

5 Varianzanalytische Modelle, komplexere lineare Modell und Random Models

Auch in diesem Kapitel werden nur wenige statistische Hintergründe geliefert. Der Fokus des Kapitels liegt in der Einübung der Arbeitsschritte zur Erstellung varianzanalytischer Modelle in GENSTAT. GENSTAT hat ein gutes Hilfesystem und bietet eine gute statistische Referenz über PDF-Dateien. Außerdem sollte sekundäre Literatur zu statistischen Modellen herangezogen werden.

Am Anfang der einzelnen Kapitel ist eine Anleitung zur Generierung von Designs für die einzelnen besprochenen Versuchsanlagen aufgezeichnet. GENSTAT ist durchaus als gutes Versuchsplanungsprogramm zu bezeichnen. Allerdings ist an einigen Stellen was die Anzahl möglicher Prüfglieder angeht limitiert. Gerade wenn komplexe Designs benötigt werden sollte Spezialsoftware angewendet werden.

Das anlegen eines Designs liefert allerdings einen guten Hinweis wie ein Auswertmodell für entsprechende Designs auszusehen hat. In jedem Fall sollten Sie einen Biometriker bzw. eine Biometrikerin konsultieren.

5.1 Syntax zum Formulieren von ANOVA Modellen

Tabelle 5: Regeln zur Formulierung von ANOVA Modellen

+	=	A +B	=	Haupteffekt von A und B
.	=	A.B	=	Interaction von A und B
*	=	A*B	=	A+B+A.B Faktorielle Struktur
/	=	A/B	=	A+A.B B nicht als Haupteffekt sondern nur als Effekt innerhalb von A

Tabelle 6: Beispiele Formulierung von ANOVA Modellen

A*B*C	=	A+B+C+A.B+A.C+B.C+A.B.C volles faktorielles Design
(A+B)*(C+D)	=	(A+B)+(C+D)+(A+B).(C+D)
	=	A+B+C+D+A.C+A.D+B.C+B.D
Block/Plot/Subplot	=	Block+Block.Plot+Block.Plot.Subplot
A/(B*C)	=	A+A.B+A.C+A.B.C

Empfehlung: Syntax von ANOVA Modellen

Sie können das über das Menü formulierte Modell aus dem Input Log übernehmen, in ein eigenes Skriptfenster einfügen und editieren. Falls Sie sich nicht ganz sicher sind sollten Sie das Modell mit Hilfe von „ + “ und „ . “ in Langform formulieren.

5.2 Beispielauswertung: Kartoffelertrag – Lateinisches Quadrat

Bitte setzen Sie den GENSTAT Server mit „Restart Server“ (vergl. 1.3.2) zurück

5.2.1 Anlegen eines Designs

Um in GENSTAT ein Design anzulegen müssen Sie das Menü „Stats -> Design -> Generate Standard Design“ auswählen.

Im sich öffnenden Dialog wählen Sie bitte über das Pulldown Menü das Design vom Typ „Latin Square“ Folgende Einstellungen sollten Sie vornehmen:

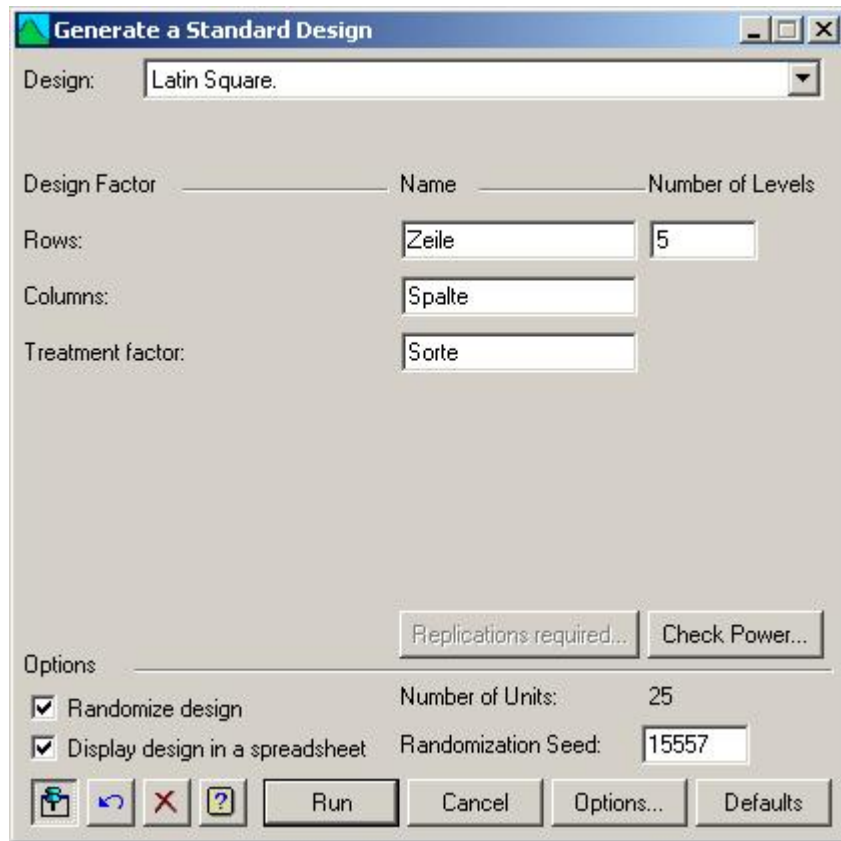


Abbildung 102: Anlegen eines Latin Square Designs

Mit einem Klick auf „Run“ wird das Design in Form eines Spreadsheets erzeugt. Die Besonderheit des Spreadsheet ist, dass die Modellinformation für die Varianzanalyse in der Tabelle hinterlegt ist. Wenn Sie in der Tabelle die Zielgröße „Ertrag“ erzeugen, Daten eingeben und die Menüoption zum Rechnen einer Varianzanalyse aufrufen können Sie dieses Überprüfen.

Nachdem Sie auf „Run“ geklickt haben können Sie auch die Power Ihres Modells prüfen. Natürlich benötigen Sie Informationen über das gewünschte Delta der Mittelwerte und eine

Streuungsinformationen aus vorherigen gleichartigen Versuchen. Falls die Power unter 80% liegt sollten Sie evtl. mehr Wiederholungen vorsehen.

Wenn Sie möchten können Sie das erstellte Design abspeichern indem Sie das Spreadsheet über die Maus aktivieren und über das Menü „File“ den „Save“ Dialog aufrufen. Das Spreadsheet ist z.B. als GenStat Spreadsheet oder als Excel Tabelle abspeicherbar. Schließen Sie es jedoch in jedem Fall.

5.2.2 Daten eingeben bzw. laden

Bitte setzen Sie den GENSTAT Server mit „Restart Server“ (vergl. 1.3.2) zurück und laden Sie bitte die Datei „Latin_Square_Data_Potato_Yield.xls“ mit Hilfe des Excel Import Wizards (vergl. 2.1).

Im Excel Wizard sollten Sie Zeile, Spalte und Sorte zur Faktorvariablen konvertieren.

Folgende Daten werden geladen:

Tabelle 7: Datensatz „Latin_Square_Data_Potato_Yield.xls“

Row (Zeile)	Column (Spalte)	Variety (Sorte)	Yield (Ertrag)
1	1	C	22
1	2	B	20
1	3	A	39
1	4	D	27
1	5	E	34
2	1	E	29
2	2	D	29
2	3	C	25
2	4	A	30
2	5	B	23
3	1	A	29
3	2	E	25
3	3	D	34
3	4	B	26
3	5	C	27
4	1	B	23
4	2	A	27
4	3	E	27
4	4	C	32
4	5	D	41
5	1	D	33
5	2	C	21
5	3	B	24
5	4	E	30
5	5	A	33

5.2.3 Auswertung

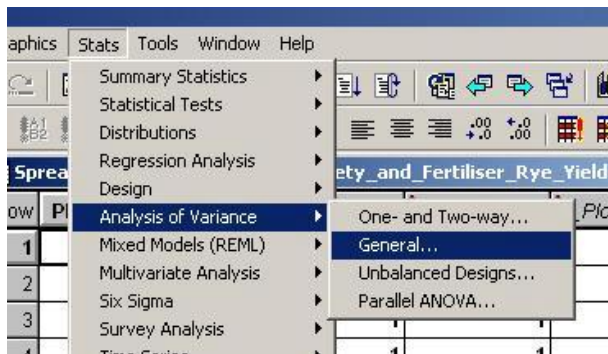


Abbildung 103: General Analysis of Variances

Die varianzanalytische Auswertung starten Sie über den Menüpunkt „Stats -> Analysis of Variances -> General“

Im sich öffnenden Dialog müssen Sie das Auswertmodell formulieren.

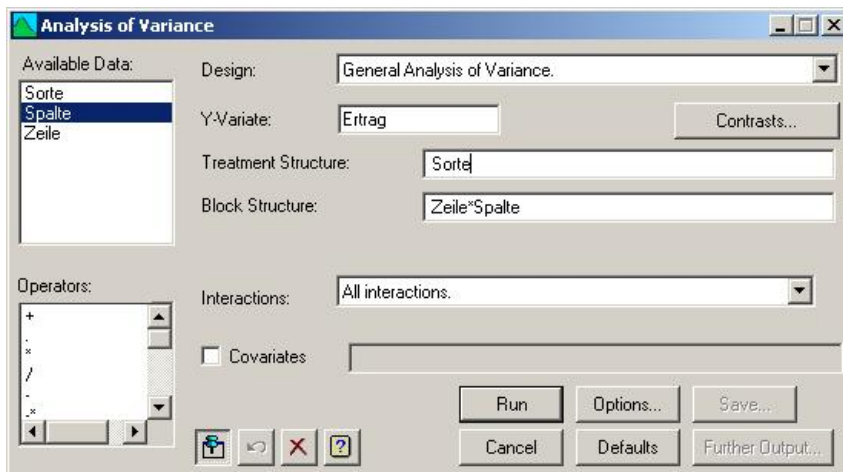


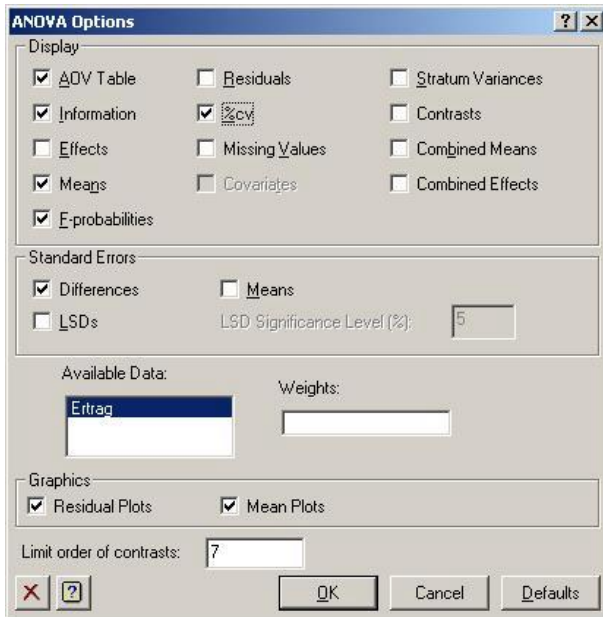
Abbildung 104: Einstellungen zum ANOVA Auswertmodell – Lateinisches Quadrat

Zielgröße des Auswertmodells ist die Variable „Ertrag“.

Als Treatment bzw. als unabhängige Größe wird die Variable Sorte genutzt.

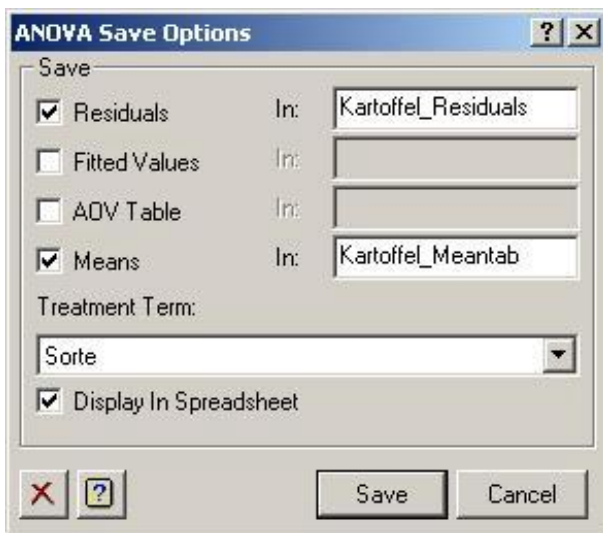
Vorgegeben durch das Modell des Lateinischen Quadrats müssen Zeile und Spalte als Blockstruktur definiert werden. Auch die Interaktion von Zeile und Spalte wird mit Hilfe des * zwischen den beiden Variablen in das Modell aufgenommen.

Bitte nehmen Sie weitere Einstellungen wie in Abbildung 105 und Abbildung 106 vor und klicken Sie auf „Run“



In jedem Fall sollten Sie als Zusatzoption die graphische Ausgabe aktivieren.

Abbildung 105: Einstellungen zum ANOVA Auswertmodell – Lateinisches Quadrat – Optionen



Um einen graphischen Test auf Normalverteilung und einen Paarvergleich der Mittelwerte vornehmen zu können sollten Sie die nebenstehenden numerischen Ergebnisse abspeichern.

Abbildung 106: Einstellungen zum ANOVA Auswertmodell – Lateinisches Quadrat – Speichern von Ergebnissen

Nachfolgende Skriptlistings werden als Befehl im Input Log generiert (vergl. Skriptlisting 4 und Skriptlisting 5). Das Input Log können Sie zur späteren Verwendung abspeichern.

Skriptlisting 4: Erzeugen des One Way ANOVA Output

```
"General Analysis of Variance."
BLOCK "No Blocking"
TREATMENTS Sorte+Spalte+Zeile
COVARIATE "No Covariate"
ANOVA [PRINT=aovtable,information,means,%cv; FACT=1; CONTRASTS=7; FPROB=yes; PSE=diff,\
means] Ertrag
APLOT [RMETHOD=simple] fitted,normal,halfnormal,histogram
AGRAPH [METHOD=means]
```

Skriptlisting 5: Speichern von notwendigen Ergebnissen der One Way ANOVA

```
DELETE [REDEFINE=yes] Kartoffel_Meantab
AKEEP [RESIDUAL=Kartoffel_Residuals; FACT=32]Sorte; MEANS=Kartoffel_Meantab
FSPREADSHEET [SHEET=29548864; METHOD=replace] Kartoffel_Residuals
FSPREADSHEET Kartoffel_Meantab
```

GENSTAT erzeugt als graphischen Output einige Diagnostik Plots und eine Darstellung der Mittelwerte. Die Diagnostik Plots lassen auf ein recht gutes Modell schließen. Im Normal Plot und im Half Normal Plot fallen ein paar Werte auf. Dieses sollte sich mit der Information über Large Residuals in der numerischen Ausgabe decken.

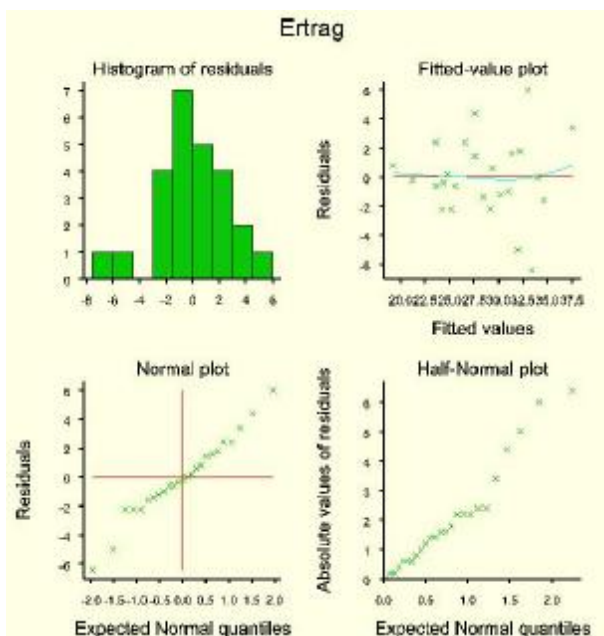


Abbildung 107: Kartoffelertrag – Diagnostikplots der ANOVA

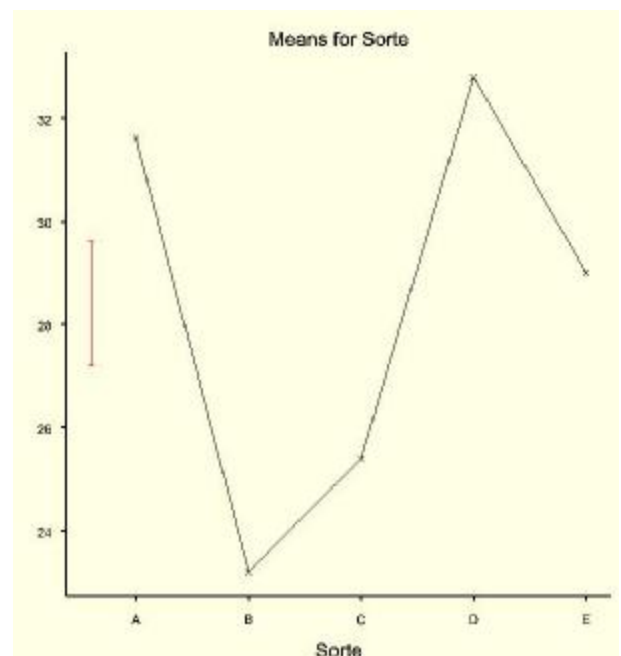


Abbildung 108: Kartoffelertrag – Means

Das numerische Ergebnis der Varianzanalyse ist nachfolgend aufgelistet.

Ergebnisausgabe 5: Ergebnis der One Way ANOVA

Analysis of variance

Variate: Ertrag

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Sorte	4	330.00	82.50	5.64	0.009
Spalte	4	150.00	37.50	2.56	0.093
Zeile	4	20.40	5.10	0.35	0.840
Residual	12	175.60	14.63		
Total	24	676.00			

Message: the following units have large residuals.

units 3	6.00
units 4	-6.40

Tables of means

Variate: Ertrag

Grand mean 28.40

Sorte	A	B	C	D	E
	31.60	23.20	25.40	32.80	29.00
Spalte	1	2	3	4	5
	27.20	24.40	29.80	29.00	31.60
Zeile	1	2	3	4	5
	28.40	27.20	28.20	30.00	28.20

Standard errors of means

Table	Sorte	Spalte	Zeile
rep.	5	5	5
d.f.	12	12	12
e.s.e.	1.711	1.711	1.711

Standard errors of differences of means

Table	Sorte	Spalte	Zeile
rep.	5	5	5
d.f.	12	12	12
s.e.d.	2.419	2.419	2.419

Stratum standard errors and coefficients of variation

Variate: Ertrag

d.f.	s.e.	cv%
12	3.825	13.5

5.2.4 Paarvergleich

Leider ist über das Menüsystem kein Paarvergleich der LS Means verfügbar. Entsprechender Code muss selbst generiert werden. Nachfolgende finden Sie ein Beispiel. Leider müssen Sie sich auch einige Werte aus dem Output der ANOVA „fischen“. Im Skript muss der Name der Tabelle übergeben werden, in der ich die zu vergleichenden Mittelwerte befinden. Außerdem wird eine Information über die Streuung und die Freiheitsgrade benötigt. Neben der konservativen Bonferroni Methode sind auch Tukey und Sidak verfügbar.

VSN hat für die Version 10 des GENSTAT Systems Besserung gelobt.

Skriptlisting 6: Erzeugen eines Paarvergleichs der Modellmittelwerte

Daten aus Tabelle „Stratum standard errors and coefficients of variation“

VARIANCE = s.e. aus Output zum Quadrat

DF = d.f. aus Output

```
ALLPAIRWISE [METHOD=Bonferroni; DIRECTION=descending; PROBABILITY=0.05]\
MEANS=Kartoffel_Meantab; REPLICATION=5; VARIANCE=14.631; DF=12
```

Ergebnisausgabe 6: Ergebnis des Paarvergleichs auf Basis der One Way ANOVA

All pairwise comparisons are tested.

Variance = 14.6310 with 12 degrees of freedom

Bonferroni test

Experimentwise error rate = 0.0500

Comparisonwise error rate = 0.0050

Mean	vs Mean	t	significant
D	A	0.496	No
D	E	1.571	No
D	C	3.059	No
D	B	3.968	Yes
A	E	1.075	No
A	C	2.563	No
A	B	3.472	Yes
E	C	1.488	No
E	B	2.398	No
C	B	0.909	No

Identifier	Mean
D	32.80
A	31.60
E	29.00
C	25.40
B	23.20

5.2.5 Tests auf Normalverteilung

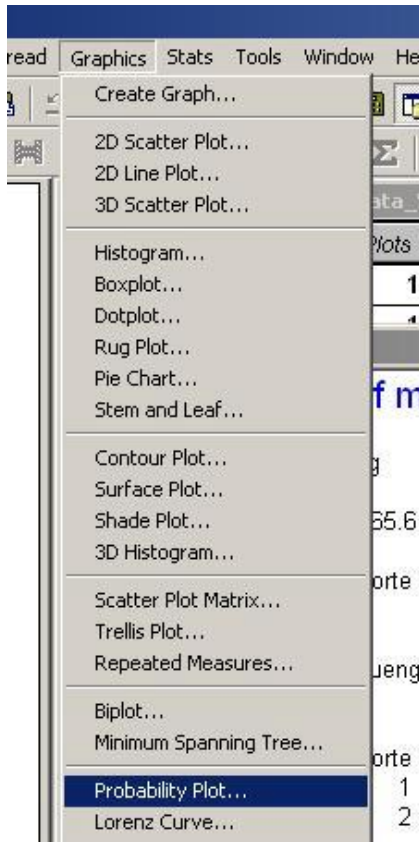


Abbildung 109: Menüpunkt Test auf Normalverteilung

Der Test auf Normalverteilung ist über das Graphik Menü aufrufbar. Nehmen Sie bitte Einstellungen wie in Abbildung 110 vor.

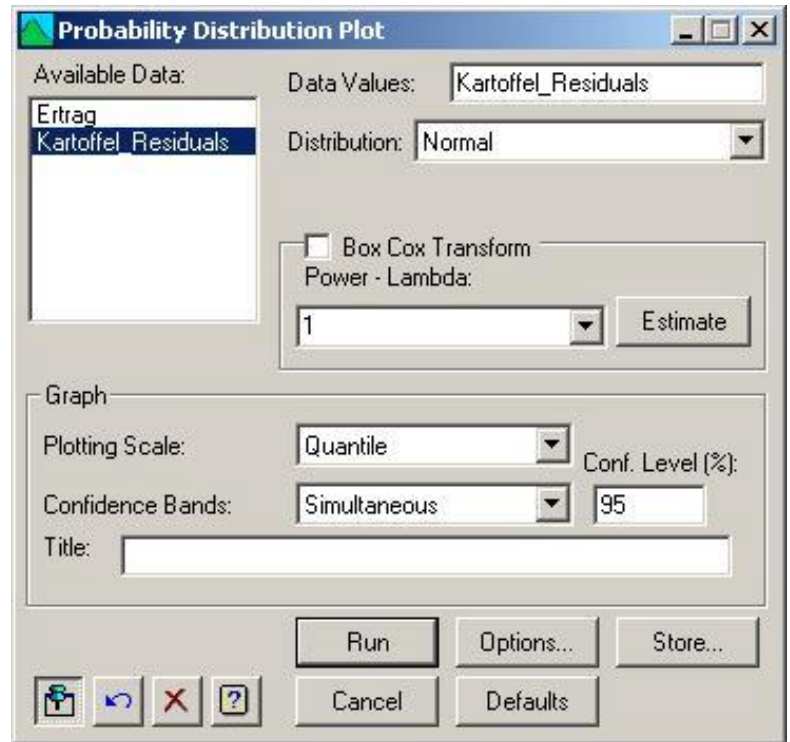


Abbildung 110: Einstellungen des Normalverteilungstests

Der Skriptbefehl sieht folgendermaßen aus:

Skriptlisting 7: Graphischer Test auf Normalverteilung

```
DPROBABILITY [PRINT=parameters,tests;DISTRIBUTION=NORMAL;METHOD=quantile;QMETHOD=standardized;\
BANDS=simultaneous;ALPHA=0.95;PLOT=reference] Kartoffel_Residuals
```

Ergebnisausgabe 7: Numerisches Ergebnis des Tests auf Normalverteilung

Critical values of test statistics (marginal tests)

Test statistic	15%	10%	5%	2.5%	1%
Anderson-Darling	0.576	0.656	0.787	0.918	1.092
Cramer-von Mises	0.091	0.104	0.126	0.148	0.178
Watson	0.085	0.096	0.116	0.136	0.163

Marginal tests

Variate	Anderson-Darling	Cramer-von Mises	Watson
1	0.3176	0.0415	0.0415

?, *, ** indicate significance at 10%, 5% and 1% levels respectively

In der graphischen Ausgabe ist erkennbar, dass sich alle Werte innerhalb der Konfidenzbänder befinden. Es ist sehr wahrscheinlich, dass eine Normalverteilung vorliegt. Dieses deckt sich mit den numerischen Ergebnissen.

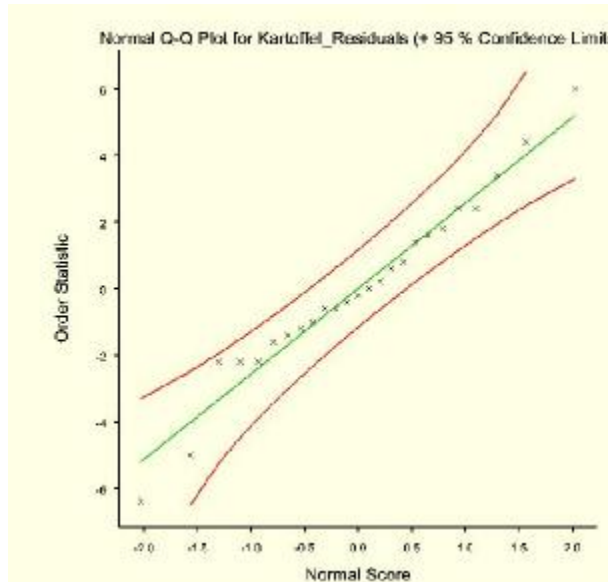


Abbildung 111: Graphische Ausgabe des Normalverteilungstests

5.2.6 Test auf Varianzhomogenität (Bartlett's)

Der Test auf Varianzhomogenität ist leider nicht über das Menü verfügbar. Allerdings benötigen Sie nur einen kurzen Skriptbefehl:

Skriptlisting 8: Test auf Varianzhomogenität

```
VHOMOGENEITY [GROUP=Variety] DATA=Yield
```

Ergebnisausgabe 8: Ergebnis des Tests auf Varianzhomogenität

Bartlett's test for homogeneity of variances

Chi-square 3.08 on 4 degrees of freedom: probability 0.544

Die Nullhypothese des Bartlett Tests lautet: „Gleichheit der Varianzen“. Im vorliegenden Fall kann also von Varianzhomogenität ausgegangen werden.

Falls Sie Versuchsdesigns mit mehr als einer Einflussgröße haben müssen Sie sich für die Durchführung des Homogenitätstests eine neue Gruppierungsvariable aus den verschiedenen Einflussgrößen generieren. Einen Hinweis hierzu liefert das Kapitel 2.9.