellt, fallen die Entscheide, eine bestimmte Nullhypothese abzulehnen, für alle Testverfahren gleich aus. Es gibt deshalb in dieser Darstellung keine Unterschiede zwischen den drei Testverfahren.

- f) Im Tukey-Anscombe Plot ist ein leichter Trichter zu erkennen. Die Varianz nimmt für grössere geschätzte Werte etwas zu. Es ist allerdings zu berücksichtigen, dass wir nur wenige Punkte haben. Eine Logarithmus-Transformation von YIELD lässt den Trichter auch nicht verschwinden. Der Normal Plot zeigt, dass die Residuen annähernd normalverteilt sind.



(a) Tukey-Anscombe-Plot



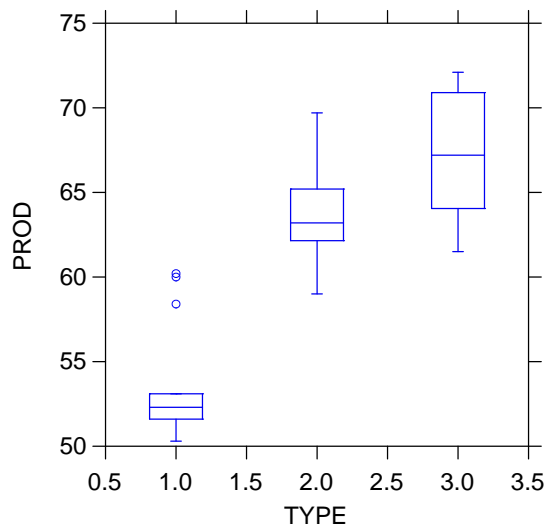
(b) Normal-Plot

g) ESTIMATE:  $\hat{y}_{\text{BCM,SE}} = \hat{\mu} + \hat{\alpha}_{\text{BCM}} + \hat{\beta}_{\text{SE}} = 62.56 + (64.70 - 62.56) + (63.00 - 62.56) = 65.14$

RESIDUAL:  $y_{\text{BCM,SE}} - \hat{y}_{\text{BCM,SE}} = 66.8 - 65.14 = 1.66$

Der Punkt (65.14, 1.66) rechts oben im Tukey-Anscombe Plot gehört also zum Ertrag der Hybrid-Sorte BCM am Ort SE.

2. a) Der Boxplot zeigt deutliche Unterschiede zwischen den Maschinentypen. Beim Typ 1 ist die Produktivität am geringsten, beim Typ 3 am höchsten.



- b) Das lineare Modell sieht folgendermassen aus:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk},$$

wobei

- $y_{ijk}$  : Produktivität
- $\mu$  : Gesamtmittel
- $\alpha_i$  : Effekt des Maschinentyps (**fix**)
- $\beta_j$  : Personen-Effekt (**zufällig**)
- $(\alpha\beta)_{ij}$  : WW-Effekt Type\*Person (**zufällig**)
- $e_{ijk}$  : zufälliger Fehler

Die beiden Faktoren TYPE (Faktor A, fix) und PERSON (Faktor B, zufällig) sind gekreuzt. Es gibt keine hierarchischen Faktoren.

c) DEP VAR: PROD N: 45 MULTIPLE R: 0.995 SQUARED MULTIPLE R: 0.989

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TYPE	1536.760	2	768.380	1072.158	0.000
PERSON	352.999	4	88.250	123.139	0.000
TYPE*PERSON	112.644	8	14.081	19.647	0.000
ERROR	21.500	30	0.717		

WARNING: CASE 23 IS AN OUTLIER (STUDENTIZED RESIDUAL = -3.934)

Da wir hier mit einem gemischten Modell arbeiten, das fixe und zufällige Effekte enthält, sind die von Systat ausgegebenen F- und P-Werte zum Teil falsch. Wir können die Resultate deshalb an dieser Stelle noch nicht interpretieren.

- d) Das oben formulierte Modell wird im Skript (siehe Kapitel 4.3.1) behandelt. Die zur korrekten Durchführung der Signifikanztests benötigten Erwartungswerte der Durchschnittsquadrate können dort nachgeschlagen werden. Die Varianzanalyse-tabelle hat folgende Struktur:

	Ursache	df	Erwartungswert(MS)
A:	TYPE	2	$\sigma_e^2 + 3\sigma_{AB}^2 + 15 \cdot Q_A$
B:	PERSON	4	$\sigma_e^2 + 9\sigma_B^2$
A*B:	TYPE*PERSON	8	$\sigma_e^2 + 3\sigma_{AB}^2$
	Fehler	30	$\sigma_e^2$
	Total	44	

Die Nullhypothese  $H_0$  für den Test des **fixen** Faktors TYPE (Faktor A) lautet:

$H_0 : \alpha_i = 0$ , d. h. der **fixe** Faktor TYPE hat keinen Einfluss auf die Produktivität.

Aufgrund der oben angegebenen Erwartungswerte der Durchschnittsquadrate muss der korrekte F-Wert folgendermassen berechnet werden:

$$F = \frac{MS(A)}{MS(A*B)} \sim F_{2,8}$$

Der dazugehörige P-Wert kann mit Systat berechnet werden, indem man den korrekten Fehlerterm (hier die Wechselwirkung TYPE\*PERSON) spezifiziert.

```
>HYPOTHESIS
>EFFECT=type
>ERROR=type*person
```

TEST OF HYPOTHESIS					
SOURCE	SS	DF	MS	F	P
HYPOTHESIS	1536.760	2	768.380	54.571	0.000
ERROR	112.644	8	14.081		

Der korrekte F-Wert (54.571) ist also erheblich kleiner als der F-Wert (1072.158) in der ursprünglichen Varianzanalysetabelle (Modell mit ausschliesslich fixen Effekten).

Die Nullhypothese  $H_0$  für den Test des **zufälligen** Faktors PERSON (Faktor B) lautet:

$H_0 : \sigma_B^2 = 0$ , d. h. die Varianz des **zufälligen** Faktors PERSON ist gleich null.

Aufgrund der oben angegebenen Erwartungswerte der Durchschnittsquadrate muss der korrekte F-Wert folgendermassen berechnet werden:

$$F = \frac{\text{MS(B)}}{\text{MS(Fehler)}} \sim F_{4,30}$$

Diese Formel entspricht der Formel des Modelles mit ausschliesslich fixen Effekten. Die in der Varianzanalysetabelle der Teilaufgabe c) angegebenen F- und P-Werte für den zufälligen Faktor Person sind deshalb korrekt.

Aufgrund der Erwartungswerte der Durchschnittsquadrate sind F- und P-Wert auch für den Test der zufälligen Wechselwirkung korrekt.

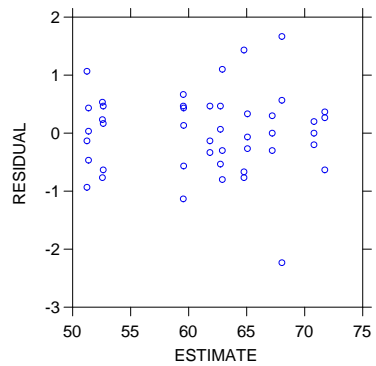
**Fazit:** Es sind also alle Faktoren, inkl. der Wechselwirkung signifikant. Dies bedeutet einerseits, dass zwischen den Maschinentypen gesicherte Unterschiede bestehen. Andererseits heisst dies aber auch, dass sich die beobachtete Gesamtvariabilität der Daten nicht nur mit dem Effekt der verschiedenen Maschinentypen und einem Zufallsfehler erklären lässt, sondern dass auch die Personen und die zufällige Wechselwirkung Maschinentyp\*Person signifikant zur beobachteten Gesamtvariabilität beitragen.

- e) Mit Hilfe der Tabellen in c) und d) können wir die Varianzkomponenten berechnen:

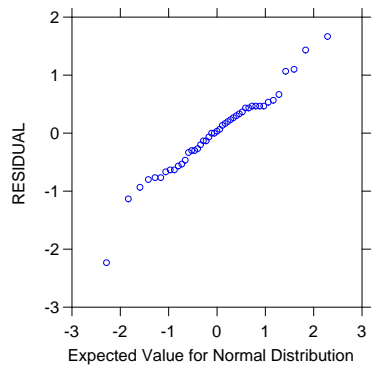
$$\begin{aligned}\hat{\sigma}_B^2 &= (\text{MS(B)} - \text{MS(Fehler)})/9 = (88.250 - 0.717)/9 = 9.725 \\ \hat{\sigma}_{AB}^2 &= (\text{MS(A*B)} - \text{MS(Fehler)})/3 = (14.081 - 0.717)/3 = 4.455 \\ \hat{\sigma}_e^2 &= \text{MS(Fehler)} = 0.717\end{aligned}$$

Mit  $\hat{\sigma}_B^2 = 9.725$  (65.3% der Totalvarianz) liefert der Faktor Person den grössten Anteil an die Variabilität der Produktivität.

- f) Bis auf einen extremer Wert (die 23<sup>te</sup> Beobachtung) zeigt der Tukey-Anscombe Plot keine Abweichungen von der Voraussetzung konstanter Fehlervarianz. Im Normal-Plot fällt wieder der extreme Wert der 23<sup>ten</sup> Beobachtung auf, ansonsten zeigen sich keine unerwünschten Strukturen. Es liegen keine eindeutigen Hinweise auf eine Verletzung der Modellvoraussetzungen vor.



(c) Tukey-Anscombe-Plot



(d) Normal-Plot

## Kurz-Übung 4

- a) Man vergleicht die Labors im Gebiet I mit denjenigen im Gebiet II.  
 b) Ja: die Koeffizienten des Kontrastes  $L_7$  sind ein Vielfaches (mal 12) der Koeffizienten des Kontrastes  $L_1$  (allgemein: die zwei Kontraste sind linear abhängig).  
 c)  $L_2$ .  
 d) Ja, denn

$$\sum_i \frac{1}{n_i} c_{1i} \cdot c_{2i} = \frac{1}{2} \cdot \left( \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot (-1) + \frac{1}{4} \cdot (-1) + \left(-\frac{1}{3}\right) \cdot 0 + \left(-\frac{1}{3}\right) \cdot 0 + \left(-\frac{1}{3}\right) \cdot 0 \right) = 0$$

2. a) Aus der Tabelle lesen wir:

$$\hat{\mu}_1 = 33.3; \quad \hat{\mu}_2 = 36.8; \quad \hat{\mu}_3 = 39.5; \quad \hat{\mu}_4 = 40.7; \quad \hat{\mu}_5 = 43.7.$$

Es folgt (als Hinweis gegeben):

$$\hat{\mu} = \frac{1}{5} \sum_{i=1}^5 \hat{\mu}_i = \frac{1}{5} \cdot (33.3 + 36.8 + 39.5 + 40.7 + 43.7) = 38.8$$

und

$$\begin{aligned} \hat{\alpha}_1 &= 33.3 - 38.8 = -5.5 \\ \hat{\alpha}_2 &= 36.8 - 38.8 = -2 \\ \hat{\alpha}_3 &= 39.5 - 38.8 = 0.7 \\ \hat{\alpha}_4 &= 40.7 - 38.8 = 1.9 \\ \hat{\alpha}_5 &= 43.7 - 38.8 = 4.9. \end{aligned}$$

Deshalb:

$$\begin{aligned} SS_{\text{zwischen Gruppen}} &= \sum_{i=1}^g n_i (\bar{y}_{i+} - \bar{y}_{++})^2 = \sum_{i=1}^g n_i (\hat{\alpha}_i)^2 = 6 \cdot \sum_{i=1}^5 (\hat{\alpha}_i)^2 = 374.16 \\ SS_{\text{innerhalb Gruppen}} &= \sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i+})^2 = \sum_{i=1}^g (n_i - 1) s_i^2 \\ &= 5 \cdot (63 + 42 + 23 + 48 + 36) = 1060. \end{aligned}$$

Alles zusammen:

Ursache	SS	df	MS	F-Wert	E(MS)
zwischen Gruppen	374.16	4	93.54	2.21	$\sigma_e^2 + 6 Q_{\text{zwischen}}$
innerhalb Gruppen	1060	25	42.40		$\sigma_e^2$
Total	1434.16	29			

Test für den Düngereffekt:

$$H_0 : \alpha_i = 0 \text{ (der Dünger hat keinen Einfluss auf den Ertrag)}$$

$$H_A : \text{mindestens ein } \alpha_i \neq 0$$

$$F\text{-Wert} = \frac{MS(\text{zwischen Gruppen})}{MS(\text{innerhalb Gruppen})} = \frac{93.54}{42.40} = 2.21$$

$$F\text{-Tabelle: } F_{4,25;95\%} \approx 2.76,$$

d.h.  $H_0 : \alpha_i = 0$  wird beibehalten.



b) Linearer Trend:

$$L_1 = -2\mu_1 - \mu_2 + 0\mu_3 + \mu_4 + 2\mu_5 .$$

Quadratischer Trend:

$$L_2 = 2\mu_1 - \mu_2 - 2\mu_3 - \mu_4 + 2\mu_5 .$$

c) Allgemein:  $SS(L)$  ist definiert als:

$$SS(L) = \frac{(\hat{L} - a)^2}{\sum_i (c_i^2/n_i)} .$$

Hier ist  $a = 0$  zu setzen.

Wir kriegen für  $\hat{L}_1$ :

$$\begin{aligned} \hat{L}_1 &= -2 \cdot 33.3 - 36.8 + 0 + 40.7 + 2 \cdot 43.7 = 24.7 \\ \sum_i (c_{1i}^2/n_i) &= \frac{1}{6} \cdot (4 + 1 + 1 + 4) = \frac{10}{6} = \frac{5}{3} \end{aligned}$$

und für  $\hat{L}_2$

$$\begin{aligned} \hat{L}_2 &= 2 \cdot 33.3 - 36.8 - 2 \cdot 39.5 - 40.7 + 2 \cdot 43.7 = -2.5 \\ \sum_i (c_{2i}^2/n_i) &= \frac{1}{6} \cdot (4 + 1 + 4 + 1 + 4) = \frac{14}{6} = \frac{7}{3} \end{aligned}$$

Zusammenfassung:

Kontrast	Mittelwerte					$\hat{L}$	$\sum_i (c_i^2/n_i)$	SS(L)
	33.3	36.8	39.5	40.7	43.7			
$L_1$	-2	-1	0	1	2	24.7	5/3	366.05
$L_2$	2	-1	-2	-1	2	-2.5	7/3	2.67

F-Tests:

- Für  $L_1$  :

$$H_0 : L_1 = 0 \text{ (kein linearer Trend)}$$

$$H_A : L_1 \neq 0$$

$$\text{F-Wert} = \frac{MS(L_1)}{MS(\text{Fehler})} = \frac{SS(L_1)}{MS(\text{Fehler})} = \frac{366.05}{42.40} = 8.63$$

$$\text{F-Tabelle: } F_{1,25;95\%} \approx 4.24,$$

d.h.  $H_0: L_1 = 0$  wird verworfen (es gibt einen linearen Trend).

- Für  $L_2$  :

$$H_0 : L_2 = 0 \text{ (kein quadratischer Trend)}$$

$$H_A : L_2 \neq 0$$

$$\text{F-Wert} = \frac{MS(L_2)}{MS(\text{Fehler})} = \frac{SS(L_2)}{MS(\text{Fehler})} = \frac{2.67}{42.40} = 0.063$$

$$\text{F-Tabelle: } F_{1,25;95\%} \approx 4.24,$$

d.h.  $H_0: L_2 = 0$  wird beibehalten (kein Hinweis auf quadratischen Trend).

## Kurz-Übung 5

a) Die Faktoren A, A\*B und C(B) werden gegen den Fehler getestet, der Faktor B gegen C(B). (Weil sich beim F-Test die Erwartungswerte des Zählers und des Nenners genau um die zu testende Komponente unterscheiden müssen.) Die zugehörigen Tests werden hier am Test für den Sorteneffekt A und den Ortseffekt B illustriert; die anderen Tests gehen analog.

- **Test von A:** Der F-Wert ergibt sich als  $F = \frac{MS(A)}{MS(Fehler)} = \frac{4.282}{0.537} = 7.97$ . Das 95%-Quantil der F-Verteilung mit 12 und 144 Freiheitsgraden beträgt  $F_{12,144} = 1.82$ . Da  $F > 1.82$  wird die Nullhypothese  $H_0$ : "Sorteneffekt spielt keine Rolle" auf dem 5%-Niveau verworfen.
- **Test von B:** Der F-Wert beträgt  $F = \frac{MS(B)}{MS(C(B))} = \frac{40.676}{1.366} = 29.77$ . Der Vergleich mit dem 95%-Quantil der entsprechenden F-Verteilung  $F_{3,12} = 3.49$  ergibt, dass auch dieser Effekt auf dem 5%-Niveau signifikant ist.

Insgesamt ergibt sich die folgende Varianzanalysetabelle:

Source		df	MS	E(MS)	F-Wert	P-Wert
A	Sorte	12	4.282	$\sigma_\varepsilon^2 + 16 \cdot Q_A$	7.97	< 0.05
B	Ort	3	40.676	$\sigma_\varepsilon^2 + 13 \cdot \sigma_C^2 + 52 \cdot Q_B$	29.77	< 0.05
A*B	Sorte*Ort	36	0.672	$\sigma_\varepsilon^2 + 4 \cdot Q_{AB}$	1.25	> 0.05
C(B)	Feld(Ort)	12	1.366	$\sigma_\varepsilon^2 + 13 \cdot \sigma_C^2$	2.54	< 0.05
	Error	144	0.537	$\sigma_\varepsilon^2$		

Ausser der Interaktion sind also alle Effekte signifikant.

b) Der Kontrast L1 entspricht der Nullhypothese

$$L1: H_0: -s_1 - s_2 - s_3 + 3s_4 = 0$$

$$\Leftrightarrow s_4 = (s_1 + s_2 + s_3)/3$$

In Worten: Er vergleicht Sorte 4 mit dem Durchschnitt der Standardsorten (Sorte 1, 2 und 3).

Entsprechend vergleichen die beiden anderen Kontraste die Sorten 5 und 13 mit dem Durchschnitt der Standardsorten.

Die Tests besagen, dass sich Sorte 4 und 13, nicht aber Sorte 5, signifikant vom Durchschnitt der Standardsorten unterscheiden.