SAS Standardprozeduren zur nichtparametrischen Datenanalyse

Edgar Brunner

Abt. Medizinische Statistik, Universität Göttingen,

Humboldt Allee 32, D-37073 Göttingen

Übersicht

1 Unverbundene Stichproben

1.1 Beispiel Relative Nierengewichte

TABELLE 1.1 *Relative Nierengewichte von 45 Wistar-Ratten einer Toxizitätsstudie.*

Relative Nierengewichte				
Placebo	Verum			
	Dosis 1	Dosis 2	Dosis 3	Dosis 4
7.11	6.23	7.40	6.65	9.26
7.08	7.93	6.51	8.11	8.62
5.95	7.59	6.85	7.37	7.72
7.36	7.14	7.17	8.43	8.54
7.58	8.03	6.76	8.21	7.88
7.39	7.31	7.69	7.14	8.44
8.25	6.91	8.18	8.25	8.02
6.95	7.52	7.05		7.72
	7.32	8.75		8.27
		7.53		7.91
				8.31

Beispiel: Relative Nierengewichte

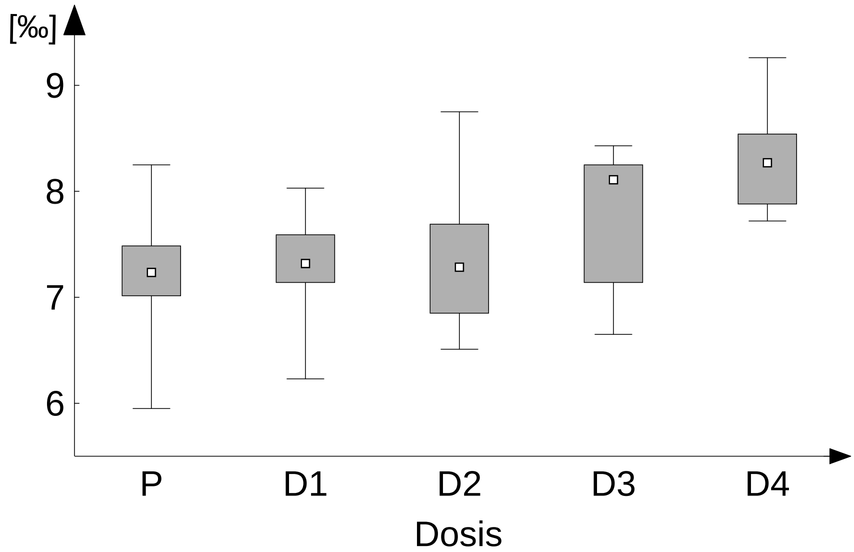

ABBILDUNG 1.1 *Box-Plots für die relativen Nierengewichte von 45 Wistar-Ratten.*

1.2 Beispiel: Corpora Lutea

TABELLE 1.2 *Anzahl der Corpora Lutea unter Placebo und einer Substanz in zwei Dosisstufen. Die Studie wurde in zwei Jahren durchgeführt.*

Gruppe	Jahr 1	Jahr 2
Placebo	13, 12, 11, 11, 14, 14, 13 13, 13	12, 16, 9, 14, 15, 12, 12 11, 13, 14, 12, 13, 12
Dosis 1	15, 12, 11, 11, 14, 13, 14 14, 12	9, 12, 11, 15, 11, 10, 13 11
Dosis 2	15, 12, 13, 14, 11, 14, 17 15	15, 13, 17, 14, 14, 13, 13 13, 9, 12, 15, 14

1.3 Nichtparametrisches Modell

$$
\begin{aligned}
X_{ik} \;\sim\; F_i(x) \;&=\; \frac{1}{2}\left[F_i^-(x) + F_i^+(x)\right] \\
& i = 1, \ldots, d, \quad k = 1, \ldots, n_i, \quad N = \sum_{i=1}^{d} n_i \,, \\
& \boldsymbol{F} = (F_1, \ldots, F_d)',
\end{aligned}
$$

$d = a$		(ein Faktor)
$d = a \cdot b$		(zwei (gekreuzte) Faktoren)
$d = a \cdot b \cdot c$		(drei (gekreuzte) Faktoren)
$\vdots$		$\vdots$

Hypothesen: $H_0^F : \boldsymbol{CF} = \boldsymbol{0}$,
$\boldsymbol{C}$: Kontrastmatrix zur Formulierung der Hypothese

Effekte und Schätzer

$$
\begin{aligned}
p_i &= \frac{1}{N}\sum_{j=1}^{d} n_j \left[P(X_{j1} < X_{i1}) + \frac{1}{2} P(X_{j1} = X_{i1}) \right] \\
\boldsymbol{p} &= (p_1, \ldots, p_d)', \\
\widehat{p}_i &= \frac{1}{N}\left(\overline{R}_{i\cdot} - \frac{1}{2} \right), \quad \overline{R}_{i\cdot} \quad \text{Mittel der Ränge in Stichprobe } i \\
\widehat{\boldsymbol{p}} &= (\widehat{p}_1, \ldots, \widehat{p}_d)'
\end{aligned}
$$

Statistik:

$$
\sqrt{N}\, \boldsymbol{C}\, \widehat{\boldsymbol{p}} = \frac{1}{\sqrt{N}} \boldsymbol{C} \begin{pmatrix} \overline{R}_{1\cdot} - \frac{1}{2} \\ \vdots \\ \overline{R}_{d\cdot} - \frac{1}{2} \end{pmatrix} = \frac{1}{\sqrt{N}} \boldsymbol{C} \begin{pmatrix} \overline{R}_{1\cdot} \\ \vdots \\ \overline{R}_{d\cdot} \end{pmatrix}
$$

Kovarianzmatrix von $\sqrt{N}\,\boldsymbol{C}\,\widehat{\boldsymbol{p}}$ unter H_0^F

$$
\begin{aligned}
\operatorname{Cov}\left(\sqrt{N}\,\boldsymbol{C}\,\widehat{\boldsymbol{p}}\right) &= \boldsymbol{C}\boldsymbol{V}_N\boldsymbol{C}', \\
\boldsymbol{V}_N &= N \cdot \operatorname{diag}\{\sigma_1^2/n_1, \ldots, \sigma_d^2/n_d\} \\
\sigma_i^2 &= \operatorname{Var}(H(X_{i1})), \quad H = \frac{1}{N}\sum_{i=1}^{d} n_i F_i\,, \\
\widehat{\sigma}_i^2 &= \frac{1}{N^2(n_i-1)}\sum_{k=1}^{n_i}\left(R_{ik} - \overline{R}_{i\cdot}\right)^2
\end{aligned}
$$

Asymptotische Verteilung unter H_0^F

$$
\sqrt{N}\,\boldsymbol{C}\,\widehat{\boldsymbol{p}} \quad \dot{\sim} \quad N(\boldsymbol{0}, \boldsymbol{C}\boldsymbol{V}_N\boldsymbol{C}')
$$

Statistiken und Verteilung unter H_0^F

1. Wald-Typ (allgemeine Alternativen)

$$Q_N \quad = \quad N \cdot \widehat{\boldsymbol{p}}' \boldsymbol{C}' [\boldsymbol{C} \widehat{\boldsymbol{V}}_N \boldsymbol{C}']^- \boldsymbol{C} \widehat{\boldsymbol{p}} \ \dot{\sim}\ \chi^2_{r(\boldsymbol{C})}$$

2. ANOVA-Typ (allgemeine Alternativen)

$$F_N \quad = \quad \frac{N}{\mathrm{Sp}(\boldsymbol{T}\widehat{\boldsymbol{V}}_N)} \widehat{\boldsymbol{p}}' \, \boldsymbol{T} \, \widehat{\boldsymbol{p}} \ \dot{\sim}\ F(\widehat{f}_1, \widehat{f}_0),$$

$$\boldsymbol{T} \ = \ \boldsymbol{C}'[\boldsymbol{C}\boldsymbol{C}']^- \boldsymbol{C},$$

$$\widehat{f}_1 = [\mathrm{Sp}(\boldsymbol{T}\widehat{\boldsymbol{V}}_N)]^2 \big/ \mathrm{Sp}(\boldsymbol{T}\widehat{\boldsymbol{V}}_N \boldsymbol{T}\widehat{\boldsymbol{V}}_N),$$

$$\widehat{f}_0 = [\mathrm{Sp}(\boldsymbol{T}\widehat{\boldsymbol{V}}_N)]^2 \big/ \mathrm{Sp}(\boldsymbol{D}_T^2 \widehat{\boldsymbol{V}}_N^2 \boldsymbol{\Lambda})$$

(Satterthwaite / Smith / Welch für σ_i^2)

3. Linearform (gemusterte Alternativen)

Hettmansperger-Norton Statistik

$$L_N \;=\; \sqrt{N}\boldsymbol{w}'\boldsymbol{C}\widehat{\boldsymbol{p}}/\widehat{\sigma}_N \;\dot{\sim}\; N(0,1),$$
$$\widehat{\sigma}_N^2 = \boldsymbol{w}'\boldsymbol{C}[\boldsymbol{C}\widehat{\boldsymbol{V}}_N\boldsymbol{C}']^-\boldsymbol{C}\boldsymbol{w}$$

Ein-Faktor Modell: Jonckheere-Terpstra Statistik

Eigenschaften

- Q_N, F_N und L_N haben
  - * unter H_0^F
  - * die *Rangtransformations- (RT-)* Eigenschaft
  - * bezüglich eines heteroskedastischen Modells.

1.4 Auswertungsschema in SAS

Daten einlesen, Ränge zuweisen

```
DATA name;
INPUT f1$ f2$ f3$ ... fk$ x;
CARDS;
f1 f2 f3 . . . x
f1 f2 f3 . . . x
.  .
.  .
;
RUN;

PROC RANK DATA=name OUT=name;
VAR x;
RANKS r;
RUN;
```

Statistiken und p-Werte

```
PROC MIXED DATA=name ANOVAF METHOD=MIVQUE0;
CLASS f1 f2 ... fk;
MODEL r = feste Faktoren / CHISQ;
REPETAED / TYPE=UN(1) GRP=f1*f2*... ;
CONTRAST 'Muster' Faktor w1...wd;
RUN;
```

Achtung:

1. $\sum_{i=1}^{d} w_d = 0$
2. bei CONTRAST: L_N^2 und der p-Wert ist zweiseitig
3. bei ESTIMATE: mit UPPER ist der p-Wert einseitig, keine ATS

Erklärung der Statements:

- **PROC MIXED DATA=**name **ANOVAF METHOD=MIVQUE0;**
- **MODEL** r = Designstruktur **/ CHISQ;**
- **REPEATED** ... **/ TYPE= ... SUB= ... GRP= ... ;**

ANOVAF	ANOVA-Typ Statistik
METHOD=MIVQUE0	verhindert Konvergenzprobleme gegenüber REML
CHISQ	Wald-Typ Statistik
TYPE = ...	Struktur der Kovarianzmatrix
SUB = ...	Faktorname der Subjects
GRP = ...	neue Schätzung der Varianz
CONTRAST	Gewichte für gemusterte Alternative $w_q, \ldots, w_d$

1.5 Auswertung der Beispiele

1.5.1 Relative Nierengewichte

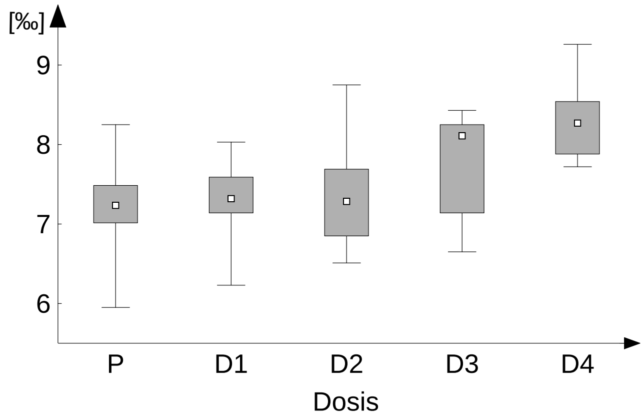

ABBILDUNG 1.1. Box-Plots für die relativen Nierengewichte von 45 Wistar-Ratten.

1.5.2 Modell, Hypothese und Statistik

Modell

$$X_{ik} \quad \sim \quad F_i(x), \quad i = 1, \ldots, a, \; k = 1, \ldots, n_i, \; N = \sum_{i=1}^{a} n_i$$

$$\boldsymbol{F} = (F_1, \ldots, F_a)', \quad \boldsymbol{p} = (p_1, \ldots, p_a)'$$

Hypothese

$$H_0^F: \qquad F_1 = \cdots = F_a \iff F_i = \overline{F}_{\cdot}\,, \quad i = 1, \ldots, a$$

$$\boldsymbol{C}_A \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_A = \boldsymbol{P}_a = \boldsymbol{I}_a - \frac{1}{a} \boldsymbol{J}_a$$

Statistik

$$\sqrt{N}\, \boldsymbol{P}_a \, \widehat{\boldsymbol{p}} \quad = \quad \frac{1}{\sqrt{N}} \, \boldsymbol{P}_a \begin{pmatrix} \overline{R}_{1\cdot} \\ \vdots \\ \overline{R}_{a\cdot} \end{pmatrix}$$

Kovarianzmatrix unter H_0^F

$$\begin{aligned}
\boldsymbol{V}_N &= N \cdot \operatorname{diag}\left\{\sigma_i^2/n_i\right\} \to N \cdot \sigma^2 \operatorname{diag}\left\{1/n_i\right\} \\
\widehat{\sigma}^2 &= \frac{1}{N^2(N-a)} \sum_{i=1}^{a} \sum_{k=1}^{n_i} (R_{ik} - \overline{R}_{i\cdot})^2 \quad - \text{ RT Test} \\
\sigma_0^2 &= \frac{1}{N^2(N-1)} \sum_{i=1}^{a} \sum_{k=1}^{n_i} [R_{ik} - (N+1)/2]^2 \\
& \qquad\qquad\qquad\qquad\qquad - \text{ Kruskal-Wallis Test}
\end{aligned}$$

SAS Prozeduren

```
DATA nierel;
INPUT gr grel;
DATALINES;
1 7.11
1 7.08
.  .
.  .
5 7.91
5 8.31
;
RUN;

PROC NPAR1WAY DATA= nierel WILCOXON; CLASS gr; VAR grel; RUN;

PROC FREQ DATA = nierel;  TABLES gr * grel / JT;  RUN;
```

RT-Test, Hettmansperger/Norton Test

```
PROC RANK DATA = nierel OUT = nierel;
VAR grel;
RANKS r;
RUN;

PROC MIXED DATA = nierel ANOVAF;
CLASS gr;
MODEL r = gr / CHISQ;
REPEATED / TYPE = simple ;
LSMEANS gr;
CONTRAST '1 2 3 4 5' gr -2 -1 0 1 2;
RUN;
```

```
The Mixed Procedure

Type 3 Tests of Fixed Effects
        Num  Den
Effect  DF   DF  Chi-Square  F Value  Pr > ChiSq  Pr > F
gr       4   40       21.30     5.32  0.0003      0.0016

         ANOVA  ANOVA
         Num    Den    ANOVA F  ANOVA
Effect   DF      DF    Value    Pr > F
gr       3.92    40    4.65     0.0037

          Contrasts
                                    ANOVA
          Num  Den                 Num Den  ANOVA F ANOVA
Label     DF   DF  F Value  Pr > F DF  DF   Value   Pr > F
1 2 3 4 5  1   40    16.10  0.0003  1  40   16.10   0.0003
```

Ergebnisse

Kruskal-Wallis-Test			
$Q_N = 15.29$	$f = 4$		$p = 0.0041$
F_N**-Test**			
$F_N = 4.65$	$\widehat{f}_1 = 3.92$	$\widehat{f}_0 = 40$	$p = 0.0037$
Jonckheere-Terpstra-Test			
$Z^{JT} = 3.56$			$p = 0.0002$
Hettmansperger-Norton-Test			
Muster $\boldsymbol{w} = (1, 2, 3, 4, 5)'$			
$L_N^2 = 16.10$	(zweiseitig)	$\widehat{f}_0 = 40$	$p = 0.0003$
$L_N = 4.01$	(einseitig)	$\widehat{f}_0 = 40$	$p = 0.00013$

1.5.3 Corpora Lutea

TABELLE 1.3 *Anzahl der Corpora Lutea unter Placebo und einer Substanz in zwei Dosisstufen. Die Studie wurde in zwei Jahren durchgeführt.*

Gruppe	Jahr 1	Jahr 2
Placebo	13, 12, 11, 11, 14, 14, 13 13, 13	12, 16, 9, 14, 15, 12, 12 11, 13, 14, 12, 13, 12
Dosis 1	15, 12, 11, 11, 14, 13, 14 14, 12	9, 12, 11, 15, 11, 10, 13 11
Dosis 2	15, 12, 13, 14, 11, 14, 17 15	15, 13, 17, 14, 14, 13, 13 13, 9, 12, 15, 14

1.5.4 Modell, Hypothesen und Statistiken

Modell

$$X_{ijk} \sim F_{ij}(x), \quad i = 1, \ldots, a, \; j = 1, \ldots, b, \; k = 1, \ldots, n_{ij},$$

$$\boldsymbol{F} = (F_{11}, \ldots, F_{ab})', \quad \boldsymbol{p} = (p_{11}, \ldots, p_{ab})', \quad N = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}$$

Hypothesen

$$\begin{aligned}
H_0^F(A): \quad & \overline{F}_{1\cdot} = \cdots = \overline{F}_{a\cdot} \iff \overline{F}_{i\cdot} = \overline{F}_{\cdot\cdot}\,, \quad i = 1, \ldots, a \\
& \boldsymbol{C}_A \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_A = \boldsymbol{P}_a \otimes \tfrac{1}{b}\boldsymbol{1}_b' \\
H_0^F(B): \quad & \overline{F}_{\cdot 1} = \cdots = \overline{F}_{\cdot b} \iff \overline{F}_{\cdot j} = \overline{F}_{\cdot\cdot}\,, \quad j = 1, \ldots, b \\
& \boldsymbol{C}_B \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_B = \tfrac{1}{a}\boldsymbol{1}_a' \otimes \boldsymbol{P}_b \\
H_0^F(AB): \quad & F_{ij} = \overline{F}_{i\cdot} + \overline{F}_{\cdot j} - \overline{F}_{\cdot\cdot}, \quad i = 1, \ldots, a, \; j = 1, \ldots, b \\
& \boldsymbol{C}_{AB} \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_{AB} = \boldsymbol{P}_a \otimes \boldsymbol{P}_b
\end{aligned}$$

Statistiken

$$\begin{aligned}
\sqrt{N}\,\boldsymbol{C}_A\,\widehat{\boldsymbol{p}} &= \frac{1}{\sqrt{N}}\left(\boldsymbol{P}_a \otimes \tfrac{1}{b}\boldsymbol{1}_b'\right)(\overline{R}_{11\cdot},\ldots,\overline{R}_{ab\cdot})', \\
\sqrt{N}\,\boldsymbol{C}_B\,\widehat{\boldsymbol{p}} &= \frac{1}{\sqrt{N}}\left(\tfrac{1}{a}\boldsymbol{1}_a' \otimes \boldsymbol{P}_b\right)(\overline{R}_{11\cdot},\ldots,\overline{R}_{ab\cdot})', \\
\sqrt{N}\,\boldsymbol{C}_{AB}\,\widehat{\boldsymbol{p}} &= \frac{1}{\sqrt{N}}\left(\boldsymbol{P}_a \otimes \boldsymbol{P}_b\right)(\overline{R}_{11\cdot},\ldots,\overline{R}_{ab\cdot})'
\end{aligned}$$

Kovarianzmatrix unter H_0^F

$$\begin{aligned}
\boldsymbol{V}_N &= N \cdot \mathrm{diag}\left\{\frac{\sigma_{11}^2}{n_{11}},\ldots,\frac{\sigma_{ab}^2}{n_{ab}}\right\} \\
\widehat{\sigma}_{ij}^2 &= \frac{1}{N^2(n_{ij}-1)}\sum_{k=1}^{n_{ij}}(R_{ijk}-\overline{R}_{ij\cdot})^2
\end{aligned}$$

SAS Prozeduren

Dateneingabe

```
DATA cl;
INPUT jahr beh clut;
CARDS;
     1  1 13
     1  1 12
     .  . .
     .  . .
     2  3 15
     2  3 14
;
RUN;
```

Ränge, Statistiken und p-Werte

```
PROC RANK DATA=cl OUT=cl;
VAR clut;
RANKS r;
RUN;

PROC MIXED DATA=cl ANOVAF METHOD=MIVQUE0;
CLASS jahr beh;
MODEL r = jahr | beh / CHISQ;
REPEATED / TYPE=UN(1) GRP=jahr*beh;
CONTRAST ’1 2 3’ beh -1 0 1;
RUN;
```

Ergebnisse

The Mixed Procedure

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
jahr	1	43.2	1.28	1.28	0.2573	0.2635
beh	2	29.5	6.91	3.38	0.0316	0.0477
jahr*beh	2	29.5	1.78	0.87	0.4116	0.4302

Effect	ANOVA Num DF	ANOVA Den DF	ANOVA F Value	ANOVA Pr > F
jahr	1	43.2	1.28	0.2635
beh	1.96	43.2	3.80	0.0310
jahr*beh	1.96	43.2	0.92	0.4026

Contrasts

Label	ANOVA Num DF	Den DF	F Value	Pr > F	Num DF	Den DF	ANOVA F Value	ANOVA Pr > F
1 2 3	1	29.7	3.93	0.0567	1	29.7	3.93	0.0567

2 Repeated Measures

2.1 Beispiel Panik-Anfall Studie

- 16 Patienten mit Panikstörung und Agoraphobie (PDA)
- 8 Wochen mit dem Antidepressivum Imipramin behandelt
- Zielgröße: Clinical Global Impression (CGI)
  - Skala mit diskreten Scores 2 bis 8
- beobachtet: zu 5 Zeitpunkten: 0, 2, 4, 6, 8 (Wochen)

TABELLE 2.1 *Score-Werte der 'Clinical Global Impression' von 16 Patienten mit einer Panik-Störung.*

CGI-Scores											
	Woche						Woche				
Patient	0	2	4	6	8	Patient	0	2	4	6	8
1	8	6	5	5	4	9	5	4	3	3	2
2	8	6	5	4	2	10	8	6	5	5	4
3	6	5	5	4	2	11	7	6	5	4	2
4	6	6	6	5	5	12	6	5	5	4	2
5	7	6	6	6	6	13	6	6	6	5	5
6	8	7	3	2	2	14	8	6	6	6	6
7	7	6	7	3	3	15	8	7	4	2	2
8	6	4	5	3	3	16	7	6	7	3	3

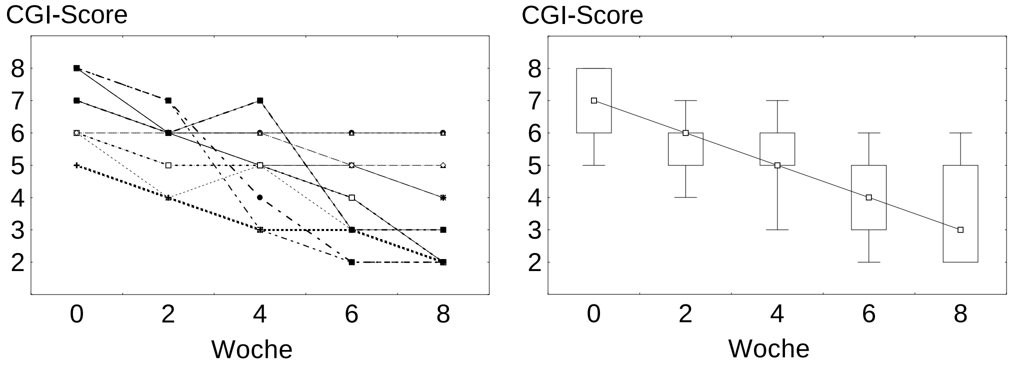

ABBILDUNG 2.1 *Verlaufskurven und Box-Plots der CGI-Scores für die 16 Patienten der Panik-Anfall Studie.*

2.2 Beispiel Schulter-Schmerz Studie

- 41 Patienten
- 2 Behandlungen (Y / N) - Faktor A
- 2 Geschlechter (M / F) - Faktor B
- 3 Tage (1 / 2 / 3) - Faktor T
- 2 Zeitpunkte (M / A) - Faktor Z
- Beobachtung: Schmerz-Score (1,2,3,4,5) nach laparoskopischer OP

TABELLE 2.2 *Schmerz-Scores von 41 Patienten nach laparoskopischer OP.*

Schmerz-Score															
Behandlung Y								Behandlung N							
		Tag								Tag					
Pat.		1		2		3		Pat.		1		2		3	
Nr.	Geschl.	M	A	M	A	M	A	Nr.	Geschl.	M	A	M	A	M	A
1	F	1	1	1	1	1	1	23	F	5	2	3	5	5	4
3	F	3	2	2	2	1	1	24	F	1	5	3	4	5	3
4	F	1	1	1	1	1	1	25	F	4	4	4	4	1	1
5	F	1	1	1	1	1	1	28	F	3	4	3	3	3	2
8	F	2	2	1	1	1	1	30	F	1	1	1	1	1	1
9	F	1	1	1	1	1	1	33	F	1	3	2	2	1	1
10	F	3	1	1	1	1	1	34	F	2	2	3	4	2	2
12	F	2	1	1	1	1	2	35	F	2	2	1	3	3	2
16	F	1	1	1	1	1	1	36	F	1	1	1	1	1	1
18	F	2	1	1	1	1	1	38	F	5	5	5	4	3	3
19	F	4	4	2	4	2	2	40	F	5	4	4	4	2	2
20	F	4	4	4	2	1	1								
21	F	1	1	1	2	1	1								
22	F	1	1	1	2	1	2								
2	M	3	2	1	1	1	1	26	M	4	4	4	4	4	3
6	M	1	2	1	1	1	1	27	M	2	3	4	3	3	2
7	M	1	3	2	1	1	1	29	M	3	3	4	4	4	3
11	M	1	1	1	1	1	1	31	M	1	1	1	1	1	1
13	M	1	2	2	2	2	2	32	M	1	5	5	5	4	3
14	M	3	1	1	1	3	3	37	M	1	1	1	1	1	1
15	M	2	1	1	1	1	1	39	M	3	3	3	3	1	1
17	M	1	1	1	1	1	1	41	M	1	3	3	3	2	1

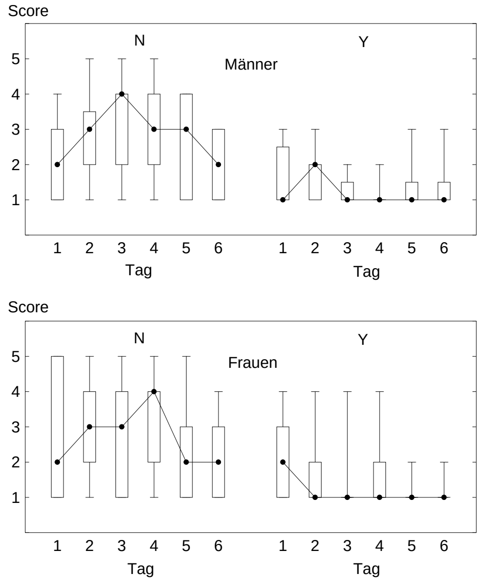

ABBILDUNG 2.2 *Verlauf und Box-Plots der Schmerz-Scores.*

2.3 Nichtparametrisches Marginal Modell

2.3.1 Modell, Hypothesen und Statistiken

Modell
Unabhängige Zufallsvektoren (t-dimensional)

$$\begin{aligned}
\boldsymbol{X}_{ik} \quad &= \quad (X_{ik1}, \ldots, X_{ikt})', \\
& \quad k = 1, \ldots, n_i \quad \text{- Individuen}, \\
& \quad i = 1, \ldots, a \quad \text{- Versuchsgruppen}, \\
& \quad n = \sum_{i=1}^{a} n_i \quad \text{- Anzahl aller Individuen}, \\
& \quad N = n \cdot t \quad \text{- Anzahl aller Beobachtungen}.
\end{aligned}$$

Randverteilungen

$$X_{iks} \;\sim\; F_{is}(x), \quad \boldsymbol{F} = (F_{11}, \ldots, F_{1t}, \ldots, F_{a1}, \ldots, F_{at})'$$

Hypothesen

$$\begin{aligned}
H_0^F(A): &\quad \overline{F}_{1\cdot} = \cdots = \overline{F}_{a\cdot} \iff \overline{F}_{i\cdot} = \overline{F}_{\cdot\cdot}\,, \quad i = 1, \ldots, a \\
&\quad \boldsymbol{C}_A \boldsymbol{F} = \mathbf{0}, \quad \boldsymbol{C}_A = \boldsymbol{P}_a \otimes \tfrac{1}{t}\mathbf{1}_t' \\
H_0^F(B): &\quad \overline{F}_{\cdot 1} = \cdots = \overline{F}_{\cdot t} \iff \overline{F}_{\cdot j} = \overline{F}_{\cdot\cdot}\,, \quad j = 1, \ldots, t \\
&\quad \boldsymbol{C}_T \boldsymbol{F} = \mathbf{0}, \quad \boldsymbol{C}_T = \tfrac{1}{a}\mathbf{1}_a' \otimes \boldsymbol{P}_t \\
H_0^F(AB): &\quad F_{ij} = \overline{F}_{i\cdot} + \overline{F}_{\cdot j} - \overline{F}_{\cdot\cdot}, \quad i = 1, \ldots, a,\; j = 1, \ldots, t \\
&\quad \boldsymbol{C}_{AT} \boldsymbol{F} = \mathbf{0}, \quad \boldsymbol{C}_{AT} = \boldsymbol{P}_a \otimes \boldsymbol{P}_t
\end{aligned}$$

Effekte und Schätzer

$$
\begin{aligned}
p_{is} &= \frac{1}{N}\sum_{j=1}^{a}\sum_{r=1}^{t} n_j \left[P(X_{j1r} < X_{i2s}) + \frac{1}{2}P(X_{j1r} = X_{i2s})\right] \\
&\quad \boldsymbol{p}_i \;=\; (p_{i1}, \ldots, p_{it})', \quad \boldsymbol{p} \;=\; (\boldsymbol{p}_1', \ldots, \boldsymbol{p}_a')', \\
\widehat{p}_{is} &= \frac{1}{N}\left(\overline{R}_{i\cdot s} - \frac{1}{2}\right), \quad \text{Mittel der Ränge in Gruppe } i \text{ zum Zeitpunt } s, \\
&\quad \widehat{\boldsymbol{p}}_i \;=\; (\widehat{p}_{i1}, \ldots, \widehat{p}_{it})' \;=\; \frac{1}{N}\left(\overline{\boldsymbol{R}}_{i\cdot} - \frac{1}{2}\boldsymbol{1}_t\right), \quad \overline{\boldsymbol{R}}_{i\cdot} \;=\; (\overline{R}_{i\cdot 1}, \ldots, \overline{R}_{i\cdot t})', \\
&\quad \widehat{\boldsymbol{p}} \;=\; (\widehat{\boldsymbol{p}}_1', \ldots, \widehat{\boldsymbol{p}}_a')'
\end{aligned}
$$

Statistik

$$
\sqrt{n}\,\boldsymbol{C}\,\widehat{\boldsymbol{p}} \;=\; \frac{1}{\sqrt{n}}\,\boldsymbol{C}\begin{pmatrix} \overline{\boldsymbol{R}}_{1\cdot} \\ \vdots \\ \overline{\boldsymbol{R}}_{a\cdot} \end{pmatrix}
$$

Schätzung der Kovarianzmatrix von $\sqrt{n}\,\boldsymbol{C}\,\widehat{\boldsymbol{p}}$ unter H_0^F

$$\widehat{\boldsymbol{V}}_n \;=\; \bigoplus_{i=1}^{a} \frac{n}{n_i}\widehat{\boldsymbol{V}}_{n,i} \;=\; n \cdot \begin{pmatrix} \frac{1}{n_1}\widehat{\boldsymbol{V}}_{n,1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \frac{1}{n_a}\widehat{\boldsymbol{V}}_{n,a} \end{pmatrix}$$

$$\widehat{\boldsymbol{V}}_{n,i} \;=\; \frac{1}{N^2(n_i-1)} \sum_{k=1}^{n_i} (\boldsymbol{R}_{ik} - \overline{\boldsymbol{R}}_{i\cdot})(\boldsymbol{R}_{ik} - \overline{\boldsymbol{R}}_{i\cdot})' \quad \text{- empirische Kovarianzmatrix}$$

$$\boldsymbol{R}_{ik} \;=\; (R_{ik1}, \ldots, R_{ikt})', \quad \overline{\boldsymbol{R}}_{i\cdot} \;=\; \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{R}_{ik}$$

Asymptotische Verteilung unter H_0^F

$$\sqrt{n}\,\boldsymbol{C}\,\widehat{\boldsymbol{p}} \quad \dot{\sim} \quad N(\mathbf{0}, \boldsymbol{C}\boldsymbol{V}_n\boldsymbol{C}')$$

Statistiken und Verteilung unter H_0^F

1. Wald-Typ (allgemeine Alternativen)

$$Q_n \;=\; n \cdot \widehat{\boldsymbol{p}}'\boldsymbol{C}'[\boldsymbol{C}\widehat{\boldsymbol{V}}_n\boldsymbol{C}']^-\boldsymbol{C}\widehat{\boldsymbol{p}} \;\dot{\sim}\; \chi^2_{r(\boldsymbol{C})}$$

2. ANOVA-Typ (allgemeine Alternativen)

$$F_n \;=\; \frac{n}{\mathrm{Sp}(\boldsymbol{T}\widehat{\boldsymbol{V}}_n)}\,\widehat{\boldsymbol{p}}'\,\boldsymbol{T}\,\widehat{\boldsymbol{p}} \;\dot{\sim}\; F(\widehat{f},\infty),$$

$$\boldsymbol{T} \;=\; \boldsymbol{C}'[\boldsymbol{C}\boldsymbol{C}']^-\boldsymbol{C}, \quad \widehat{f} \;=\; [\mathrm{Sp}(\boldsymbol{T}\widehat{\boldsymbol{V}}_n)]^2 \big/ \mathrm{Sp}(\boldsymbol{T}\widehat{\boldsymbol{V}}_n\boldsymbol{T}\widehat{\boldsymbol{V}}_n)$$

3. Linearform (gemusterte Alternativen)

$$L_n \;=\; \sqrt{n}\boldsymbol{w}'\boldsymbol{C}\widehat{\boldsymbol{p}}/\widehat{\sigma}_n \;\dot{\sim}\; N(0,1), \qquad \widehat{\sigma}_n^2 \;=\; \boldsymbol{w}'\boldsymbol{C}[\boldsymbol{C}\widehat{\boldsymbol{V}}_n\boldsymbol{C}']^-\boldsymbol{C}\boldsymbol{w}$$

Eigenschaften

- Q_n, F_n und L_n haben die RT-Eigenschaft unter H_0^F bezüglich eines multivariaten **heteroskedastischen** Modells.

2.4 Schema zur Repeated Measures Analyse

Daten einlesen, Ränge zuweisen und sortieren

```
DATA name;
INPUT indiv f1$ f2$ f3$ ... fk$ x1 ... xt;
      ARRAY xx{t} x1-xt;
         DO time=1 to t;  x = xx{time};  OUTPUT;  END;
      DROP x1-xt;
CARDS;
1 f1 f2 f3 . . . x1 x2 ... xt
2 f1 f2 f3 . . . x1 x2 ... xt
.  .
.  .
;
RUN;

PROC RANK DATA=name OUT=name; VAR x; RANKS r; RUN;
PROC SORT DATA=name OUT=name; BY indiv time;  RUN;
```

Statistiken und p-Werte

```
PROC MIXED DATA=name ANOVAF METHOD=MIVQUE0;
CLASS f1 f2 ... fk;
MODEL r = feste Faktoren / CHISQ;
REPETAED time / TYPE=UN SUB=indiv GRP=f1*f2*... ;
LSMEANS f1*f2* ... ;
CONTRAST ’Muster’ Faktor(en) w1...wd;
RUN;
```

Anmerkung

- Bei Verwendung von PROC MIXED immer zuerst
  - nach den Individuen (subjects)
  - dann nach den Sub-Plot Faktoren (Zeit)
- **sortieren**. - Ergebnisse hängen von der Reihenfolge ab, in der die Daten eingegeben werden.

Drei wichtige Statements:

- **PROC MIXED DATA=**name **ANOVAF METHOD=MIVQUE0;**
- **MODEL** r = Designstruktur **/ CHISQ;**
- **REPEATED** time **/ TYPE=... SUB=... GRP=... ;**

ANOVAF	ANOVA-Typ Statistik
METHOD=MIVQUE0	verhindert Konvergenzprobleme gegenüber REML
CHISQ	Wald-Typ Statistik
TYPE=UN	keine Struktur der Kovarianzmatrix (unstructured)
SUB=...	Faktorname der Subjects
GRP=...	neue Schätzung der Kovarianzmatrix

Unterschiede

- Bei allen Effekten, welche die Zeit (mit mehr als zwei Zeitpunkten, d,h. Num DF > 1) beinhalten, liefert PROC MIXED konservative p-Werte, da
  - 2. Freiheitsgrad aus ddfm=kr übernommen
  - anstatt ∞ gesetzt
  - kann durch Ausrechnen des p-Wertes in einem getrennten DATA-Step behoben werden.

Beispiel:

	Value	Num DF	Den DF
time	4.08	2.89	$103 \to \infty$
beh * time	3.77	2.89	$103 \to \infty$

Anmerkung $F(x|f_1, \infty) = \chi^2(f_1 \cdot x|f_1)$

```
DATA pval;
INPUT value numdf;
p = 1 - PROBCHI(numdf*value,numdf);
DATALINES;
4.08 2.89
3.77 2.89
;
RUN;

PROC PRINT DATA=pval; RUN;

Obs    value    numdf        p
1      4.08     2.89    0.007319
2      3.77     2.89    0.011114
```

2.5 Auswertung der Beispiele

2.5.1 Panik-Anfall Studie

CGI-Scores											
	Woche						Woche				
Patient	0	2	4	6	8	Patient	0	2	4	6	8
1	8	6	5	5	4	9	5	4	3	3	2
2	8	6	5	4	2	10	8	6	5	5	4
3	6	5	5	4	2	11	7	6	5	4	2
4	6	6	6	5	5	12	6	5	5	4	2
5	7	6	6	6	6	13	6	6	6	5	5
6	8	7	3	2	2	14	8	6	6	6	6
7	7	6	7	3	3	15	8	7	4	2	2
8	6	4	5	3	3	16	7	6	7	3	3

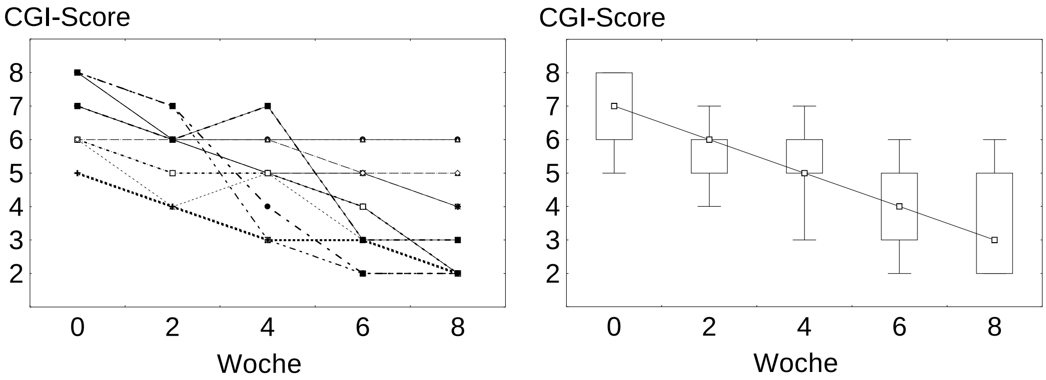

ABBILDUNG 2.1 *Verlaufskurven und Box-Plots für die Daten der Panik-Score Studie.*

Modell: eine Gruppe von Individuen, t Zeitpunkte (LD-F1)

Unabhängige Zufallsvektoren

$$\boldsymbol{X}_k \quad = \quad (X_{k1}, \ldots, X_{kt})', \quad k = 1, \ldots, n$$

Randverteilungen

$$X_{ks} \quad \sim \quad F_s(x), \qquad \boldsymbol{F} = (F_1, \ldots, F_t)'$$

Hypothese

$$\begin{aligned} H_0^F: \quad & F_1 = \cdots = F_a \iff F_i = \overline{F}_\cdot\,, \quad i = 1, \ldots, a \\ & \boldsymbol{C}_A \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_A = \boldsymbol{P}_a = \boldsymbol{I}_a - \tfrac{1}{a}\boldsymbol{J}_a \end{aligned}$$

Statistik $\sqrt{n}\, \boldsymbol{P}_a\, \widehat{\boldsymbol{p}}$

Kovarianzmatrix unter H_0^F: unstrukturiert

$$\widehat{\boldsymbol{V}}_n \quad = \quad \frac{1}{N^2(n-1)} \sum_{k=1}^{n} (\boldsymbol{R}_k - \overline{\boldsymbol{R}}_\cdot)(\boldsymbol{R}_k - \overline{\boldsymbol{R}}_\cdot)', \quad \text{- empirische Kovarianzmatrix}$$

SAS Prozeduren
Dateneingabe

```
DATA pssi;
INPUT pat s1-s5;
ARRAY tt{5} s1-s5;
 DO time = 1 TO 5; score=tt{time}; OUTPUT; END;
DROP s1-s5;
DATALINES;
 1 8 6 5 5 4
 2 8 6 5 4 2
 . . . . . .
15 8 7 4 2 2
16 7 6 7 3 3
;
RUN;
```

Sortieren, Ränge zuweisen, Statistiken

```
PROC SORT DATA=pssi out=pssi; BY pat time; RUN;

PROC RANK DATA=pssi OUT=pssi; VAR score; RANKS r; RUN;

PROC MIXED DATA=pssi ANOVAF METHOD=MIVQUE0;
CLASS time;
MODEL r = time / CHISQ;
REPEATED time / TYPE=UN SUB=pat;
CONTRAST ’5 4 3 2 1’ time 2 1 0 -1 -2;
RUN;
```

Ergebnisse

```
The Mixed Procedure
Type 3 Tests of Fixed Effects
       Num Den
Effect DF  DF  Chi-Square  F Value  Pr>ChiSq  Pr>F
time    4  15  126.69      31.67    <.0001    <.0001

       ANOVA  ANOVA
       Num    Den     ANOVA F   ANOVA
Effect DF     DF      Value     Pr > F
time   2.23   33.5    36.94     <.0001 -->> 0.00000

             Contrasts
                                          ANOVA
             Num  Den                    Num Den  ANOVA F    ANOVA
Label        DF   DF  F Value  Pr > F   DF  DF      Value   Pr > F
5 4 3 2 1     1   15    70.17  <.0001    1  15      70.17   0.0000
```

```
DATA pvals;
INPUT fn numdf confn conddf;
pfn = 1 - PROBCHI(numdf*fn,numdf);
lin = SQRT(confn);
pcon = 1 - PROBT(lin,conddf);
DATALINES;
36.94  2.23  70.17  15
;
RUN;

PROC PRINT DATA=pvals;
VAR fn numdf pfn lin conddf pcon; RUN;

Obs       fn  numdf      pfn     lin  conddf          pcon
1      36.94   2.23   0.0000    8.38      15   0.000000243
```

2.5.2 Schulter-Schmerz Studie (F1-LD-F1)

- 41 Patienten
- 2 Behandlungen (Y / N) - Faktor A (whole-plot)
  d.h. nicht nach 'Geschlecht' geschichtet
- 6 Zeitpunkte (1 / 2 / 3 / 4 / 5 / 6) - Faktor T (sub-plot)
  d.h. keine Strukturierung nach Tag / Tageszeit
- Beobachtung: Schmerz-Score (1,2,3,4,5) nach laparoskopischer OP

Schmerz-Score													
Behandlung Y							Behandlung N						
Pat.- Nr.	Zeitpunkt 1	2	3	4	5	6	Pat.- Nr.	Zeitpunkt 1	2	3	4	5	6
1	1	1	1	1	1	1	23	5	2	3	5	5	4
3	3	2	2	2	1	1	24	1	5	3	4	5	3
4	1	1	1	1	1	1	25	4	4	4	4	1	1
5	1	1	1	1	1	1	28	3	4	3	3	3	2
8	2	2	1	1	1	1	30	1	1	1	1	1	1
9	1	1	1	1	1	1	33	1	3	2	2	1	1
10	3	1	1	1	1	1	34	2	2	3	4	2	2
12	2	1	1	1	1	2	35	2	2	1	3	3	2
16	1	1	1	1	1	1	36	1	1	1	1	1	1
18	2	1	1	1	1	1	38	5	5	5	4	3	3
19	4	4	2	4	2	2	40	5	4	4	4	2	2
20	4	4	4	2	1	1							
21	1	1	1	2	1	1							
22	1	1	1	2	1	2							
2	3	2	1	1	1	1	26	4	4	4	4	4	3
6	1	2	1	1	1	1	27	2	3	4	3	3	2
7	1	3	2	1	1	1	29	3	3	4	4	4	3
11	1	1	1	1	1	1	31	1	1	1	1	1	1
13	1	2	2	2	2	2	32	1	5	5	5	4	3
14	3	1	1	1	3	3	37	1	1	1	1	1	1
15	2	1	1	1	1	1	39	3	3	3	3	1	1
17	1	1	1	1	1	1	41	1	3	3	3	2	1

2.5.3 Modell, Hypothesen und Statistiken

Modell
Unabhängige Zufallsvektoren (t-dimensional, $t = 6$)

$$\begin{aligned} \boldsymbol{X}_{ik} &= (X_{ik1}, \ldots, X_{ikt})', \\ & k = 1, \ldots, n_i \quad \text{- Individuen} \\ & i = 1, 2 \quad \text{- Versuchsgruppen} \\ & n = n_1 + n_2 \quad \text{- Anzahl aller Individuen} \\ & N = n \cdot t \quad \text{- Anzahl aller Beobachtungen} \end{aligned}$$

Randverteilungen

$$X_{iks} \sim F_{is}(x), \quad \boldsymbol{F} = (F_{11}, \ldots, F_{16}, F_{21}, \ldots, F_{26})'$$

Hypothesen

$$
\begin{aligned}
H_0^F(A): \quad & \overline{F}_{1\cdot} = \overline{F}_{2\cdot} \iff \overline{F}_{i\cdot} = \overline{F}_{\cdot\cdot}\,, \quad i = 1,2 \\
& \boldsymbol{C}_A \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_A = \boldsymbol{P}_2 \otimes \tfrac{1}{6}\boldsymbol{1}_6' \\
H_0^F(B): \quad & \overline{F}_{\cdot 1} = \cdots = \overline{F}_{\cdot 6} \iff \overline{F}_{\cdot j} = \overline{F}_{\cdot\cdot}\,, \quad j = 1, \ldots, 6 \\
& \boldsymbol{C}_T \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_T = \tfrac{1}{2}\boldsymbol{1}_2' \otimes \boldsymbol{P}_6 \\
H_0^F(AB): \quad & F_{ij} = \overline{F}_{i\cdot} + \overline{F}_{\cdot j} - \overline{F}_{\cdot\cdot}, \quad i = 1,2,\ j = 1, \ldots, 6 \\
& \boldsymbol{C}_{AT} \boldsymbol{F} = \boldsymbol{0}, \quad \boldsymbol{C}_{AT} = \boldsymbol{P}_2 \otimes \boldsymbol{P}_6
\end{aligned}
$$

Statistik $\sqrt{n}\,\boldsymbol{C}\,\widehat{\boldsymbol{p}} \;=\; \dfrac{1}{2\sqrt{n}}\boldsymbol{C} \begin{pmatrix} \overline{R}_{1\cdot 1} \\ \vdots \\ \overline{R}_{2\cdot 6} \end{pmatrix}$

Kovarianzmatrix unter H_0^F: $\boldsymbol{C}\boldsymbol{V}_n\boldsymbol{C}'$

$$\begin{aligned}
\widehat{\boldsymbol{V}}_n &= \bigoplus_{i=1}^{2} \frac{n}{n_i}\widehat{\boldsymbol{V}}_{n,i} = \begin{pmatrix} \frac{n}{n_1}\widehat{\boldsymbol{V}}_{n,1} & \mathbf{0} \\ \mathbf{0} & \frac{n}{n_2}\widehat{\boldsymbol{V}}_{n,2} \end{pmatrix} \\
\widehat{\boldsymbol{V}}_{n,i} &= \frac{1}{N^2(n_i-1)}\sum_{k=1}^{n_i}(\boldsymbol{R}_{ik} - \overline{\boldsymbol{R}}_{i\cdot})(\boldsymbol{R}_{ik} - \overline{\boldsymbol{R}}_{i\cdot})'
\end{aligned}$$

– empirische Kovarianzmatrix - unstrukturiert

SAS Prozeduren

Dateneingabe

```
DATA sss;
INPUT pat beh$ sex$ alt t1-t6;
ARRAY ss{6} t1-t6;
 DO time=1 TO 6; score=ss{time}; OUTPUT; END;
DROP t1-t6;
DATALINES;
 1  Y  F  64  1  1  1  1  1  1
 3  Y  F  77  3  2  2  2  1  1
.   .  .   .  .  .  .  .  .  .
39  N  M  72  3  3  3  3  1  1
41  N  M  61  1  3  3  3  2  1
;
RUN;
```

Sortieren, Ränge bilden, Statistiken

```
PROC RANK DATA=sss OUT=sss; VAR score; RANKS r; RUN;

PROC SORT DATA=sss OUT=sss; BY pat time; RUN;

PROC MIXED DATA=sss ANOVAF METHOD=MIVQUE0;
CLASS beh time;
MODEL r = beh|time / CHISQ;
REPEATED time / TYPE=UN SUB=pat GRP=beh;
RUN;
```

Ergebnisse

```
The Mixed Procedure
         Type 3 Tests of Fixed Effects
         Num Den
Effect   DF  DF     Chi-Square  F Value  Pr>ChiSq  Pr>F
beh      1   28.6   17.86       17.86    <.0001    0.0002
time     5   33.5   18.91        3.39    0.0020    0.0139
beh*time 5   33.5   25.46        4.56    0.0001    0.0028

         ANOVA  ANOVA
         Num    Den    ANOVA F  ANOVA
Effect   DF     DF     Value    Pr>F
beh      1      28.6   17.86    0.0002
time     2.89  103      4.08    0.0096 -->> 0.007319
beh*time 2.89  103      3.77    0.0139 -->> 0.011114
```

```
DATA pval;
INPUT value numdf;
p = 1 - PROBCHI(numdf*value,numdf);
DATALINES;
4.08 2.89
3.77 2.89
;
RUN;

PROC PRINT DATA=pval; RUN;

Obs    value    numdf        p
1      4.08     2.89    0.007319
2      3.77     2.89    0.011114
```

2.5.4 Schulter-Schmerz Studie (F2-LD-F1)

- 41 Patienten
- 2 Behandlungen (Y / N) - Faktor A (whole-plot)
- 2 Geschlechter (M / F) - Faktor B (whole-plot)
- 6 Zeitpunkte (1 / 2 / 3 / 4 / 5 / 6) - Faktor T (sub-plot)
  d.h. keine Strukturierung nach Tag / Tageszeit
- Beobachtung: Schmerz-Score (1,2,3,4,5) nach laparoskopischer OP

Schmerz-Score															
Behandlung Y								Behandlung N							
Pat.-		Zeitpunkt						Pat.-		Zeitpunkt					
Nr.	Geschl.	1	2	3	4	5	6	Nr.	Geschl.	1	2	3	4	5	6
1	F	1	1	1	1	1	1	23	F	5	2	3	5	5	4
3	F	3	2	2	2	1	1	24	F	1	5	3	4	5	3
4	F	1	1	1	1	1	1	25	F	4	4	4	4	1	1
5	F	1	1	1	1	1	1	28	F	3	4	3	3	3	2
8	F	2	2	1	1	1	1	30	F	1	1	1	1	1	1
9	F	1	1	1	1	1	1	33	F	1	3	2	2	1	1
10	F	3	1	1	1	1	1	34	F	2	2	3	4	2	2
12	F	2	1	1	1	1	2	35	F	2	2	1	3	3	2
16	F	1	1	1	1	1	1	36	F	1	1	1	1	1	1
18	F	2	1	1	1	1	1	38	F	5	5	5	4	3	3
19	F	4	4	2	4	2	2	40	F	5	4	4	4	2	2
20	F	4	4	4	2	1	1								
21	F	1	1	1	2	1	1								
22	F	1	1	1	2	1	2								
2	M	3	2	1	1	1	1	26	M	4	4	4	4	4	3
6	M	1	2	1	1	1	1	27	M	2	3	4	3	3	2
7	M	1	3	2	1	1	1	29	M	3	3	4	4	4	3
11	M	1	1	1	1	1	1	31	M	1	1	1	1	1	1
13	M	1	2	2	2	2	2	32	M	1	5	5	5	4	3
14	M	3	1	1	1	3	3	37	M	1	1	1	1	1	1
15	M	2	1	1	1	1	1	39	M	3	3	3	3	1	1
17	M	1	1	1	1	1	1	41	M	1	3	3	3	2	1

SAS Prozeduren

Dateneingabe

```
DATA sss;
INPUT pat beh$ sex$ alt t1-t6;
ARRAY ss{6} t1-t6;
 DO time=1 TO 6; score=ss{time}; OUTPUT; END;
DROP t1-t6;
DATALINES;
 1  Y  F  64  1  1  1  1  1  1
 3  Y  F  77  3  2  2  2  1  1
.   .  .   .  .  .  .  .  .  .
39  N  M  72  3  3  3  3  1  1
41  N  M  61  1  3  3  3  2  1
;
RUN;
```

Sortieren, Ränge bilden, Statistiken

```
PROC RANK DATA=sss OUT=sss; VAR score; RANKS r; RUN;

PROC SORT DATA=sss OUT=sss; BY pat time; RUN;

PROC MIXED DATA=sss ANOVAF METHOD=MIVQUE0;
CLASS beh sex time;
MODEL r = beh|sex|time / CHISQ;
REPEATED time / TYPE=UN SUB=pat GRP=beh*sex;
RUN;
```

Ergebnisse

```
           The Mixed Procedure
           Type 3 Tests of Fixed Effects
            Num Den
Effect       DF   DF  Chi-Square  F Value  Pr>ChiSq   Pr>F
beh           1  37    16.40       16.40    <.0001    0.0003
sex           1  37     0.05        0.05    0.8297    0.8308
beh*sex       1  37     0.04        0.04    0.8499    0.8509
time          5  185  16.34         3.27    0.0059    0.0075
beh*time      5  185  27.51         5.50    <.0001    <.0001
sex*time      5  185  12.38         2.48    0.0299    0.0337
beh*sex*time  5  185   5.12         1.02    0.4016    0.4050

          ANOVA  ANOVA
            Num    Den    ANOVA F  ANOVA
Effect       DF     DF     Value    Pr>F
beh           1    22.5   16.40    0.0005
sex           1    22.5    0.05    0.8316
beh*sex       1    22.5    0.04    0.8516
time          2.7  79.6    3.38    0.0262 -->> 0.02127
beh*time      2.7  79.6    3.71    0.0180 -->> 0.01400
sex*time      2.7  79.6    1.14    0.3337 -->> 0.32893
beh*sex*time  2.7  79.6    0.44    0.7063 -->> 0.70366
```

```
DATA pval;
INPUT value numdf;
p = 1 - PROBCHI(numdf*value,numdf);
DATALINES;
2.7  3.38
2.7  3.71
2.7  1.14
2.7  0.44
;
RUN;

PROC PRINT DATA=pval; RUN;

 Obs    value    numdf       p
  1      3.38     2.7     0.02127
  2      3.71     2.7     0.01400
  3      1.14     2.7     0.32893
  4      0.44     2.7     0.70366
```

2.5.5 Schulter-Schmerz Studie (F2-LD-F2)

- 41 Patienten
- 2 Behandlungen (Y / N) - Faktor A
- 2 Geschlechter (M / F) - Faktor B
- 3 Tage (1 / 2 / 3) - Faktor T
- 2 Zeitpunkte (M / A) - Faktor Z
- Beobachtung: Schmerz-Score (1,2,3,4,5) nach laparoskopischer Operation

TABELLE 2.3 *Schmerz-Scores von 41 Patienten nach laparoskopischer OP.*

Schmerz-Score															
Behandlung Y								Behandlung N							
		Tag								Tag					
Pat.		1		2		3		Pat.		1		2		3	
Nr.	Geschl.	M	A	M	A	M	A	Nr.	Geschl.	M	A	M	A	M	A
1	F	1	1	1	1	1	1	23	F	5	2	3	5	5	4
3	F	3	2	2	2	1	1	24	F	1	5	3	4	5	3
4	F	1	1	1	1	1	1	25	F	4	4	4	4	1	1
5	F	1	1	1	1	1	1	28	F	3	4	3	3	3	2
8	F	2	2	1	1	1	1	30	F	1	1	1	1	1	1
9	F	1	1	1	1	1	1	33	F	1	3	2	2	1	1
10	F	3	1	1	1	1	1	34	F	2	2	3	4	2	2
12	F	2	1	1	1	1	2	35	F	2	2	1	3	3	2
16	F	1	1	1	1	1	1	36	F	1	1	1	1	1	1
18	F	2	1	1	1	1	1	38	F	5	5	5	4	3	3
19	F	4	4	2	4	2	2	40	F	5	4	4	4	2	2
20	F	4	4	4	2	1	1								
21	F	1	1	1	2	1	1								
22	F	1	1	1	2	1	2								
2	M	3	2	1	1	1	1	26	M	4	4	4	4	4	3
6	M	1	2	1	1	1	1	27	M	2	3	4	3	3	2
7	M	1	3	2	1	1	1	29	M	3	3	4	4	4	3
11	M	1	1	1	1	1	1	31	M	1	1	1	1	1	1
13	M	1	2	2	2	2	2	32	M	1	5	5	5	4	3
14	M	3	1	1	1	3	3	37	M	1	1	1	1	1	1
15	M	2	1	1	1	1	1	39	M	3	3	3	3	1	1
17	M	1	1	1	1	1	1	41	M	1	3	3	3	2	1

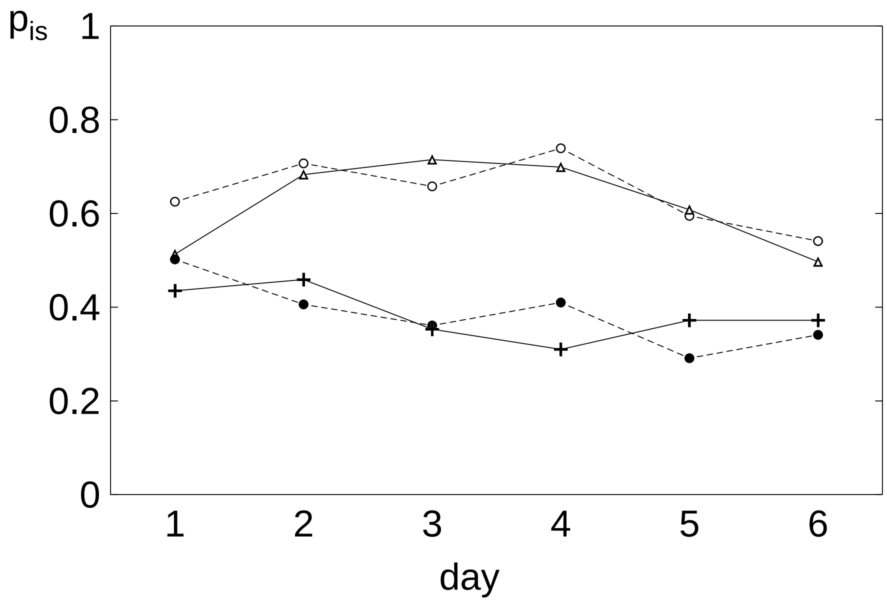

ABBILDUNG 2.2 *Verlaufskurven der relativen Effekte für die Daten der Schulzer-Schmerz Studie.*

SAS Prozeduren

Dateneingabe

```
DATA sss;
INPUT pat beh$ sex$ alt t1-t6;
ARRAY s1{2} t1-t2;
 DO zeit=1 TO 2; score=s1{zeit}; tag = 1; OUTPUT; END;
ARRAY s2{2} t3-t4;
 DO zeit=1 TO 2; score=s2{zeit}; tag = 2; OUTPUT; END;
ARRAY s3{2} t5-t6;
 DO zeit=1 TO 2; score=s3{zeit}; tag = 3; OUTPUT; END;
DROP t1-t6;
DATALINES;
 1  Y  F  64  1  1  1  1  1  1
 3  Y  F  77  3  2  2  2  1  1
 .  .  .   .  .  .  .  .  .  .
41  N  M  61  1  3  3  3  2  1
;
RUN;
```

Sortieren, Ränge bilden, Statistiken

```
PROC RANK DATA=sss OUT=sss; VAR score; RANKS r; RUN;

PROC SORT DATA=sss OUT=sss; BY pat tag zeit; RUN;

PROC MIXED DATA=sss ANOVAF METHOD=MIVQUE0;
CLASS beh sex tag zeit;
MODEL r = beh|sex|tag|zeit / CHISQ;
REPEATED tag*zeit / TYPE=UN SUB=pat GRP=beh*sex;
RUN;
```

```
                    The Mixed Procedure
           Type 3 Tests of Fixed Effects
                 ANOVA    ANOVA
                   Num      Den    ANOVA F     ANOVA
Effect              DF       DF      Value     Pr > F
beh                  1     22.5      16.40     0.0005
sex                  1     22.5       0.05     0.8316
beh*sex              1     22.5       0.04     0.8516
tag               1.67     68.9       5.09     0.0125 --> 0.00965
beh*tag           1.67     68.9       3.94     0.0307 --> 0.02611
sex*tag           1.67     68.9       0.48     0.5847 --> 0.58446
beh*sex*tag       1.67     68.9       0.75     0.4540 --> 0.45018
zeit                 1     31.9       0.58     0.4530
beh*zeit             1     31.9       0.88     0.3545
sex*zeit             1     31.9       0.23     0.6325
beh*sex*zeit         1     31.9       0.05     0.8161
tag*zeit          1.33     32.7       1.28     0.2780 --> 0.26881
beh*tag*zeit      1.33     32.7       4.18     0.0384 --> 0.02966
sex*tag*zeit      1.33     32.7       2.56     0.1100 --> 0.09853
beh*sex*tag*zeit  1.33     32.7       0.01     0.9515 --> 0.96059
```

```
DATA pval;
INPUT value numdf;
p = 1 - PROBCHI(numdf*value,numdf);
DATALINES;
5.09  1.67
3.94  1.67
0.48  1.67
0.75  1.67
1.28  1.33
4.18  1.33
2.56  1.33
0.01  1.33
;
RUN;

PROC PRINT DATA=pval; RUN;
```

Obs	numdf	value	p
1	1.67	5.09	0.00965
2	1.67	3.94	0.02611
3	1.67	0.48	0.58446
4	1.67	0.75	0.45018
5	1.33	1.28	0.26881
6	1.33	4.18	0.02966
7	1.33	2.56	0.09853
8	1.33	0.01	0.96059

3 Literatur

BRUNNER, E., DOMHOF.S. und LANGER, F. (2001). *Nonparametric Analysis of Longitudinal Data in Factorial Designs.* Wiley, New York (in preparation).

BRUNNER, E. AND LANGER, F. (1999). *Nichtparametrische Analyse longitudinaler Daten,* Oldenbourg, München.

BRUNNER, E. and Puri, M. L. (2001). Nonparametric methods in factorial designs. *Statist. Papers* **42**, 1–52.

HETTMANSPERGER, T. P. AND NORTON, R. M. (1987). Tests for patterned alternatives in k–sample problems. *J. Amer. Statist. Assoc.* **82**, 292–299.