



**JKI**



..... **Berichte**

Eckard Moll<sup>1</sup> und Thomas Stauber<sup>2</sup>

## **Datensichtung, Unterstützung bei der Problemanalyse, erste Schritte einer Datenanalyse**

**143  
2008**

<sup>1</sup> Julius Kühn-Institut  
Bundesforschungsinstitut für Kulturpflanzen  
- Zentrale Datenverarbeitung -

<sup>2</sup> Johann Heinrich von Thünen-Institut  
Bundesforschungsinstitut für Ländliche Räume, Wald und Fischerei  
- Institut für Waldökologie und Waldinventuren -

## **Kontaktadresse**

Dr. Eckard Moll

Julius Kühn-Institut (JKI) – Bundesforschungsinstitut für Kulturpflanzen

Zentrale Datenverarbeitung

Stahnsdorfer Damm 81

14532 Kleinmachnow

Telefon +49 (0)33203 48-0

Telefax +49 (0)33203 48-425

Internet <http://www.jki.bund.de>

Der Forschungsbereich des Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (BMELV) hat seit dem 1. Januar 2008 eine neue Struktur.

Die Biologische Bundesanstalt für Land- und Forstwirtschaft (BBA), die Bundesanstalt für Züchtungsforschung an Kulturpflanzen (BAZ) sowie zwei Institute der Bundesforschungsanstalt für Landwirtschaft (FAL) wurden zum Julius Kühn-Institut - Bundesforschungsinstitut für Kulturpflanzen zusammengeschlossen. Das Johann Heinrich von Thünen-Institut (vTI) wurde aus der Bundesforschungsanstalt für Fischerei, der Bundesforschungsanstalt für Forst- und Holzwirtschaft und aus Teilen der Bundesforschungsanstalt für Landwirtschaft errichtet.

The research branch of the Federal Ministry of Food, Agriculture and Consumer Protection (BMELV) has been reorganized. The former Biological Research Centre for Agriculture and Forestry (BBA) has been merged with other institutions. The newly established Julius Kühn Institute (JKI), Federal Research Centre for Cultivated Plants, is working on plant protection, plant breeding, crop and soil science. The Johann Heinrich von Thünen Institute (vTI) was created from the German Federal Research Centre for Fisheries, the German Federal Research Centre for Forestry and Forest Products and part of the German Federal Agricultural Research Centre.

**Wir unterstützen den offenen Zugang zu wissenschaftlichem Wissen.**

**Die Berichte aus dem Julius Kühn-Institut erscheinen daher als OPEN ACCESS-Zeitschrift.**

**Alle Ausgaben stehen kostenfrei im Internet zur Verfügung:**

**<http://www.jki.bund.de> Bereich Veröffentlichungen – Berichte.**

We advocate open access to scientific knowledge. Reports from the Julius Kühn Institute are therefore published as open access journal. All issues are available free of charge under <http://www.jki.bund.de> (see Publications – Reports).

## **Herausgeber / Editor**

Julius Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen, Braunschweig, Deutschland

Julius Kühn Institute, Federal Research Centre for Cultivated Plants, Braunschweig, Germany

## **Verlag**

Eigenverlag

## **Vertrieb**

Saphir Verlag, Gutsstraße 15, 38551 Ribbesbüttel

Telefon +49 (0)5374 6576

Telefax +49 (0)5374 6577

## **ISSN 1866-590X**

© Julius Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen, 2008

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersendung, des Nachdrucks, des Vortrages, der Entnahme von Abbildungen, der Funksendung, der Wiedergabe auf fotomechanischem oder ähnlichem Wege und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten.

## Inhaltsverzeichnis

Vorwort	5
1 Berechnen statistischer Maßzahlen und Veranschaulichung der Daten	7
1.1 Eine Stichprobe	7
1.1.1 Beispiel Ährenlängen, Übernahme der in MS-Excel gespeicherten Daten nach SAS	7
1.1.2 Berechnen statistischer Maßzahlen	8
1.1.3 Box-Whisker-Plot	11
1.1.4 Häufigkeitsverteilung für nicht klassifizierte Daten	12
1.1.5 Sind die Daten normalverteilt?	13
1.1.6 Aufgabenstellungen für zwei weitere Beispiele	16
1.1.7 Maßzahlen der 3 Beispiele Ährenlängen, Ct-Gehalte, Mäusegewichte	16
1.2 Zwei Stichproben	18
1.2.1 Beispiel Mäusegewichte	18
1.2.2 Bemerkungen zu abhängigen und unabhängigen Stichproben sowie zu fachlichen Aufgabenstellungen, die entsprechende Auswertungsverfahren zur Folge haben	21
1.3 Zusammenfassung	22
2 Tests von Hypothesen über Parameter, zwei Stichproben	23
2.1 Fachliche Aufgabenstellungen und Ableitung der statistischen Fragestellungen	23
2.2 t-Test zweier unabhängiger Stichproben	24
2.2.1 Die Voraussetzungen des t-Tests	24
2.2.2 Bemerkungen zur Normalverteiltheit	24
2.2.3 Der Levene-s-Test zur Überprüfung der Gleichheit der Varianzen	24
2.2.4 Die SAS-Prozedur TTEST	26
2.2.5 Bemerkungen zur Testentscheidung und grafische Interpretation	27
2.2.6 Durchführung des t-Tests mit dem Analyst	27
2.2.7 Bemerkungen zur Versuchsplanung mit dem Analyst	28
2.3 t-Test zweier abhängiger Stichproben	30
2.3.1 Aufgabenstellung und Daten	30
2.3.2 Rückführung auf das Einstichprobenproblem	30
2.4 Zusammenfassung	31
3 Gruppierte Diagramme für mehr als zwei Stichproben	32
3.1 Beispiel Stammhöhen von Eichen in den Jahren 2001 und 2005	32
3.2 Boxplots und Punktediagramm (Scatterplot)	33
3.3 Zusammenfassung	37
4 Skalierung der Faktoren und Merkmale und Wahl des statistischen Auswertungsverfahrens	38
4.1 Faktor(en)	38
4.2 Datenkomprimierung in Abhängigkeit von der Skalierung eines Prüfmerkmals	38
4.3 Die Daten zweier Versuche	39
4.4 Statistische Auswertungsverfahren, die aus den Eigenschaften der Faktoren und Prüfmerkmale resultieren	41
4.5 Versuch V1 - Varianzanalyse	41
4.6 Versuch V2 - Korrelationsanalyse, Regressionsanalyse	44

4.7	Hinweise für die statistische Auswertung eines ordinal skalierten Merkmals	47
4.7.1	Beispiel Arzneimittelpflanzen, verschiedene Behandlungen gegen eine bestimmte Wurzelkrankheit	47
4.7.2	Boxplots und Zusammenhang zwischen zwei Variablen	47
4.7.3	Vergleich mittlerer Werte der 5 Behandlungen	49
4.7.3.1	Verfahren der statistischen Analyse	49
4.7.3.2	Merkmal Anzahl gesunder Pflanzen	50
4.7.3.3	Merkmal Anzahl Pflanzen mit Neuaustrieb	55
4.7.4	Analyse von Häufigkeiten	58
4.8	Zusammenfassung	63
	Literatur	64
	Verzeichnis der SAS-Programme	65
	Abbildungsverzeichnis	65
	Tabellenverzeichnis	66

## Vorbemerkungen

„Politiker benutzen die Statistik wie ein Betrunkener einen Laternenpfahl: nicht, um eine Sache zu beleuchten, sondern um sich daran festzuhalten.“

**Gerd Bosbach**, Remagener Mathematikprofessor, in einem dpa-Gespräch, in dem er zum kritischen Umgang mit statistischen Daten auffordert.

Dem lässt sich hinzu fügen:

Manche Wissenschaftler benutzen die Statistik, weil sie von Referees der Fachzeitschriften dazu aufgefordert wurden.

Sollten sie doch einmal aus eigenem Antrieb statistische Verfahren und Methoden einsetzen, dann häufig nur, um *Sterne* zeichnen oder das Wort *signifikant* verwenden zu können.

Oft wird die fachliche Versuchsfrage erst formuliert, wenn die Experimente durchgeführt wurden. Hinzu kommt, dass unter der fachlichen Versuchsfrage viel zu häufig nur das Aufzählen zu beobachtender/beobachteter (Prüf-)Merkmale verstanden wird. Statistik wird in der Regel eingesetzt, wenn die Experimente durchgeführt wurden. Und das hauptsächlich, weil die Zeitschrift oder andere es fordern. Dann allerdings werden biometrisch ungeplante Versuche ausgewertet und methodische Fehler wie z.B. keine Zufallsauswahl sind nicht mehr korrigierbar.

Es werden in diesem Heft zwar Aspekte der Versuchsplanung gestreift, sie sind aber nicht Hauptgegenstand der Betrachtungen. Wir wollen Hinweise, Anregungen und vor allem Unterstützung geben, wie (erste) Schritte der Datensichtung und -analyse aussehen *könn(t)en*. Es bietet sich dabei an, auf die im Auftrag des Senats der Bundesforschungsanstalten herausgegebenen Hefte „Einführung in die Biometrie“ Bezug zu nehmen. Wir halten uns auch an die Symbolik dieser Hefte.

Die softwareseitige Basis ist SAS<sup>®</sup> 9. Wegen der z.T. ungünstigen deutschen „Übersetzung“ wird die englische Version bevorzugt. Die SAS-Programme werden schrittweise aufgebaut. Den SAS-Programmen sind immer die folgenden Anweisungen voranzustellen:

```
OPTIONS NOCENTER NODATE NONUMBER; TITLE;
```

Die Ausgaben werden erläutert. Da eine Datensichtung ohne eine Visualisierung der Daten kaum zu machen ist, werden auch verschiedene grafische Darstellungen eine Rolle spielen.

Im Kapitel 4 wird bewusst über erste Schritte der Datenanalyse hinaus gegangen, um grafische Analysen mit statistischen zu verbinden und neben der Anwendung klassischer Testverfahren einen Blick auf parameterfreie zu richten.

Dank sagen die Autoren Herrn Dr. Erich Schumacher (Universität Hohenheim) für die Durchsicht von Kapitel 4 und seine Anregungen.

Kleinmachnow, Waldsiefersdorf, Juni 2007

Eckard Moll  
Thomas Stauber



# 1 Berechnen statistischer Maßzahlen und Veranschaulichung der Daten

## 1.1 Eine Stichprobe

### 1.1.1 Beispiel Ährenlängen, Übernahme der in MS-Excel gespeicherten Daten nach SAS

Als Ausgangspunkt wählen wir das Ährenlängenbeispiel aus RICHTER (2004, S. 41). Die Daten sind noch einmal in der Tab. 1 angegeben.

**Tab. 1:** Urliste für die Längen von  $n = 240$  Ähren ( $x_i$  in cm,  $i = 1, 2, \dots, 240$ )

1	6,0	31	7,5	61	5,7	91	7,4	121	7,3	151	6,6	181	6,0	211	6,5
2	9,0	32	8,0	62	8,1	92	8,3	122	9,2	152	7,5	182	6,0	212	7,0
3	6,8	33	6,6	63	8,5	93	10,0	123	6,3	153	6,6	183	7,6	213	8,8
4	8,2	34	6,6	64	9,6	94	10,0	124	8,9	154	8,0	184	7,8	214	4,5
5	9,6	35	11,3	65	8,6	95	5,3	125	7,3	155	10,2	185	6,9	215	9,3
6	7,9	36	7,5	66	7,8	96	6,3	126	9,6	156	8,0	186	8,2	216	8,3
7	7,3	37	8,0	67	7,3	97	8,3	127	7,8	157	7,8	187	6,7	217	5,6
8	6,7	38	9,0	68	9,0	98	7,6	128	4,8	158	6,5	188	7,0	218	8,4
9	8,6	39	4,0	69	9,3	99	8,8	129	7,0	159	6,0	189	5,0	219	7,3
10	7,1	40	6,0	70	7,3	100	7,6	130	7,4	160	9,5	190	8,8	220	8,6
11	5,0	41	6,5	71	6,4	101	6,4	131	5,0	161	6,6	191	7,0	221	9,1
12	7,0	42	11,0	72	7,7	102	8,6	132	7,5	162	6,8	192	7,0	222	7,2
13	7,4	43	8,3	73	9,5	103	6,0	133	8,8	163	8,5	193	7,5	223	5,4
14	5,8	44	5,2	74	9,3	104	6,1	134	7,4	164	9,1	194	10,5	224	6,3
15	9,5	45	7,0	75	7,0	105	6,6	135	8,5	165	4,4	195	7,8	225	6,2
16	6,7	46	9,0	76	4,5	106	7,0	136	6,8	166	4,8	196	10,0	226	7,7
17	6,0	47	5,8	77	9,0	107	6,8	137	8,0	167	5,6	197	7,1	227	9,0
18	8,3	48	7,8	78	7,4	108	9,0	138	8,9	168	10,2	198	10,0	228	7,0
19	8,0	49	8,2	79	5,1	109	6,5	139	8,8	169	6,0	199	4,5	229	8,0
20	7,6	50	8,5	80	4,7	110	8,2	140	8,1	170	7,0	200	10,0	230	5,3
21	7,3	51	7,8	81	7,8	111	8,0	141	6,0	171	8,3	201	9,8	231	8,7
22	9,2	52	5,2	82	8,0	112	8,0	142	5,9	172	7,8	202	7,4	232	8,7
23	9,0	53	7,2	83	8,0	113	11,0	143	5,5	173	6,7	203	5,8	233	6,8
24	6,4	54	7,8	84	7,5	114	5,3	144	8,0	174	6,3	204	8,2	234	7,5
25	6,7	55	9,0	85	6,8	115	5,0	145	5,0	175	6,4	205	5,4	235	7,5
26	9,0	56	7,6	86	10,0	116	9,4	146	9,0	176	8,2	206	6,8	236	7,0
27	8,5	57	6,9	87	5,0	117	7,0	147	9,2	177	6,7	207	6,0	237	4,3
28	9,8	58	5,1	88	6,9	118	8,7	148	5,6	178	8,0	208	8,4	238	5,4
29	5,6	59	7,5	89	6,3	119	5,9	149	8,5	179	7,3	209	7,4	239	5,4
30	7,8	60	7,0	90	8,2	120	8,4	150	7,2	180	5,3	210	5,6	240	8,2

Daten werden oft mit MS-Excel erfasst. Das könnte dann wie in Abb. 1 aussehen. Dabei wurde obige Struktur übernommen.

The screenshot shows a Microsoft Excel spreadsheet titled 'Aehrenlaengen.xls'. The data is entered into columns A through H, with rows 1 through 30. The values in the spreadsheet match the data provided in Tab. 1.

**Abb. 1:** Mit MS-Excel erfasste Ährenlängen

Gemessen wurde das Merkmal Ährenlänge an 204 verschiedenen, zufällig ausgewählten Halmen. Die zufällige Auswahl ist notwendig, wenn statistische Verfahren und Methoden angewendet werden sollen.

### 1.1.2 Berechnen statistischer Maßzahlen

Das Merkmal Ährenlänge ist ein metrisches, quantitatives und stetiges Merkmal, das uns alle Möglichkeiten der Datenkomprimierung und –analyse bei der Betrachtung der Häufigkeitsverteilung, den Maßzahlen der Lage und der Streuung erlaubt (s.a. RICHTER 2004). Deshalb beginnen wir mit einer Liste statistischer Maßzahlen. Auch wenn sie nicht vollständig ist, wird selten so umfangreich gebraucht.

- Anzahl der Beobachtungen,
- (arithmetischer) Mittelwert,
- Minimum,
- Maximum,
- Spannweite,
- Modalwert,
- Varianz,
- Standardabweichung,
- Schiefe,
- Exzeß,
- Quartile,
- Quartilabstand,
- Quantile ( $P_{0,05}, P_{0,9}, P_{0,95}$ ).

Diese Maßzahlen werden aus einer Stichprobe (genauer: Zufallsstichprobe) für das Merkmal Ährenlänge berechnet und mit Symbolen gekennzeichnet. Für die Berechnung der meisten statistischen Maßzahlen gibt es in MS-Excel Funktionen. Sowohl die Symbole als auch die MS-Excel-Funktionen (deutsche Version) sind in Tab. 2 zu finden. Dabei steht  $M$  für die Datenmatrix,  $i = 1, 2, 3$  für das Quartil und  $p$  für die Wahrscheinlichkeit.

**Tab. 2:** Liste der statistischen Maßzahlen der Stichprobe, ihre Symbole und die MS-EXCEL-Funktionen

Maßzahl	Symbol	Funktion (MS-Excel)
Anzahl der Daten	n	Anzahl( $M$ )
(arithmetischer) Mittelwert	$\bar{x}$	Mittelwert( $M$ )
Minimum	Min	Min( $M$ )
Maximum	Max	Max( $M$ )
Spannweite	w	-
Modalwert	Mo	Modalwert( $M$ )
Varianz	$s^2$	Varianz( $M$ )
Standardabweichung	s	Stabw( $M$ )
Schiefe	$g_1$	Schiefe( $M$ )
Exzeß	$g_2$	Kurt( $M$ )
Quartile	$Q_1, Q_2 = Me = P_{0,5}, Q_3$	Quartile( $M; i$ )
Quartilsabstand	$I_{0,5}$	-
Quantile ( $P_{0,05}, P_{0,9}, P_{0,95}$ )	$P_{0,05}, P_{0,9}, P_{0,95}$	Quantile( $M; p$ )

Die Spannweite  $w$  ist die Differenz zwischen dem Maximum und dem Minimum und der Quartilsabstand  $I_{0,5}$  die Differenz zwischen  $Q_3 (= P_{0,75})$  und  $Q_1 (= P_{0,25})$ . Mit Hilfe der Funktionsaufrufe und diesen Differenzen lassen sich in MS-Excel die aufgeführten Maßzahlen berechnen.

Um die Maßzahlen aus Tab. 2 in SAS zu berechnen, braucht es etwas Vorbereitung, denn die Daten sollen ja nicht noch einmal erfasst werden. Der erste Schritt ist, die Daten der Ährenlängen aus der MS-Excel-Datei in eine SAS-Datei zu übertragen. Dazu nutzen wir den Import von SAS:

1. Aufruf des Import-Wizard: File → Import Data...
2. Auswahl/Bestätigung der Datenquelle MS-Excel:
3. Workbook angeben: MS-Excel-Datei eingeben

4. gewünschte(s) Tabelle/Arbeitsblatt auswählen  
 unter **Options...** besonders die erste Zeile beachten  
 → wird der Haken gesetzt (Standard), dann sollen die in der ersten Zeile stehenden Spaltenbezeichnungen als Variablenbezeichnungen in SAS übernommen werden (d.h. keine Umlaute, Sonder-, Operator- und Leerzeichen verwenden!)

Für unser Beispiel (s. Abb. 1) darf der Haken nicht gesetzt werden, denn in der ersten Zeile stehen Daten und keine Spaltenbezeichnungen. **OK** **Next >**

5. Speichern als SAS-Datei:

Library

als temporäre Datei in die Library Work (Standard)

als permanente Datei in eine vorher benannte Library (z.B.: lib, l, meier, sasuser)

Member

SAS-Dateinamen eintragen (keine Sonder- oder Leerzeichen)

Mit **Finish** wird der Import sofort vorgenommen, mit **Next >** und dann **Finish** wird zuvor das SAS-Programm zum Import der Daten gespeichert.

Zuerst ein Blick auf das SAS-Programm (Programm 1) zum Import der Daten:

**Programm 1:** Import\_Aehrenlaenge.sas

```
PROC IMPORT OUT= WORK.Aehrenlaengen
            DATAFILE= "G:\Aehrenlaengen.xls"
            DBMS=EXCEL REPLACE;
            SHEET="Ährenlängen$";
            GETNAMES=NO;
            MIXED=NO;
            SCANTEXT=YES;
            USEDATE=YES;
            SCANTIME=YES;
RUN;
```

Dieses Programm kann wiederholt genutzt werden, auch nach Änderungen der Daten in der MS-Excel-Datei.

Abb. 2 zeigt den Inhalt der so angelegten (temporären) Datei Aehrenlaengen. Es fällt auf, dass sieben verschiedene Variablen F1 bis F7 benannt wurden. Ziel ist eine Variable: Aehrenlaenge. Das wird im Programm 2 realisiert.

	F1	F2	F3	F4	F5	F6	F7
1	6	7.5	6.4	7	6	8.2	6.5
2	9	8	7.7	6.8	5.9	6.7	7
3	6.8	9	9.5	9	5.5	8	8.8
4	8.2	4	9.3	6.5	8	7.3	4.5
5	9.6	6	7	8.2	5	5.3	9.3
6	7.9	6.5	4.5	8	9	6	8.3
7	7.3	11	9	8	9.2	6	5.6
8	6.7	8.3	7.4	11	5.6	7.6	8.4
9	8.6	5.2	5.1	5.3	8.5	7.8	7.3
10	7.1	7	4.7	5	7.2	6.9	8.6
11	5	9	7.8	9.4	6.6	8.2	9.1
12	7	5.8	8	7	7.5	6.7	7.2
13	7.4	7.8	8	8.7	6.6	7	5.4
14	5.8	8.2	7.5	5.9	8	5	6.3
15	9.5	8.5	6.8	8.4	10.2	8.8	6.2
16	6.7	7.8	10	7.3	8	7	7.7
17	6	5.2	5	9.2	7.8	7	9
18	8.3	7.2	6.9	6.3	6.5	7.5	7

**Abb. 2:** Mit MS-Excel erfasste Ährenlängen

**Programm 2:** Variable\_Aehrenlaenge.sas

```

DATA Kopie;
  SET aehrenlaengen;
  Aehrenlaenge = F1; OUTPUT;
  Aehrenlaenge = F2; OUTPUT;
  Aehrenlaenge = F3; OUTPUT;
  Aehrenlaenge = F4; OUTPUT;
  Aehrenlaenge = F5; OUTPUT;
  Aehrenlaenge = F6; OUTPUT;
  Aehrenlaenge = F7; OUTPUT;
RUN;
DATA aehrenlaengen;
  SET Kopie (WHERE=(Aehrenlaenge ^= .));
  DROP F1-F7;
RUN;

```

Die statistischen Maßzahlen berechnen wir mit Hilfe des Programms 3.

**Programm 3:** Aehrenlaenge1.sas

```

PROC UNIVARIATE DATA= aehrenlaengen;
  VAR Aehrenlaenge;
RUN;

```

In Tab. 3 werden die berechneten statistischen Maßzahlen aufgelistet.

**Tab. 3:** Statistischen Maßzahlen, berechnet mit MS-Excel-Funktionen und SAS

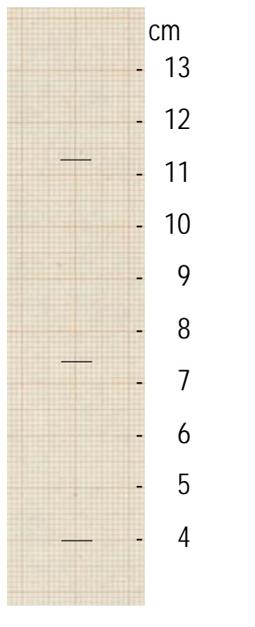
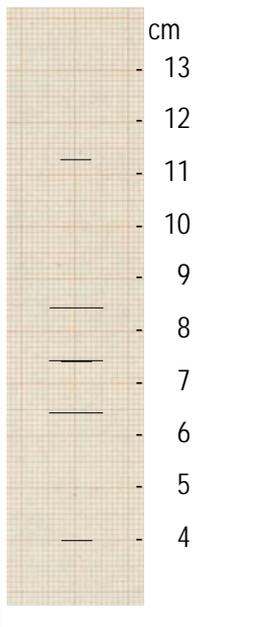
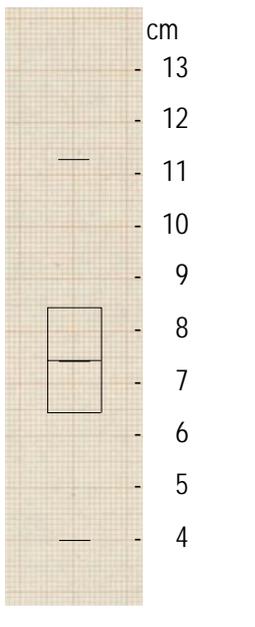
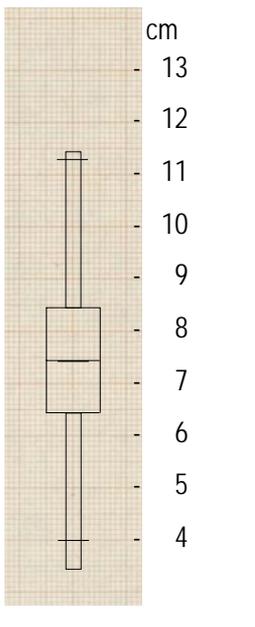
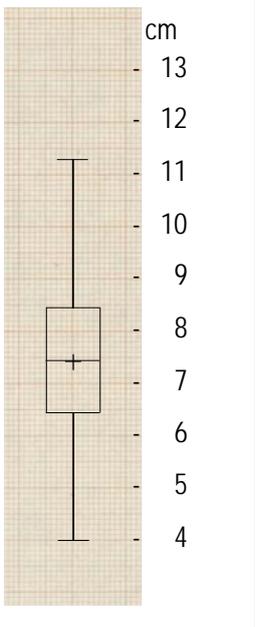
Maßzahl	Symbol	Ergebnisse	
		MS-EXCEL	SAS
Anzahl der Daten	n	240	240
(arithmetischer) Mittelwert	$\bar{x}$	7,42	7.42
Minimum	Min	4,00	4.00
Maximum	Max	11,30	11.30
Spannweite	w	7,30	7.30
Modalwert	Mo	7,00	7.00
Varianz	$s^2$	2,12	2.12
Standardabweichung	s	1,46	1.46
Schiefe	$g_1$	0,010	0.010
Exzeß	$g_2$	-0,394	-0.394
Quartile	$Q_1 = P_{0,25}$	6,48	6.45
	$Q_2 = Me = P_{0,5}$	7,45	7.45
	$Q_3 = P_{0,75}$	8,43	8.45
Quartilabstand	$I_{0,5}$	1,95	1.95
Quantile	$P_{0,05}$	5,00	5.00
	$P_{0,9}$	9,21	9.25
	$P_{0,95}$	9,81	9.90

Es gibt numerische Unterschiede in den Schätzungen der Quantile. Der Hintergrund dafür sind unterschiedliche Definitionen und Algorithmen, auf die RICHTER (2004, S. 60) hinweist. Bei genauem Lesen der Ausgabe des Programms 3 fällt auf, dass die Überschrift zu den Quantilen einen Zusatz auf die (standardseitig) verwendete Berechnung hat: Quantiles (Definition 5). Diese Definition soll auch hier verwendet werden.

### 1.1.3 Box-Whisker-Plot

Die Struktur nüchterner Zahlen lässt sich besser in einer Grafik erfassen. In Tab. 4 wird auf Millimeterpapier schrittweise ein Box-Plot entwickelt.

**Tab. 4:** Wege zum Box-Whisker-Plot

Min, Max, $\bar{x}$	$Q_1, Q_2 (= Me), Q_3$	Markieren des Quartilabstands	beidseitiges „Anheften“ des 1,5-fachen Quartilabstands	Begrenzen der Whisker bis zu den Extremwerten, spezielle Markierung des Mittelwerts
				
	Der geringe Abstand zwischen $Q_2$ und $\bar{x}$ veranschaulicht die geringe Schiefe. Das positive Vorzeichen weist auf eine rechtsschiefe Verteilung ( $Q_2 < \bar{x}$ ).	Innerhalb dieser Box liegen 50 % aller Werte.		Werte, die außerhalb des 1,5-fachen der Quartilabstände liegen, werden mitunter mit Symbolen markiert, um so eventuelle Ausreißer kenntlich zu machen.

Der Box-Whisker-Plot veranschaulicht in einer einfachen Grafik fast alle Maßzahlen aus Tab. 3 und lässt Rückschlüsse auf die Symmetrie der Verteilung zu. Deshalb sind Box-Plots so informativ und sollten viel häufiger eingesetzt werden. Zunehmend findet man sie auch in Veröffentlichungen, weil sie viel mehr über die Daten aussagen, als eine Tabelle. Auf keinen Fall ist es hinreichend, nur den Mittelwert mit evtl. der Anzahl der Daten anzugeben, aus denen er berechnet wurde. Es **muss** die Variabilität der Daten in geeigneter Form mitgeteilt werden. Das ist durch die geschätzte Varianz  $s^2$  und bevorzugt bei symmetrischer Verteilung durch die Angabe der Form  $\bar{x} \pm s$  möglich. Bei nichtmetrischen Daten und/oder stärkerer Asymmetrie bieten sich die Quantile an.

Ein Box-Whisker-Plot kann in MS-Excel nicht gezeichnet werden. In SAS steht die Prozedur *BOXPLOT* zur Verfügung (Programm 4).

**Programm 4:** Grundstruktur der Prozedur *BOXPLOT*

```
PROC BOXPLOT DATA=<daten> ;
  PLOT y * x ;
RUN;
```

Anstelle der allgemeinen Variablen y kann die Variable Aehrenlaenge gesetzt werden. Aber was steht für x, eine Variable, die immer gebraucht wird? Sie könnte einen konstanten Wert haben, numerische oder Zeichenkettenvariable sein. Dieser Wert muss aber in einem DATA-Schritt vorher zugewiesen werden:

**Programm 5:** Aehrenlaenge2.sas

```
DATA boxplot;
  SET aehrenlaengen;
  x = 'n = 204' ;           * oder: x = 'in cm';           * oder: x = 1 ;
RUN;
PROC BOXPLOT DATA=boxplot ;
  LABEL x = 'Box-Whisker-Plot';
  PLOT Aehrenlaenge * x /CBOXES=black ;
RUN; QUIT;
```

Wenn man sich nun noch daran gewöhnt, vor die erste Prozedur, die eine Grafik erzeugt, grafische Optionen z.B.

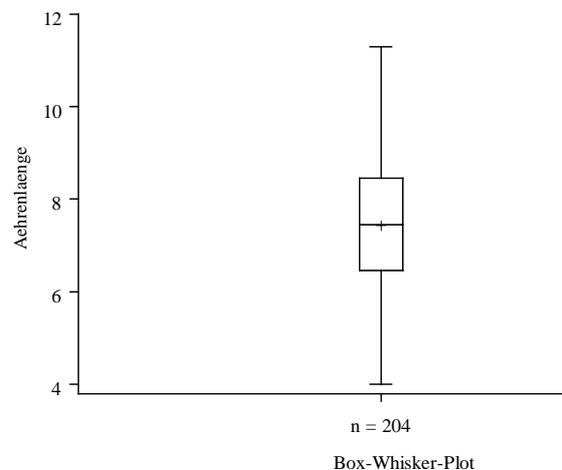
```
GOPTIONS device=emf
          ftext="Times"           hier kann als Schriftart z.B. auch „Arial“ stehen
          htext=1.2 ;
```

zu vereinbaren, ist das Ergebnis auch sehenswert (Abb. 3). Diese Grafik kann als Datei gespeichert werden:

File → Export as Image...

Als Dateityp **muss** unbedingt EMF Files (\*.emf) gewählt werden!

Der Vorteil dieser EMF-Grafiken ist, dass sie hochauflösende Vektorgrafiken sind und in MS-Word bearbeitet werden können (Modus: Grafik bearbeiten).



**Abb. 3:** Box-Whisker-Plot der Ährenlängen

Mit einer weiteren Anweisung vor dem Aufruf von *PROC BOXPLOT* (am besten in Verbindung mit der *GOPTIONS*-Anweisung):  
`SYMBOL c=blue h=0.4 v=dot;`  
 wird das den Mittelwert kennzeichnende + -Zeichen durch einen (blauen) Punkt ersetzt.

**1.1.4 Häufigkeitsverteilung für nicht klassifizierte Daten**

Einen schnellen Eindruck über die Häufigkeitsverteilung der 240 Ährenlängen liefert das Programm

**Programm 6:** Aehrenlaenge3.sas

```
GOPTIONS device=emf ftext="Times" htext=1.2;
PROC GCHART DATA= Aehrenlaengen;
  VBAR Aehrenlaenge;
RUN; QUIT;
```

Die Häufigkeitsverteilung (Abb. 4) kann nur ein erster Eindruck sein. Eine Klasseneinteilung – besonders unter Berücksichtigung der Reduktionslagen – wie sie RICHTER (2004, S. 45 bis 56) beschreibt, wird hierbei nicht verwendet.

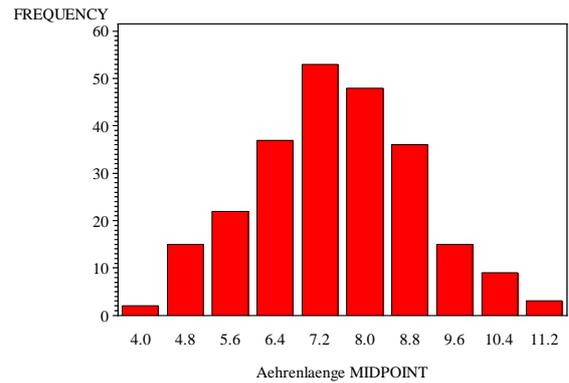


Abb. 4: Häufigkeitsverteilung der Ährenlängen

Hilfreich zum Erkennen möglicher Ausreißer, Schreibfehler oder dgl. kann eine weitere einfache grafische Darstellung sein (Programm 7).

Programm 7: Aehrenlaenge4.sas

```
DATA plot;
  SET aehrenlaengen;
  x = 1 ;
RUN;
SYMBOL c=black h=0.6 v=dot;
GOPTIONS device=emf ftext="Times" htext=1.2;
PROC GPLOT DATA=plot ;
  LABEL x = 'Ährenlänge'
        Aehrenlaenge = 'in cm';
  PLOT Aehrenlaenge*x;
RUN; QUIT;
```

Das Ergebnis (Abb. 5) zeigt keine Auffälligkeiten.

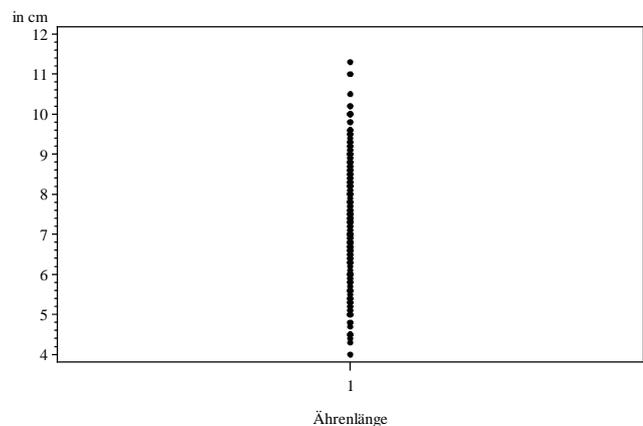


Abb. 5: Verteilung der Ährenlängen

### 1.1.5 Sind die Daten normalverteilt?

Eine häufig gestellte Frage – seltener wird nach dem Warum gefragt – ist, ob die Daten normalverteilt sind.

Die Normalverteilung ist eine stetige Verteilung. Sie kommt folglich für diskrete oder qualitative Merkmale nicht in Frage bzw. nur approximativ. Das Programm 8 liefert Maßzahlen, Tests (Ausgabe 8.1) und grafische Darstellungen (Ausgabe 8.2), um sich für oder gegen die Annahme der Normalverteilung zu entscheiden.

Bemerkung:

Beim Testen bezieht sich die Voraussetzung der Normalverteilung unter der Nullhypothese im Allgemeinen auf die Residuen. Viel wichtiger als ein Test der Residuen auf Normalverteilung ist bei mehreren Stichproben die Einhaltung und Überprüfung der Varianzhomogenität. Auch ist es sinnvoll von den Eigenschaften des auszuwertenden Merkmals auszugehen, ob die Berechnung statistischer Maßzahlen und darauf aufbauende Auswertungsverfahren zulässig sind. Besonders im einfaktoriellen

Fall sind parametrische Tests robust gegenüber Abweichungen der Residuen von der Normalverteilung.

**Programm 8:** Aehrenlaenge5.sas

```
GOPTIONS reset=all;
GOPTIONS device=emf ftext="Times" htext=1.2;
SYMBOL v=dot h=0.6 i=none;
PROC UNIVARIATE DATA= Aehrenlaengen NORMAL;
  VAR Aehrenlaenge;
  HISTOGRAM / normal ;
  PROBLOT / normal ;
  QQPLOT / normal ;
RUN; QUIT;
```

**Ausgabe 8.1:** zu Aehrenlaenge4.sas - Textausgabe

```
...
Tests for Normality
Test          --Statistic--      -----p Value-----
Shapiro-Wilk      W      0.993641      Pr < W      0.4021
Kolmogorov-Smirnov D      0.040167      Pr > D      >0.1500
Cramer-von Mises  W-Sq   0.041912      Pr > W-Sq   >0.2500
Anderson-Darling  A-Sq   0.318816      Pr > A-Sq   >0.2500
...

The UNIVARIATE Procedure
Fitted Distribution for Aehrenlaenge
Parameters for Normal Distribution
Parameter      Symbol      Estimate
Mean           Mu          7.42375
Std Dev        Sigma       1.455554

Goodness-of-Fit Tests for Normal Distribution
Test          ---Statistic---      -----p Value-----
Kolmogorov-Smirnov D      0.04016651      Pr > D      >0.150
Cramer-von Mises  W-Sq   0.04191176      Pr > W-Sq   >0.250
Anderson-Darling  A-Sq   0.31881555      Pr > A-Sq   >0.250

Quantiles for Normal Distribution
Percent      -----Quantile-----
Observed      Estimated
1.0          4.40000      4.03762
5.0          5.00000      5.02958
10.0         5.40000      5.55838
25.0         6.45000      6.44199
50.0         7.45000      7.42375
75.0         8.45000      8.40551
90.0         9.25000      9.28912
95.0         9.90000      9.81792
99.0        11.00000     10.80988
```

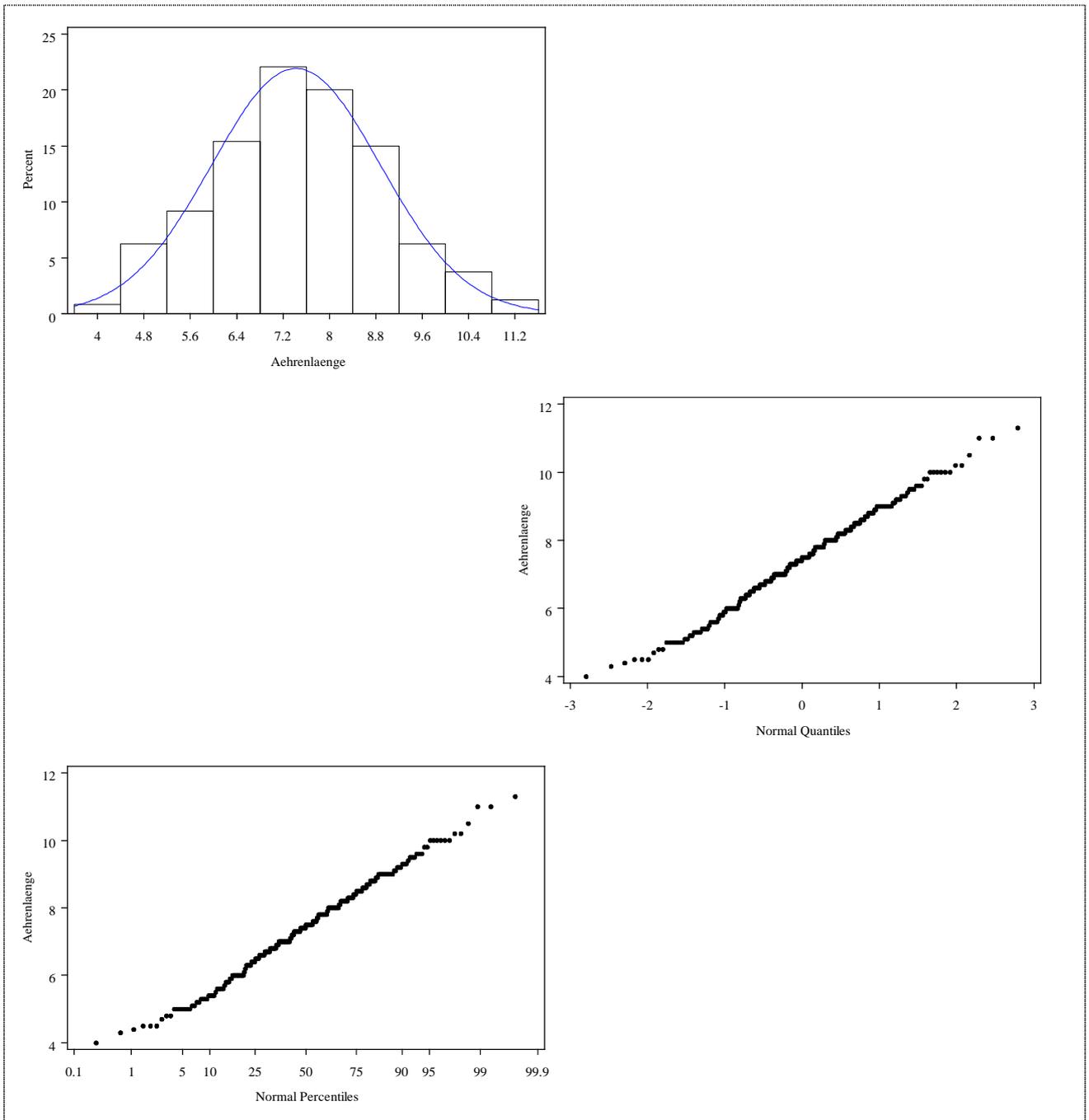
Wird nur mit der Anweisung HISTOGRAM ausgegeben

Die Quantile der gemessenen Ährenlängen und der geschätzten Normalverteilung ähneln sich. Für eine Testentscheidung zur Normalverteilung werden 4 Tests angeboten. Mit der Orientierung auf den Shapiro-Wilk-Test würde die Annahme der Nullhypothese „die Ährenlängen sind annähernd normalverteilt“ nicht abgelehnt werden, egal welche Philosophie man vertritt: ein Signifikanzniveau  $\alpha \approx 0,20$ , um Abweichungen frühzeitig zu erkennen, oder ein Signifikanzniveau  $\alpha < 0,05$ , um die Annahme der Normalverteilung sehr lange beizubehalten. Bei der Entscheidung sollten auch die Grafiken (Ausgabe 8.2) herangezogen werden

- Häufigkeitsverteilung (Histogramm) mit angepasster Normalverteilung,
- Wahrscheinlichkeitsdarstellung und

- Quantile-Quantile-Plot.

**Ausgabe 8.2:** zu Aehrenlaenge4.sas - Grafiken



Die geschätzte Normalverteilung passt sich der Häufigkeitsverteilung (vgl. Abb. 4) gut an.

Die Punkte der Wahrscheinlichkeitsdarstellung und der Q-Q-Plot (Quantile-Quantile-Plot) liegen fast auf einer Geraden, so dass auch von dieser Seite der Annahme der Normalverteilung in der Grundgesamtheit nichts im Wege steht. Diese Geraden können mit einer zusätzlichen Option mit den aus den Daten geschätzten Werten für den Erwartungswert  $\mu$  und die Standardabweichung  $\sigma$  eingezeichnet werden:

```
PROBPLOT / normal(mu=est sigma=est) ;
QQPLOT / normal(mu=est sigma=est) ;
```

Sehr gut fällt in den beiden unteren Grafiken (Ausgabe 8.2) auch das von RICHTER (2004, S. 44) beschriebene Phänomen auf, dass Ährenlängen mit Nachkommastellen „1“ bzw. „9“ häufig dem nächstgelegenen ganzzahligen Wert zugerechnet werden.

### 1.1.6 Aufgabenstellungen für zwei weitere Beispiele

Berechnen Sie geeignete statistische Maßzahlen für nachfolgende Daten und erstellen Sie Grafiken, die diese am besten widerspiegeln.

**A1:** Ct-Gehalt in mg/100 g Boden an kondensierten Tanninen (phenologische Inhaltsstoffe, Gerbstoffe), der von 90 Probenahmestellen einer Ackerfläche ermittelt wurde (RICHTER 2004, S. 75):

0.70	0.79	0.73	0.63	0.71	0.65	0.65	0.85	0.78	0.55	0.88	0.74	0.72	0.47	0.68
0.66	0.71	0.76	0.70	0.76	0.86	0.72	0.81	0.68	0.69	0.81	0.61	0.72	0.64	0.61
0.84	0.56	0.70	0.56	0.77	0.59	0.68	0.68	0.55	0.51	0.81	0.6	0.85	0.73	0.60
0.52	0.68	0.57	0.82	0.59	0.62	0.65	0.72	0.60	0.71	0.61	0.69	0.66	0.60	0.56
0.68	0.49	0.65	0.70	0.76	0.59	0.88	0.62	0.68	0.71	0.67	0.61	0.67	0.72	0.71
0.52	0.79	0.78	0.50	0.63	0.64	0.64	0.65	0.68	0.57	0.59	0.74	0.64	0.67	0.59

**A2:** Gewicht von Mäusen in g im Alter von 125 Tagen:

26,0	16,5	20,0	17,0	18,0	16,0	28,5	21,0	23,5	23,0	20,0	19,5	22,5
18,0	24,0	18,5	24,0	20,0	25,0	28,0	22,0	27,5	24,0	20,5	22,5	23,0

### 1.1.7 Maßzahlen der 3 Beispiele Ährenlängen, Ct-Gehalte, Mäusegewichte

Für jede der drei voneinander unabhängigen Datenreihen Ährenlängen (Tab. 1), Ct-Gehalte (A1), Mäusegewichte (A2) gilt, dass durch die Zufallsauswahl auch die Daten innerhalb jeder Datenreihe voneinander unabhängig sind – eine Voraussetzung für die Berechnung statistischer Maßzahlen und statistische Analysen. Eine Auswahl der geschätzten Parameter ist:

**Tab. 5a:** Statistische Maßzahlen

	Ährenlängen	Ct-Gehalte	Mäusegewichte
Maßeinheit	mm	mg	g
n	240	90	26
$\bar{x}$	7,42	0,674	21,87
s	1,46	0,093	3,503
$\bar{x} \pm s$	7,42 ± 1,46	0,674 ± 0,093	21,87 ± 3,503

**Tab. 5b:** Konfidenzintervall und Quartile

	Ährenlängen	Ct-Gehalte	Mäusegewichte
Freiheitsgrade FG für die Schätzung des 0,95-Konfidenzintervalls	240 - 1 = 239	90 - 1 = 89	26 - 1 = 25
$t_{1-\alpha/2; FG}$	1,96994	1,98698	2,05954
0,95-Konfidenzintervall $\left[ \bar{x} - t_{1-\alpha/2; FG} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{1-\alpha/2; FG} \cdot \frac{s}{\sqrt{n}} \right]$	[7,23 , 7,61]	[0,654 , 0,694]	[20,46 , 23,28]
Q <sub>1</sub>	6,45	0,61	19,50
Q <sub>2</sub> = Me	7,45	0,68	22,25

$Q_3$	8,45	0,72	24,00
-------	------	------	-------

Die miteinander verknüpfte Angabe von Mittelwert und Standardabweichung  $\bar{x} \pm s$  sollte ebenso wie die eines Konfidenzintervalls bevorzugt bei symmetrischer Verteilung des Merkmals genutzt werden. Die Quartile spiegeln dagegen besser eine asymmetrische (eingipflige) Verteilung wider. Sie gilt es auch bei einem diskreten Merkmal anzugeben, zumal dann im Allgemeinen ein Mittelwert keine Aussagekraft und Zulässigkeit hat.

Die Werte in Tab. 5a und b sind bis auf Anzahl, Freiheitsgrade und t-Quantil alles Werte, zu denen die angegebene Maßeinheit gehört.

Ein Vergleich der Standardabweichungen (Tab. 5a) zeigt, dass die Ct-Gehalte am geringsten und die Mäusegewichte am stärksten streuen. Die Maßzahl Variationskoeffizient  $s\%$  ist eine bezüglich des Mittelwerts relativierte Streuungsangabe  $s\% = \frac{s}{\bar{x}} \cdot 100\%$ . Die Maßeinheit ist nun Prozent. Die

Bezugnahme auf den Mittelwert macht es möglich, verschiedene Merkmale hinsichtlich ihrer Streuung zu vergleichen (Tab. 6). Bezogen auf den jeweiligen Mittelwert oder anders gesagt in Einheiten des jeweiligen Mittelwerts streuen die Ährenlängen am stärksten.

**Tab. 6:** Variationskoeffizienten verschiedener Merkmale

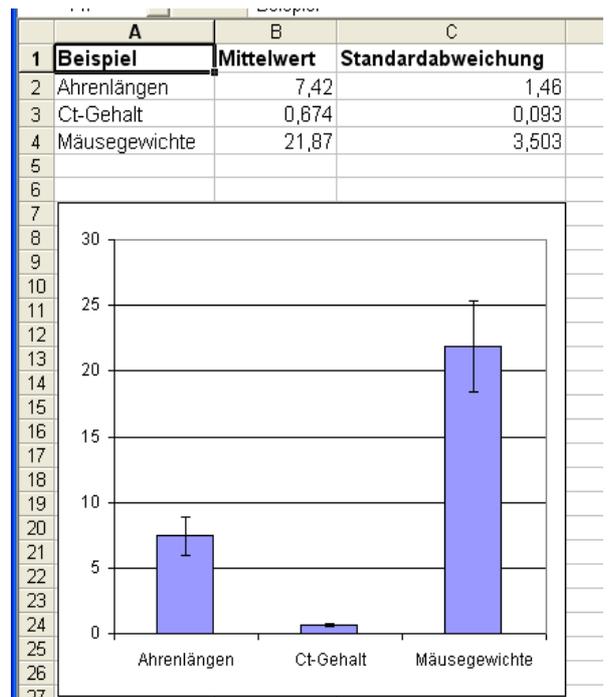
	Ährenlängen	Ct-Gehalte	Mäusegewichte
$s\%$	19,68 %	13,80 %	16,02 %

Grundsätzlich gilt, dass zu einem mittleren Wert (Mittelwert oder Median) ein Streuungsmaß anzugeben ist. Bei einem Boxplot ist die Streuung der Daten gut erkennbar, bei einem Balkendiagramm, das nur die mittleren Werte ohne zusätzlich Informationen charakterisiert, dagegen nicht. Vergleichbar der Angabe  $\bar{x} \pm s$  sollte es auch Standard sein, in einem Balkendiagramm die Standardabweichung vom Mittelwert ausgehend nach beiden Seiten abzutragen.

In MS-Excel gibt es die Möglichkeit, Balkendiagramme – die Höhe eines Balkens entspricht dem Mittelwert des Merkmals – zusätzlich mit Fehlerindikatoren zu versehen. Als Fehlerindikatoren können die Standardabweichung (Abb. 6) oder das  $(1-\alpha)$ -Konfidenzintervall genutzt werden; seltener trifft (ausgehend von der Aufgabenstellung) der Standardfehler zu.

Ein Hinweis an dieser Stelle ist vielleicht angebracht: Dreidimensionale Balken mögen „hübsch“ sein, die dritte Dimension ist aber häufig fachlich und sachlich Unsinn.

In SAS kann eine ähnliche Grafik nur unter Verwendung der Einzeldaten gezeichnet werden. Dafür wird die Prozedur *GCHART* mit der *VBAR*-Option *ERRORBAR=* verwendet. Diese Grafiken sind nicht so ansehnlich wie die von MS-Excel.



**Abb. 6:** Balkendiagramme mit Standardabweichungen

## 1.2 Zwei Stichproben

### 1.2.1 Beispiel Mäusegewichte

Leider kommt es in der biometrischen Beratung vor, dass ein Versuchsansteller nicht nur nicht weiß, mit welcher Fragestellung er etwas untersucht hat, sondern auch wichtige Informationen als zu vernachlässigen eingestuft werden. So könnte es bei den Gewichten in g von Mäusen im Alter von 125 Tagen (s. Kap. 1.1.6, A2) auch gewesen sein. Die beiden Spalten sind geschlechtspezifisch aufgestellt. Die erste Spalte umfasst männliche und die zweite weibliche Tiere. Beide Datenreihen enthalten voneinander unabhängig ermittelte Messwerte von für jede Stichprobe zufällig ausgewählten Mäusen.

männlich	26,0	16,5	20,0	17,0	18,0	16,0	28,5	21,0	23,5	23,0	20,0	19,5	22,5
weiblich	18,0	24,0	18,5	24,0	20,0	25,0	28,0	22,0	27,5	24,0	20,5	22,5	23,0

Wir haben es nun mit zwei voneinander unabhängigen Stichproben zu tun. Die Berechnung der statistischen Maßzahlen und das Erstellen von Grafiken geht auf dem gleichen, bereits beschriebenen Weg. Liegen die Mäusegewichte in einer MS-Excel-Datei ohne Spaltenbezeichnung (s. Abb. 7) vor, dann kann eine SAS-Datei wie auf den Seiten 8 und 9 aufgeführt erstellt werden. Der Dateiname ist *Maus*. [Wenn in der ersten Zeile eine Bezeichnung steht, darf sie keine Umlaute, Operatoren, Leer- oder Sonderzeichen enthalten, da sie als Variablenbezeichnungen in SAS übernommen werden: gemäß Punkt 4, S. 9 entsprechenden Haken setzen.]

	A	B
1	26,0	18,0
2	16,5	24,0
3	20,0	18,5
4	17,0	24,0
5	18,0	20,0
6	16,0	25,0
7	28,5	28,0
8	21,0	22,0
9	23,5	27,5
10	23,0	24,0
11	20,0	20,5
12	19,5	22,5
13	22,5	23,0
14		

Abb. 7: Mäusegewichte in MS-Excel-Datei

Das SAS-Programm zur Berechnung statistischer Maßzahlen und zum Zeichnen der Boxplots und anderer Grafiken ist im Programm 9 angegeben, die Ergebnisse in Ausgabe 9.

#### Programm 9: Maus.sas

```

DATA MausGewicht;
  SET Maus;
  Geschlecht = 'männlich';
  Gewicht = F1;
  OUTPUT;
  Geschlecht = 'weiblich';
  Gewicht = F2;
  OUTPUT;
  DROP F1 F2;
RUN;

PROC SORT DATA=MausGewicht;
  BY Geschlecht;
RUN;

GOPTIONS device=emf
          ftext="Times"
          htext=1.2;

PROC BOXPLOT DATA=MausGewicht;
  PLOT Gewicht * Geschlecht /CBOXES=black ;
RUN;
PROC UNIVARIATE DATA=MausGewicht NORMAL;
  VAR Gewicht;
  BY Geschlecht;
  HISTOGRAM / normal ;
  PROBLOT / normal ;
RUN;

```

Die SAS-Datei Maus hat 2 Variablen F1 und F2, die das Gewicht der Mäuse beinhalten. Das auszuwertende Merkmal steht folglich in 2 Variablen. Daraus muss eine Variable werden. Zusätzlich wird eine Variable gebraucht, die die Gruppierung, das Geschlecht, charakterisiert.

Die Berechnung der statistischen Maßzahlen und einiger Grafiken soll für die Geschlechtergruppen getrennt vorgenommen werden. Das erfordert eine Sortierung.

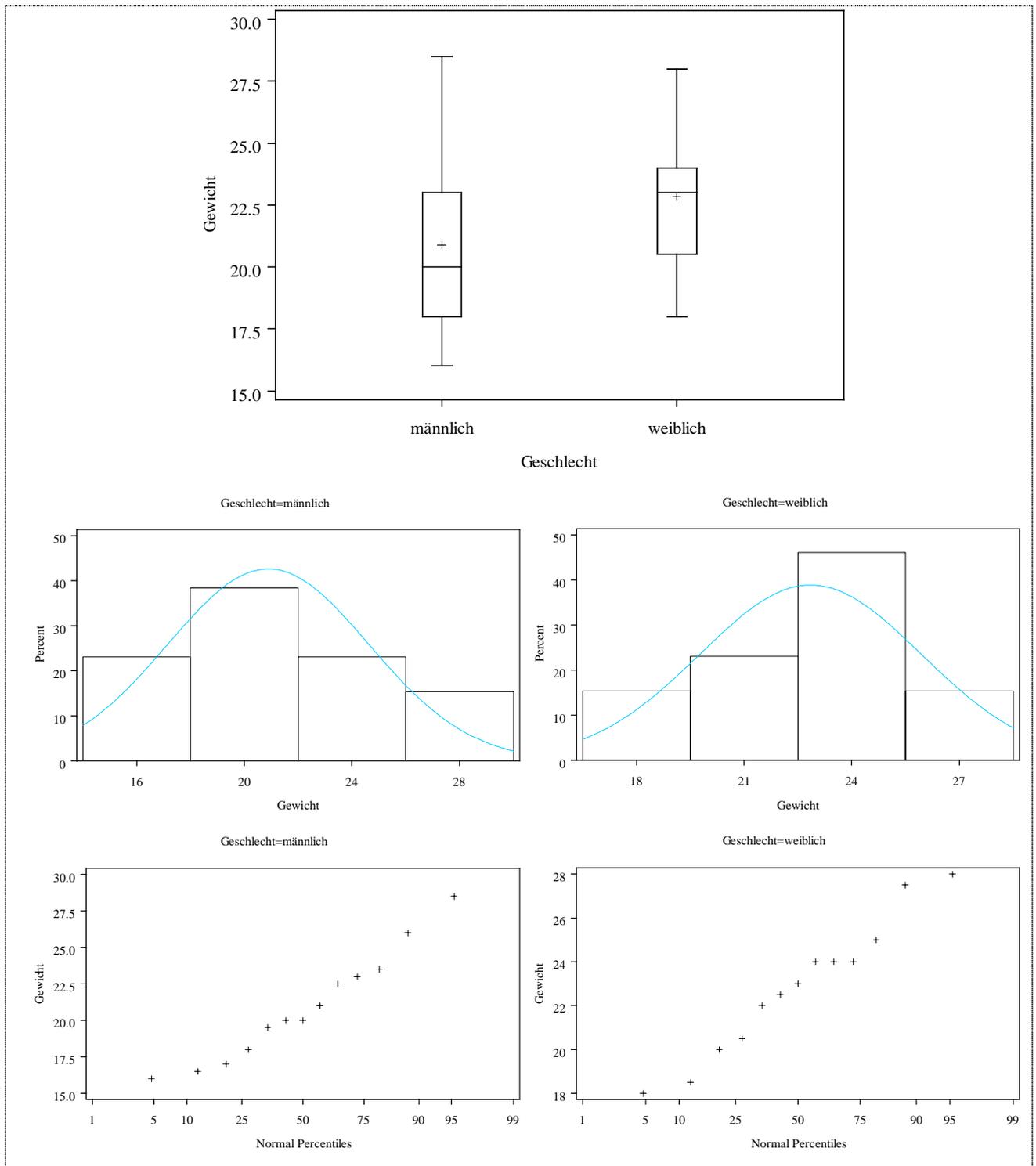
Diese oder ähnliche grafischen Optionen sollten Standard sein (s.S.12).

Box-Whisker-Plot  
y: Gewicht, x: Geschlecht

statistische Maßzahlen und Grafiken für jedes Geschlecht getrennt

**QUIT;**

**Ausgabe 9.1:** zu Maus.sas - Grafiken



**Ausgabe 9.2:** zu Maus.sas - Textausgabe

Geschlecht=männlich				Geschlecht=weiblich			
The UNIVARIATE Procedure				The UNIVARIATE Procedure			
Variable: Gewicht				Variable: Gewicht			
Moments				Moments			
N	13	Sum Weights	13	N	13	Sum Weights	13
Mean	20.8846154	Sum Observations	271.5	Mean	22.8461538	Sum Observations	297
Std Deviation	3.7425139	Variance	14.0064103	Std Deviation	3.07804467	Variance	9.47435897
Skewness	0.60041642	Kurtosis	-0.1584876	Skewness	0.06676778	Kurtosis	-0.5384005
Uncorrected SS	5838.25	Corrected SS	168.076923	Uncorrected SS	6899	Corrected SS	113.692308
Coeff Variation	17.919956	Std Error Mean	1.0379866	Coeff Variation	13.4729228	Std Error Mean	0.85369599

# Kapitel 1: Berechnen statistischer Maßzahlen und Veranschaulichung der Daten

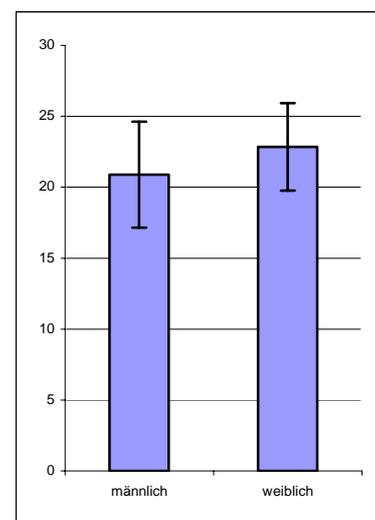
Basic Statistical Measures				Basic Statistical Measures			
Location		Variability		Location		Variability	
Mean	20.88462	Std Deviation	3.74251	Mean	22.84615	Std Deviation	3.07804
Median	20.00000	Variance	14.00641	Median	23.00000	Variance	9.47436
Mode	20.00000	Range	12.50000	Mode	24.00000	Range	10.00000
		Interquartile Range	5.00000			Interquartile Range	3.50000
Tests for Location: Mu0=0				Tests for Location: Mu0=0			
Test	-Statistic-	-----p Value-----		Test	-Statistic-	-----p Value-----	
Student's t	t 20.12031	Pr >  t	<.0001	Student's t	t 26.76146	Pr >  t	<.0001
Sign	M 6.5	Pr >=  M	0.0002	Sign	M 6.5	Pr >=  M	0.0002
Signed Rank	S 45.5	Pr >=  S	0.0002	Signed Rank	S 45.5	Pr >=  S	0.0002
Tests for Normality				Tests for Normality			
Test	--Statistic--	-----p Value-----		Test	--Statistic--	-----p Value-----	
Shapiro-Wilk	W 0.953902	Pr < W	0.6584	Shapiro-Wilk	W 0.961533	Pr < W	0.7771
Kolmogorov-Smirnov	D 0.131888	Pr > D	>0.1500	Kolmogorov-Smirnov	D 0.123112	Pr > D	>0.1500
Cramer-von Mises	W-Sq 0.031346	Pr > W-Sq	>0.2500	Cramer-von Mises	W-Sq 0.03104	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq 0.229829	Pr > A-Sq	>0.2500	Anderson-Darling	A-Sq 0.217269	Pr > A-Sq	>0.2500
Quantiles (Definition 5)				Quantiles (Definition 5)			
Quantile	Estimate			Quantile	Estimate		
100% Max	28.5			100% Max	28.0		
99%	28.5			99%	28.0		
95%	28.5			95%	28.0		
90%	26.0			90%	27.5		
75% Q3	23.0			75% Q3	24.0		
50% Median	20.0			50% Median	23.0		
25% Q1	18.0			25% Q1	20.5		
10%	16.5			10%	18.5		
5%	16.0			5%	18.0		
1%	16.0			1%	18.0		
0% Min	16.0			0% Min	18.0		
Extreme Observations				Extreme Observations			
----Lowest----		----Highest----		----Lowest----		----Highest----	
Value	Obs	Value	Obs	Value	Obs	Value	Obs
16.0	6	22.5	13	18.0	14	24.0	17
16.5	2	23.0	10	18.5	16	24.0	23
17.0	4	23.5	9	20.0	18	25.0	19
18.0	5	26.0	1	20.5	24	27.5	22
19.5	12	28.5	7	22.0	21	28.0	20
Geschlecht=männlich				Geschlecht=weiblich			
The UNIVARIATE Procedure				The UNIVARIATE Procedure			
Fitted Distribution for Gewicht				Fitted Distribution for Gewicht			
Parameters for Normal Distribution				Parameters for Normal Distribution			
Parameter	Symbol	Estimate		Parameter	Symbol	Estimate	
Mean	Mu	20.88462		Mean	Mu	22.84615	
Std Dev	Sigma	3.742514		Std Dev	Sigma	3.078045	
Goodness-of-Fit Tests for Normal Distribution				Goodness-of-Fit Tests for Normal Distribution			
Test	---Statistic---	-----p Value-----		Test	---Statistic---	-----p Value-----	
Kolmogorov-Smirnov	D 0.13188848	Pr > D	>0.150	Kolmogorov-Smirnov	D 0.12311182	Pr > D	>0.150
Cramer-von Mises	W-Sq 0.03134577	Pr > W-Sq	>0.250	Cramer-von Mises	W-Sq 0.03104028	Pr > W-Sq	>0.250
Anderson-Darling	A-Sq 0.22982857	Pr > A-Sq	>0.250	Anderson-Darling	A-Sq 0.21726855	Pr > A-Sq	>0.250
Quantiles for Normal Distribution				Quantiles for Normal Distribution			
Percent	-----Quantile-----	-----		Percent	-----Quantile-----	-----	
	Observed	Estimated			Observed	Estimated	
1.0	16.0000	12.1782		1.0	18.0000	15.6856	
5.0	16.0000	14.7287		5.0	18.0000	17.7832	
10.0	16.5000	16.0884		10.0	18.5000	18.9015	
25.0	18.0000	18.3603		25.0	20.5000	20.7700	
50.0	20.0000	20.8846		50.0	23.0000	22.8462	
75.0	23.0000	23.4089		75.0	24.0000	24.9223	
90.0	26.0000	25.6808		90.0	27.5000	26.7908	
95.0	28.5000	27.0405		95.0	28.0000	27.9091	
99.0	28.5000	29.5910		99.0	28.0000	30.0068	

Zur besseren Übersicht werden die wichtigsten Maßzahlen in Tab. 7 zusammen gefasst.

**Tab. 7:** Statistische Maßzahlen für die männlichen und weiblichen Mäusegewichte

	männlich	weiblich
n	13	13
$\bar{x}$	20,88	22,85
s	3,74	3,08
$s^2$	14,01	9,47
$Q_1$	18,0	20,5
$Q_2 = \text{Me}$	20,0	23,0
$Q_3$	23,0	24,0
$\bar{x} \pm s$	$20,88 \pm 3,74$	$22,85 \pm 3,08$
$\left[ \bar{x} - t_{1-\alpha/2; \text{FG}} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{1-\alpha/2; \text{FG}} \cdot \frac{s}{\sqrt{n}} \right]$	[18,62 , 23,15]	[20,99 , 24,71]

Die Darstellung der mittleren geschlechts-abhängigen Mäusegewichte erfolgt natürlich unter Angabe der jeweiligen Standardabweichungen (Abb. 8).



**Abb. 8:** Mäusegewichte mit Standardabweichung

### 1.2.2 Bemerkungen zu abhängigen und unabhängigen Stichproben sowie zu fachlichen Aufgabenstellungen, die entsprechende Auswertungsverfahren zur Folge haben

Für jede Stichprobe gilt, dass seine Elemente voneinander unabhängig sind und jedes Element der Grundgesamtheit dieselbe Chance hat, für die Stichprobe ausgewählt zu werden. Wir sprechen deshalb von Zufallsstichproben: Die Objekte, an denen die Messung oder Beobachtung des Merkmals vorgenommen werden sollen, **müssen** aus der Grundgesamtheit/den Grundgesamtheiten zufällig ausgewählt werden. Dieser Grundsatz kann nicht häufig genug wiederholt werden.

Für weitere Betrachtungen ist es aber wichtig zu wissen, ob die Stichproben abhängig oder unabhängig sind. Neben dem Begriff der Abhängigkeit wird häufig auch von gepaarten und ungepaarten Stichproben gesprochen. In unserem Beispiel der Mäusegewichte sind aus einer Population gleichaltriger Tiere 13 männliche zufällig und unabhängig davon 13 weibliche zufällig ausgewählt worden. Eine jeweilige Paarbildung oder sonstige Abhängigkeit ist nicht erkennbar. Anders sieht es bei der Bewertung eine Behandlung (z.B. vorher, nachher) aus, wobei das Merkmal am selben Tier erhoben wird. Von abhängigen oder gepaarten Stichproben spricht man auch, wenn von Zwillingspaaren einer in die eine und der andere in die andere Gruppe zufällig eingeordnet wird. Da die Ähnlichkeit zwischen den Paaren größer ist als zu den anderen Stichprobenelementen, hat das auch Auswirkungen auf die Auswertung, wenn es z.B. darum geht, mittlere Wirkungen miteinander zu vergleichen.

Für die Mäusegewichte könnten mögliche Fragen sein: „Unterscheiden sich die mittleren Gewichte und/oder Varianzen männlicher und weiblicher Tiere statistisch?“ Um das zu entscheiden, brauchen wir die Schätzwerte aus (unabhängigen) Zufalls-Stichproben und geeignete Verfahren mit ihren Voraussetzungen.

### 1.3 Zusammenfassung

Der erste Schritt **muss** immer sein, sich ein „Bild“ von den Daten zu machen. Das kann mit Hilfe geeigneter statistischer Maßzahlen, die aus den Daten der Zufallsstichprobe berechnet werden, und/oder Grafiken erfolgen. Zur Berechnung statistischer Maßzahlen stehen die SAS-Prozeduren *UNIVARIATE* (auch zu nutzen zur Überprüfung der Normalverteilung) und *MEANS* zur Verfügung. Neben dem mittleren Wert, Mittelwert oder Median, ist unbedingt ein geeignetes Streuungsmaß anzugeben. Für grafische Darstellungen ist an vorderster Stelle die SAS-Prozedur *BOXPLOT* zu nennen. Genutzt werden können auch *GLOT* und *GCHART*. In Grafiken wie Balkendiagrammen ist auch die Variabilität, z.B. mit Hilfe der Standardabweichung, mit darzustellen. Das kann in MS-Excel unter Nutzung der Fehlerindikatoren vorgenommen werden.

Bei zwei Stichproben ist es für weitere Auswertungsverfahren entscheidend, ob es sich um unabhängige oder abhängige, gepaarte Stichproben handelt.

## 2 Tests von Hypothesen über Parameter, zwei Stichproben

### 2.1 Fachliche Aufgabenstellungen und Ableitung der statistischen Fragestellungen

Für die Gewichte der Mäuse (in g) nach 125 Tagen (s. Tab. 7) ist festzuhalten:

- $\bar{x}_{\text{männlich}} < \bar{x}_{\text{weiblich}}$  (20,88 < 22,85)
- $s_{\text{männlich}}^2 > s_{\text{weiblich}}^2$  (14,01 > 9,47).

Diese Beziehungen gelten – für die Stichproben. Da aber Aussagen für die Grundgesamtheit zu treffen sind, müssen wir ein Hypothesenpaar für Parameter der Grundgesamtheiten formulieren, das dann getestet wird. Voraussetzung ist, dass die Stichprobenelemente, d.h. die Objekte, an denen die Merkmale beobachtet wurden, zufällig aus den Grundgesamtheiten ausgewählt wurden. Nur dann sind die Stichproben für die Grundgesamtheiten repräsentativ.

Die erste Hypothese, die Nullhypothese, wird dahingehend formuliert, dass (in den Grundgesamtheiten) kein Unterschied zwischen den Parametern besteht. Die zweite, notwendige Hypothese ist von der fachlichen Zielstellung abgängig. Nur jeweils eine der 3 in Tab. 8 aufgeführten Alternativhypothesen kann zutreffen.

**Tab. 8:** Auswahl des Hypothesenpaares

Hypothese		Erwartungswert	Parameter Varianz	allgemein
Nullhypothese $H_0$		$\mu_{\text{männlich}} = \mu_{\text{weiblich}}$	$\sigma_{\text{männlich}}^2 = \sigma_{\text{weiblich}}^2$	$\Theta_{\text{männlich}} = \Theta_{\text{weiblich}}$
Alternativhypothese	$H_A^1$	Es soll ausschließlich betrachtet werden, ob der zu vergleichende Parameter in einer Grundgesamtheit größer als in der anderen ist. $\mu_{\text{männlich}} > \mu_{\text{weiblich}}$	$\sigma_{\text{männlich}}^2 > \sigma_{\text{weiblich}}^2$	$\Theta_{\text{männlich}} > \Theta_{\text{weiblich}}$
	$H_A^2$	Es soll ausschließlich betrachtet werden, ob der zu vergleichende Parameter in einer Grundgesamtheit kleiner als in der anderen ist. $\mu_{\text{männlich}} < \mu_{\text{weiblich}}$	$\sigma_{\text{männlich}}^2 < \sigma_{\text{weiblich}}^2$	$\Theta_{\text{männlich}} < \Theta_{\text{weiblich}}$
	$H_A^3$	Der zu vergleichende Parameter unterscheidet sich in den Grundgesamtheiten: er kann größer oder kleiner sein. $\mu_{\text{männlich}} \neq \mu_{\text{weiblich}}$	$\sigma_{\text{männlich}}^2 \neq \sigma_{\text{weiblich}}^2$	$\Theta_{\text{männlich}} \neq \Theta_{\text{weiblich}}$

Sollte man allgemein erwarten können, dass das Gewicht der männlichen Tiere (fast immer) größer als das der weiblichen ist, dann könnte das Hypothesenpaar

$$H_0: \mu_{\text{männlich}} = \mu_{\text{weiblich}}$$

$$H_A: \mu_{\text{männlich}} > \mu_{\text{weiblich}}$$

zugrunde gelegt werden. Dafür gibt es wohl keine fachliche Begründung. Ebenso kann man auch nicht davon ausgehen, dass das Gewicht der männlichen Tiere kleiner als das der weiblichen ist.

Folglich kann das zu testende Hypothesenpaar nur lauten:

$$H_0^\mu: \mu_{\text{männlich}} = \mu_{\text{weiblich}}$$

$$H_A^\mu : \mu_{\text{männlich}} \neq \mu_{\text{weiblich}} .$$

Für den Parameter Varianz kann ebenso wenig davon ausgegangen werden, dass die Variabilität eines Geschlechts größer als die des anderen ist. Es kommt also auch nur die zweiseitige Fragestellung mit dem Hypothesenpaar

$$H_0^{\sigma^2} : \sigma_{\text{männlich}}^2 = \sigma_{\text{weiblich}}^2$$

$$H_A^{\sigma^2} : \sigma_{\text{männlich}}^2 \neq \sigma_{\text{weiblich}}^2$$

in Frage.

## 2.2 t-Test zweier unabhängiger Stichproben

### 2.2.1 Die Voraussetzungen des t-Tests

Der t-Test hat bekanntlich die Voraussetzungen

- gleiche Varianzen in den Grundgesamtheiten und
- Normalverteilung unter der Nullhypothese.

Es ist robust gegenüber Abweichungen von der Normalverteilung (RASCH u. TIKU 1985).

### 2.2.2 Bemerkungen zur Normalverteilttheit

Das Gewicht der Mäuse in g ist ein stetiges, metrisches Merkmal. Die Boxplots und die (orientierenden) Histogramme (Ausgabe 9.1) lassen eine leichte Asymmetrie erkennen. Der Shapiro-Wilk-Test lässt für jedes Geschlecht die Annahme der Nullhypothese nicht ablehnen (egal, ob zum Signifikanzniveau  $\alpha \approx 0,20$  (Abweichungen frühzeitig erkennen) oder  $\alpha < 0,05$  (Annahme der Nullhypothese lange beibehalten)).

Die schwächste der beiden Voraussetzungen, die Normalverteilungsannahme, wird am liebsten überprüft - wenn überhaupt. Sehr selten dagegen wird auf die Varianzhomogenität geschaut. Das Mindeste ist ein Blick auf die Variabilität in beiden Gruppen (s. Tab. 7):

	männlich	weiblich
s	3,74	3,08
s <sup>2</sup>	14,01	9,47

### 2.2.3 Der Levene-s-Test zur Überprüfung der Gleichheit der Varianzen

Riesige Unterschiede in den Variabilitäten (s.o.) sind das nicht. Das zeigen auch die Box-Whisker-Plots (Ausgabe 9.1). Das mag bei hinreichender Erfahrung für ein „Gefühl“ reichen, ein Test ist das nicht. Manch einer kennt als Test zum Überprüfen der Gleichheit der beiden Varianzen den F-Test als Homogenitätstest. Dieser F-Test ist nicht robust gegenüber Abweichungen von der Normalverteilung (NÜRNBERG 1982). Auch wenn SAS ihn standardmäßig ausgibt: Er sollte keinesfalls eingesetzt werden! Auf LEVENE (1960) basierend gibt es mehrere robuste Tests (s. SUMPFF u. MOLL 2004, S. 86-87, SCHUMACHER 2004, S. 22-23).

Im Programm 10 wird der Levene-s-Test mit der quadratische Transformation  $z_{ij} = (x_{ij} - \bar{x}_{i\bullet})^2$  ( $i = \text{männlich, weiblich}; j = 1, 2, \dots, 13$ ) vorgestellt.

**Programm 10:** Levene.sas

```

* Levene-s-Test zum Vergleich der Varianzen;
DATA levene1;
  SET MausGewicht;
  Gewicht2 = Gewicht*Gewicht;
  IF Geschlecht = 'männlich' then do;
    mSUM + Gewicht;
    mSUM2 + Gewicht2;
  END;
  IF Geschlecht = 'weiblich' then do;
    wSUM + Gewicht;
    wSUM2 + Gewicht2;
  END;
RUN;

DATA _NULL_;
  SET levene1;
  CALL SYMPUT ('mSUM' , mSUM);
  CALL SYMPUT ('mSUM2' , mSUM2);
  CALL SYMPUT ('wSUM' , wSUM);
  CALL SYMPUT ('wSUM2' , wSUM2);
RUN;

DATA _NULL_;
  SET levene1 (where=(Geschlecht = 'männlich'));
  n = _n_;
  CALL SYMPUT ('mn' , n);
DATA _NULL_;
  SET levene1 (where=(Geschlecht = 'weiblich'));
  n = _n_;
  CALL SYMPUT ('wn' , n);
RUN;

DATA levene_s;
  W = (((&mSUM*&mSUM/&mn + &wSUM*&wSUM/&wn -
        (&mSUM+&wSUM) * (&mSUM+&wSUM) / (&mn+&wn)) * (&mn+&wn-2))
        / (&mSUM2+&wSUM2 - (&mSUM*&mSUM/&mn + &wSUM*&wSUM/&wn)));
  FWert = FINV(0.95,1,&mn+&wn-2);
  IF W < FWert
    THEN TestEntscheidung = ' n.s. : t-Test ';
    ELSE TestEntscheidung = 'signifikant : Welch-Test';
RUN;

TITLE 'Levene-s-Test zum Vergleich der Varianzen';
PROC PRINT NOOBS;
RUN;
TITLE ;

```

Aufbauend auf die SAS-Datei MausGewicht (Programm 9) werden die Gewichte und deren Quadrate geschlechtsabhängig aufsummiert.

Die Summen werden auf Macrovariable gelegt.

Die Anzahlen in den beiden Gruppen werden auf Macrovariable gelegt.

Der Wert W für die Teststatistik **W** wird berechnet, ebenso der F-Wert. (s. SUMPFF u. MOLL 2004, S. 86)

Die Testentscheidung basiert auf dem Vergleich von W mit  $F_{1-\alpha; 1, n_1+n_2-2}$ .

Das Ergebnis steht in Ausgabe 10.

**Ausgabe 10:** zu Levene.sas

```

Levene-s-Test zum Vergleich der Varianzen

      W           FWert      TestEntscheidung
-----
2.13022      4.25968      n.s. : t-Test

```

Das Quantil der F-Verteilung ist größer als der berechnete Wert W der Teststatistik. Folglich kann die Nullhypothese nicht abgelehnt werden. Mit Hilfe des Levene-Tests wird die Gleichheit der Varianzen getestet, also das Hypothesenpaar

$$H_0^{\sigma^2} : \sigma_{\text{männlich}}^2 = \sigma_{\text{weiblich}}^2$$

$$H_A^{\sigma^2} : \sigma_{\text{männlich}}^2 \neq \sigma_{\text{weiblich}}^2$$

Damit ist Antwort auf eine der möglichen Aufgabenstellungen gegeben und gleichzeitig die Voraussetzung für den t-Test zweier unabhängiger Stichproben überprüft worden.

### 2.2.4 Die SAS-Prozedur TTEST

Kann von gleichen Varianzen in den Grundgesamtheiten ausgegangen werden, ist die wichtigste Voraussetzung für den Vergleich zweier Mittelwerte erfüllt. Sollte diese Voraussetzung nicht erfüllt sein, wird der Welch-Test (WELCH 1947), t-Test für ungleiche Varianzen, empfohlen. Diese Entscheidung ist im Programm 10 bereits verankert. Der t- bzw. Welch-Test wird mit Hilfe des Programms 11 durchgeführt. Die Ergebnisse stehen in Ausgabe 11.

**Programm 11:** tTest.sas

```

PROC TTEST DATA= MausGewicht;
  CLASS Geschlecht;
  VAR Gewicht;
RUN;
    
```

Klassifizierungsvariable zur Kennzeichnung  
beider Gruppen ist notwendig

**Ausgabe 11:** zu tTest.sas

```

The TTEST Procedure
    
```

		Statistics				
Variable	Geschlecht	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev
Gewicht	männlich	13	18.623	20.885	23.146	2.6837
Gewicht	weiblich	13	20.986	22.846	24.706	2.2072
Gewicht	Diff (1-2)		-4.735	-1.962	0.8122	2.6754

		Statistics				
Variable	Geschlecht	Std Dev	Upper CL Std Dev	Std Err	Minimum	Maximum
Gewicht	männlich	3.7425	6.1779	1.038	16	28.5
Gewicht	weiblich	3.078	5.081	0.8537	18	28
Gewicht	Diff (1-2)	3.4264	4.7667	1.344		

		T-Tests			
Variable	Method	Variances	DF	t Value	Pr >  t
Gewicht	Pooled	Equal	24	-1.46	0.1574
Gewicht	Satterthwaite	Unequal	23.1	-1.46	0.1579

		Equality of Variances			
Variable	Method	Num DF	Den DF	F Value	Pr > F
Gewicht	Folded F	12	12	1.48	0.5085

Im ersten Teil der Ausgabe 11 finden wir einige statistische Maßzahlen:

Anzahl, Mittelwert mit 0,95-Konfidenzintervall, Standardabweichung mit 0,95-Konfidenzintervall, Standardfehler sowie Minimum und Maximum

in der Regel für die beiden Gruppen männlich und weiblich getrennt, aber auch einige für die Differenz aus dem i-ten Gewicht der männlichen und dem i-ten Gewicht (i = 1, 2, ... 13) der weiblichen Tiere.

Unter der Überschrift T-Tests finden wir sowohl das Ergebnis des t-Tests als auch das des Welch-Tests. Was welcher Test ist, wird durch die Methode gekennzeichnet. ‚Pooled‘ weist auf die Berechnung einer gemeinsamen Varianz (bei gleichen Varianzen) und damit auf den t-Test und ‚Satterthwaite‘ auf die auf SATTERTHWAITTE (1946) zurück gehende Wichtung der Freiheitsgrade mit den unterschiedlichen Varianzen und damit auf den Welch-Test hin. Sie sind im Allgemeinen reell. Bei vorher festgelegtem Signifikanzniveau von  $\alpha = 0,05$  und zweiseitiger Fragestellung (das zu testende Hypothesenpaar lautet:

$$H_0^{\sigma^2} : \sigma_{\text{männlich}}^2 = \sigma_{\text{weiblich}}^2 \text{ vs. } H_A^{\sigma^2} : \sigma_{\text{männlich}}^2 \neq \sigma_{\text{weiblich}}^2 )$$

kann die Nullhypothese zum Niveau  $\alpha = 0,05$  nicht abgelehnt werden, weil die berechnete Überschreitungswahrscheinlichkeit größer als das Signifikanzniveau ist:  $(Pr>|t|) = 0,1574 > \alpha = 0,05$ .

Die letzte Information der Ausgabe 11 bezieht sich auf den F-Test. Der soll als Test auf Gleichheit der Varianzen nicht herangezogen werden (s.o.). Er wird deshalb auch nicht weiter beachtet, zumal wir uns auf die Ergebnisse des robusten Levene-s-Tests berufen können.

### 2.2.5 Bemerkungen zur Testentscheidung und grafische Interpretation

Die Testentscheidung kann entweder durch den Vergleich der Überschreitungswahrscheinlichkeit mit dem vorgegebenen Signifikanzniveau (s.o.) oder durch den Vergleich des berechneten t-Werts der Teststatistik mit dem t-Quantil zur Wahrscheinlichkeit  $1-\alpha/2$  (zweiseitige Fragestellung!) erfolgen. Das t-Quantil  $t_{1-\alpha/2; n_1+n_2-2}$  hat den Wert  $t_{0,975; 24} = 2,064$ . Es ist größer als der Absolutwert des berechneten t-Wert.

Die Testsituation unter der Nullhypothese zeigt die Abb. 9 recht gut. Die Einheiten auf der Abszissenachse sind Einheiten der standardisierten Standardabweichung, der Mittelwert liegt bei Null. Der berechnete t-Wert liegt mit  $t = -1,46$  innerhalb des Bereichs der Nullhypothese und nicht im farblich ausgefüllten Bereich außerhalb der Nullhypothese, der durch das vorgegebene Signifikanzniveau  $\alpha$  bestimmt wird. Die beidseitige Aufteilung des Signifikanzniveaus aufgrund der zweiseitigen Fragestellung wird deutlich.

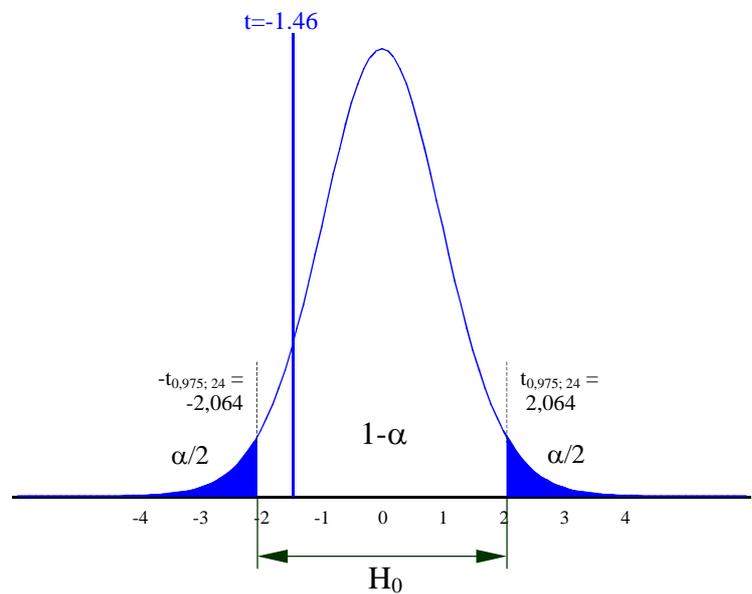


Abb. 9: Veranschaulichung der Testsituation zum Vergleich der mittleren Mäusegewichte (zweiseitige Fragestellung)

Folglich kann die Testaussage nur sein, die Nullhypothese nicht abzulehnen.

### 2.2.6 Durchführung des t-Tests mit dem Analyst

Der SAS/Analyst bietet schnelle Möglichkeiten, den t-Test durchzuführen und Grafiken zu erstellen. Entweder wird er über den Aufruf `analyst` im Command-Fenster oder durch

Solutions → Analysis ▶ → Analyst  
aktiviert.

Die Daten liegen in der notwendigen Struktur (Klassifikationsvariable und Variable für das auszuwertende Merkmal) in der Datei WORK.MausGewicht. Diese Datei wird über

File → Open By SAS Name  
angesprochen und deren Inhalt wird in der Tabellenstruktur des Analyst sichtbar.

Mit

Statistics → Hypothesis Tests ▶ → Two-Sample t-test for Means ...  
wird der t-Test angesprochen.

Die Zuweisung der Variablen (Depend: Gewicht, Group: Geschlecht) und die Auswahl der Alternativhypothese (vgl. Tab. 8) kann der Abb. 10 entnommen werden.

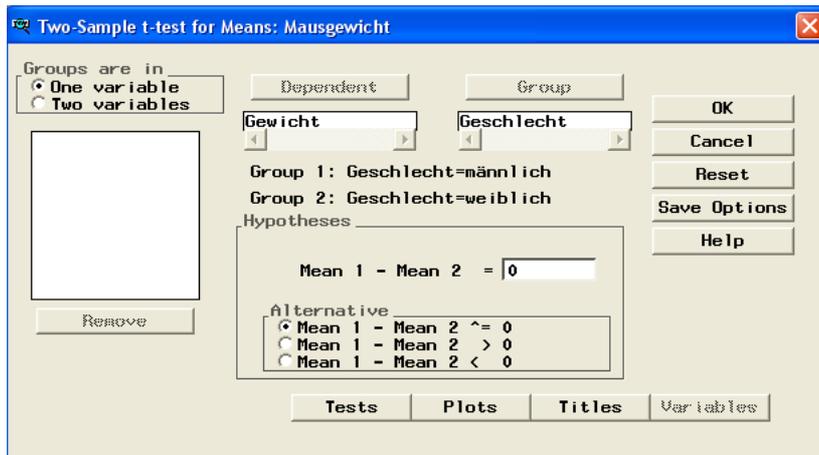


Abb. 10: Eingaben für den t-Test

Der Schalter ‚Tests‘ gestattet beispielsweise die zusätzliche Berechnung eines Konfidenzintervalls für die Differenz der beiden Mittelwerte. Der Schalter ‚Plots‘ liefert wahlweise

- Box-Whisker-Plots,
- Balkendiagramme,
- Mittelwertdarstellungen und
- Darstellung der t-Verteilung.

Die letzte Darstellungsform richtet sich nach der Wahl der Alternativhypothese und entspricht bei zweiseitiger Fragestellung der Abb. 9.

Die Ergebnisse entsprechen denen der Ausgabe 11 (ohne die des F-Tests). Wurde beispielsweise zusätzlich die Berechnung des Konfidenzintervalls gewählt, erweitert sich die Ausgabe um:

95% Confidence Interval for the Difference between Two Means

Lower Limit	Upper Limit
----- -4.74	----- 0.81

Erkennbar ist, dass das berechnete 0,95-Konfidenzintervall der Differenz der Mittelwerte die Null einschließt. Diese Aussage kann auch als Testaussage für die Gleichheit der Mittelwerte herangezogen werden: Die Nullhypothese kann zum Signifikanzniveau  $\alpha = 0,05$  nicht abgelehnt werden.

## 2.2.7 Bemerkungen zur Versuchsplanung mit dem Analyst

Der SAS/Analyst bietet auch Berechnungen von Stichprobenumfängen und zum Risiko 2. Art an. Das erfolgt für den t-Test über: Statistics → Sample Size ► → Two-Sample t-test ...

Bei der Versuchsplanung für einen Test sind zwei Wahrscheinlichkeiten zu berücksichtigen: das Signifikanzniveau  $\alpha$  oder das Risiko 1. Art und das Risiko 2. Art  $\beta$ . Die Differenz  $1-\beta$  heißt Güte, engl. Power. Die Testentscheidung, die Entscheidung für die Nullhypothese  $H_0$  oder gegen sie (für die Alternativhypothese  $H_A$ ), wird zum Signifikanzniveau  $\alpha$  vorgenommen. Folglich unterscheiden sich geplante und ungeplante Versuche darin, dass geplante Versuche durch die Anzahl Wiederholungen beide Risiken berücksichtigen. Bei einem ungeplanten Versuch passiert es sehr häufig, dass aufgrund der zu geringen Wiederholung keine Unterschiede zwischen den Mittelwerten der Varianten erkannt werden können, weil die im Versuch erkennbare Differenz viel zu groß ist.

Bekanntlich ist das Risiko 1. Art  $\alpha$  die Wahrscheinlichkeit dafür, die (in der Grundgesamtheit) richtige Nullhypothese abzulehnen, d.h. eine falsche Entscheidung für Signifikanz zu treffen, und das Risiko 2. Art  $\beta$  die Wahrscheinlichkeit dafür, eine (in der Grundgesamtheit) nicht zutreffende Nullhypothese anzunehmen, d.h. signifikante Differenzen nicht zu erkennen.

Für die Daten des Beispiels der Mäusegewichte werden die Gruppenmittelwerte, die (gemeinsame) Standardabweichung (s. Tab 5a) und das Signifikanzniveau (Risiko 1. Art)  $\alpha = 0.05$  eingetragen. Von Vorteil ist es, sich für verschiedene Vorgaben der Güte  $1-\beta$  (Power) die Stichprobenumfänge berechnen zu lassen und das auch grafisch darzustellen (Tab. 9).

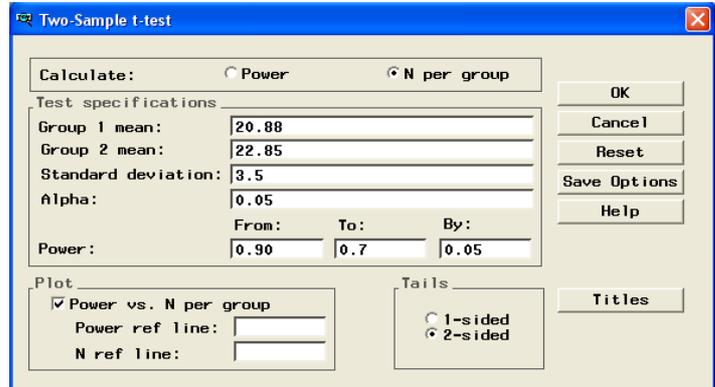
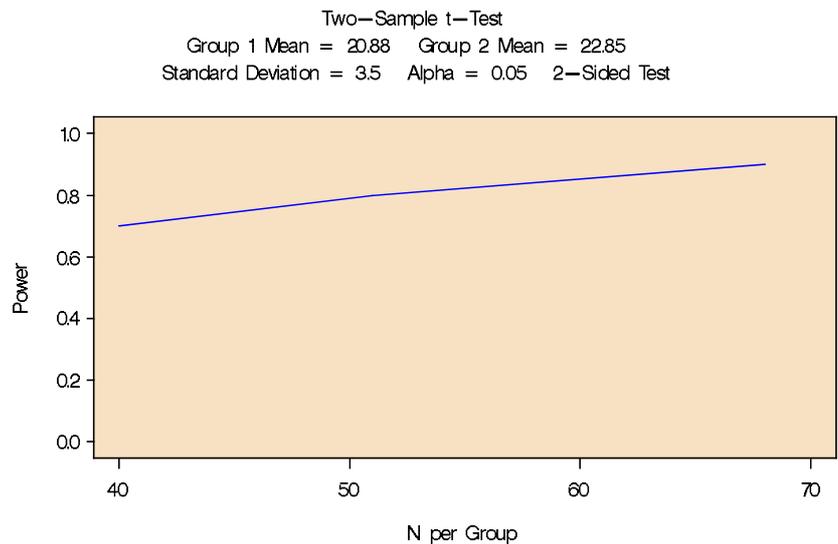


Abb. 11: Eingaben für Stichprobenumfangsberechnungen

Tab. 9: Mindeststichprobenumfang in Abhängigkeit von der Güte  $1-\beta$

Group 1 Mean = 20.88  
 Group 2 Mean = 22.85  
 Standard Deviation = 3.5  
 Alpha = 0.05 2-Sided  
 Test

Power	N per Group
0.700	40
0.750	45
0.800	51
0.850	58
0.900	68



Um einen Unterschied zwischen den beiden Mittelwerten von  $22,85 \text{ g} - 20,88 \text{ g} = 1,97 \text{ g}$  bei einer Standardabweichung von  $3,5 \text{ g}$  und einem Risiko 1. Art  $\alpha = 0,05$  braucht man für einen Wiederholungsversuch pro Gruppe

- bei einem Risiko 2. Art  $\beta = 0,25$  ( $1-\beta = 0,75$ ) mindestens 45 Tiere
- bei einem Risiko 2. Art  $\beta = 0,20$  ( $1-\beta = 0,80$ ) mindestens 51 Tiere.

Wird als Zielgröße nicht der Stichprobenumfang je Gruppe (Abb. 11), sondern die Güte (Power)  $1 - \beta$  gewählt, ist das Ergebnis für  $n = 13$  in jeder der beiden Gruppen

Group 1 Mean	Group 2 Mean
20.88	22.85
Standard Deviation = 3.5	Alpha = 0.05 2-Sided Test
N per Group	Power
13	0.280

Eine Power  $1 - \beta = 0,28$  bedeutet, dass das Risiko 2. Art  $\beta = 0,72$  ist. Mit einem Risiko 2. Art von 72 % ist ein Versuch untauglich, eine Differenz von 1,97 g zwischen den Mittelwerten nachweisen zu wollen; der Stichprobenumfang ist zu gering.

Nun bietet es sich natürlich an, mit den Möglichkeiten des Analyst zu „spielen“. Trauen Sie sich!

## 2.3 t-Test zweier abhängiger Stichproben

### 2.3.1 Aufgabenstellung und Daten

Wir nehmen einmal an, dass 13 Mäuse aus einer Population zufällig ausgewählt und zu einem bestimmten Zeitpunkt ihr Gewicht gewogen wurde. Diese Wägung wurde nach wenigen Tagen (das ist natürlich genau festgelegt, z.B. 4 Tage) bei gleich bleibender Ernährung und sonstigen Haltungsbedingungen für jedes der Tiere wiederholt. Die Versuchsfrage könnte sein, ob es nach diesen 4 Tagen einen signifikanten Gewichtszuwachs gibt. Diese Aufgabe zeigt eine Richtung an, nämlich Zuwachs. Wir haben es mit einer einseitigen Fragestellung zu tun. Wesentlich ist, dass die Messungen am gleichen Tier erfolgen (oder am gleichen Paar, wobei ein Tier desselben Pairs zufällig der einen und das andere der anderen Gruppe zugeordnet wird). Wir haben es folglich nicht mehr mit unabhängigen Stichproben zu tun, sondern mit abhängigen oder gepaarten. Die Formulierung der Versuchsfrage zeigt bereits den Weg: Bilden der Differenzen.

### 2.3.2 Rückführung auf das Einstichprobenproblem

Zwischen den Messungen am selben Tier oder des gleichen Pairs werden die Differenzen gebildet. Das Ergebnis ist eine Wertereihe aus Differenzen. Es werden nicht mehr zwei Stichproben, sondern eine betrachtet. Gehen wir der Einfachheit halber von denselben Zahlen aus, ergibt sich folgender Stand:

Tier	1	2	3	4	5	6	7	8	9	10	11	12	13
erste Wägung	26,0	16,5	20,0	17,0	18,0	16,0	28,5	21,0	23,5	23,0	20,0	19,5	22,5
zweite Wägung	18,0	24,0	18,5	24,0	20,0	25,0	28,0	22,0	27,5	24,0	20,5	22,5	23,0
Differenz	-8,0	7,5	-1,5	7,0	2,0	9,0	-0,5	1,0	4,0	1,0	0,5	3,0	0,5

Die Datenerfassung entspräche der der Abb. 7, so dass wir im Programm 12 von der SAS-Datei *Maus* (mit den Variablen F1 und F2) ausgehen können.

#### Programm 12: tTest\_.sas

```

DATA Differenz;
  SET Maus;
  Differenz = F2 - F1;
RUN;
PROC TTEST DATA=Differenz;
  VAR Differenz;
RUN;
    
```

keine Klassifizierungsvariable

#### Ausgabe 12: zu tTest\_.sas

```

The TTEST Procedure

                               Statistics
Variable      N      Lower CL      Mean      Upper CL      Lower CL      Std Dev      Upper CL
                Mean      Mean      Mean      Std Dev      Std Dev      Std Dev
Differenz     13      -0.708  1.9615     4.6313     3.1681     4.4181     7.293
    
```

Statistics			
Variable	Std Err	Minimum	Maximum
Differenz	1.2253	-8	9

T-Tests			
Variable	DF	t Value	Pr >  t
Differenz	12	1.60	0.1354

Nach einigen statistischen Maßzahlen mit Konfidenzschätzung folgen die Testergebnisse. Für die Testentscheidung ist zu beachten, dass wir eine einseitige Fragestellung mit dem zu testenden Hypothesenpaar haben:

$$H_0: \mu_1. \text{ Wägung} = \mu_2. \text{ Wägung} \quad (\delta = 0)$$

$$H_A: \mu_1. \text{ Wägung} < \mu_2. \text{ Wägung} \Rightarrow \mu_2. \text{ Wägung} - \mu_1. \text{ Wägung} = \delta > 0 .$$

Folglich ist bei festgelegtem Signifikanzniveau  $\alpha = 0,05$  bei einseitiger Fragestellung entweder der berechnete t-Wert 1,60 mit dem t-Quantil  $t_{1-\alpha; n-1} = t_{0,95; 12} = 1,782$  [ $t_{0,95; 12} = 1,782 > 1,60$ ] oder die halbe berechnete Überschreitungswahrscheinlichkeit mit dem Signifikanzniveau  $[(Pr>|t|)/2 = 0,1354/2 > \alpha = 0,05]$  zu vergleichen. Die Testentscheidung lautet, dass die Nullhypothese nicht abgelehnt werden kann. Der mittlere Gewichtszuwachs von 1,96 g (s. Ausgabe 12) liegt bei dieser Wiederholungsanzahl und der geschätzten Streuung im Zufallsbereich. Das geschätzte 0,95-Konfidenzintervall aus Ausgabe 12 ist ein zweiseitiges und kann somit nicht für die Testentscheidung der einseitigen Fragestellung herangezogen werden.

## 2.4 Zusammenfassung

Eine Entscheidung für welchen Test geht einher mit der Formulierung des Hypothesenpaars aus Null- und Alternativhypothese über die zu testenden Parameter. Der t-Test zum Vergleich zweier Mittelwerte unabhängiger Stichproben geht von der Gleichheit der Varianzen beider Grundgesamtheiten aus. Die sollte mit Hilfe des Levene-Tests oder dessen Modifikationen überprüft werden, nicht mit dem F-Test. Bei ungleichen Varianzen kann der Welch-Test verwendet werden. Beide Tests werden mit der SAS-Prozedur *TTEST* realisiert. Abhängige Stichproben werden durch Differenzenbildung auf das Einstichprobenproblem zurück geführt. Auch hier ist die SAS-Prozedur *TTEST* – allerdings ohne die *CLASS*-Anweisung.

Mit Hilfe des SAS-Analyst kann neben der Durchführung des t-Tests unter anderem auch das Ergebnis grafisch veranschaulicht werden. Stichprobenumfänge und Güte  $1-\beta$  können berechnet werden.

### 3 Gruppierete Diagramme für mehr als zwei Stichproben

#### 3.1 Beispiel Stammhöhen von Eichen in den Jahren 2001 und 2005

Als Beispiel werden aus einem größeren Merkmalspool eines Nachkommenschaftsversuchs mit Eichen nur die mittleren Stammhöhen in den Jahren 2001 und 2005 (H01 und H05) betrachtet:

Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05
6	1	1	482,8	701,8	6	21	2	748,5	7	9	1	325,2	508,2	7	24	3	256,5	433,9	8	9	4	240,7	574,3	
6	1	2	743,3		6	21	3	807,8	7	9	2			7	24	4	237,1	366,0	8	11	1	259,9	542,3	
6	1	3	753,2		6	23	1	541,5	838,0	7	9	3	292,6	513,0	7	26	1	282,5	517,9	8	11	2	285,2	593,1
6	2	1	525,1	770,8	6	23	2	832,8	7	9	4	238,3	372,3	7	26	2		443,9	8	11	3	263,9	553,7	
6	2	2	837,6		6	23	3	948,1	7	11	1	308,8	529,6	7	26	3	282,2	458,4	8	11	4	315,7	638,3	
6	2	3	917,4		6	24	1	508,9	740,0	7	11	2		509,2	7	26	4	250,9	389,4	8	13	1	223,3	512,6
6	3	1	549,6	788,1	6	24	2	803,4	7	11	3	223,0	426,2	7	28	1	302,9	555,6	8	13	2	191,6	439,6	
6	3	2	812,0		6	24	3	872,6	7	11	4	264,6	391,4	7	28	2		530,9	8	13	3	260,1	557,9	
6	3	3	933,2		6	26	1	570,0	839,2	7	13	1	269,1	454,3	7	28	3			8	13	4	286,2	614,5
6	6	1	496,5	802,0	6	26	2	786,2	7	13	2		543,6	7	30	2		453,6	8	14	1	240,9	583,7	
6	6	2	884,7		6	26	3	539,5	716,6	7	13	3	325,1	530,4	7	36	1	320,6	536,9	8	14	2	210,3	529,6
6	6	3	898,2		6	28	1	564,8	880,3	7	13	4	258,9	447,1	7	36	2		563,1	8	14	3	299,1	605,6
6	7	1	609,7	866,1	6	28	2	835,7	7	14	1	290,7	553,6	7	36	3	313,0	522,3	8	14	4	276,7	605,3	
6	7	2	778,0		6	28	3	817,2	7	14	2	263,4	438,3	7	36	4	249,8	356,6	8	15	1	221,8	440,2	
6	7	3	748,0		6	30	1	586,4	900,0	7	14	3			7	37	1	334,6	548,4	8	15	2	247,8	513,2
6	8	1	604,1	794,4	6	30	2	762,1	7	14	4	211,6	315,5	7	37	2		537,8	8	15	3	262,6	539,6	
6	8	2	811,2		6	30	3	559,0	777,8	7	15	1	329,1	577,4	7	37	3	358,7	590,2	8	15	4	262,5	594,5
6	8	3	905,3		6	36	1	497,7	725,3	7	15	3	209,4	379,3	7	37	4	241,2	365,2	8	16	1	260,7	547,6
6	9	1	551,3	818,9	6	36	2	950,7	7	15	4	235,3	361,2	7	38	1	300,2	508,8	8	16	2	213,8	528,4	
6	9	2	787,9		6	36	3	679,3	946,9	7	16	1	294,0	495,1	7	38	2		521,3	8	16	3	308,1	585,9
6	9	3	626,3	917,3	6	37	1	556,7	824,3	7	16	2			7	38	3	325,2	570,7	8	16	4	240,8	541,9
6	11	1	578,9	882,1	6	37	2	897,0	7	16	3	293,9	517,9	7	38	4	252,7	418,6	8	17	1	232,7	519,1	
6	11	2	789,1		6	37	3	940,6	7	16	4	251,3	374,9	8	1	1	270,6	543,6	8	17	2	222,6	480,9	
6	11	3	909,7		6	38	1	581,7	923,3	7	17	1	234,6	417,3	8	1	2	247,6	531,3	8	17	3	239,5	518,9
6	13	1	549,6	872,3	6	38	2	857,3	7	17	2		366,0	8	1	3	270,1	613,1	8	17	4	244,4	538,9	
6	13	2	854,4		6	38	3	901,0	7	17	3	260,9	440,8	8	1	4	249,1	551,6	8	18	1	192,9	454,4	
6	13	3	868,1		7	1	1	285,5	544,3	7	17	4	271,3	451,8	8	2	1	232,7	494,7	8	18	2	241,5	543,1
6	14	1	573,1	895,6	7	1	2	471,3	7	18	1	273,6	481,9	8	2	2	193,9	465,0	8	18	3	263,9	550,8	
6	14	2	839,7		7	1	3	229,2	390,2	7	18	2		474,6	8	2	3	199,9	475,8	8	18	4	202,9	469,5
6	14	3	788,7		7	1	4	273,3	468,5	7	18	3	269,8	478,2	8	2	4	244,3	568,5	8	19	1	182,0	454,2
6	15	1	627,9	887,4	7	2	1	291,9	471,5	7	18	4	231,8	372,9	8	3	1	312,9	627,3	8	19	2	227,9	549,5
6	15	2	810,9		7	2	2	418,5	7	19	1	302,3	499,8	8	3	2	172,9	513,9	8	19	3	172,5	447,5	
6	15	3	883,8		7	2	3	255,4	422,2	7	19	2	262,4	394,2	8	3	3	184,6	448,5	8	19	4	259,5	614,3
6	16	1	570,8	867,2	7	2	4	206,7	324,9	7	19	3	261,0	417,3	8	3	4	300,4	607,6	8	20	1	195,2	419,4
6	16	2	652,7		7	3	1	352,7	546,5	7	19	4	212,5	288,0	8	6	1	230,9	539,2	8	20	2	189,2	434,8
6	16	3	904,4		7	3	2			7	20	1	277,7	463,4	8	6	2	281,3	606,3	8	20	3	207,4	504,3
6	17	1	584,1	845,3	7	3	3	325,1	550,4	7	20	2			8	6	3	333,2	634,3	8	20	4	263,9	539,0
6	17	2	817,6		7	3	4	235,5	350,0	7	20	3			8	6	4	295,5	583,7	8	21	1	140,0	459,4
6	17	3	865,9		7	6	1	344,9	557,1	7	20	4	211,6	287,6	8	7	1	188,6	496,8	8	21	2	218,0	509,5
6	18	1	522,2	755,7	7	6	3	360,0	621,1	7	21	1	342,4	607,6	8	7	2	212,1	505,0	8	21	3	235,5	576,7
6	18	2	785,6		7	6	4	297,2	452,5	7	21	2		380,8	8	7	3	174,3	448,9	8	21	4	185,3	489,7
6	18	3	885,5		7	7	1	273,9	518,3	7	21	3	265,6	387,6	8	7	4	252,5	590,2	8	23	1	185,9	373,0
6	19	1	470,7	800,0	7	7	2	422,1	7	21	4	224,2	349,1	8	8	1	197,7	418,2	8	23	2	246,2	529,8	
6	19	2	891,1		7	7	3	309,1	521,7	7	23	1	307,9	490,8	8	8	2	290,2	568,3	8	23	3	259,6	557,9
6	19	3	873,5		7	7	4	205,5	306,9	7	23	2		509,6	8	8	3	261,9	550,9	8	23	4	201,4	506,3
6	20	1	530,8	770,7	7	8	1	265,3	491,4	7	23	3	283,5	479,6	8	8	4	272,7	557,5	8	24	1	188,7	501,0
6	20	2	673,6		7	8	2	475,0	7	23	4	267,8	351,3	8	9	1	236,4	584,8	8	24	2	239,0	518,2	
6	20	3	864,2		7	8	3	223,7	386,7	7	24	1	301,8	498,0	8	9	2	307,6	587,0	8	24	3	278,4	545,6
6	21	1	567,5	893,9	7	8	4	270,7	444,0	7	24	2		585,0	8	9	3	253,1	554,2	8	24	4	251,3	632,5

Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05	Ort	Herkunft	Wiederholung	H01	H05
8	26	1	186,7	378,2	8	38	1	226,6	467,7	9	8	2	392,1	680,3	9	17	2	309,2	494,0	9	26	1	499,8	751,8
8	26	2	281,0	567,0	8	38	2	304,3	606,3	9	9	1	499,7	785,0	9	17	3	392,3	609,9	9	26	2	500,2	792,3
8	26	3	228,8	525,2	8	38	3	350,7	676,8	9	9	2	346,6	538,3	9	18	1	425,1	717,7	9	26	3	511,1	800,4
8	26	4	224,0	550,9	8	38	4	331,6	620,1	9	9	3	528,5	843,4	9	18	2	449,3	712,0	9	28	1	495,4	778,2
8	28	1	266,6	575,6	9	1	1	464,6	668,9	9	11	1	374,8	650,3	9	18	3	414,0	634,4	9	28	2	538,5	844,4
8	28	2	231,3	533,6	9	1	2	487,5	744,8	9	11	2	512,4	771,4	9	19	1	527,6	802,8	9	28	3	472,6	708,3
8	28	3	292,8	571,1	9	1	3	511,9	796,3	9	11	3	438,2	647,1	9	19	2	234,5	451,1	9	30	1	549,1	814,4
8	28	4	276,5	586,7	9	2	1	405,0	687,4	9	13	1	496,3	795,1	9	19	3	438,8	787,1	9	30	2	386,9	634,5
8	30	1	241,3	543,7	9	2	2	367,3	569,6	9	13	2	412,0	640,1	9	20	1	475,2	771,8	9	30	3	536,0	781,0
8	30	2	315,8	590,6	9	2	3	498,9	703,6	9	13	3	609,2	913,3	9	20	2	426,1	663,5	9	36	1	548,2	820,6
8	30	3	274,6	611,2	9	3	1	417,9	675,4	9	14	1	520,7	826,6	9	20	3	439,8	608,0	9	36	2	322,5	543,8
8	30	4	283,1	578,3	9	3	2	443,9	748,1	9	14	2	446,8	720,0	9	21	1	449,7	698,1	9	36	3	613,8	940,9
8	36	1	282,0	560,1	9	3	3	559,8	820,9	9	14	3	493,7	760,7	9	21	2	473,2	719,0	9	37	1	429,1	741,4
8	36	2	276,5	520,0	9	6	1	477,0	773,6	9	15	1	456,4	675,7	9	21	3	494,9	780,8	9	37	2	509,8	793,4
8	36	3	288,3	624,4	9	6	2	332,3	532,4	9	15	2	444,6	703,5	9	23	1	512,9	710,6	9	37	3	627,6	925,8
8	36	4	270,7	526,0	9	6	3	570,7	938,5	9	15	3	558,3	846,5	9	23	2	459,4	749,0	9	38	1	504,2	781,0
8	37	1	265,2	552,6	9	7	1	473,8	760,0	9	16	1	517,4	813,8	9	23	3	508,3	826,1	9	38	2	595,0	884,2
8	37	2	231,9	573,8	9	7	2	405,0	625,8	9	16	2	502,1	774,6	9	24	1	486,3	816,1	9	38	3	454,9	733,3
8	37	3	230,3	563,4	9	7	3	355,9	548,8	9	16	3	601,2	900,0	9	24	2	414,4	676,0					
8	37	4	268,2	575,0	9	8	1	474,8	739,6	9	17	1	383,6	712,3	9	24	3	520,0	850,5					

Weil die auf einer bestimmten Fläche stehenden Bäume als zusammen gehörig angesehen werden, wird für weitere Betrachtungen das Merkmal mittlere Stammhöhe für dieses Areal verwendet. Eine Wiederholung ist dann ein weiteres Flächenstück. Für die Betrachtungen hier soll nur dieses Merkmal an verschiedenen Standorten ohne Beachtung der Herkünfte interessieren.

### 3.2 Boxplots und Punktediagramm (Scatterplot)

Zunächst verwenden wir eine bekannte Darstellungsform, die Boxplots. Wir gehen davon aus, dass die Daten in der SAS-Datei *Eiche* vorliegen. Wenn die 5 Spalten untereinander und nicht zusätzlich noch nebeneinander (s. Datenauflistung oben) in einer MS-Excel-Datei vorliegen, dann kann ihre Übernahme in eine SAS-Datei wie auf den Seiten 8 und 9 beschrieben erfolgen. Bei Punkt 4 wäre die Standardeinstellung (die in der ersten Zeile stehenden Bezeichnungen werden als SAS-Variablen übernommen) zu wählen. Das Programm 13 erzeugt die in Ausgabe 13 vorgestellten Boxplots.

Programm 13: Eichel.sas

```

PROC SORT DATA=Eiche;
  BY Ort;
RUN;

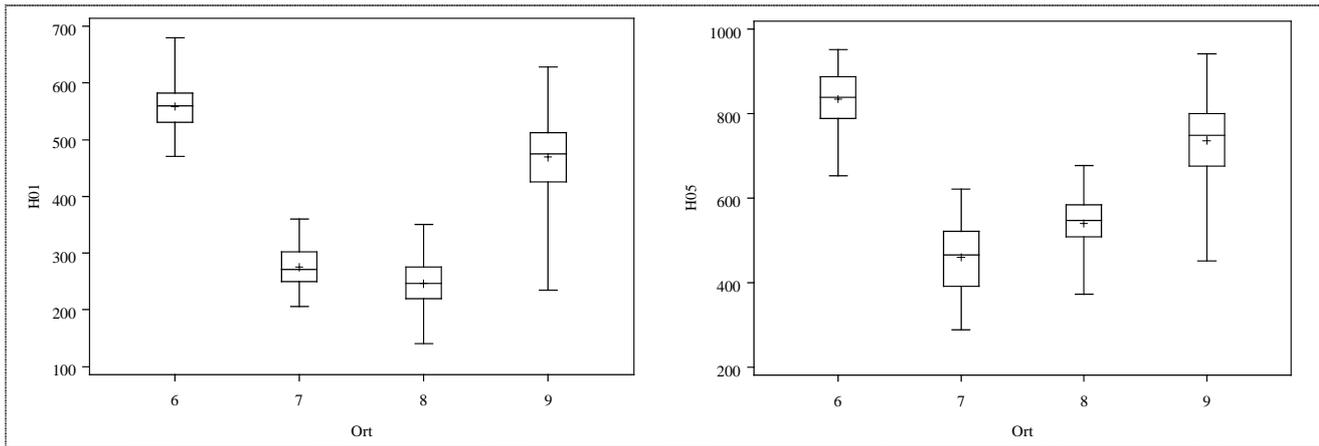
GOPTIONS
  Ftext="Times New Roman"
  Htext=1.2
  Device=emf;

PROC BOXPLOT DATA=Eiche;
  PLOT H01 *Ort /cboxes = black;
  PLOT H05 *Ort /cboxes = black;
RUN;
QUIT;

```

Es werden 2 getrennte Grafiken erzeugt.

**Ausgabe 13:** zu Eiche1.sas - Grafiken



Ein optischer Vergleich der Stammhöhen der beiden Jahre ist wegen der unterschiedlichen Achseneinteilungen nicht so einfach. Für eine nach Ort und Jahr gruppierte Darstellung innerhalb einer Grafik brauchen wir die Variablen *Hoehe*, die die Werte der Stammhöhen *H01* und *H05* enthält, und *Jahr*, eine Gruppierungsvariable für die Jahre. Das umfasst den ersten Teil des Programms 14. Die Boxplots (Ausgabe 14) der je Ort gruppierten Stammhöhen mit zusätzlich farblich markierten Mittelwerten erzeugt die Prozedur *BOXPLOT*.

**Programm 14:** Eiche2.sas

```

DATA GruppGraf;
  SET eiche;
  IF h01 ^= . THEN DO;
    Jahr = '2001';
    Hoehe= h01;
    OUTPUT;
  END;
  IF h05 ^= . THEN DO;
    Jahr = '2005';
    Hoehe= h05;
    OUTPUT;
  END;

PROC SORT DATA=GruppGraf;
  BY Ort Jahr;
RUN;

GOPTIONS
  Ftext="Times New Roman"
  Htext=1.2
  Device=emf;

SYMBOL1 h=.5 v=dot i=none c=black;
SYMBOL2 h=.5 v=dot i=none c=yellow;
SYMBOL3 h=.5 v=dot i=none c=green;
SYMBOL4 h=.5 v=dot i=none c=gray;

PROC BOXPLOT DATA=GruppGraf;
  LABEL Hoehe = 'Stammhöhe in cm';
  PLOT Hoehe * Jahr = Ort / cboxes = black;
RUN; QUIT;

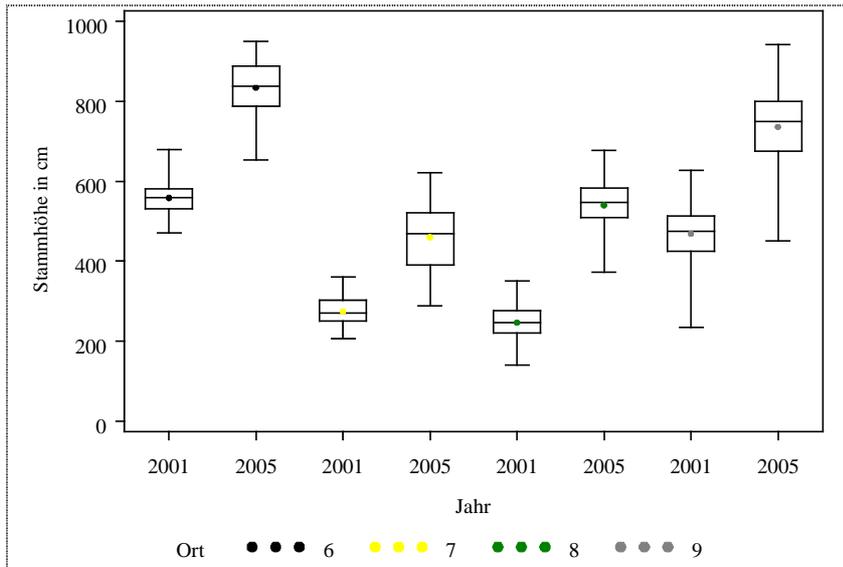
```

die Reihenfolge der Sortierungsvariablen, der BY-Variablen, ist wichtig

Symbole für die Mittelwerte: ausgefüllte Kreise halber Zellhöhe in unterschiedlichen Farben

Die ortsabhängige Struktur der beiden Merkmale *H01* und *H05* kann man auch in einem Punktediagramm (Scatterplot) veranschaulichen (Programm 15 und Ausgabe 15).

**Ausgabe 14:** zu Eiche2.sas - gruppierte Boxplots



**Programm 15:** Eiche3.sas

```

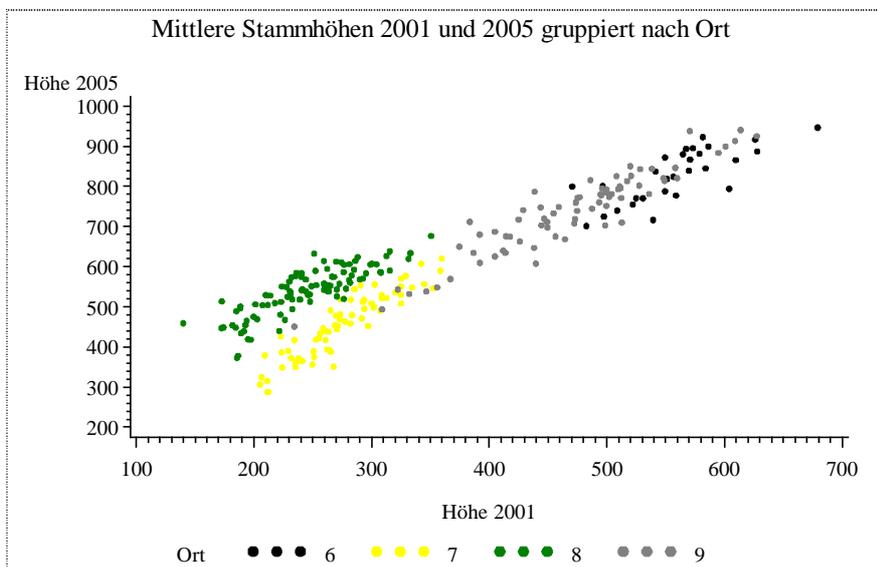
GOPTIONS RESET=all;
GOPTIONS
  Ftitle="Times New Roman"  Htitle=1.5
  Ftext="Times New Roman"   Htext=1.2
  Device=emf;
SYMBOL1 h=.5 v=dot i=none c=black;
SYMBOL2 h=.5 v=dot i=none c=yellow;
SYMBOL3 h=.5 v=dot i=none c=green;
SYMBOL4 h=.5 v=dot i=none c=gray;

TITLE "Mittlere Stammhöhen 2001 und 2005 gruppiert nach Ort";

PROC GGPLOT DATA=Eiche;
  LABEL h05='Höhe 2005' /* y-achse */
        h01='Höhe 2001'; /* x-achse */
  PLOT h05 * h01 =ort
        /noframe ;
RUN; QUIT;
TITLE;
    
```

zurücksetzen aller grafischen Einstellungen  
 grafische Einstellungen für Überschrift und  
 sonstige Beschriftung  
 Symbole: ausgefüllte Kreise halber Zellhöhe  
 in unterschiedlichen Farben

**Ausgabe 15:** zu Eiche3.sas - gruppiertes Scatterplot



Eine Zuordnung der Herkünfte innerhalb der Orte ist so weder in den gruppierten Boxplots noch in dem gruppierten Scatterplot möglich. Die Grafiken (Ausgaben 14 und 15) geben Hinweise für unterschiedliches Wachstum an den Orten. Den unterschiedlichen Einfluss der einzelnen Herkünfte an den Orten im jeweiligen Jahr kann man beispielsweise anhand der Mittelwerte mit einem anderen gruppierten Scatterplot zeigen. Das Programm 16 liefert für das Jahr 2005 eine solche Grafik (Ausgabe 16). Dass die Variable Herkünfte numerisch gewählt wurde, hat keine Auswirkung auf die Darstellung. Die Herkünfte könnten auch mit (möglichst wenig) Buchstaben bezeichnet werden. Die Verbindung der Mittelwerte der Stammhöhen der Herkünfte je Ort ist keine inhaltliche; sie dient ausschließlich der Anschaulichkeit.

**Programm 16:** Eiche4.sas

```

PROC MEANS DATA=eiche (WHERE=(h05 ^= . ));
  CLASS Ort Herkunft ;
  VAR H05;
  OUTPUT OUT=m_eiche mean=mean;
RUN;

PROC SORT DATA=m_eiche;
  BY Herkunft;
RUN;

GOPTIONS
  Ftext="Times New Roman"
  Htext=1.2
  Device=emf;

SYMBOL1 h=.5 v=dot POINTLABEL=( "#Herkunft" )i=join c=black;
SYMBOL2 h=.5 v=dot POINTLABEL=( "#Herkunft" )i=join c=yellow;
SYMBOL3 h=.5 v=dot POINTLABEL=( "#Herkunft" )i=join c=green;
SYMBOL4 h=.5 v=dot POINTLABEL=( "#Herkunft" )i=join c=gray;

TITLE "Stammhöhe 2005 in cm";
PROC GGPLOT DATA=m_eiche (WHERE=( _type_ =3 AND mean ^= . ));
  PLOT mean * Herkunft = ort
    /HMINOR=0 NOFRAME;
RUN;
QUIT;

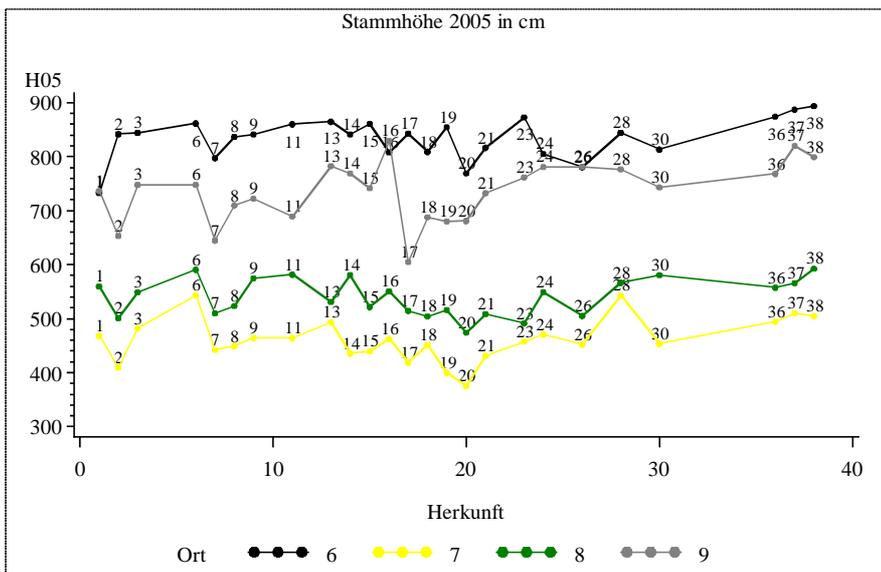
```

berechnen der Mittelwerte

mit POINTLABEL werden zusätzlich die Bezeichnungen der Herkünfte abgetragen

mittlere Stammhöhen über den Herkünften für jeden Ort

**Ausgabe 16:** zu Eiche4.sas - gruppierte Darstellung der Mittelwerte



Die mittleren Stammhöhen der Herkünfte zeigen für 2005 ein im Großen und Ganzen ähnliches Verhalten, aber auch ortspezifische Besonderheiten wie z.B. die Herkunft 17.

Welche Grafik(en) für die Beantwortung der jeweiligen Versuchsfragen geeignet ist/sind, kann und soll hier nicht diskutiert werden. Gestützt auf die Visualisierung der Daten, die immer sinnvoll ist, lassen sich Fragen zur Gleichheit von Varianzen, zum Vergleich von Mittelwerten, nach einem linearen (ortsabhängigen) Trend, nach der Gleichheit der Anstiege, nach Wechselwirkungen oder zu anderen Problemen vorteilhafter für die statistische Analyse formulieren und beantworten.

### **3.3 Zusammenfassung**

Die grafische Darstellung von Lage und Streuung des untersuchten Merkmals unter Beachtung der Strukturen ist für eine weitere statistische Analyse von Vorteil. Geeignete Grafiken können das Anliegen der untersuchten Versuchsfrage veranschaulichen und so den Weg für eine statistische Auswertung bereiten.

## **4 Skalierung der Faktoren und Merkmale und Wahl des statistischen Auswertungsverfahrens**

### **4.1 Faktor(en)**

In der Biometrie werden allgemeine Einflussgrößen, die auf die zu untersuchenden Merkmale (Zufallsgrößen) einwirken, als Faktoren bezeichnet.

Zur Analyse der fachlichen Versuchsfrage zählt nicht nur die Präzisierung der Frage nach dem biologischen Sachverhalt, der durch den Versuch zu beantworten ist, sondern auch die Analyse des Bedingungskomplexes, d.h. die Analyse der auf das Versuchsobjekt wirkenden Faktoren.

Es wird dringend empfohlen, zur Thematik der Faktoren bei RICHTER (2004) die Seiten 29-30 und bei SCHUMACHER (2004) die Seiten 100-101 nachzulesen, da die nachfolgenden Bemerkungen nur zusammenfassenden Charakter haben sollen und können.

Zu unterscheiden ist zwischen Plan- und Störfaktoren. Prüffaktor(en) und Blockfaktor(en) sind Planfaktoren. Einzig Blockfaktoren sind neben der Eigenschaft Planfaktor(en) zu sein auch Störfaktor(en) und Organisationsprinzip für die Arbeitsorganisation. Prüffaktor(en) ist/sind Faktor(en), dessen/deren Wirkung auf das Versuchsobjekt untersucht werden soll. Zum Erkennen einer Wirkung werden verschiedene Ausprägungen des Prüffaktors – Stufen genannt – in den Versuch einbezogen. Gibt es im Versuch nur einen Prüffaktor, spricht man vom einfaktoriellen oder einfachen, bei mehreren Prüffaktoren vom mehrfaktoriellen oder faktoriellen Versuch. Allgemein werden Prüffaktoren mit den ersten Großbuchstaben des Alphabets gekennzeichnet (A, B, C, ...), die Anzahl der Stufen mit den entsprechenden Kleinbuchstaben (a, b, c, ...). Die Prüfglieder – auch Varianten oder Behandlungen genannt – eines einfaktoriellen Versuchs sind die Stufen des Prüffaktors, eines mehrfaktoriellen Versuchs die Stufenkombinationen der Prüffaktoren. Auf das Versuchsobjekt, an dem die Prüfmerkmale beobachtet werden, wirken neben den Prüffaktore(n) die Konstant-, Rest- und systematische Störfaktoren. Konstantfaktoren sind die Faktoren, deren Einfluss auf das Versuchsobjekt nicht untersucht werden soll. Sie sind folglich für den Versuch konstant zu halten. Als Restfaktoren bezeichnet man die Gesamtheit der zufällig auf die Versuchsobjekte wirkenden Einflüsse. Sie umfassen unter anderem die Variabilität der Versuchsobjekte, Störfaktoren sowie Versuchsfehler und sollten deshalb möglichst gering gehalten werden. Systematische Störfaktoren sind spezielle Störfaktoren mit räumlichen, zeitlichen und/oder personellen Ursachen, deren Einfluss bei der Planung eines Versuchs zu berücksichtigen ist. Sie dadurch von den übrigen Restfaktoren zu trennen, erhöht die Wiederholungsgenauigkeit.

Die Anzahl der Prüffaktoren und die zusätzlichen Planfaktoren (z.B. Blockfaktoren) kennzeichnen den Versuch und bestimmen die Versuchsanlage. Die Prüffaktoren widerspiegeln das Versuchsziel und ihre Wirkung auf die Versuchsobjekte. Das Ziel muss sein, die Versuchsfrage vor Durchführung des Versuchs (!) so zu präzisieren, dass störende Einflüssen auf die Prüffaktoren gering sind. Dem entsprechend sind dann Versuchsanlage und statistisches Modell zu wählen.

### **4.2 Datenkomprimierung in Abhängigkeit von der Skalierung eines Prüfmerkmals**

Bestimmte Eigenschaft(en) werden am Versuchsobjekt gemessen, gezählt, geschätzt, bonitiert, ..., allgemein: beobachtet. Diese Eigenschaft(en) werden als Prüfmerkmale, mitunter auch nur als Merkmale bezeichnet. Sie werden herangezogen, um die Wirkung der Prüffaktoren einzuschätzen. Die beobachteten Ausprägungen eines Prüfmerkmals können sehr unterschiedlich sein. Ihre Skalierung ist aber von entscheidender Bedeutung für eine weitere Verdichtung der Daten. Auch für die Charakterisierung des Prüfmerkmals wird ausdrücklich auf RICHTER (2004, Kapitel 2) verwiesen und hier nur die zusammenfassende Tabelle RICHTER (2004, S. 79) übernommen (Tab. 10).

**Tab. 10:** Möglichkeiten der Datenkomprimierung in Abhängigkeit von der Skalierung (RICHTER 2004, S. 79)

Skalierung des Merkmals	Häufigkeitsverteilung (bei evtl. notwendiger Klasseneinteilung)	Maßzahlen der Lage	Maßzahlen der Streuung
metrisch	absolute und relative Häufigkeiten absolute und relative Summenhäufigkeiten	arithmetisches Mittel Median Modalwert <sup>1)</sup> empirische Quantile (z.B. Quartile)	Varianz Standardabweichung Variationskoeffizient Spannweite Quantilabstände (z.B. Hälftespielraum)
ordinal	absolute und relative Häufigkeiten absolute und relative Summenhäufigkeiten	Median Modalwert <sup>1)</sup> empirische Quantile (z.B. Quartile)	indirekte Charakterisierung durch Angabe von • Minimum und Maximum • unteres und oberes Quartil
nominal	absolute und relative Häufigkeiten	Modalwert <sup>1)</sup>	-

<sup>1)</sup> Bei Mehrgipfligkeit sind mehrere Modalwerte anzugeben.

Für nominal- und ordinalskalierte Prüfmerkmale können etliche Maßzahlen **nicht** berechnet werden!

### 4.3 Die Daten zweier Versuche

Von 2 Versuchen, nennen wir sie V1 und V2, wollen wir zunächst nur die Daten betrachten. Wir gehen davon aus, dass erforderliche Randomisationen erfolgt sind.

V1:

N1	N2	N3	N4	N5
1,25	3,25	11,07	24,16	28,32
1,61	2,68	19,12	21,53	31,72
1,79	4,37	13,81	29,49	25,51
1,98	3,73	16,79	18,98	40,91

V2:

M1	M2	M3	M4
75,5	67,4	64,4	36,6
77,2	68,3	66,0	39,3
78,6	68,5	67,3	42,0
81,4	71,3	67,8	46,2
83,8	71,6	69,2	49,7
94,9	77,5	76,4	72,2
103,9	82,1	83,5	93,0
112,0	87,4	89,3	112,0
119,8	92,1	94,8	141,8
127,2	96,7	105,0	173,6
159,5	110,4	116,0	269,9
163,8	115,2	126,1	303,4
166,8	119,6	132,0	354,4
176,3	122,7	136,6	391,5
182,0	125,2	142,5	458,3
184,4	126,2	144,5	472,2
190,5	128,5	147,6	520,5

Die Daten sehen metrisch aus. Mehr lässt sich ohne die Kenntnis der Aufgabenstellung und ohne einen Hinweis darauf, was Stufe eines Faktors oder Merkmal ist, nicht sagen. Wir gehen davon aus, dass die Daten dieselbe Maßeinheit haben. Aus MS-Excel-Dateien mit obiger Struktur und den Spaltenbezeichnung wurden - wie auf den Seiten 8 und 9 vorgestellt - SAS-Dateien mit den Namen V1 und V2 erzeugt. Mit Hilfe des Programms 17 werden die Box-Whisker-Plots (Ausgabe 17) gezeichnet.

**Programm 17:** V1\_V2.sas

```

DATA a;
  SET V1;
  Bezeichnung = 'N1'; Wert = N1; OUTPUT;
  Bezeichnung = 'N2'; Wert = N2; OUTPUT;
  Bezeichnung = 'N3'; Wert = N3; OUTPUT;
  Bezeichnung = 'N4'; Wert = N4; OUTPUT;
  Bezeichnung = 'N5'; Wert = N5; OUTPUT;
RUN;
DATA V1;
  SET a;
  DROP N1-N5;
RUN;

DATA b;
  SET V2;
  Bezeichnung = 'M1'; Wert = M1; OUTPUT;
  Bezeichnung = 'M2'; Wert = M2; OUTPUT;
  Bezeichnung = 'M3'; Wert = M3; OUTPUT;
  Bezeichnung = 'M4'; Wert = M4; OUTPUT;
RUN;
DATA V2N;
  SET b;
  DROP M1-M4;
RUN;
PROC SORT DATA=V1;
  BY Bezeichnung;
PROC SORT DATA=V2N;
  BY Bezeichnung;
RUN;

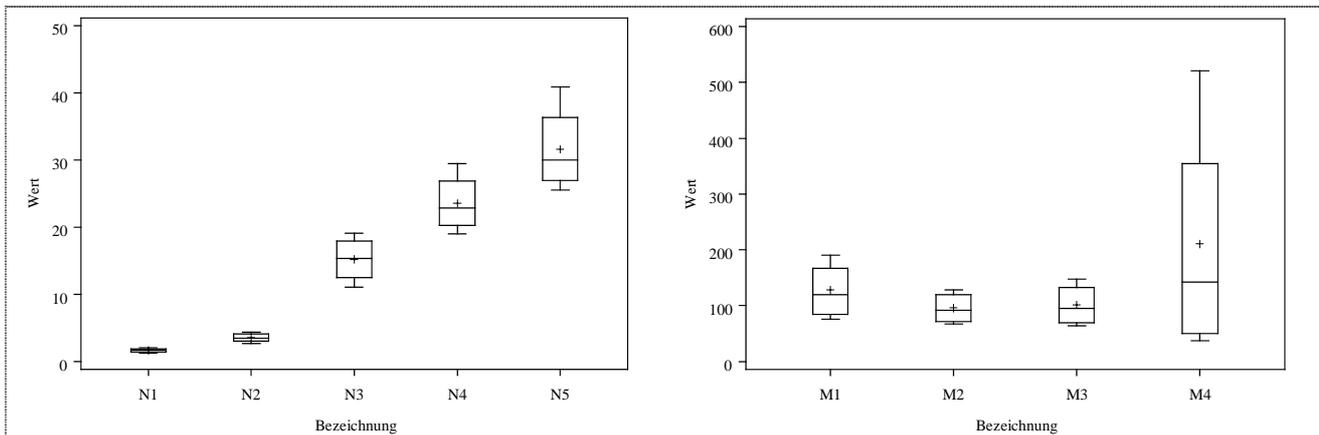
GOPTIONS device=emf
          ftext="Times"
          htext=1.2 ;

PROC BOXPLOT DATA=V1 ;
  PLOT Wert * Bezeichnung /CBOXES=black ;
RUN;
PROC BOXPLOT DATA=V2N ;
  PLOT Wert * Bezeichnung /CBOXES=black ;
RUN; QUIT;
    
```

Die Daten müssen umstrukturiert werden.

Die Daten müssen umstrukturiert werden.

**Ausgabe 17:** zu V1\_V2.sas



Deutlich zu erkennen ist, in *VI* steigt die Variabilität mit größer werdenden Mittelwerten und in *V2* ist die Variabilität unterschiedlich.

#### 4.4 Statistische Auswertungsverfahren, die aus den Eigenschaften der Faktoren und Prüfmerkmale resultieren

Bekanntlich bedingen die Wahl der Stufen des Faktors bzw. der Faktoren und die Skalierung des Merkmals bzw. der Merkmale das bzw. die statistischen Auswertungsverfahren (Tab. 11).

**Tab. 11:** Schematische Darstellung der aus der Wahl der Faktorstufen und/oder der Merkmalskalierung abzuleitenden statistischen Auswertungsverfahren

		Stufen des Faktors/der Faktoren oder Merkmal		
		nominal	ordinal	metrisch
Merkmal(e) oder Merkmal	nominal	Kontingenztafelanalyse	Rangvarianzanalyse Rangtests, parameterfreie Tests	Varianzanalyse
	ordinal	Rangvarianzanalyse Rangtests, parameterfreie Tests	Rangkorrelation	Rangkorrelation
	metrisch	Varianzanalyse	Rangkorrelation	Korrelationsanalyse, Regressionsanalyse

Ohne weiter darauf einzugehen soll an dieser Stelle nur vermerkt werden, dass die statistischen Maßzahlen, die einen Zusammenhang zwischen Merkmalen beschreiben, ebenfalls von der Skalierung der Merkmale abhängen. So beschreibt beispielsweise der bekannte Produkt-Momenten-Korrelationskoeffizient, auch Pearsonscher Korrelationskoeffizient genannt, den linearen Zusammenhang zwischen zwei (oder mehreren) quantitativen Merkmalen.

#### 4.5 Versuch V1 - Varianzanalyse

Die Daten zu *VI* sind Trockengewichte in mg der Pilzkulturen *Rhizopus oryzae* auf 5 verschiedenen künstlichen Nährböden (SCHUMACHER 2004, S. 32):

V1:

Nährboden 1	Nährboden 2	Nährboden 3	Nährboden 4	Nährboden 5
1,25	3,25	11,07	24,16	28,32
1,61	2,68	19,12	21,53	31,72
1,79	4,37	13,81	29,49	25,51
1,98	3,73	16,79	18,98	40,91

Abb. 13 zeigt deutlich die zunehmende Variabilität, eine Zunahme, die nichts mit der Bezeichnung des Nährbodens zu tun hat, sondern mit den mittleren Trockengewichten in Beziehung steht. Die SAS-Prozeduren *UNIVARIATE* und *MEANS* liefern die in Tab. 12 stehenden Maßzahlen.

**Tab. 12:** Mittelwerte, Standardabweichungen und Varianzen von V1

Trockengewicht	Mean	Standardabweichung	Varianz
Nährboden1	1.6575	0.31085	0.0966
Nährboden2	3.5075	0.71751	0.5148
Nährboden3	15.1975	3.50635	12.2945
Nährboden4	23.5400	4.49520	20.2069
Nährboden5	31.6150	6.69667	44.8454

Was auch im Balkendiagramm (Abb. 12) sichtbar ist, verdeutlicht die Beziehung der Mittelwerte und Standardabweichungen (Abb. 13): die Abhängigkeit von Variabilität und Mittelwert. Mittelwert und Standardabweichung verhalten sich annähernd proportional zueinander:  $\mu \sim \sigma$ .

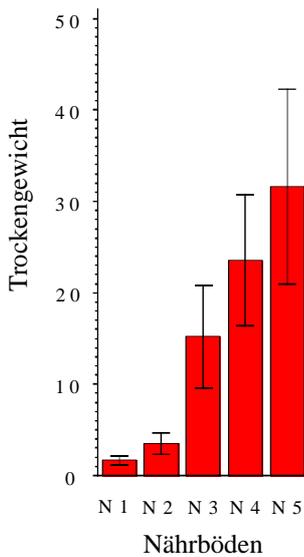


Abb. 12: Balkendiagramm der Trockengewichte

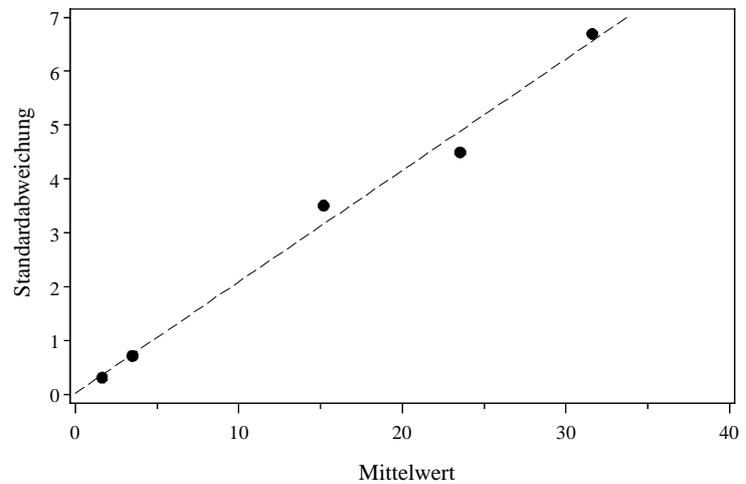


Abb. 13: Abhängigkeit von Mittelwert und Standardabweichung

Wenn so etwas auftritt, kann man im Allgemeinen davon ausgehen, dass die (neben der Additivität der Effekte im Modell) wichtigste Voraussetzung, die Gleichheit der Varianzen, **nicht** angenommen werden kann. Hauptsächlich für diese Betrachtung haben wir uns unabhängig von einer fachlichen Aufgabenstellung und der Kenntnis von Faktoren und/oder Merkmalen ausschließlich auf die Daten orientiert. Beim Merkmal Trockengewicht würde man nicht unbedingt auf eine Proportionalität zwischen Mittelwert und Standardabweichung schließen, bei der exponentiellen Entwicklung von Pilzgeflechten aber sehr wohl. Es bietet sich eine logarithmische Transformation der Daten an. Und genau die verwendet SCHUMACHER (2004). Wir haben einen Faktor Nährböden mit 5 nominalen Stufen und ein metrisches Prüfmerkmal, das (logarithmierte) Trockengewicht. Das führt uns nach Tab. 11 zur Varianzanalyse. Ein Blockfaktor ist nicht benannt, so dass wir vom Modell einer einfaktoriellen randomisierten Anlage A-R ausgehen. Die statistische Auswertung des Versuchs *V1* umfasst einen Test auf Varianzhomogenität und einen zum Vergleich der Mittelwerte (Tukey-Prozedur) (s. Programm 18). Die jeweilige Überprüfung der Residuen auf Normalverteilung ist dem SAS-Programm von SCHUMACHER (2004, S. 34) zu entnehmen. Die Ausgabe der Ergebnisse ist verkürzt in Ausgabe 18 angegeben.

**Programm 18:** V1\_GLM.sas

<pre> PROC GLM DATA=V1;   CLASS Bezeichnung;   MODEL Wert = Bezeichnung;   MEANS Bezeichnung /HOVTEST=BF; RUN;  DATA V1T;   SET V1;   logTGewicht = log(Wert); RUN;  PROC GLM DATA=V1T;   CLASS Bezeichnung;   MODEL logTGewicht = Bezeichnung;   MEANS Bezeichnung /HOVTEST=BF TUKEY; RUN; QUIT; </pre>	<p>ausgewertet wird die Datei V1 aus Programm 17 die Option HOVTEST=BF führt den Homogenitätstest nach Brown und Forsythe durch, eine Modifikation des Levene-Tests auf der Basis der Mediane (s. SCHUMACHER 2004)</p> <p>logarithmische Transformation der Daten</p> <p>Test auf Homogenität der Varianzen, Varianzanalyse und Tukey-Prozedur</p>
--	--

**Ausgabe 18:** zu V1\_GLM.sas

The GLM Procedure							Varianztabelle der Originaldaten
Dependent Variable: Wert							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	4	2636.300530	659.075132	42.27	<.0001		
Error	15	233.874525	15.591635				
Corrected Total	19	2870.175055					
Source	DF	Type III SS	Mean Square	F Value	Pr > F		
Bezeichnung	4	2636.300530	659.075132	42.27	<.0001		
Brown and Forsythe's Test for Homogeneity of Wert Variance ANOVA of Absolute Deviations from Group Medians					Test auf Homogenität der Varianzen: Empfohlen wird ein Signifikanzniveau $\alpha=0,10$ . Das bedeutet, dass von unterschiedlichen Varianzen ausgegangen werden muss.		
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Bezeichnung	4	57.2942	14.3236	2.60	<b>0.0785</b>		
Error	15	82.7088	5.5139				
Level of -----Wert-----							
Bezeichnung	N	Mean	Std Dev				
N1	4	1.6575000	0.31084562				
N2	4	3.5075000	0.71751307				
N3	4	15.1975000	3.50635019				
N4	4	23.5400000	4.49520485				
N5	4	31.6150000	6.69666833				
The GLM Procedure							Varianztabelle der logarithmierten Daten
Dependent Variable: logTGewicht							
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Model	4	26.07923460	6.51980865	151.47	<.0001		
Error	15	0.64566314	0.04304421				
Corrected Total	19	26.72489774					
Source	DF	Type III SS	Mean Square	F Value	Pr > F		
Bezeichnung	4	26.07923460	6.51980865	151.47	<.0001		
Brown and Forsythe's Test for Homogeneity of logTGewicht Variance ANOVA of Absolute Deviations from Group Medians					Test auf Homogenität der Varianzen: Es kann von gleichen Varianzen ausgegangen werden.		
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F		
Bezeichnung	4	0.00575	0.00144	0.12	<b>0.9738</b>		
Error	15	0.1819	0.0121				
Tukey's Studentized Range (HSD) Test for logTGewicht							Tukey-Prozedur
NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.							
Alpha				0.05			
Error Degrees of Freedom				15			
Error Mean Square				0.043044			
Critical Value of Studentized Range				4.36699			
Minimum Significant Difference				0.453			
Means with the same letter are not significantly different.							
Tukey Grouping							
	Mean	N	Bezeichnung				
A	3.4377	4	N5				
B	3.1454	4	N4				
B	2.7003	4	N3				
C	1.2389	4	N2				
D	0.4912	4	N1				

## 4.6 Versuch V2 - Korrelationsanalyse, Regressionsanalyse

Ein ganz anderes Bild ergibt sich für den Versuch V2, wenn fachliches Hintergrundwissen einbezogen wird. Die 4 Spalten sind verschiedene mittlere Körpermaße Brustumfang, Widerristhöhe, Rumpflänge und Körpermasse gleichaltriger Bullenkälber im Alter zwischen Geburt und 48 Monaten (RASCH und VERDOOREN 2004, S. 13):

V2:

Brustumfang (in cm)	Widerristhöhe (in cm)	Rumpflänge (in cm)	Körpermasse (in kg)
75,5	67,4	64,4	36,6
77,2	68,3	66,0	39,3
78,6	68,5	67,3	42,0
81,4	71,3	67,8	46,2
83,8	71,6	69,2	49,7
94,9	77,5	76,4	72,2
103,9	82,1	83,5	93,0
112,0	87,4	89,3	112,0
119,8	92,1	94,8	141,8
127,2	96,7	105,0	173,6
159,5	110,4	116,0	269,9
163,8	115,2	126,1	303,4
166,8	119,6	132,0	354,4
176,3	122,7	136,6	391,5
182,0	125,2	142,5	458,3
184,4	126,2	144,5	472,2
190,5	128,5	147,6	520,5

Alle 4 Merkmale sind metrisch. Ein Faktor ist nicht erkennbar. Als statistische Auswertungsverfahren kommen nach Tab. 11 die Korrelations- und/oder Regressionsanalyse in Frage. Im Allgemeinen gehen die Zusammenhangsmaße von einem linearen Zusammenhang aus. Sie zu berechnen setzt voraus, dass es sich um Zufallsvariable, Merkmale mit zufälligen Ausprägungen, handelt. Das ist eine Voraussetzung, die häufig „übersehen“ wird. Eine Gerade ist das einfachste Modell zur Beschreibung des funktionalen Zusammenhangs zwischen zwei Merkmalen. In beiden Fällen ist zu prüfen, ob wir es mit einem annähernd linearen Verlauf zu tun haben. Ein geeigneter Weg ist die grafische Darstellung der Daten. Um das zu realisieren gehen wir von der SAS-Datei V2 aus, nicht von V2N (s. Programm 17) und verwenden die Merkmalsbezeichnungen – allerdings ohne Umlaute (Programm 19). Nur die erste Grafik ist hier angegeben. Die anderen, die die Tab. 13 komplettieren, lauten für alle vier Merkmale analog.

### Programm 19: V2\_GPLOT.sas

```

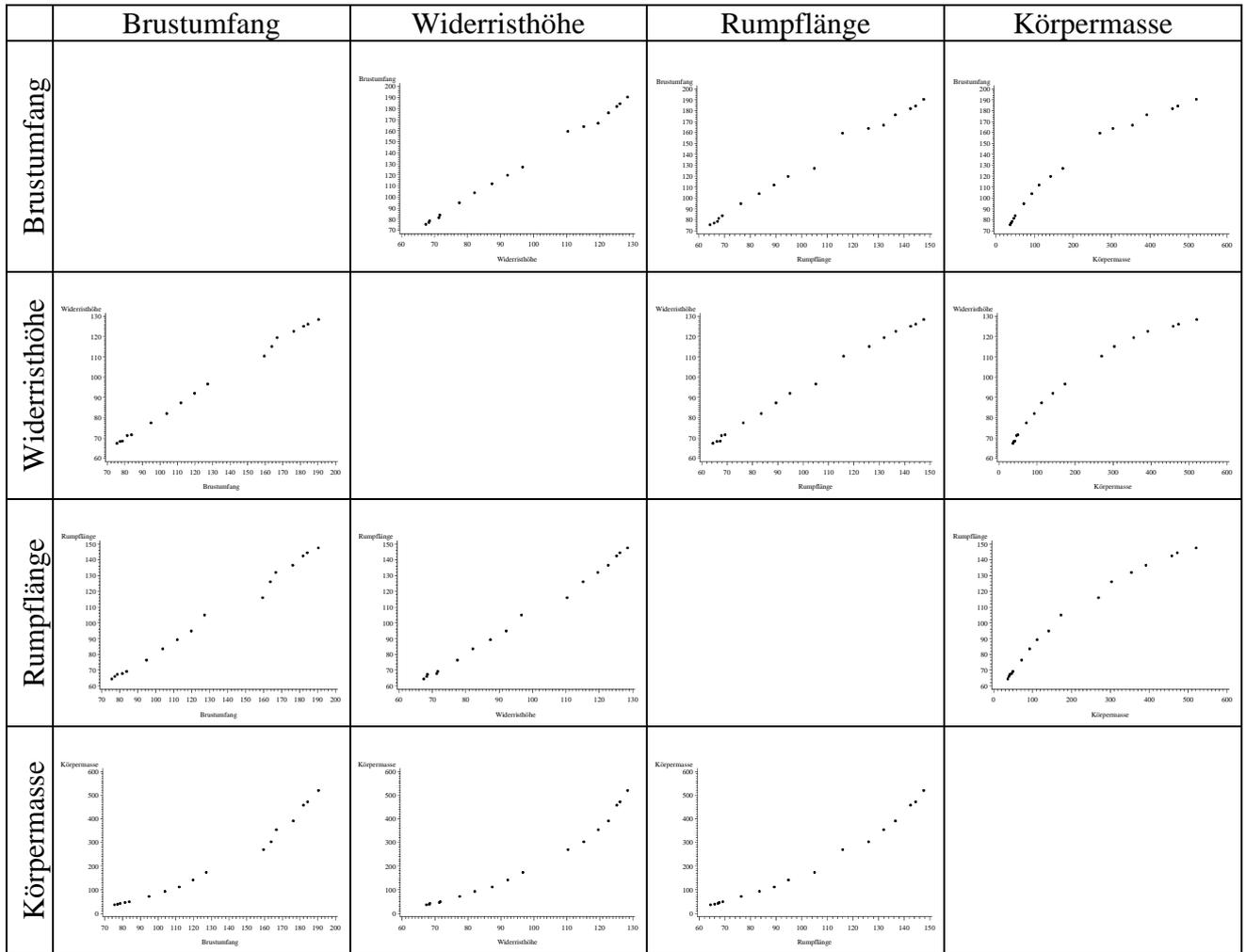
DATA V2R;
  SET V2
    (rename=(M1=Brustumfang M2=Widerristhoehe M3=Rumpflaenge M4=Koerpermasse));
RUN;

GOPTIONS FTEXT="Times" HTEXT=1.2
  DEVICE=emf;
SYMBOL h=.5 v=dot i=none c=black;

PROC GPLOT DATA=V2R;
  LABEL Brustumfang='Brustumfang'
    Widerristhoehe='Widerristhöhe';
  PLOT Brustumfang*Widerristhoehe / NOFRAME;      PLOT <y-Achse> * <x-Achse>;
RUN; QUIT;

```

**Tab. 13:** Darstellungen der Zusammenhänge zwischen den vier Merkmalen



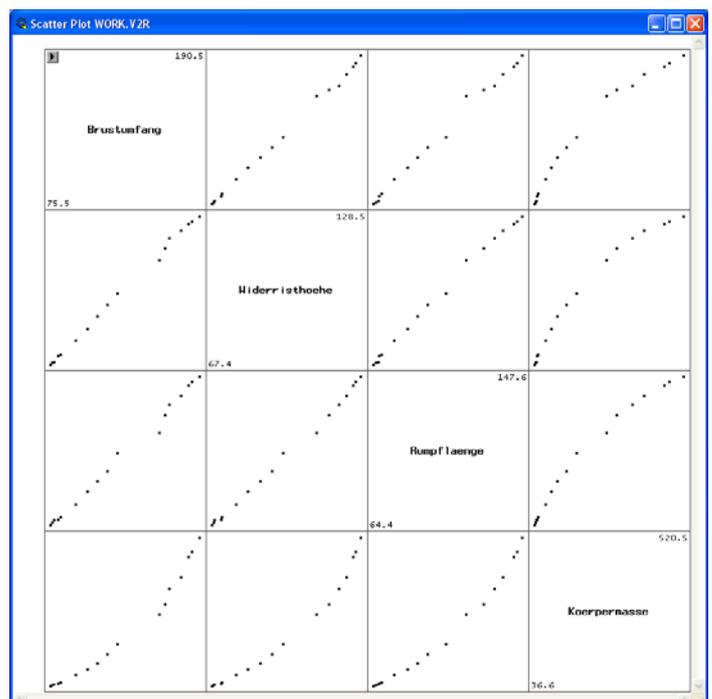
Eine ähnliche Darstellung wie die in Tab. 12 erhält man, wenn SAS/INSIGHT verwendet wird. Der Aufruf erfolgt ähnlich dem des SAS/Analyst über `insight` im Command-Fenster oder durch

Solutions → Analysis ► → Interactive Data Analysis .

Beim Aufruf werden die SAS-Library und die Datei, die bearbeitet werden soll, festgelegt. Mit

Analyze → Scatter Plot (Y X)

öffnet sich ein Fenster, in dem links jeweils alle vier Merkmale markiert werden und einmal durch Anklicken des Y-Schalters als Variablen der y-Achse und zum anderen durch Anklicken des X-Schalters als Variablen der x-Achse vereinbart werden. Der Schalter OK liefert eine gemeinsame Darstellung (Abb. 14).



**Abb.14:** paarweise Darstellung der vier Merkmale

Die paarweisen Beziehungen mit dem Merkmal Körpermasse scheinen nicht linear zu sein.

Wir haben bereits festgestellt, dass alle vier Merkmale zufällige Ausprägungen haben. Sie sind Zufallsvariablen. Die paarweisen Korrelationskoeffizienten können also berechnet werden (Programm 20, Ausgabe 20).

**Programm 20:** V2\_CORR.sas

```
PROC CORR DATA=V2R PEARSON;
VAR Brustumfang Widerristhoehe Rumpflaenge Koerpermasse;
RUN;
```

**Ausgabe 20:** zu V2\_CORR.sas

```
The CORR Procedure
  4 Variables:  Brustumfang  Widerristhoehe  Rumpflaenge  Koerpermasse

                Simple Statistics
Variable          N          Mean          Std Dev          Sum          Minimum          Maximum
Brustumfang       17    128.09412    43.31428    2178    75.50000    190.50000
Widerristhoehe    17    95.92353    23.50366    1631    67.40000    128.50000
Rumpflaenge       17    101.70588    31.40206    1729    64.40000    147.60000
Koerpermasse      17    210.38824    173.40870    3577    36.60000    520.50000

                Pearson Correlation Coefficients, N = 17
                Prob > |r| under H0: Rho=0

                Brustumfang  Widerristhoehe  Rumpflaenge  Koerpermasse
Brustumfang      1.00000          0.99890          0.99653          0.97799
                  <.0001          <.0001          <.0001          <.0001
Widerristhoehe   0.99890          1.00000          0.99804          0.97672
                  <.0001          <.0001          <.0001          <.0001
Rumpflaenge      0.99653          0.99804          1.00000          0.98403
                  <.0001          <.0001          <.0001          <.0001
Koerpermasse     0.97799          0.97672          0.98403          1.00000
                  <.0001          <.0001          <.0001          <.0001
```

Im oberen Teil der Ausgabe 20 werden einige statistische Maßzahlen ausgegeben. Dann folgen die paarweisen linearen Korrelationskoeffizienten. Es stehen - bis auf die Beziehung der Variablen mit sich selbst - immer zwei Zahlen untereinander. Die obere ist der Korrelationskoeffizient und die untere die Überschreitungswahrscheinlichkeit des Tests der Hypothesen

$$H_0: \rho = 0$$

$$H_A: \rho \neq 0.$$

Die Option *PEARSON* kann im Programm 20 fehlen. Sie ist die Standardeinstellung. Werden die Optionen *SPEARMAN* und/oder *KENDALL* verwendet, dann werden die Rangkorrelationskoeffizienten nach Spearman für monoton fallende oder monoton steigende Zusammenhänge bzw. Kendalls  $\tau$  berechnet. Kendalls  $\tau$  ist eine in der Agrarforschung selten gebrauchte Maßzahl.

Die Beschreibung des Zusammenhangs zwischen zwei (oder mehreren) Variablen mit Hilfe einer Funktion, einer Regressionsfunktion, bedarf Überlegungen zur Modellwahl. Das einfachste Modell ist die Gerade. Sie ist aber nicht immer zutreffend. RASCH und VERDOOREN (2004) geben für die Modellwahl und die Regressionsanalyse mit Hilfe von SAS die notwendige Unterstützung.

## 4.7 Hinweise für die statistische Auswertung eines ordinal skalierten Merkmals

### 4.7.1 Beispiel Arzneimittelpflanzen, verschiedene Behandlungen gegen eine bestimmte Wurzelkrankheit

Auf jeder Parzelle einer Blockanlage mit 4 Wiederholungen und 5 Behandlungen (A, B, C, D, E) wurden 45 zufällig ausgewählte Pflanzen der gleichen Anzucht einer Bärentraubenherkunft gepflanzt. Es wurden mehrere Bonituren während des Beobachtungszeitraums durchgeführt. Betrachtet wird hier die Endbonitur, bei der die gesunden und kranken oder toten Pflanzen sowie die Pflanzen mit und ohne Neuaustrieb gezählt wurden. Da sich die Anzahlen (Tab. 14) sowohl für die Charakterisierung des Krankheitsbefalls als auch den Neuaustrieb zu 45 ergänzen, genügt es, das Merkmal ‚Anzahl gesunder Pflanzen‘ und das Merkmal ‚Anzahl Pflanzen mit Neuaustrieb‘ heran zu ziehen. Alle Behandlungen sollen in ihren mittleren Wirkungen miteinander verglichen werden.

**Tab. 14:** Anzahl der gesunden und kranken bzw. toten Pflanzen sowie der Pflanzen mit und ohne Neuaustrieb

Behandlung	Block	gesund	krank_tot	mit_Neuaustrieb	ohne_Neuaustrieb
A	1	35	10	37	8
A	2	31	14	38	7
A	3	34	11	38	7
A	4	37	8	40	5
B	1	18	27	15	30
B	2	15	30	33	12
B	3	11	34	24	21
B	4	12	33	20	25
C	1	26	19	23	22
C	2	21	24	29	16
C	3	16	29	14	31
C	4	25	20	43	2
D	1	23	22	15	30
D	2	28	17	10	35
D	3	18	27	12	33
D	4	30	15	18	27
E	1	7	38	16	29
E	2	14	31	28	17
E	3	17	28	34	11
E	4	11	34	23	22

Wenn die Daten mit MS-Excel gemäß Tab. 14 erfasst wurden, dann ist nur auf die Spaltenbezeichnung zu achten, die keine Leer-, Sonderzeichen und Operatoren enthalten darf. Die Spaltenbezeichnungen in Excel werden durch Import nach SAS zu den Variablennamen. Die Übernahme der Daten in eine SAS-Datei erfolgt wie auf den Seiten 7 und 8 beschrieben. Die SAS-Datei trage den Namen *Arzneipfl*.

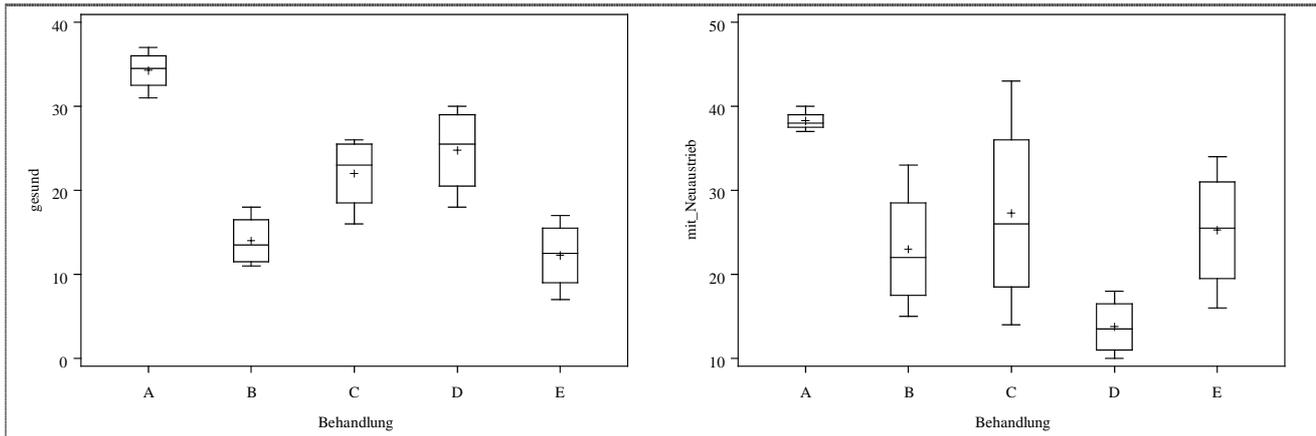
### 4.7.2 Boxplots und Zusammenhang zwischen zwei Variablen

Wir betrachten die Merkmale ‚Anzahl gesunder Pflanzen‘ und ‚Anzahl Pflanzen mit Neuaustrieb‘. Das Programm 21 liefert die Boxplots (Ausgabe 21).

**Programm 21:** Arzneipfl1.sas - Boxplots

```
GOPTIONS Device=emf
      Ftext = "Times New Roman"
      Htext = 1.2;
PROC BOXPLOT DATA=arzneipfl;
  PLOT gesund * Behandlung / CBOXES=black;
  PLOT mit_Neuaustrieb * Behandlung / CBOXES=black;
RUN;
QUIT;
```

**Ausgabe 21:** zu Arzneipfl1.sas - Boxplots



Für die Anzahl gesunder Pflanzen fällt auf, dass die Variabilität keine großen Unterschiede aufweist, dagegen aber die der Anzahl der Pflanzen mit Neuaustrieb große Variabilität zeigt. Die Verteilungen sind – wie an der guten Übereinstimmung von Median und arithmetischem Mittelwert erkennbar – annähernd symmetrisch. Bis auf die Behandlung D scheint es so zu sein, dass eine größere Anzahl gesunder Pflanzen mit einer größeren Anzahl von Pflanzen mit Neuaustrieb korrespondiert. Die Merkmale ‚Anzahl gesunder Pflanzen‘ und ‚Anzahl Pflanzen mit Neuaustrieb‘ sind Zufallsvariable. Das Programm 22 berechnet dafür die linearen Zusammenhänge. Da die Merkmale ordinal skaliert sind und folglich nach Tab. 10 kein Mittelwert berechnet werden kann, orientieren wir uns auf den Spearmanschen Rangkorrelationskoeffizienten  $r_s$ . Der Produkt-Moment-Korrelationskoeffizient, der Pearsonsche-Korrelationskoeffizient  $r$ , wird nur zum Vergleich berechnet. Wegen des auffälligen Verhaltens der Behandlung D werden die Korrelationskoeffizienten für den Versuch, für alle Behandlungen außer D und danach ausschließlich für die Behandlung D berechnet. Der eigentlich zu erwartende stärkere Zusammenhang zwischen den beiden Merkmalen, dass gesunde Pflanzen auch mehr Neuaustriebe haben, zeigt sich deutlicher ohne die Behandlung D.

**Programm 22:** Arzneipfl2.sas - Korrelationskoeffizienten

```
TITLE 'Behandlungen A, B, C, D, E';
PROC CORR DATA=arzneipfl PEARSON SPEARMAN;
  VAR gesund mit_Neuaustrieb;
RUN;

TITLE 'Behandlungen - außer D';
PROC CORR DATA=arzneipfl (WHERE=(Behandlung^='D')) PEARSON SPEARMAN;
  VAR gesund mit_Neuaustrieb;
RUN;

TITLE 'nur Behandlung D';
PROC CORR DATA=arzneipfl (WHERE=(Behandlung='D')) PEARSON SPEARMAN;
  VAR gesund mit_Neuaustrieb;
RUN;
TITLE;
```

**Ausgabe 22:** zu Arzneipfl2.sas

Behandlungen A, B, C, D, E			Behandlungen A, B, C, D, E		
Pearson Correlation Coefficients, N = 20 Prob >  r  under H0: Rho=0			Spearman Correlation Coefficients, N = 20 Prob >  r  under H0: Rho=0		
	gesund	mit_Neuaustrieb		gesund	mit_Neuaustrieb
gesund	1.00000	0.44954	gesund	1.00000	0.36836
gesund		0.0468	gesund		0.1100
mit_Neuaustrieb	0.44954	1.00000	mit_Neuaustrieb	0.36836	1.00000
mit_Neuaustrieb	0.0468		mit_Neuaustrieb	0.1100	

Behandlungen - außer D			Behandlungen - außer D		
Pearson Correlation Coefficients, N = 16 Prob >  r  unter H0: Rho=0			Spearman Correlation Coefficients, N = 16 Prob >  r  unter H0: Rho=0		
	gesund	mit_Neuaustrieb		gesund	mit_Neuaustrieb
gesund	1.00000	0.71380	gesund	1.00000	0.65954
gesund		0.0019	gesund		0.0054
mit_Neuaustrieb	0.71380	1.00000	mit_Neuaustrieb	0.65954	1.00000
mit_Neuaustrieb	0.0019		mit_Neuaustrieb	0.0054	

nur Behandlung D			nur Behandlung D		
Pearson Correlation Coefficients, N = 4 Prob >  r  unter H0: Rho=0			Spearman Correlation Coefficients, N = 4 Prob >  r  unter H0: Rho=0		
	gesund	mit_Neuaustrieb		gesund	mit_Neuaustrieb
gesund	1.00000	0.34979	gesund	1.00000	0.40000
gesund		0.6502	gesund		0.6000
mit_Neuaustrieb	0.34979	1.00000	mit_Neuaustrieb	0.40000	1.00000
mit_Neuaustrieb	0.6502		mit_Neuaustrieb	0.6000	

Die Korrelationskoeffizienten und Überschreitungswahrscheinlichkeiten sind in Tab. 15 noch einmal aufgeführt. Der Spearmansche Rangkorrelationskoeffizient  $r_s$  ist nur für den Fall der Behandlungen ohne die Behandlung D signifikant von Null verschieden (zum Niveau  $\alpha = 0,05$ ).

**Tab. 15:** Korrelationskoeffizienten und Überschreitungswahrscheinlichkeiten

	Versuch (alle Behandlungen)	Behandlungen außer D	nur Behandlung D
$r_s$	0,37	0,66	0,40
Prob > $ r_s $	0,110	0,005	0,600
$r$	0,45	0,71	0,35
Prob > $ r $	0,047	0,002	0,650

### 4.7.3 Vergleich mittlerer Werte der 5 Behandlungen

#### 4.7.3.1 Verfahren der statistischen Analyse

Die beiden Merkmale ‚Anzahl gesunder Pflanzen‘ und ‚Anzahl Pflanzen mit Neuaustrieb‘ sind ordinal skaliert. Der für metrische Merkmale mögliche Vergleich von Mittelwerten mit Hilfe der Varianzanalyse (globaler F-Test) ist durch den Friedman-Test oder den Permutationstest unter Berücksichtigung der Versuchsanlage, der einfaktoriellen randomisierten Blockanlage, zu ersetzen. An die Stelle der multiplen paarweisen Tests - beispielsweise der Tukey-Prozedur - müssen parameterfreie Tests treten. Wir haben für jeden der paarweisen Vergleiche den Wilcoxon-Test mit Berechnung der exakten Überschreitungswahrscheinlichkeiten gewählt.

Die Verteilung der Pflanzenanzahl ist für jede Behandlung weitestgehend symmetrisch (s. Ausgabe 21). Bei weniger als 4 Wiederholungen ist diese für parameterfreie Verfahren wichtige Voraussetzung schwer einzuschätzen. Zum Vergleich werden zusätzlich Varianzanalyse und Tukey-Tests gerechnet. Ein Test auf Varianzhomogenität ist derzeit nur für einfaktorielle Anlagen (keine Blockanlage) implementiert. Für den Permutationstest wird das Macro von SCHUMACHER und FRISCH (<http://www.uni-hohenheim.de/inst110/> → Personenverzeichnis → Schumacher, Dr. → SAS-Macros zu Permutationstests → *ribdperm.mac* (Download)) genutzt.

Die SAS-Programme zur Analyse werden für die beiden Merkmale getrennt betrachtet, damit die evtl. unterschiedlichen Ausgaben und Ergebnisinterpretationen eine Einheit bilden.

### 4.7.3.2 Merkmal Anzahl gesunder Pflanzen

Das Programm 23 und die dazu gehörende Ausgabe 23 sind nachfolgend aufgeführt.

**Programm 23:** Arzneipfl3.sas – Testen, Merkmal: Anzahl gesunder Pflanzen

```

%INCLUDE 'G:\SAS_MACRO\Permutation\ribdperm.mac';

TITLE1 'Merkmal: Anzahl gesunder Pflanzen';
TITLE2 'Varianzanalyse und Tukey-Prozedur';
PROC MIXED DATA=arzneipfl NOBOUND;
  CLASS Behandlung Block;
  MODEL gesund = Behandlung / DDFM=KenwardRoger;
  RANDOM Block;
  LSMEANS Behandlung;
  LSMEANS Behandlung / ADJUST=tukey ;
RUN;

TITLE2 'Permutationstest';
DATA permutation;
  set arznejpfl;
  Bezeichnung = 'Arzneipfl.';
  IF Behandlung = 'A' THEN NVariante = 1;
  IF Behandlung = 'B' THEN NVariante = 2;
  IF Behandlung = 'C' THEN NVariante = 3;
  IF Behandlung = 'D' THEN NVariante = 4;
  IF Behandlung = 'E' THEN NVariante = 5;
RUN;

%ribdperm (g,
  pw_t,
  daten =permutation,
  versuch =Bezeichnung,
  block =Block,
  beh =NVariante,
  wert =gesund,
  nperm=10000,
  alpha=0.05);

TITLE2 'Friedman-Test ';
PROC FREQ DATA=arzneipfl;
  TABLES Block*Behandlung*gesund / NOPRINT CMH2 SCORES=RANK;
RUN;

TITLE2 'Wilcoxon-Test (Mann-Whitney-Test) für jeweils 2 Varianten';
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='A' OR Behandlung='B')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='A' OR Behandlung='C')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='A' OR Behandlung='D')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='A' OR Behandlung='E')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='B' OR Behandlung='C')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='B' OR Behandlung='D')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='B' OR Behandlung='E')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='C' OR Behandlung='D')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='C' OR Behandlung='E')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
PROC NPARIWAY DATA=arzneipfl (WHERE=(Behandlung='D' OR Behandlung='E')) WILCOXON;
  CLASS Behandlung; VAR gesund; EXACT; RUN;
TITLE;

```

Das heruntergeladene Macro (SCHUMACHER u. FRISCH) zur Analyse einer einfaktoriellen Blockanlage muss aktiviert werden.

Varianzanalyse und Tukey-Prozedur  
Blocks: zufällig

Für das Macro *ribdperm* ist eine zusätzliche Versuchsbezeichnung notwendig. Die Behandlungen müssen numerisch sein (Behandlung → NVariante).

Globaltest  
paarweise Vergleiche vom Tukey-Typ  
Blockfaktor  
Merkmal  
Anzahl Permutationen  
multiples Signifikanzniveau

Parameter des Macro-Aufrufs

Erwähnt werden soll, dass PROC FREQ mit der Anweisung EXACT ... auch exakte Testgrößen und Überschreitungswahrscheinlichkeiten berechnet.

**Ausgabe 23:** zu Arzneipfl3.sas (gekürzt)

Merkmal: Anzahl gesunder Pflanzen  
 Varianzanalyse und Tukey-Prozedur

Covariance Parameter Estimates

Cov Parm Estimate  
 Block -0.9917  
 Residual 17.8083

Die geschätzte negative Blockvarianz führt zu einer Mitteilung:  
 NOTE: Estimated G matrix is not positive definite.

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Behandlung	4	12	17.70	<.0001

Least Squares Means

Effect	Behandlung	Estimate	Standard Error	DF	t Value	Pr >  t
Behandlung	A	34.2500	2.0504	14.8	16.70	<.0001
Behandlung	B	14.0000	2.0504	14.8	6.83	<.0001
Behandlung	C	22.0000	2.0504	14.8	10.73	<.0001
Behandlung	D	24.7500	2.0504	14.8	12.07	<.0001
Behandlung	E	12.2500	2.0504	14.8	5.97	<.0001

Differences of Least Squares Means

Effect	Behandlung	Behandlung	Estimate	Standard Error	DF	Adjustment	Adj P
Behandlung	A	B	20.2500	2.9840	12	Tukey-Kramer	0.0002
Behandlung	A	C	12.2500	2.9840	12	Tukey-Kramer	0.0104
Behandlung	A	D	9.5000	2.9840	12	Tukey-Kramer	0.0503
Behandlung	A	E	22.0000	2.9840	12	Tukey-Kramer	<.0001
Behandlung	B	C	-8.0000	2.9840	12	Tukey-Kramer	0.1161
Behandlung	B	D	-10.7500	2.9840	12	Tukey-Kramer	0.0246
Behandlung	B	E	1.7500	2.9840	12	Tukey-Kramer	0.9746
Behandlung	C	D	-2.7500	2.9840	12	Tukey-Kramer	0.8832
Behandlung	C	E	9.7500	2.9840	12	Tukey-Kramer	0.0436
Behandlung	D	E	12.5000	2.9840	12	Tukey-Kramer	0.0090

Merkmal: Anzahl gesunder Pflanzen  
 Permutationstest

Bezeichnung des Versuches, Anzahl der Simulationsschritte und multiples Alpha

VERS NPERM ALPHA  
 Arzneipfl. 10000 0.05

Globaler Test (Simulation) H0: "Keine Behandlungsunterschiede"

ERGEBNIS  
 Pr>Perm 0

Die approximative Überschreitungswahrscheinlichkeit mit dem Wert 0 und deren 0,99-Konfidenzintervall [0 ; 0,0005297] liegen im Bereich des vorgegebenen multiplen Signifikanzniveaus  $\alpha=0,05$ . Es kann also davon ausgegangen werden, dass es eingehalten wird.

0.99 Konfidenzintervall für Pr  
 ERGEBNIS  
 0 0.0005297

Paarweise Vergleiche, Tukey-Typ (Simulation)

1. Differenzen der Prüfgliedmittelwerte und Grenzdifferenz

ERGEBNIS

Differenzen	1	2	3	4	5
1	0.000	20.250	12.250	9.500	22.000
2		0.000	8.000	10.750	1.750
3			0.000	2.750	9.750
4				0.000	12.500
5					0.000

ERGEBNIS  
 Grenzdifferenz 17.75

2. Überschreitungswahrscheinlichkeiten

ERGEBNIS

Pr>Prob	1	2	3	4	5
1		0.009	0.404	0.693	0.001
2			0.817	0.560	1.000
3				0.998	0.666
4					0.381
5					

Überschreitungswahrscheinlichkeiten der dem klassischen Tukey-Test nahe kommenden paarweisen Permutationstests, Vergleiche 1-2 und 1-5 zum Niveau 0,05 signifikant

Merkmal: Anzahl gesunder Pflanzen  
Friedman-Test

Summary Statistics for Behandlung by gesund  
Controlling for Block

Cochran-Mantel-Haenszel Statistics (Based on Rank Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	4.0000	0.0455
2	Row Mean Scores Differ	4	13.4000	0.0095

← Teststatistik: 13,4 und  
Überschreitungswahrscheinlichkeit: 0,0095

Merkmal: Anzahl gesunder Pflanzen

Wilcoxon-Test (Mann-Whitney-Test) für jeweils 2 Varianten

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
A	4	26.0	18.0	3.464102	6.50
B	4	10.0	18.0	3.464102	2.50

Wilcoxon Two-Sample Test  
Statistic 26.0000

t Approximation

One-Sided Pr > Z 0.0336  
Two-Sided Pr > |Z| 0.0671

zweiseitige Überschreitungswahrscheinlichkeit (t-Approximation)

Exact Test

One-Sided Pr >= S 0.0143  
Two-Sided Pr >= |S - Mean| 0.0286

zweiseitige Überschreitungswahrscheinlichkeit (exakter Test)

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
A	4	26.0	18.0	3.464102	6.50
C	4	10.0	18.0	3.464102	2.50

Wilcoxon Two-Sample Test  
Statistic 26.0000

t Approximation

One-Sided Pr > Z 0.0336  
Two-Sided Pr > |Z| 0.0671

Exact Test

One-Sided Pr >= S 0.0143  
Two-Sided Pr >= |S - Mean| 0.0286

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
A	4	26.0	18.0	3.464102	6.50
D	4	10.0	18.0	3.464102	2.50

Wilcoxon Two-Sample Test  
Statistic 26.0000

t Approximation

One-Sided Pr > Z 0.0336  
Two-Sided Pr > |Z| 0.0671

Exact Test

One-Sided Pr >= S 0.0143  
Two-Sided Pr >= |S - Mean| 0.0286

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
A	4	26.0	18.0	3.464102	6.50
E	4	10.0	18.0	3.464102	2.50

Wilcoxon Two-Sample Test  
Statistic 26.0000

t Approximation  
One-Sided Pr > Z 0.0336  
Two-Sided Pr > |Z| 0.0671

Exact Test  
One-Sided Pr >= S 0.0143  
Two-Sided Pr >= |S - Mean| 0.0286

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
B	4	11.0	18.0	3.464102	2.750
C	4	25.0	18.0	3.464102	6.250

Wilcoxon Two-Sample Test  
Statistic 11.0000

t Approximation  
One-Sided Pr < Z 0.0514  
Two-Sided Pr > |Z| 0.1027

Exact Test  
One-Sided Pr <= S 0.0286  
Two-Sided Pr >= |S - Mean| 0.0571

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
B	4	10.50	18.0	3.443420	2.6250
D	4	25.50	18.0	3.443420	6.3750

Wilcoxon Two-Sample Test  
Statistic 10.5000

t Approximation  
One-Sided Pr < Z 0.0408  
Two-Sided Pr > |Z| 0.0816

Exact Test  
One-Sided Pr <= S 0.0286  
Two-Sided Pr >= |S - Mean| 0.0571

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
B	4	20.50	18.0	3.443420	5.1250
E	4	15.50	18.0	3.443420	3.8750

Wilcoxon Two-Sample Test  
Statistic 20.5000

t Approximation  
One-Sided Pr > Z 0.2898  
Two-Sided Pr > |Z| 0.5796

Exact Test  
One-Sided Pr >= S 0.2714  
Two-Sided Pr >= |S - Mean| 0.5429

Wilcoxon Scores (Rank Sums) for Variable gesund  
Classified by Variable Behandlung

Behandlung	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
C	4	15.0	18.0	3.464102	3.750
D	4	21.0	18.0	3.464102	5.250

Wilcoxon Two-Sample Test  
Statistic 15.0000

```

t Approximation
One-Sided Pr < Z      0.2469
Two-Sided Pr > |Z|    0.4939

Exact Test
One-Sided Pr <= S      0.2429
Two-Sided Pr >= |S - Mean| 0.4857

          Wilcoxon Scores (Rank Sums) for Variable gesund
          Classified by Variable Behandlung
Behandlung      N      Sum of      Expected      Std Dev      Mean
-----
C                4        25.0        18.0        3.464102     6.250
E                4        11.0        18.0        3.464102     2.750

Wilcoxon Two-Sample Test
Statistic                25.0000

t Approximation
One-Sided Pr > Z      0.0514
Two-Sided Pr > |Z|    0.1027

Exact Test
One-Sided Pr >= S      0.0286
Two-Sided Pr >= |S - Mean| 0.0571

          Wilcoxon Scores (Rank Sums) for Variable gesund
          Classified by Variable Behandlung
Behandlung      N      Sum of      Expected      Std Dev      Mean
-----
D                4        26.0        18.0        3.464102     6.500
E                4        10.0        18.0        3.464102     2.500

Wilcoxon Two-Sample Test
Statistic                26.0000

t Approximation
One-Sided Pr > Z      0.0336
Two-Sided Pr > |Z|    0.0671

Exact Test
One-Sided Pr >= S      0.0143
Two-Sided Pr >= |S - Mean| 0.0286
    
```

Trotz bereits vorgenommener Kürzung umfasst die Ausgabe 23 noch sehr viele Informationen. Davon werden nur die für die globalen und paarweisen Testentscheidungen notwendigen Daten in die Tab. 16 und 17 übernommen. Das Signifikanzniveau ist  $\alpha = 0,05$ .

**Tab. 16:** Merkmal: Anzahl gesunde Pflanzen, globale Tests

Test	F-Test	Permutationstest	Friedman-Test
Schätzwert der Teststatistik	17,70		13,40
Überschreitungswahrscheinlichkeit	<0,0001	0	0,0095
Testentscheidung	signifikant	signifikant	signifikant

**Tab. 17:** Merkmal: Anzahl gesunde Pflanzen, paarweise Vergleiche

Test		A-B	A-C	A-D	A-E	B-C	B-D	B-E	C-D	C-E	D-E
Tukey-Test	Adj P	0,0002	0,0104	0,0503	<,0001	0,1161	0,0246	0,9746	0,8832	0,0436	0,0090
	Entscheidung	signif.	signif.		signif.		signif.			signif.	signif.
Permutationstest Tukey-Type Test	Pr>Prob	0,009	0,404	0,693	0,001	0,817	0,560	1,0	0,998	0,666	0,381
	Entscheidung	signif.			signif.						
Wilcoxon-Test (exakt)	Pr> S-Mean	0,0286	0,0286	0,0286	0,0286	0,0571	0,0571	0,5429	0,4857	0,0571	0,0433
	Entscheidung	signif.	signif.	signif.	signif.						signif.

Für die Signifikanzentscheidung bei den paarweisen Vergleichen ist der Hinweis notwendig, dass das Signifikanzniveau  $\alpha$  für den Tukey-Test und für den Permutationstest (Tukey-Type Test) ein

multiples ist. Das gilt natürlich auch für die globalen Tests. Für den Wilcoxon-Test ist es ein vergleichsbezogenes Risiko. Hier liegt ein Problem. Um auch hier eine versuchsbezogene Aussage vorzunehmen, ist eine Bonferroni-Korrektur des Signifikanzniveaus anzuraten. Das Signifikanzniveau  $\alpha_{\text{vergleichsbezogen}} = 0,05$  für die Wilcoxon-Tests ist bei  $m = 10$  Vergleiche für versuchsbezogene Aussage durch 10 zu dividieren:  $\alpha_{\text{Bonferroni}} = \frac{\alpha_{\text{vergleichsbezogen}}}{m} = 0,005$ .

Allerdings ist diese Korrektur eher eine konservative Überkorrektur, die zu weniger Signifikanzen führt. Dann gibt es bezüglich der Wilcoxon-Tests für keinen der Vergleiche einen signifikanten Unterschied.

Weitere Probleme sind, dass

- die korrekte Randomisation der Behandlungen innerhalb der Blocks wichtigste Voraussetzung für die Varianzanalyse und Tukey-Test, Permutationstests, Friedman- und Wilcoxon-Test ist,
- die auf den Daten basierenden Permutationstests in der Regel trennschärfer als Friedman- und Wilcoxon-Test sind und folglich den rangbasierten Test vorgezogen werden sollten,
- für die Varianzanalyse und multiplen Mittelwertvergleiche notwendige Voraussetzungen die Varianzhomogenität und Normalverteiltheit der Residuen sind.

Nach der Verteilung der Anzahlen gesunder Pflanzen (Ausgabe 21) wäre die Annahme der Varianzhomogenität kaum abzulehnen und wir würden bezüglich deren mittlerer Werte (Mittelwerte oder Mediane) signifikante Unterschiede der Behandlungen von A zu E und B sowie zu C erwarten und D zu E vermuten. Diese liefern zum multiplen Niveau  $\alpha = 0,05$  die parametrischen Tukey-Tests und zusätzlich noch die signifikanten Vergleiche B zu D und C zu E. Die Permutationstests vom Tukey-Type zeigen nur das als signifikant an, was offensichtlich ist: A zu B und A zu E.

Die Varianzanalyse und die Tukey-Test-Prozedur sind robust gegenüber Abweichungen von der Normalverteilung der Residuen. Varianzhomogenität ist eine wesentlich schärfere Voraussetzung. Der Anwender muss nun von Fall zu Fall entscheiden, ob und wie er Mittelwerte von Anzahlen interpretieren kann, denn (formal) berechnen kann man sie immer. Das fällt bei großen Anzahlen leichter. Der Anwender entscheidet, ob er vom Grundsatz abweichen darf/will, dass für ordinal skalierte Daten die Berechnung von Mittelwerten nicht zulässig ist und folglich auch Verfahren, die auf den Vergleich von Mittelwerten basieren, nicht eingesetzt werden können.

Für das Merkmal ‚Anzahl gesunder Pflanzen‘ (die Zählwerte sind von Null hinreichend weit entfernt und annähernd symmetrisch) scheinen Mittelwerte interpretierbar zu sein und Varianzanalyse sowie Tukey-Prozedur akzeptable Ergebnisse zu liefern. Das kann und soll aber keine Empfehlung sein.

Ein Vorschlag wäre, den Permutationstest wie beschrieben anzuwenden oder das Simulate-Verfahren einzusetzen. Dieses Simulate-Verfahren simuliert auf der Basis häufig gezogener Zufallsstichproben die Verteilung der kritischen Quantile zur Einhaltung des versuchsbezogenen Signifikanzniveaus für das zugrundegelegte Modell. Es ist auch für unbalancierte Daten und selbst bei Varianzheterogenität anwendbar. Dieses Verfahren ist zudem nicht so konservativ wie die Tukey-Prozedur. (WESTFALL u.a. 1999, PIEPHO 2006, SCHUMACHER u. WEIMER 2006, HIGH oA)

#### 4.7.3.3 Merkmal Anzahl Pflanzen mit Neuaustrieb

Das als Empfehlung (Permutationstest oder Simulate-Verfahren) genannte Vorgehen wird für beide Verfahren im Programm 24 umgesetzt. Dabei gehen wir davon aus, dass das Macro für den Permutationstest per `%include` aktiviert und die Datei *Permutation* gebildet wurde, d.h. das Programm 23 abgearbeitet wurde. Auch für dieses Merkmal ist die Ausgabe 24 gekürzt.

**Programm 24:** Arzneipfl4.sas – Testen, Merkmal: Anzahl Pflanzen mit Neuaustrieb

```

TITLE1 'Merkmal: Anzahl Pflanzen mit Neuaustrieb';
TITLE2 'Simulate-Verfahren';
PROC MIXED DATA=arzneipfl NOBOUND;
  CLASS Behandlung Block;
  MODEL mit_Neuaustrieb = Behandlung / DDFM=KenwardRoger;
  RANDOM Block;
  LSMEANS Behandlung;
  LSMEANS Behandlung / ADJUST=simulate ;
RUN;

TITLE2 'Permutationstest';
%ribdperm (g,
           pw_t,
           daten =permutation,
           versuch =Bezeichnung,
           block =Block,
           beh =NVariante,
           wert =mit_Neuaustrieb,
           nperm=10000,
           alpha=0.05);

TITLE;
    
```

**Ausgabe 24:** zu Arzneipfl4.sas (gekürzt)

Merkmal: Anzahl Pflanzen mit Neuaustrieb  
 Simulate-Verfahren

Covariance Parameter Estimates

Cov Parm	Estimate
Block	0.6667
Residual	55.0000

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Behandlung	4	12	5.64	0.0086

Least Squares Means

Effect	Behandlung	Estimate	Standard Error	DF	t Value	Pr >  t
Behandlung	A	38.2500	3.7305	15	10.25	<.0001
Behandlung	B	23.0000	3.7305	15	6.17	<.0001
Behandlung	C	27.2500	3.7305	15	7.30	<.0001
Behandlung	D	13.7500	3.7305	15	3.69	0.0022
Behandlung	E	25.2500	3.7305	15	6.77	<.0001
Behandlung	A	38.2500	3.7305	15	10.25	<.0001
Behandlung	B	23.0000	3.7305	15	6.17	<.0001
Behandlung	C	27.2500	3.7305	15	7.30	<.0001
Behandlung	D	13.7500	3.7305	15	3.69	0.0022
Behandlung	E	25.2500	3.7305	15	6.77	<.0001

Differences of Least Squares Means

Effect	Behandlung	Behandlung	Estimate	Standard Error	DF	t Value	Pr >  t	Adjustment	Adj P
Behandlung	A	B	15.2500	5.2440	12	2.91	0.0131	Simulate	0.0772
Behandlung	A	C	11.0000	5.2440	12	2.10	0.0578	Simulate	0.2797
Behandlung	A	D	24.5000	5.2440	12	4.67	0.0005	Simulate	0.0044
Behandlung	A	E	13.0000	5.2440	12	2.48	0.0290	Simulate	0.1569
Behandlung	B	C	-4.2500	5.2440	12	-0.81	0.4335	Simulate	0.9197
Behandlung	B	D	9.2500	5.2440	12	1.76	0.1032	Simulate	0.4302
Behandlung	B	E	-2.2500	5.2440	12	-0.43	0.6755	Simulate	0.9919
Behandlung	C	D	13.5000	5.2440	12	2.57	0.0243	Simulate	0.1347
Behandlung	C	E	2.0000	5.2440	12	0.38	0.7096	Simulate	0.9942
Behandlung	D	E	-11.5000	5.2440	12	-2.19	0.0487	Simulate	0.2460

```

Merkmal: Anzahl Pflanzen mit Neuaustrieb
Permutationstest
-----
Bezeichnung des Versuches, Anzahl der Simulationsschritte und
multiples Alpha
VERS      NPERM      ALPHA
Arzneipl. 10000      0.05

Globaler Test (Simulation) H0:"Keine Behandlungsunterschiede"
ERGENNIS
Pr>Perm   0.0119

0.99 Konfidenzintervall für Pr
ERGENNIS
0.0091023 0.0146977

Paarweise Vergleiche, Tukey-Typ (Simulation)

1. Differenzen der Prüfgliedmittelwerte und Grenzdifferenz
ERGENNIS
Differenzen      1      2      3      4      5
1      0.000 15.250 11.000 24.500 13.000
2      0.000  4.250  9.250  2.250
3      0.000 13.500  2.000
4      0.000 11.500
5      0.000

Grenzdifferenz      ERGENNIS
20.25

2. Überschreitungswahrscheinlichkeiten
ERGENNIS
Pr>Prob      1      2      3      4      5
1      0.289 0.667 0.002 0.463
2      0.981 0.827 0.998
3      0.413 0.999
4      0.618
5

Merkmal: Anzahl Pflanzen mit Neuaustrieb
Friedman-Test

Summary Statistics for Behandlung by mit_Neuaustrieb
Controlling for Block
Cochran-Mantel-Haenszel Statistics (Based on Rank Scores)
Statistic  Alternative Hypothesis  DF      Value      Prob
-----
1      Nonzero Correlation      1      4.0506      0.0442
2      Row Mean Scores Differ    4      11.2911      0.0235
    
```

Die wichtigsten Daten für die Testentscheidungen finden wir in Tab. 18 und 19 (multiples Signifikanzniveau:  $\alpha = 0,05$ ).

**Tab. 18:** Merkmal: Anzahl Pflanzen mit Neuaustrieb, globale Tests

Test	F-Test	Permutationstest
Schätzwert der Teststatistik	5,64	
Überschreitungswahrscheinlichkeit	0,0086	0,0119
Testentscheidung	signifikant	signifikant

**Tab. 19:** Merkmal: Anzahl Pflanzen mit Neuaustrieb, paarweise Vergleiche

Test		A-B	A-C	A-D	A-E	B-C	B-D	B-E	C-D	C-E	D-E
Simulate-Verfahren	Adj P	0,0772	0,2797	0,0044	0,1569	0,9197	0,4302	0,9919	0,1347	0,9942	0,2460
	Entscheidung			signif.							
Permutationstest Tukey-Type Test	Pr>Prob	0,289	0,667	0,002	0,463	0,981	0,827	0,998	0,413	0,999	0,618
	Entscheidung			signif.							

Die Probleme sind dieselben wie sie oben für das andere ordinal skalierte Merkmal genannt wurden.

Für die Verteilungen der Pflanzenanzahlen mit Neutrieben (vgl. Ausgabe 21) ist auch ohne Test eher nicht von gleichen Variabilitäten (in den Grundgesamtheiten) auszugehen, was besonders deutlich an den Wirkungen der Behandlungen A und C zu sehen ist. Damit sind Varianzanalyse (Varianzhomogenität!) und Tukey-Test nicht anzuwenden. Der Permutationstest vom Tukey-Type liefert zum Niveau  $\alpha = 0,05$  einen signifikanten Unterschied zwischen den Medianen der Behandlungen A und D. Zur gleichen Testentscheidung kommt auch das Simulate-Verfahren.

#### 4.7.4 Analyse von Häufigkeiten

Für die Erfassung der Krankheit einer Pflanze haben wir die beiden Möglichkeiten gesund oder krank (bzw. tot). Ein solches dichotomes oder bivariates Merkmal mit zwei Ausprägungen liegt auch für die Charakterisierung des Neuaustriebs vor (vgl. Tab. 14). Jeweils beide Ausprägungen sind nicht unabhängig voneinander, da insgesamt 45 Pflanzen auf jeder Parzelle gepflanzt wurden. Wir haben es also bei 5 Behandlungen für jedes Merkmal mit einer (5x2)-Kontingenztafel mit festen Randsummen, dem Modell 2 der Kontingenztafelanalyse, zu tun. Mit den Variablen *gesund* und *krank\_tot* sowie *mit\_Neuaustrieb* und *ohne\_Neuaustrieb* bilden wir im Programm 25 zuerst entsprechende Klassifikationsmerkmale wie wir sie beispielsweise vom t-Test her kennen. Analysiert werden die Häufigkeiten. Die mit no... beginnenden Optionen unterdrücken zusätzliche Angaben in der Kontingenztafel. Neben den Häufigkeiten sollen die erwartete Zellhäufigkeiten berechnet werden (Option *expected*). Die eigentlichen Teststatistiken liefert die Option *chisq*. Die Überschreitungswahrscheinlichkeiten werden exakt (*EXACT chisq ;*) berechnet und die damit verbundene beträchtlich erhöhten Laufzeiten in Kauf genommen. Die Ergebnisdarstellung ist in Ausgabe 25 auf die Kontingenztafel (nur für die (5x2)-Tafel), den  $\chi^2$ -Wert und seine Überschreitungswahrscheinlichkeit begrenzt.

#### Programm 25: Arzneipfl5.sas - Kontingenztafelanalyse

```

OPTIONS formchar= '|----|+|----+=|';

DATA KTA;
  SET arzneipfl;
  Krankheit      = 'gesund  ';
  ANZ_krank      = gesund;
  Neuaustrieb    = 'mit_Neuaustrieb';
  ANZ_Neuaustr  = mit_Neuaustrieb;
  OUTPUT;
  Krankheit      = 'krank/tot';
  ANZ_krank      = krank_tot;
  Neuaustrieb    = 'ohne_Neuaustrieb';
  ANZ_Neuaustr  = ohne_Neuaustrieb;
  OUTPUT;
;
TITLE1 "Kontingenztafelanalyse";
TITLE2 "Krankheit";
TITLE3 "(5*2) - Tafel";
PROC FREQ DATA=KTA;
  TABLE Behandlung * Krankheit / nocol norow nopercent expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank;
RUN;

TITLE "(2*2) - Untertafel A, B";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='B'));
  TABLE Behandlung * Krankheit / nocol norow nopercent expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank;  RUN;

TITLE "(2*2) - Untertafel A, C";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='C'));
  TABLE Behandlung * Krankheit / nocol norow nopercent expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank;  RUN;

```

```

TITLE "(2*2) - Untertafel A, D";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='D'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank; RUN;
TITLE "(2*2) - Untertafel A, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='E'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank; RUN;
TITLE "(2*2) - Untertafel B, C";
PROC FREQ DATA=KTA (WHERE=(Behandlung='B' OR Behandlung='C'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank; RUN;
TITLE "(2*2) - Untertafel B, D";
PROC FREQ DATA=KTA (WHERE=(Behandlung='B' OR Behandlung='D'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank;
RUN;
TITLE "(2*2) - Untertafel B, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='B' OR Behandlung='E'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank; RUN;
TITLE "(2*2) - Untertafel C, D";
PROC FREQ DATA=KTA (WHERE=(Behandlung='C' OR Behandlung='D'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank; RUN;
TITLE "(2*2) - Untertafel C, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='C' OR Behandlung='E'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank; RUN;
TITLE "(2*2) - Untertafel D, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='D' OR Behandlung='E'));
  TABLE Behandlung * Krankheit / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_krank; RUN;

TITLE1 "Kontingenztafelanalyse";
TITLE2 "Neuaustrieb";
TITLE3 "(5*2) - Tafel";
PROC FREQ DATA=KTA;
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr;
RUN;

TITLE "(2*2) - Untertafel A, B";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='B'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel A, C";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='C'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel A, D";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='D'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel A, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='A' OR Behandlung='E'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel B, C";
PROC FREQ DATA=KTA (WHERE=(Behandlung='B' OR Behandlung='C'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;

```

```
TITLE "(2*2) - Untertafel B, D";
PROC FREQ DATA=KTA (WHERE=(Behandlung='B' OR Behandlung='D'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel B, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='B' OR Behandlung='E'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel C, D";
PROC FREQ DATA=KTA (WHERE=(Behandlung='C' OR Behandlung='D'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel C, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='C' OR Behandlung='E'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE "(2*2) - Untertafel D, E";
PROC FREQ DATA=KTA (WHERE=(Behandlung='D' OR Behandlung='E'));
  TABLE Behandlung * Neuaustrieb / nocol norow nopercnt expected chisq ;
  EXACT chisq ;
  WEIGHT ANZ_Neuaustr; RUN;
TITLE;
```

**Ausgabe 25:** zu Arzneipfl5.sas

Kontingenztafelanalyse							
Krankheit							
(5*2) - Tafel							
Table of Behandlung by Krankheit							
Behandlung(Behandlung)							
Krankheit							
Frequency	Expected		Total				
	gesund	krank/to					
		t					
A	137 85.8	43 94.2	180				
B	56 85.8	124 94.2	180				
C	88 85.8	92 94.2	180				
D	99 85.8	81 94.2	180				
E	49 85.8	131 94.2	180				
Total	429	471	900				
				Pearson Chi-Square Test			
(2*2) - Untertafel A, B				Chi-Square		112.3068	
				DF		4	
				Asymptotic Pr > ChiSq		<.0001	
				Exact Pr >= ChiSq		2.856E-24	
				Pearson Chi-Square Test			
(2*2) - Untertafel A, C				Chi-Square		73.2822	
				DF		1	
				Asymptotic Pr > ChiSq		<.0001	
				Exact Pr >= ChiSq		7.471E-18	
				Pearson Chi-Square Test			
(2*2) - Untertafel A, D				Chi-Square		28.4563	
				DF		1	
				Asymptotic Pr > ChiSq		<.0001	
				Exact Pr >= ChiSq		1.382E-07	
				Pearson Chi-Square Test			
(2*2) - Untertafel A, D				Chi-Square		17.7638	
				DF		1	
				Asymptotic Pr > ChiSq		<.0001	
				Exact Pr >= ChiSq		3.692E-05	

(2*2) - Untertafel A, E	<pre> Pearson Chi-Square Test ----- Chi-Square          86.1402 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  7.227E-21                     </pre>																																																
(2*2) - Untertafel B, C	<pre> Pearson Chi-Square Test ----- Chi-Square          11.8519 DF                  1 Asymptotic Pr &gt; ChiSq  0.0006 Exact      Pr &gt;= ChiSq  8.249E-04                     </pre>																																																
(2*2) - Untertafel B, D	<pre> Pearson Chi-Square Test ----- Chi-Square          20.9485 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  7.049E-06                     </pre>																																																
(2*2) - Untertafel B, E	<pre> Pearson Chi-Square Test ----- Chi-Square           0.6588 DF                  1 Asymptotic Pr &gt; ChiSq  0.4170 Exact      Pr &gt;= ChiSq  0.4867                     </pre>																																																
(2*2) - Untertafel C, D	<pre> Pearson Chi-Square Test ----- Chi-Square           1.3465 DF                  1 Asymptotic Pr &gt; ChiSq  0.2459 Exact      Pr &gt;= ChiSq  0.2915                     </pre>																																																
(2*2) - Untertafel C, E	<pre> Pearson Chi-Square Test ----- Chi-Square          17.9228 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  3.406E-05                     </pre>																																																
(2*2) - Untertafel D, E	<pre> Pearson Chi-Square Test ----- Chi-Square          28.6843 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  1.247E-07                     </pre>																																																
<p>Kontingenztafelanalyse  Neuaustrieb  (5*2) - Tafel  Table of Behandlung by Neuaustrieb  Behandlung (Behandlung)  Neuaustrieb</p> <table border="1"> <thead> <tr> <th>Frequency</th> <th>mit_Neuaustrieb</th> <th>ohne_Neuaustrieb</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>153</td> <td>27</td> <td>180</td> </tr> <tr> <td></td> <td>102</td> <td>78</td> <td></td> </tr> <tr> <td>B</td> <td>92</td> <td>88</td> <td>180</td> </tr> <tr> <td></td> <td>102</td> <td>78</td> <td></td> </tr> <tr> <td>C</td> <td>109</td> <td>71</td> <td>180</td> </tr> <tr> <td></td> <td>102</td> <td>78</td> <td></td> </tr> <tr> <td>D</td> <td>55</td> <td>125</td> <td>180</td> </tr> <tr> <td></td> <td>102</td> <td>78</td> <td></td> </tr> <tr> <td>E</td> <td>101</td> <td>79</td> <td>180</td> </tr> <tr> <td></td> <td>102</td> <td>78</td> <td></td> </tr> <tr> <td>Total</td> <td>510</td> <td>390</td> <td>900</td> </tr> </tbody> </table>		Frequency	mit_Neuaustrieb	ohne_Neuaustrieb	Total	A	153	27	180		102	78		B	92	88	180		102	78		C	109	71	180		102	78		D	55	125	180		102	78		E	101	79	180		102	78		Total	510	390	900
Frequency	mit_Neuaustrieb	ohne_Neuaustrieb	Total																																														
A	153	27	180																																														
	102	78																																															
B	92	88	180																																														
	102	78																																															
C	109	71	180																																														
	102	78																																															
D	55	125	180																																														
	102	78																																															
E	101	79	180																																														
	102	78																																															
Total	510	390	900																																														
	<pre> Pearson Chi-Square Test ----- Chi-Square          112.2172 DF                  4 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  3.995E-24                     </pre>																																																

(2*2) - Untertafel A, B	<pre> Pearson Chi-Square Test ----- Chi-Square          47.5443 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  4.935E-12                     </pre>
(2*2) - Untertafel A, C	<pre> Pearson Chi-Square Test ----- Chi-Square          27.1444 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  2.519E-07                     </pre>
(2*2) - Untertafel A, D	<pre> Pearson Chi-Square Test ----- Chi-Square         109.3573 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  1.328E-26                     </pre>
(2*2) - Untertafel A, E	<pre> Pearson Chi-Square Test ----- Chi-Square          36.1551 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  2.117E-09                     </pre>
(2*2) - Untertafel B, C	<pre> Pearson Chi-Square Test ----- Chi-Square           3.2554 DF                  1 Asymptotic Pr &gt; ChiSq  0.0712 Exact      Pr &gt;= ChiSq  0.0893                     </pre>
(2*2) - Untertafel B, D	<pre> Pearson Chi-Square Test ----- Chi-Square          15.7402 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  1.068E-04                     </pre>
(2*2) - Untertafel B, E	<pre> Pearson Chi-Square Test ----- Chi-Square           0.9047 DF                  1 Asymptotic Pr &gt; ChiSq  0.3415 Exact      Pr &gt;= ChiSq  0.3979                     </pre>
(2*2) - Untertafel C, D	<pre> Pearson Chi-Square Test ----- Chi-Square          32.6580 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  1.594E-08                     </pre>
(2*2) - Untertafel C, E	<pre> Pearson Chi-Square Test ----- Chi-Square           0.7314 DF                  1 Asymptotic Pr &gt; ChiSq  0.3924 Exact      Pr &gt;= ChiSq  0.4543                     </pre>
(2*2) - Untertafel D, E	<pre> Pearson Chi-Square Test ----- Chi-Square          23.9367 DF                  1 Asymptotic Pr &gt; ChiSq  &lt;.0001 Exact      Pr &gt;= ChiSq  1.488E-06                     </pre>

Verglichen werden Wahrscheinlichkeiten. Sowohl die Überschreitungswahrscheinlichkeit für den  $\chi^2$ -Wert für die Krankheit als auch die für den Neuaustrieb weisen mit  $<0,0001$  zum Niveau  $\alpha=0,05$  auf signifikante Unterschiede zwischen den Wahrscheinlichkeiten hin. Für die (5x2)-Tafeln heißt das, dass die 5 Behandlungen in mindestens einem Fall signifikant unterschiedliche Wahrscheinlichkeiten für die Zuordnung der Pflanzen zu gesund oder krank/tot bzw. mit oder ohne Neuaustrieb bewirken.

Die jeweils 10 (2x2)-Untertafeln zum paarweisen Vergleich der Behandlungen zählen zu einem Versuch, für das das Signifikanzniveau versuchsbezogen betrachtet werden soll. Das bedeutet, dass auch hier – wie bereits auf S. 55 beschrieben – eine Bonferroni-Korrektur vorgenommen werden sollte. Der Vergleich wird folglich nicht zu 0,05, sondern zu  $\alpha_{\text{Bonferroni}} = 0,005$  vorgenommen. Die Überschreitungswahrscheinlichkeiten und Testentscheidungen für die beiden Merkmale sind in Tab. 20 und 21 zusammen gestellt.

**Tab. 20:** Überschreitungswahrscheinlichkeiten der (2x2)-Untertafeln für die  $\chi^2$ -Werte, Merkmal: Krankheit

	A-B	A-C	A-D	A-E	B-C	B-D	B-E	C-D	C-E	D-E
P <sub>asymptotisch</sub>	<0,0001	<0,0001	<0,0001	<0,0001	0,0006	<0,0001	0,4170	0,2459	<0,0001	<0,0001
P <sub>exakt</sub>	7,171 *10 <sup>-18</sup>	1,382 *10 <sup>-7</sup>	3,692 *10 <sup>-5</sup>	7,227 *10 <sup>-21</sup>	8,249 *10 <sup>-4</sup>	7,049 *10 <sup>-6</sup>	0,4867	0,2915	3,406 *10 <sup>-5</sup>	1,247 *10 <sup>-7</sup>
Entscheidung	signif.	signif.	signif.	signif.	signif.	signif.			signif.	signif.

**Tab. 21:** Überschreitungswahrscheinlichkeiten der (2x2)-Untertafeln für die  $\chi^2$ -Werte, Merkmal: Neuaustrieb

	A-B	A-C	A-D	A-E	B-C	B-D	B-E	C-D	C-E	D-E
P <sub>asymptotisch</sub>	<0,0001	<0,0001	<0,0001	<0,0001	0,0712	<0,0001	0,3415	<0,0001	0,3924	<0,0001
P <sub>exakt</sub>	4,935 *10 <sup>-12</sup>	2,519 *10 <sup>-7</sup>	1,328 *10 <sup>-26</sup>	2,117 *10 <sup>-9</sup>	0,0893	1,068 *10 <sup>-4</sup>	0,3979	1,594 *10 <sup>-8</sup>	0,4543	1,488 *10 <sup>-6</sup>
Entscheidung	signif.	signif.	signif.	signif.		signif.		signif.		signif.

Für die (2x2)-Untertafeln sind es für das Merkmal Krankheit nur die Vergleiche B und E sowie C und D für die keine signifikanten Unterschiede aufgezeigt werden konnten, für das Merkmal Neuaustrieb trifft das für die Vergleiche B und C, B und E sowie C und E zu.

## 4.8 Zusammenfassung

Die Eigenschaften - besonders das Skalenniveau - der Faktoren und/oder Merkmale sind entscheidend für die Wahl eines geeigneten Auswertungsverfahrens. Anhand zweier verschiedener Versuche stellen wir einfache parametrische Auswertungsverfahren vor. Dabei gehen wir beim Vergleich der Mittelwerte auf die wichtige Eigenschaft der Varianzhomogenität ein. Bei der Formulierung des Zusammenhangs zwischen zwei Merkmalen mit zufälligen Ausprägungen muss darauf geachtet werden, dass Zusammenhangsmaße für den linearen Zusammenhang definiert sind. Wenn andere Regressionsmodelle als das lineare zugrunde gelegt werden sollen, ist besonderer Wert auf die Modellwahl zu legen.

Bei der Analyse ordinal skalierten Merkmale haben wir das Problem, dass bevorzugte Maßzahlen wie beispielsweise Mittelwert und Varianz nicht definiert sind. Für die Versuchsanlage, eine einfaktorielles randomisierte Blockanlage, werden parameterfreie Verfahren auf der Grundlage der Daten (Permutationstest) und der Ränge der Daten (Friedman- und Wilcoxon-Test) für zwei qualitative Merkmale genutzt. Varianzanalyse und Tukey-Testprozedur werden zum Vergleich gerechnet. Der Anwender könnte sich bei einem Merkmal auch für diese parametrischen Verfahren entscheiden, wenn er die Mittelwerte als interpretierbar ansieht. Für das andere Merkmal sind diese Verfahren aufgrund offensichtlich unterschiedlicher Variabilitäten nicht geeignet. Grundsätzlich gilt, dass für ordinal skalierte Merkmale keine Mittelwerte berechnet und folglich auch keine Analyseverfahren eingesetzt werden können, die auf den Vergleich von Mittelwerten basieren.

Zum Vergleich der Häufigkeiten werden für die (m x n)-Tafeln und (2x2)-Untertafeln die  $\chi^2$ -Werte berechnet. Diese Maßzahlen können Aussagen über signifikant unterschiedliche Zuordnungen zu den Ausprägungen liefern.

## Literatur

- HIGHT, R. (oA): [www.uoregon.edu/~robinh/glm10\\_homog\\_var.txt](http://www.uoregon.edu/~robinh/glm10_homog_var.txt) , Universität Oregon
- LEVENE, H. (1960): Robust Tests for Equality of Variances  
In: OLKIN, I. et al. (Hrsg.): Contributions to Probability and Statistics  
Essays in Honor of H. Hotelling  
Stanford University Press, S. 278-292
- NÜRNBERG, G. (1982): Beiträge zur Versuchsplanung für die Schätzung von Varianzkomponenten und Robustheitsuntersuchungen zum Vergleich zweier Varianzen  
Probleme der angewandten Statistik, Heft 6  
Forschungszentrum für Tierproduktion Dummerstorf-Rostock
- PIEPHO, H.-P. (2006): Auswertung linearer Modelle mit den SAS Prozeduren GLM, MIXED, GENMOD und GLIMMIX  
Skript zum Fortbildungskurs 153 des Senats der Bundesforschungsanstalten des BMELV
- RASCH, D. und M.L. TIKU (1985): Robustness of Statistical Methods and Nonparametric Statistics  
Reidel Publ. Co., Dordrecht Boston Lancaster Tokyo
- RASCH, D. und R. VERDOOREN: Grundlagen der Korrelationsanalyse und der Regressionsanalyse  
In: MOLL, E., J. GRÖGER, M. LIESEBACH, P.E. RUDOLPH, T. STAUBER und M. ZILLER (Hrsg.): Einführung in die Biometrie, Heft 4  
2. Aufl. 2004, ISBN 3-930037-18-1
- RICHTER, Ch. (2004): Grundbegriffe und Datenanalyse  
In: MOLL, E., J. GRÖGER, M. LIESEBACH, P.E. RUDOLPH, T. STAUBER und M. ZILLER (Hrsg.): Einführung in die Biometrie, Heft 1  
2. Aufl., ISBN 3-930037-15-7
- SATTERTHWAITE, F.E. (1946): An approximate distribution of estimates of variance components  
Biometrics **2**, S. 110-114
- SCHUMACHER, E. (2004): Vergleich von mehr als zwei Parametern  
In: MOLL, E., J. GRÖGER, M. LIESEBACH, P.E. RUDOLPH, T. STAUBER und M. ZILLER (Hrsg.): Einführung in die Biometrie, Heft 3  
2. Aufl. 2004, ISBN 3-930037-17-3
- SCHUMACHER, E. und M. WEIMER (2006): Multiple Vergleiche mit der SAS<sup>®</sup>-Prozedur Mixed  
In: KAISER, K. und R.-H. BÖDEKER (Hrsg.): Statistik und Datenanalyse mit SAS<sup>®</sup>  
Proceedings der 10. Konferenz der SAS-Anwender in Forschung und Entwicklung (KSFE), 23.-24. Februar 2006, Universität Hamburg, S. 171-187
- SUMPF, D. und E. MOLL (2004): Schätzen eines Parameters und Vergleich von bis zu zwei Parametern  
In: MOLL, E., J. GRÖGER, M. LIESEBACH, P.E. RUDOLPH, T. STAUBER und M. ZILLER (Hrsg.): Einführung in die Biometrie, Heft 2  
2. Aufl. 2004, ISBN 3-930037-16-5
- WELCH, B.L. (1947): The Generalization of Student's Problem when Several Different Population Variances are Involved  
Biometrika **34**, S. 28-35
- WESTFALL, P. H., T. D. RANDALL, R. DROR, R. D. WOLFINGER, and Y. HOCHBERG (1999): Multiple Comparisons and Multiple Tests Using the SAS System<sup>®</sup>, Cary, NC: SAS Institute Inc.,

## Verzeichnis der SAS-Programme

1:	Import_Aehrenlaenge.sas	9
2:	Variable_Aehrenlaenge.sas	10
3:	Aehrenlaenge1.sas	10
4:	Grundstruktur der Prozedur <i>BOXPLOT</i>	12
5:	Aehrenlaenge2.sas	12
6:	Aehrenlaenge3.sas	12
7:	Aehrenlaenge4.sas	13
8:	Aehrenlaenge5.sas	14
9:	Maus.sas	18
10:	Levene.sas	25
11:	tTest.sas	26
12:	tTest_.sas	30
13:	Eiche1.sas	33
14:	Eiche2.sas	34
15:	Eiche3.sas	35
16:	Eiche4.sas	36
17:	V1_V2.sas	40
18:	V1_GLM.sas	42
19:	V2_GPLOT.sas	44
20:	V2_CORR.sas	46
21:	Arzneipfl1.sas – Boxplots	47
22:	Arzneipfl2.sas – Korrelationskoeffizienten	48
23:	Arzneipfl3.sas – Testen, Merkmal: Anzahl gesunder Pflanzen	50
24:	Arzneipfl4.sas – Testen, Merkmal: Anzahl Pflanzen mit Neuaustrieb	56
25:	Arzneipfl5.sas – Kontingenztafelanalyse	58

## Abbildungsverzeichnis

1:	Mit MS-Excel erfasste Ährenlängen	7
2:	Mit MS-Excel erfasste Ährenlängen	9
3:	Box-Whisker-Plot der Ährenlängen	12
4:	Häufigkeitsverteilung der Ährenlänge	13
5:	Verteilung der Ährenlängen	13
6:	Balkendiagramme mit Standardabweichungen	17
7:	Mäusegewichte in MS-Excel-Datei	18
8:	Mäusegewichte mit Standardabweichung	21
9:	Veranschaulichung der Testsituation zum Vergleich der mittleren Mäusegewichte (zweiseitige Fragestellung)	27
10:	Eingaben für den t-Test	28
11:	Eingaben für Stichprobenumfangsberechnungen	29
12:	Balkendiagramm der Trockengewichte	42
13:	Abhängigkeit von Mittelwert und Standardabweichung	42
14:	paarweise Darstellung der vier Merkmale	45

## Tabellenverzeichnis

<b>1:</b>	Urliste für die Längen von $n = 240$ Ähren ( $x_i$ in cm, $i = 1, 2, \dots, 240$ )	7
<b>2:</b>	Liste der statistischen Maßzahlen der Stichprobe, ihre Symbole und die MS-EXCEL-Funktionen	8
<b>3:</b>	Statistischen Maßzahlen, berechnet mit MS-Excel-Funktionen und SAS	10
<b>4:</b>	Wege zum Box-Whisker-Plot	11
<b>5a:</b>	Statistische Maßzahlen	16
<b>5b:</b>	Konfidenzintervall und Quartile	16
<b>6:</b>	Variationskoeffizienten verschiedener Merkmale	17
<b>7:</b>	Statistische Maßzahlen für die männlichen und weiblichen Mäusegewichte	21
<b>8:</b>	Auswahl des Hypothesenpaares	23
<b>9:</b>	Mindeststichprobenumfang in Abhängigkeit von der Güte $1-\beta$	29
<b>10:</b>	Möglichkeiten der Datenkomprimierung in Abhängigkeit von der Skalierung (RICHTER 2004, S. 79)	39
<b>11:</b>	Schematische Darstellung der aus der Wahl der Faktorstufen und/oder der Merkmalskalierung abzuleitenden statistischen Auswertungsverfahren	41
<b>12:</b>	Mittelwerte, Standardabweichungen und Varianzen von V1	41
<b>13:</b>	Darstellungen der Zusammenhänge zwischen den vier Merkmalen	45
<b>14:</b>	Anzahl der gesunden und kranken bzw. toten Pflanzen sowie der Pflanzen mit und ohne Neuaustrieb	47
<b>15:</b>	Korrelationskoeffizienten und Überschreitungswahrscheinlichkeiten	49
<b>16:</b>	Merkmal: Anzahl gesunde Pflanzen, globale Tests	54
<b>17:</b>	Merkmal: Anzahl gesunde Pflanzen, paarweise Vergleiche	54
<b>18:</b>	Merkmal: Anzahl Pflanzen mit Neuaustrieb, globale Tests	57
<b>19:</b>	Merkmal: Anzahl Pflanzen mit Neuaustrieb, paarweise Vergleiche	57
<b>20:</b>	Überschreitungswahrscheinlichkeiten der (2x2)-Untertafeln für die $\chi^2$ -Werte, Merkmal: Krankheit	63
<b>21:</b>	Überschreitungswahrscheinlichkeiten der (2x2)-Untertafeln für die $\chi^2$ -Werte, Merkmal: Neuaustrieb	63