



universität  
wien

# MAGISTERARBEIT

Titel der Magisterarbeit

„Gemischte Modelle in der Biostatistik“

Verfasser

Raimund Möser

angestrebter akademischer Grad

Magister der Sozial- und Wirtschaftswissenschaften  
(Mag. rer. soc. oec.)

Wien, im Oktober 2010

Studienkennzahl lt. Studienblatt: A 066 951

Studienrichtung lt. Studienblatt: Statistik

Betreuer: Ao. Univ.-Prof. Mag. Dr. Andreas Futschik

# Vorwort

Gemischte Modelle bieten eine flexible Möglichkeit, gruppierte Datenstrukturen zu handhaben. Die Abhängigkeit von wiederholten Messungen an den Versuchsobjekten wird durch das Modell berücksichtigt. Es werden nicht nur Mittelwerte, wie bei linearen Modellen, sondern auch Varianzen innerhalb der Daten modelliert.

Modelle dieser Art beinhalten nicht nur feste, sondern auch zufällige Effekte, welche durch die zufällige Auswahl der Versuchsobjekte entstehen. Diese werden als Stichprobe aus einer Grundgesamtheit angesehen. Varianzstrukturen innerhalb der Daten können mit geeigneten Schätzmethoden geschätzt und auf deren Signifikanz getestet werden.

In der Biostatistik gibt es eine Vielzahl von Anwendungsgebieten für gemischte Modelle, etwa im medizinischen oder veterinärmedizinischen Bereich, in der psychologischen und soziologischen Forschung, bis hin zur Landwirtschaft.

In der Arbeit werden die theoretischen Grundlagen des Modells vorgestellt und anschließend anhand eines realen Datensatzes angewendet. Zu Beginn werden einführende Beispiele gebracht, welche die Anwendungsmöglichkeiten der gemischten Modelle näher bringen. In den weiteren Abschnitten des theoretischen Teils wird das Modell exakt formuliert, Schätzmethoden besprochen und auf das Testen von Hypothesen sowie die Konstruktion von Konfidenzintervallen eingegangen.

Die Anwendung der Modelle wird anhand eines aus dem Bereich der Tierhaltung stammenden Datensatzes vorgeführt. Zu Beginn werden einfache Modelle angepasst, im weiteren Verlauf wird die Komplexität der Modelle stufenweise erhöht. Durch Modellselektion werden sinnvolle Modelle ausgewählt, und abschließend die Bedeutung verschiedener Einflussgrößen auf die abhängige Variable diskutiert.

Ein weiteres Ziel der Arbeit ist der Vergleich der durch die Softwarepakete R, SPSS und SAS ausgegebenen Ergebnisse.

Besonderer Dank gilt Prof. Dr. Andreas Futschik, welcher mir das Thema näher gebracht, und mich im Rahmen der Arbeit betreut hat. Weiters möchte ich meinen Eltern danken, welche mich während meines Studiums unterstützt haben.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
1.1	Einleitende Beispiele . . . . .	4
1.2	Modell mit zufälligem Intercept . . . . .	6
1.3	Modell mit zufälliger Steigung . . . . .	9
1.4	Modell mit zufälliger Wechselwirkung . . . . .	11
1.5	Verschachtelte Modelle . . . . .	12
<b>2</b>	<b>Das lineare gemischte Modell</b>	<b>14</b>
2.1	Das lineare Modell . . . . .	14
2.2	Das lineare gemischte Modell . . . . .	15
2.3	Beispiele . . . . .	18
2.4	Das mehrstufige gemischte Modell . . . . .	20
<b>3</b>	<b>Schätzen der Parameter</b>	<b>23</b>
3.1	ANOVA-Methode . . . . .	23
3.2	Maximum-Likelihood-Methode . . . . .	24
3.3	Restringierte ML-Schätzung . . . . .	30
3.4	Optimierungsalgorithmen . . . . .	32
<b>4</b>	<b>Inferenz im linearen gemischten Modell</b>	<b>34</b>
4.1	Verteilung der ML- bzw. REML-Schätzer . . . . .	35
4.2	Testen der festen Effekte . . . . .	36
4.3	Testen der Varianzkomponenten . . . . .	38
4.4	Konfidenzintervalle für Varianzkomponenten . . . . .	39
4.5	Vorhersage im gemischten Modell . . . . .	40
<b>5</b>	<b>Der Datensatz</b>	<b>42</b>
<b>6</b>	<b>Deskriptive Analyse</b>	<b>43</b>
<b>7</b>	<b>Zufälliges Intercept-Modell</b>	<b>50</b>
7.1	Analyse in R . . . . .	51
7.2	Analyse in SPSS . . . . .	58
7.3	Analyse in SAS . . . . .	61
<b>8</b>	<b>Marginale Effekte</b>	<b>64</b>
8.1	Analyse in R . . . . .	65
8.2	Analyse in SPSS . . . . .	70
8.3	Analyse in SAS . . . . .	72

<b>9</b>	<b>Verschachtelte zufällige Effekte</b>	<b>75</b>
9.1	Analyse in R . . . . .	76
9.2	Analyse in SPSS . . . . .	79
9.3	Analyse in SAS . . . . .	80
<b>10</b>	<b>Zufällige Wechselwirkungen</b>	<b>80</b>
10.1	Analyse in R . . . . .	81
10.2	Analyse in SPSS . . . . .	84
10.3	Analyse in SAS . . . . .	85
<b>11</b>	<b>Modellselektion</b>	<b>86</b>
11.1	Haupteffektmodell . . . . .	86
11.2	Analyse in R . . . . .	87
11.3	Analyse in SPSS . . . . .	89
11.4	Analyse in SAS . . . . .	90
11.5	Modell mit Wechselwirkungstermen . . . . .	91
11.6	Analyse in R . . . . .	92
11.7	Analyse in SPSS . . . . .	97
11.8	Analyse in SAS . . . . .	99

# 1 Einleitung

## 1.1 Einleitende Beispiele

Lineare Modelle werden in der Statistik verwendet, um den Zusammenhang zwischen einer abhängigen Variable, auch Responsevariable genannt, und ein oder mehreren unabhängigen Variablen zu untersuchen. Die unabhängigen Variablen können nominales, ordinales wie auch metrisches Skalenniveau aufweisen. Falls es sich um nominales Skalenniveau handelt, wird die Variable Faktor genannt. Die einzelnen Ausprägungen des Faktors sind die Faktorstufen, welche die Einteilung der Responsevariable in Gruppen darstellen. Die Analyse eines Datensatzes dieser Art wird als Varianzanalyse bezeichnet.

Es ist auch möglich, mehrere Gruppierungsvariablen simultan zu untersuchen und deren Interaktionen zu analysieren. Je nach Wahl der Designmatrix  $X$  werden die Effekte geschätzt. Um die Signifikanz des Faktors zu bestimmen, werden die Parameter gemeinsam mit dem F-Test getestet.

Es wird für jede Ausprägung des Faktors ein eigener Effekt bestimmt. Handelt es sich bei den Faktorstufen zum Beispiel um drei verschiedene Medikamente, deren Wirkung nachgewiesen werden soll, um vier verschiedene Futtermittel, deren Unterschiede von Interesse sind, oder um verschiedene Joghurtsorten, deren Geschmack in einem Test untersucht werden soll, das Interesse gilt den speziellen Effekten der Faktorausprägungen. Ist Medikament A besser als B, oder lässt sich Futtermittel C von den beiden herkömmlichen Futtermitteln A und B unterscheiden. Jede Faktorstufe stellt eine ganz spezielle, einzigartige Ausprägung dar, deren Wirkung von Interesse ist.

Es gibt Situationen, wo die Einzigartigkeit der Faktorstufen nicht in diesem Sinne gegeben ist, an den Effekten der einzelnen Faktorstufen kein Interesse besteht. In so einem Fall wird von einer Gruppierungsvariable gesprochen. Gruppierungen von Daten können durch eine Vielzahl von Möglichkeiten entstehen. Zum Beispiel soll ein neues Medikament getestet werden, dafür werden einzelne Krankenhäuser ausgewählt, an welchen jeweils Patienten das Präparat verabreicht bekommen. Die verschiedenen Krankenhäuser stellen eine Gruppierungsvariable dar, an deren speziellen Effekten kein wirkliches Interesse besteht, vielmehr ist die Wirkung des Medikaments von Bedeutung.

Weiters können die Krankenhäuser auch nicht als einzigartige Faktorausprägungen betrachtet werden, da Krankenhäuser im Grunde dieselben Aufgaben und Eigenschaften besitzen. Ein spezieller Vergleich von Krankenhaus A mit Krankenhaus B ist uninteressant, die Krankenhäuser, die in dem Datensatz vorkommen, sind eine zufällige Stichprobe aus allen vergleichbaren Krankenhäusern im Land.

Was aber durchaus auftreten kann, sind Unterschiede zwischen den Kran-

kenhäusern in Bezug auf das Ergebnis des Medikamentenversuches. Die Unterschiede können aufgrund verschiedener Merkmale des Krankenhauses bzw. der Patienten entstehen, wie z.B. die Größe des Krankenhauses, oder ob es sich um eine Stadt- oder Landregion handelt, in welcher sich das Krankenhaus befindet. Die Unterschiede können sich auch durch Personen ergeben, die in irgendeiner Weise in den Versuch eingreifen.

Da die krankenhausspezifischen Effekte nicht vorhergesehen werden können, muss mit einer zufälligen Streuung zwischen den Krankenhäusern gerechnet werden, also mit zufälligen Effekten. Um der Streuung gerecht zu werden, wird eine Varianzkomponente in das Modell aufgenommen, welche die Streuung zwischen den Gruppen misst. Solche Modelle werden gemischte Modelle genannt, da sie aus festen sowie aus zufälligen Effekten bestehen.

Ein anderer Anwendungsbereich für ein gemischtes Modell sieht folgendermaßen aus. Biologen sind am Vorkommen einer bestimmten Pflanze auf Hochalmen interessiert. Dazu werden einige Berge ausgewählt, welche in ihrer Lage und Höhe vergleichbar sind, und Areale auf den entsprechenden Almen abgesteckt. Da Berge im Allgemeinen eine nichthomogene Oberfläche besitzen, werden mehrere Areale pro Berg ausgewählt, um ein allgemeines Bild der Bepflanzung zu gewährleisten. Dann wird die Anzahl der entsprechenden Pflanze in jedem Areal aufgenommen, es sind mehrere Messungen für jeden Berg vorhanden. Die Frage, ob es zwischen den Bergen einen Unterschied gibt, soll beantwortet werden.

Die Frage lässt sich mit einer einfachen Varianzanalyse beantworten, indem der Effekt für jeden Berg als fest betrachtet wird. Jeder Berg wird als einzigartige Ausprägung angesehen, dessen Effekt geschätzt wird. Soll aber eine Aussage über alle vergleichbaren Berge dieser Art gemacht werden, ist das Modell ungeeignet. Die Forscher sind vielleicht nicht an den speziellen Bergen interessiert, sondern an der Grundgesamtheit aller Berge, die in dem Datensatz vorkommenden Berge können als Stichprobe aus der Grundgesamtheit betrachtet werden.

In so einem Fall bietet es sich an, den Effekt der Berge als zufällig zu betrachten und die Streuung zwischen den Bergen zu messen. Falls diese in Bezug auf die Streuung innerhalb der Berge eher groß ist, spricht das dafür, dass die Berge eine unterschiedliche Bepflanzung aufweisen. Falls die Streuung innerhalb der Berge überwiegt, ist kein Effekt vorhanden. Die Frage kann mit einem Test auf Signifikanz der Varianzkomponente beantwortet werden.

In dem Beispiel handelt es sich um wiederholte Messungen an einem Objekt, welches durch einen Berg gegeben ist. Das gemischte Modell ist geeignet, um Datensätze mit wiederholten Messungen zu analysieren, da die Korrela-

tion der Messungen innerhalb eines Objektes „heraus gerechnet“ wird und die Vergleichbarkeit zwischen den Objekten gewährleistet ist. Wenn alle Beobachtungen „in einen Topf geschmissen“ werden, führt das zu falschen Ergebnissen, da die Abhängigkeitsstruktur nicht berücksichtigt wird.

Die Frage, ob ein Faktor als fest oder zufällig angenommen werden soll, ist nicht immer leicht zu beantworten. Wenn es sich bei den Faktorausprägungen um einzigartige, spezielle Objekte handelt, an deren speziellen Effekten Interesse besteht, ist der Faktor als fest zu behandeln. Handelt es sich bei den Ausprägungen um Objekte, die den Charakter des Vertauschbaren oder willkürlich Ausgewählten besitzen, oder gar um eine Strichprobe aus einer Grundgesamtheit, sind zufällige Effekte angebracht.

## 1.2 Modell mit zufälligem Intercept

Die zwei oben angeführten Beispiele lassen sich durch ein Modell mit zufälligem Intercept modellieren. Das Modell wird auch *Random-Effects-Model* genannt. Im Beispiel mit den Bergen kann der Effekt der Berge als zufällig betrachtet und ein Modell auf folgende Weise formuliert werden,

$$y_{ij} = \beta + b_i + \epsilon_{ij}.$$

Die Laufvariable  $i = 1, \dots, M$  steht für den  $i$ -ten Berg, die Laufvariable  $j = 1, \dots, n_i$  für die  $j$ -te Messung innerhalb des  $i$ -ten Berges, welche durch die Anzahl einer Pflanze im  $j$ -ten Areal gegeben ist. Die Responsevariable  $y_{ij}$  bezeichnet die jeweilige Messung. Der Koeffizient  $\beta$  ist der Erwartungswert über alle Beobachtungen. Obwohl das Modell den Namen *Random-Effects-Model* trägt, handelt es sich um einen fixen Effekt. Falls es sich um ein balanciertes Design handelt, entspricht  $\beta$  dem Stichprobenmittelwert  $\sum_{i=1}^M \sum_{j=1}^{n_i} y_{ij} / N$ , bei unbalancierten Designs ist das nicht der Fall, da die Gruppen ein unterschiedliches Gewicht auf den Mittelwert haben.  $N$  bezeichnet die gesamte Stichprobenzahl. Bei  $b_i$  handelt es sich um den Effekt des  $i$ -ten Berges, und die Variable  $\epsilon_{ij}$  ist der gewöhnliche Fehlerterm für jede einzelne Beobachtung. Falls es sich um ein balanciertes Design handelt, gibt es pro Berg dieselbe Anzahl an Beobachtungen  $n$ . Wenn die Anzahl der Beobachtungen pro Berg variiert, gibt es für jeden  $n_i$  Beobachtungen. Die Gesamtanzahl der Beobachtungen ist dann  $N = \sum_{i=1}^M n_i$ .

Das Modell ist in seiner Gleichung nicht von einer normalen Varianzanalyse zu unterscheiden. Es scheint ein Koeffizient  $\beta$  für den Gesamtmittelwert auf,  $b_i$  stellt die festen Effekte der Berge dar, welche geschätzt werden. Der große Unterschied zwischen der Varianzanalyse und dem Modell mit zufälligen Effekten ist folgender:

Die  $b_i$ s sind keine festen Effekte sondern Zufallsvariablen und können daher nicht geschätzt werden. Vielmehr wird den  $b_i$ s ein Wahrscheinlichkeitsmodell unterlegt, dessen Parameter geschätzt werden. In den meisten Fällen ist das Wahrscheinlichkeitsmodell die Normalverteilung mit den Parametern  $\mu$  und  $\sigma^2$ , kurz geschrieben als  $N(\mu, \sigma^2)$ . Es können auch andere Verteilungsfamilien herangezogen werden, die Komplexität des Modells in Bezug auf die Parameterschätzung wird dadurch aber noch zusätzlich erhöht. In so gut wie allen Fällen wird eine Normalverteilung verwendet. Durch das Ansehen der  $b_i$ s als Zufallsvariablen bekommt das Modell einen bayesianischen Charakter, es kann von den  $b_i$ s auch als Zufallsvariablen einer Aprioriverteilung gesprochen werden.

Der Grund warum der feste Effekt  $\beta$  für den Mittelwert im Modell aufscheint ist, dass dadurch die Zufallseffekte um Null zentriert werden können. Der Mittelwert kann genauso in die Verteilungsannahme der  $b_i$ s mit einbezogen werden, dann wird kein fester Effekt  $\beta$  im Modell benötigt. Die Modellformulierung hat sich aber nicht durchgesetzt.

Jedes  $b_i$  kann als

$$b_i \sim N(0, \sigma_B^2) \quad \text{für } i = 1, \dots, M$$

betrachtet werden. Das Subskript  $B$  bei  $\sigma_B^2$  steht dafür, dass es von der herkömmlichen Fehlervarianz  $\sigma^2$  unterschieden werden kann.  $\sigma_B^2$  stellt eine zusätzliche Varianzkomponente dar. Im Beispiel ist es die Varianz zwischen den Bergen. Allgemein wird von  $\sigma_B^2$  als der Varianz zwischen den Gruppen, von  $\sigma^2$  als der Varianz innerhalb der Gruppen gesprochen.

Eine weitere Annahme für die Zufallsvariablen  $b_i$  ist, dass sie für verschiedene  $i$  unabhängig sind,

$$\text{Cov}(b_i, b_j) = 0 \quad \text{für alle } i \neq j.$$

Das bedeutet, dass die zufälligen Effekte für die jeweiligen Gruppen voneinander unabhängig sind. Im Beispiel, dass die verschiedenen Berge voneinander unabhängige Objekte darstellen, welche in keinem Zusammenhang stehen. Würden zwei benachbarte Berge untersucht werden, die in irgendeiner Weise miteinander verbunden sind, etwa einer größeren Luftbelastung ausgesetzt sind, wäre die Annahme nicht gewährleistet. Im Allgemeinen kann nicht vermieden werden, dass es zwischen den Subjekten Ähnlichkeiten gibt, deshalb ist eine randomisierte Auswahl wichtig.

Eine weitere Standardannahme betrifft die Fehlerterme  $\epsilon_{ij}$ . Diese sind normalverteilt mit  $N(0, \sigma^2)$  und ebenfalls unabhängig voneinander,

$$\text{Cov}(\epsilon_{ij}, \epsilon_{i'j'}) = 0 \quad \text{außer für } i = i', j = j'.$$



Die Annahme setzt voraus, dass die Beobachtungen auf gleiche Art um den Mittelwert der jeweiligen Gruppe schwanken, und es keine Muster innerhalb der Gruppen gibt. Das kann adaptiert werden, indem zum Beispiel heteroskedastische Fehlervarianz oder Autokorrelationen eingebunden werden, für das einführende Beispiel ist das nicht der Fall.

Zusätzlich wird angenommen, dass die zufälligen Effekte  $b_i$  unabhängig von den Fehlertermen  $\epsilon_{ij}$  sind,

$$\text{Cov}(b_i, \epsilon_{i'j}) = 0 \quad \text{für alle } i, i', j.$$

Die Annahme bedeutet, dass der Wert des zufälligen Effektes nichts mit dem Wert der Fehlerterme zu tun hat. Bei der einfachen Varianzanalyse ist das die Annahme der Gleichheit der Varianzen zwischen den Gruppen. Im Modell ist es nicht möglich, dass zum Beispiel bei einem Berg mit einem größeren Effekt auch mit einer größeren Schwankung der Beobachtungen innerhalb des Berges zu rechnen ist.

Jetzt kann die Verteilung der Responsevariablen  $y_{ij}$  genauer betrachtet werden. Die Linearkombination unabhängiger normalverteilter Zufallsvariablen  $b_i$  und  $\epsilon_{ij}$  ergibt, dass

$$y_{ij} \sim N(\beta, \sigma_B^2 + \sigma^2).$$

Die Kovarianz zwischen den Responsewerten der gleichen Gruppe  $i$  ist

$$\begin{aligned} \text{Cov}(y_{ij}, y_{ij'}) &= E[(y_{ij} - E[y_{ij}])(y_{ij'} - E[y_{ij'}])] \\ &= E[(y_{ij} - \beta)(y_{ij'} - \beta)] \\ &= E[y_{ij}y_{ij'} - \beta y_{ij} - \beta y_{ij'} + \beta^2] \\ &= E[y_{ij}y_{ij'}] - \beta^2 \\ &= E[(\beta + b_i + \epsilon_{ij})(\beta + b_i + \epsilon_{ij'})] - \beta^2 \\ &= \beta^2 + E[b_i^2] - \beta^2 \\ &= \sigma_B^2 \quad \text{für alle } j \neq j'. \end{aligned}$$

Die Terme  $E(b_i \epsilon_{ij'})$  sowie  $E(\epsilon_{ij} \epsilon_{ij'})$  sind Null, da sie aus unabhängigen Zufallsvariablen bestehen, welche um Null zentriert sind. Die Kovarianz der

Responsevariable für verschiedene Gruppen ist Null da

$$\begin{aligned}
Cov(y_{ij}, y_{i'j'}) &= E[(y_{ij} - E[y_{ij}])(y_{i'j'} - E[y_{i'j'}])] \\
&= E[(y_{ij} - \beta)(y_{i'j'} - \beta)] \\
&= E[y_{ij}y_{i'j'} - \beta y_{ij} - \beta y_{i'j'} + \beta^2] \\
&= E[y_{ij}y_{i'j'}] - \beta^2 \\
&= E[(\beta + b_i + \epsilon_{ij})(\beta + b_{i'} + \epsilon_{i'j'})] - \beta^2 \\
&= \beta^2 - \beta^2 \\
&= 0 \quad \text{für alle } i \neq i'.
\end{aligned}$$

Hier wird zusätzlich verwendet, dass  $E(b_i b_{i'}) = 0$ , da verschiedene  $b_i$ s voneinander unabhängig sind.

Jetzt kann die Intra-Klassen-Korrelation zwischen zwei verschiedenen Werten derselben Gruppe berechnet werden mit

$$\begin{aligned}
Corr(y_{ij}, y_{ij'}) &= \frac{Cov(y_{ij}, y_{ij'})}{\sqrt{Var(y_{ij})}\sqrt{Var(y_{ij'})}} \\
&= \frac{\sigma_B^2}{\sigma_B^2 + \sigma^2} \quad \text{für alle } j \neq j'.
\end{aligned}$$

Das ist ein Maß, welches mit einer Varianzanalyse nicht berechnet werden kann, da keine Varianzkomponente zwischen den Gruppen geschätzt wird. In einigen Anwendungsbereichen ist die Größe von Interesse.

Das Modell mit zufälligem Intercept ist das einfachste Modell mit zufälligen Effekten, da es nur einen zufälligen Faktor gibt und keine festen Effekte. Wenn eine kontinuierliche Kovariable in den Daten zur Verfügung steht, kann eine weitere Form des gemischten Modells angewendet werden.

### 1.3 Modell mit zufälliger Steigung

Eine weitere Form eines einfachen gemischten Modells ist das Modell mit zufälliger Steigung. Handelt es sich bei den Daten zum Beispiel um Wachstumskurven einzelner Subjekte oder Messeinheiten, ist das Modell adäquat. In dem Fall ist für jeden Beobachtungswert eine kontinuierliche Kovariable vorhanden, welche den Zeitpunkt der Messung beinhaltet.

Ziel einer solchen Analyse ist es, einen Gesamt-Intercept und eine Gesamt-Steigung über alle Gruppen zu schätzen. Diese stellen die Populationsmittelwerte dar. Die Schwankung der einzelnen Gruppen um den Intercept sowie die Steigung wird mit den zufälligen Effekten modelliert. Eine interessante Hypothese ist, ob alle Gruppen ein ähnliches Wachstumsmuster aufweisen.

Die Frage kann mit der Prüfung auf Signifikanz der Varianzkomponente für die Steigung beantwortet werden.

Mit dem einfachen linearen Modell können solche Datenstrukturen nicht optimal analysiert werden, da die Daten gruppiert sind. Es besteht die einzige Möglichkeit, separate Regressionen zu rechnen und daraus Schätzer für den Gesamt-Intercept sowie die Gesamt-Steigung zu berechnen. Die Vorgangsweise ist aber unbefriedigend, da bei unbalancierten Datensätzen die Frage der Gewichtung der Gruppen eine Rolle spielt. Weiters kann es der Fall sein, dass sich die Zeitfenster der einzelnen Gruppen nicht genau überdecken, was bei einfachen Regressionen ebenfalls zu Problemen führt. Das gemischte Modell bietet die Möglichkeit, solche Datenstrukturen auf elegante Weise zu handhaben und alle zur Verfügung stehenden Datenpunkte einfließen zu lassen.

Angenommen, Forscher sind am Wachstum von Städten in einer Zeitperiode interessiert. Die Städte weisen in ihrer Struktur und Größe eine Ähnlichkeit auf. Es werden zufällig einige der Städte aus der Grundgesamtheit ausgewählt und deren Bevölkerungsstand zu verschiedenen Zeitpunkten notiert. Falls für jede Stadt dieselbe Anzahl an Beobachtungen vorhanden ist, könnte das Modell folgendermaßen aussehen,

$$y_{ij} = \beta_1 + \beta_2 x_j + b_i + \epsilon_{ij}.$$

Die Variable  $i$  stellt die  $i$ -te Stadt dar und  $j$  die  $j$ -te Beobachtung innerhalb der Stadt. Die Einwohnerzahlen der Städte werden mit  $y_{ij}$  bezeichnet. Die Kovariable  $x_j$ , welche den Zeitpunkt beinhaltet, ist für alle Städte gleich, da es sich um ein balanciertes Design handelt. Die Parameter des Modells sind  $\beta_1$  für den Gesamt-Intercept,  $\beta_2$  für eine Gesamt-Steigung aller Gruppen und  $b_i$  für einen zufälligen Intercept für jede Stadt. Die Verteilungsannahme für die Zufallsvariable  $b_i$  ist  $N(0, \sigma_B^2)$ , und  $\epsilon_{ij} \sim N(0, \sigma^2)$ .

Das Modell beinhaltet einen zufälligen Effekt  $b_i$  für den Intercept jeder Stadt. Es lässt zu, dass die Städte ein unterschiedliches Einwohnerniveau aufweisen, die Regressionslinie vertikal verschoben werden kann. Die Schwankung wird durch  $\sigma_B^2$  ausgedrückt. Das Modell beinhaltet keinen zufälligen Effekt für die Steigung, da nur der feste Parameter  $\beta_2$  vorhanden ist. Die Regressionslinien der Gruppen sind in der Steigung nicht variabel. Falls die Städte in ihrem Wachstum ein unterschiedliches Verhalten aufweisen, ist das Modell nicht adäquat. Ein weiterer zufälliger Parameter muss in das Modell aufgenommen werden,

$$y_{ij} = \beta_1 + \beta_2 x_j + b_{1i} + b_{2i} x_j + \epsilon_{ij}.$$

Der zusätzliche Parameter ist  $b_{2i}$ , welcher für jede Stadt eine eigene Steigung zulässt, deren Schwankung durch  $\sigma_S^2$  gemessen wird. Dieser wird mit der Kovariablen  $x_j$  multipliziert. Die Verteilung von  $b_{2i}$  ist  $N(0, \sigma_S^2)$ . Da das Modell

zwei zufällige Parameter beinhaltet, kann ein weiterer Parameter geschätzt werden, die Kovarianz zwischen den beiden Effekten  $\sigma_{BS}$ . Der Parametervektor der zufälligen Effekte  $[b_{1i}, b_{2i}]^T$  entspricht einer multivariaten Normalverteilung mit Erwartungsvektor Null und Varianzkovarianzmatrix

$$\Psi = \begin{bmatrix} \sigma_B^2 & \sigma_{BS} \\ \sigma_{BS} & \sigma_S^2 \end{bmatrix}.$$

Die Interpretation des Kovarianzparameters ist der Zusammenhang zwischen der Steigung und dem Intercept der Städte. Ist allgemein das Wachstum geringer je größer die Einwohnerzahl ist, sind die Parameter negativ korreliert.

Die Hypothese, ob die Wachstumsmuster der Städte ähnlich sind oder nicht, kann mit einem Test der Varianzkomponente  $\sigma_S^2$  beantwortet werden. Das entspricht der Formulierung der Nullhypothese  $H_0: \sigma_S^2 = 0$ , beziehungsweise der Alternative  $H_1: \sigma_S^2 > 0$ . In dem Fall gilt das Ergebnis für die Grundgesamtheit, bei Modellen mit festen Parametern ist so eine Schlussfolgerung nicht möglich.

## 1.4 Modell mit zufälliger Wechselwirkung

Wenn ein Datensatz zwei Gruppierungsvariablen beinhaltet, und für jede Kombination der Ausprägungen mehr als eine Beobachtung vorhanden ist, kann ein Wechselwirkungsterm in das Modell aufgenommen werden. Dieser misst die Wechselwirkung zwischen den Gruppierungsvariablen zusätzlich zum linearen Trend. Bei einem linearen Modell mit festen Termen werden bei einem  $2 \times 3$ -Modell, also drei Faktorstufen pro Faktor, vier Wechselwirkungsterme nötig sein. Insgesamt sind dann zehn Parameter zu schätzen, was bei kleinen Datensätzen eine beträchtliche Anzahl ist. Je mehr Faktorstufen, Gruppen, oder Blöcke es gibt, desto mehr Parameter werden geschätzt.

Mit gemischten Modellen ist es möglich, die Wechselwirkungen als zufällige Effekte zu betrachten und deren Schwankung um den linearen Trend zu messen. Dadurch reduzieren sich die zusätzlichen Parameter zu einer Varianzkomponente, unabhängig von der Anzahl der Faktorstufen. Das kann bei kleinen Datensätzen von Vorteil sein. Die Signifikanz einer solchen Varianzkomponente kann als Evidenz für das Vorhandensein einer Wechselwirkung herangezogen werden.

Ein Anwendungsbeispiel für ein Modell mit zufälligen Wechselwirkungstermen könnte folgendermaßen aussehen. Gemessen wird der Produktivitäts-Score von zufällig ausgewählten Arbeitern auf drei verschiedenen Maschinen. Es sind pro Arbeiter und Maschine jeweils drei Messungen vorhanden. Dadurch kann ein Wechselwirkungsterm berechnet werden. Die Arbeiter stellen einen Blockfaktor dar, welcher nicht von vorrangigem Interesse ist. Es wird

angenommen, dass es natürliche Schwankungen zwischen den Arbeitern gibt. Es bietet sich an, einen zufälligen Intercept-Effekt  $b_i$  für jeden Arbeiter zu modellieren, da die Arbeiter eine zufällige Auswahl aus der Grundgesamtheit aller Arbeiter darstellen. Der Vergleich zwischen den drei Maschinen ist von Interesse, deshalb werden feste Effekte verwendet, um die Maschinen zu vergleichen. Die Wechselwirkung zwischen den Maschinen und den Arbeitern kann als zufällig angenommen werden, da der Arbeitereffekt zufällig ist. Das Modell hat folgende Form,

$$y_{ijk} = \beta + \beta_j + b_i + b_{ij} + \epsilon_{ijk}.$$

Der Produktivitäts-Score stellt die Responsevariable dar und wird mit  $y_{ijk}$  bezeichnet. Die Laufvariablen sind  $i$  für den  $i$ -ten Arbeiter,  $j$  für die  $j$ -te Maschine und  $k$  für die  $k$ -te Beobachtung innerhalb einer Maschine bzw. eines Arbeiters. Die festen Effekte  $\beta$  und  $\beta_j$  modellieren den Gesamt-Intercept sowie den Unterschied zwischen den Maschinen, die zufälligen Effekte  $b_i$  stellen einen Intercept für jeden Arbeiter dar, und der zufällige Term  $b_{ij}$  misst den Unterschied des  $i$ -ten Arbeiters zum linearen Trend der  $j$ -ten Maschine, also den Wechselwirkungsterm. Die Terme  $\epsilon_{ijk}$  sind die Fehlerterme jeder einzelnen Beobachtung.

Bei dem gemischten Modell handelt es sich um zufällige Effekte auf zwei Ebenen, der Ebene „Arbeiter“ und der Ebene „Maschine innerhalb Arbeiter“. Das Modell beinhaltet verschachtelte zufällige Effekte. Die Zufallsvariablen  $b_i$  und  $b_{ij}$  sind voneinander unabhängig und stellen verschiedene Ebenen der Streuung dar. In so einem Fall macht das Messen einer Kovarianz zwischen den Parametern keinen Sinn. Die jeweiligen Varianzkomponenten werden mit  $\sigma_B^2$  und  $\sigma_W^2$  bezeichnet.  $\sigma_W^2 > 0$  ist ein Indiz für Wechselwirkungen zwischen den Arbeitern und den Maschinen. Der Effekt der Maschinen, welcher mit festen Parametern geschätzt wird, kann auf gewöhnliche Weise auf Signifikanz getestet werden, wobei der Arbeitereffekt bereits heraus gerechnet ist.

## 1.5 Verschachtelte Modelle

Im vorigen Beispiel werden verschachtelte zufällige Effekte verwendet, um Wechselwirkungen zu modellieren. Es gibt auch andere Situationen, wo verschachtelte zufällige Effekte auftreten. Das ist der Fall, wenn die Daten eine mehrstufige Gruppierungsstruktur aufweisen, die Gruppen ineinander verschachtelt sind. Solche Modelle werden auch hierarchische Modelle genannt.

Die Frage soll beantwortet werden, ob zwischen den Schulen in einer Stadt Unterschiede in Bezug auf das Ausbildungsniveau bestehen. Es werden einige Schulen aus der Grundgesamtheit aller vergleichbaren Schulen ausgewählt,

welche einem Test unterzogen werden. In jeder der ausgewählten Schulen gibt es mehrere Schulklassen. Falls Kinder im Alter von vierzehn Jahren getestet werden, wird die entsprechende Schulstufe gewählt. Dann werden zufällig Schüler aus den einzelnen Klassen ausgewählt, welche einen Test absolvieren müssen. Der Testscore wird als Responsevariable angesehen, anhand derer die Schulen verglichen werden.

Im Datensatz gibt es zwei Ebenen der Gruppierung, die der Schulen und die der Klassen innerhalb der Schulen. Diese haben eine verschachtelte Struktur, da jede Klasse einer bestimmten Schule zugewiesen ist. Der Effekt zwischen den Schulen und der Effekt der Klassen kann mit einem zufälligen Intercept modelliert werden. Somit besteht die Möglichkeit, dass es zwischen den Schulen einen Unterschied gibt und dass es zwischen den Klassen innerhalb einer Schule Unterschiede gibt. Das Modell könnte so aussehen,

$$y_{ijk} = \beta + b_i + b_{ij} + \epsilon_{ijk}.$$

Die Laufvariable  $i$  steht für die  $i$ -te Schule, die Laufvariable  $j$  für die  $j$ -te Klasse innerhalb der  $i$ -ten Schule. Die Responsevariable ist  $y_{ijk}$ , der Testscore des jeweiligen Schulkindes. In dem Modell gibt es nur einen festen Effekt, den des Gesamt-Interceptes  $\beta$ . Die zufälligen Effekte sind  $b_i$  für die Schulen und  $b_{ij}$  für die Klassen innerhalb der Schulen.  $\epsilon_{ijk}$  stellt den gewöhnlichen Fehlerterm dar, dessen Verteilung eine  $N(0, \sigma^2)$  ist. Die Laufvariable  $k$  zählt die Schulkinder innerhalb einer Klasse. Das Modell ist eigentlich kein gemischtes Modell, sondern ein *Random-Effects-Model*, da es bis auf den Interceptparameter keine festen Effekte beinhaltet.

Die Varianzkomponenten sind  $\sigma_B^2$  für die Varianz zwischen den Schulen,  $\sigma_K^2$  für die Varianz zwischen den Klassen innerhalb der Schulen und  $\sigma^2$  für die Varianz der Schulkinder innerhalb der Klassen. Varianzen werden auf drei verschiedenen, ineinander verschachtelten Ebenen gemessen.

Die Verteilung der Zufallsvariablen  $b_i$  ist  $N(0, \sigma_B^2)$  und  $b_{ij}$  ist  $N(0, \sigma_K^2)$ . Die Zufallsvariablen sind voneinander unabhängig. Das hat zu bedeuten, dass die Streuung der einzelnen Klassen um den Mittelwert der Schule nicht von dem Bildungsgrad der Schule abhängig ist. Die Varianzkomponente  $\sigma_K^2$  misst die Streuung der Klassen innerhalb der Schulen und kann nicht mit einer speziellen Schule in Verbindung gebracht werden.

In dem Beispiel macht es Sinn, für die Schulen zufällige Effekte zu wählen, da Vergleiche zwischen speziellen Schulen nicht von Interesse sind, sondern Aussagen über die Grundgesamtheit der Schulen getroffen werden sollen. Im gemischten Modell wird nur eine Stichprobe benötigt, das Testen aller Schulen bedeutet einen größeren Aufwand.

Die Varianzkomponenten  $\sigma_B^2$  und  $\sigma_K^2$  können auf deren Signifikanz getestet werden. Falls die Hypothese  $\sigma_B^2 > 0$  zutrifft, dann kann davon aus-

gegangen werden, dass es Unterschiede im Ausbildungsniveau der Schulen gibt. Falls  $\sigma_K^2 > 0$ , gibt es Unterschiede zwischen den Klassen innerhalb der Schulen.

## 2 Das lineare gemischte Modell

In dem Abschnitt wird das lineare gemischte Modell in Matrixschreibweise formuliert. Die Responsevariable  $y$  ist eine Linearkombination der festen wie auch zufälligen Koeffizienten mit den Regressormatrizen. Zusätzlich ist noch ein gewöhnlicher Fehlerterm in der Modellgleichung zu finden. Die allgemeine Formulierung des linearen gemischten Modells in der Form geht auf Laird und Ware (1982) zurück. Da das gemischte Modell auf dem linearen Modell mit festen Parametern aufbaut, werden die wichtigsten Eigenschaften dieses Modells zuerst formuliert.

### 2.1 Das lineare Modell

Das lineare Modell in Matrixschreibweise wird definiert durch

$$y = X\beta + \epsilon,$$

wobei  $y$  ein  $n \times 1$ ,  $\beta$  ein  $k \times 1$  und  $\epsilon$  ein  $n \times 1$  Vektor ist. Die Dimension der Regressormatrix  $X$  ist  $n \times k$ . Jeder einzelne Wert der Responsevariable  $y_i$ ,  $i = 1, \dots, n$ , kann als Linearkombination der  $\beta_j$ ,  $j = 1, \dots, k$ , mit der  $i$ -ten Zeile der Matrix  $X$  sowie einem zufälligen Fehlerterm  $\epsilon_i$  identifiziert werden,

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \dots + x_{ik}\beta_k + \epsilon_i.$$

Die Fehlerterme  $\epsilon_i$  sind Zufallsvariablen mit der Eigenschaft  $E(\epsilon_i) = 0$  und  $Var(\epsilon_i) = \sigma^2$  für alle  $i = 1, \dots, n$ . Eine daraus resultierende Konsequenz ist

$$E(y) = X\beta.$$

Die Koeffizienten  $\beta_j$  werden als feste unbekannte Größen angesehen, deren Werte mit der Technik der kleinsten Quadrate (OLS) geschätzt werden. Die Schätzer können durch folgende Formel errechnet werden,

$$\hat{\beta} = (X^T X)^{-1} X^T y.$$

Der Schätzer ist der beste lineare unverzerrte Schätzer (b.l.u.e.), er hat die geringste Varianz aller linearen unverzerrten Schätzer. Das Resultat wird als

das Gauß-Markov-Theorem bezeichnet. Die Varianzkovarianzmatrix von  $\hat{\beta}$  ist

$$VC(\hat{\beta}) = \sigma^2(X^T X)^{-1}.$$

Die Residuen  $\hat{\epsilon}$  können durch

$$\hat{\epsilon} = y - X\hat{\beta}$$

errechnet werden, und die Varianz der Fehlerterme  $\sigma^2$  lässt sich durch

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{n - k}$$

unverzerrt schätzen. Unter der Standardannahme  $\epsilon_i \sim N(0, \sigma^2)$  für alle  $i$ , lassen sich Inferenzmethoden für die Parameterschätzer herleiten. Es können sowohl die marginalen Effekte einzelner Parameter, wie auch eine Kombination mehrerer Parameter auf Signifikanz getestet werden. Die Signifikanz der gesamten Regression lässt sich durch die Statistik

$$F = \frac{ESS/(k - 1)}{RSS/(n - k)},$$

welche unter der Nullhypothese einer  $F_{k-1, n-k}$ -Verteilung folgt, testen. ESS beschreibt die erklärte Quadratsumme und RSS die Quadratsumme der Residuen.

Die Varianzkovarianzmatrix der Fehlerterme  $I\sigma^2$  ist die einfachste Form der Modellierung. Es können auch Abhängigkeiten der Fehler in der Matrix modelliert werden, das führt zu einem allgemeineren Modell, welches als GLS(Generalized-Least-Squares)-Modell bezeichnet wird. Nun aber zur Formulierung des gemischten Modells, welches sich um einen wesentlichen Teil vom linearen Modell unterscheidet.

## 2.2 Das lineare gemischte Modell

Das lineare gemischte Modell mit einer Gruppierungsvariable wird formuliert durch

$$y_i = X_i\beta + Z_ib_i + \epsilon_i \quad i = 1, \dots, M,$$

wobei  $y_i$  den  $n_i \times 1$  Responsevektor der  $i$ -ten Gruppe bezeichnet,  $X_i$  die  $n_i \times k$  Regressormatrix der festen Effekte, und  $Z_i$  die  $n_i \times q$  Regressormatrix der zufälligen Effekte darstellt.  $\beta$  ist der gewöhnliche  $k \times 1$  Parametervektor der festen Effekte und  $b_i$  der  $q \times 1$  Vektor der zufälligen Effekte.  $q$  ist die



Anzahl der zufälligen Parameter. Die Fehlerterme für jede Beobachtung werden durch den  $n_i \times 1$  Vektor  $\epsilon_i$  beschrieben. Es gibt  $i = 1, \dots, M$  Gruppen und  $n_i$  ist die Anzahl der Beobachtungen für die  $i$ -te Gruppe. Falls es sich um ein balanciertes Design handelt, kann die Größe für jede Gruppe mit  $n$  festgesetzt werden. Die Gesamtanzahl der Beobachtungen  $N = \sum_{i=1}^M n_i$ , und für das balancierte Design gilt  $N = Mn$ .

Das Modell lässt sich kompakter formulieren, indem die verschiedenen Gruppen in einer Matrix zusammengefasst werden. Das hat folgende Gestalt,

$$y = X\beta + Zb + \epsilon,$$

die Vektoren bzw. Matrizen stellen sich zusammen durch,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}, \quad X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_M \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_M \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix},$$

$$Z = \begin{bmatrix} Z_1 & 0 & 0 & \cdots & 0 \\ 0 & Z_2 & 0 & \cdots & 0 \\ 0 & 0 & Z_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & Z_M \end{bmatrix}.$$

Die Dimension von  $y$  und  $\epsilon$  ist  $N \times 1$ ,  $X$  ist eine  $N \times k$ ,  $Z$  eine  $N \times Mq$  Matrix,  $b$  ein  $Mq \times 1$  Vektor, und  $\beta$  hat dieselben Dimensionen wie zuvor. Das ist eine äquivalente Möglichkeit, das Modell zu formulieren. Um auf weitere Eigenschaften des Modells einzugehen, wird die Formulierung verwendet, die zu Beginn des Abschnitts eingeführt wurde.

$b_i$  ist ein Vektor bestehend aus Zufallsvariablen, er kann durch Verteilungseigenschaften beschrieben werden. Die erste Annahme betrifft den Erwartungswert,

$$E(b_i) = 0 \quad \text{für alle } i = 1, \dots, M.$$

0 steht für den Nullvektor der Dimension  $n_i \times 1$ . Die Annahme setzt voraus, dass der Populationsmittelwert der zufälligen Effekte durch einen der festen Parameter geschätzt wird. Das bedeutet, dass die Spalten der Matrix  $Z_i$  eine Teilmenge der Spalten von  $X_i$  sind. Da die Matrix  $X_i$  aus linear unabhängigen Spalten besteht,  $\text{rank}(X_i) = k$ , gilt die lineare Unabhängigkeit der Spalten auch für  $Z_i$ . Eine lineare Abhängigkeit der Spalten kann durch eine Reparametrisierung des Modells vermieden werden.

Weiters wird für den Vektor  $b_i$  eine Varianzkovarianzmatrix  $\Psi$  spezifiziert. Die Matrix ist von der Form  $q \times q$ , und hat im allgemeinen Fall folgende Gestalt,

$$\Psi = \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_1 b_2} & \sigma_{b_1 b_3} & \cdots \\ \sigma_{b_1 b_2} & \sigma_{b_2}^2 & \sigma_{b_2 b_3} & \cdots \\ \sigma_{b_1 b_3} & \sigma_{b_2 b_3} & \sigma_{b_3}^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

$\sigma_{b_1}^2$  misst die Varianz des ersten Elementes der  $b_i$  über alle Gruppen,  $\sigma_{b_2}^2$  misst die des zweiten Elementes, usw. Die Parameter in den Nebendiagonalen stehen für die Kovarianzen zwischen den zufälligen Effekten.

Die einfachste Form der Parametrisierung ist, für jeden Koeffizienten in der Hauptdiagonale der Varianzkovarianzmatrix einen Varianzparameter festzulegen, und die Kovarianzen auf Null zu setzen,

$$\Psi = \begin{bmatrix} \sigma_{b_1}^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_{b_2}^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_{b_3}^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \sigma_{b_q}^2 \end{bmatrix}.$$

Es gibt eine Vielzahl von Möglichkeiten, die Kovarianzstruktur zu modellieren, die oben angeführten Matrizen stellen die beiden Extreme dar. Die Matrix  $\Psi$  muss jedenfalls symmetrisch und positiv-semidefinit sein. Die zweite Eigenschaft ist mit der Bedingung nichtnegativer Eigenwerte gleichzusetzen.

In der Praxis ist es sinnvoller, die Matrix  $\Psi$  positiv-definit zu definieren, das bedeutet, alle Eigenwerte müssen größer als Null sein. Die Restriktion der Matrix  $\Psi$  ist deshalb möglich, da ein indefinites Modell durch ein positiv-definites Modell mit geringerer Ordnung reparametrisiert werden kann.

Die zweite Eigenschaft der  $b_i$ s ist, dass sie für verschiedene Gruppen unabhängig sind,

$$Cov(b_i, b_{i'}) = 0 \quad \text{für } i \neq i'.$$

Da es sich bei  $b_i$  um einen Zufallsvektor handelt, ist jedes Element von  $b_i$  unabhängig von jedem Element  $b_{i'}$  für  $i \neq i'$ .

Eine weitere Annahme ist die Unabhängigkeit zwischen  $b_i$  und den Vektoren der Fehlerterme  $\epsilon_i$ , für verschiedene Gruppen wie auch innerhalb der Gruppe,

$$Cov(b_i, \epsilon_{i'}) = 0 \quad \text{für alle } i, i'.$$

Wiederum bedeutet das die Unabhängigkeit jedes Elementes von  $b_i$  mit jedem Element von  $\epsilon_i$ .

Nun zu den Eigenschaften des Vektors  $\epsilon_i$  der Fehlerterme, welche dieselben sind wie im linearen Modell,

$$E(\epsilon_i) = 0 \quad \text{für alle } i = 1, \dots, M,$$

und deren Unabhängigkeit für verschiedene Gruppen,

$$Cov(\epsilon_i, \epsilon_{i'}) = 0 \quad \text{für alle } i \neq i'.$$

Die Form der Varianzkovarianzmatrix der Fehlerterme innerhalb einer Gruppe wird wie folgt definiert,

$$VC(\epsilon_i) = \sigma^2 I \quad \text{für alle } i = 1, \dots, M,$$

wobei  $I$  die  $n_i \times n_i$  Einheitsmatrix ist. Diese Form der  $VC(\epsilon_i)$  ist die einfachste und gilt als die Standardannahme. Es ist möglich, heteroskedastische Varianzstrukturen oder Autokorrelation in der Matrix zu modellieren.

Der Erwartungswert des Responsevektors  $y_i$  ist

$$E(y_i) = X_i \beta \quad \text{für alle } i = 1, \dots, M,$$

das ist dasselbe Resultat wie im linearen Modell. Die Koeffizienten  $\beta$  stellen die festen Parameter über die gesamte Population dar. Der bedingte Erwartungswert  $E(y_i|b_i)$  kann ebenfalls angegeben werden.  $b_i$  wird als Realisation und nicht als Zufallsvektor betrachtet.

$$E(y_i|b_i) = X_i \beta + Z_i b_i \quad \text{für alle } i = 1, \dots, M.$$

Das sind die gruppenspezifischen Erwartungswerte, welche bei der Vorhersage von Interesse sind.

## 2.3 Beispiele

Im Folgenden werden die Modelle der Beispiele aus der Einleitung mittels Matrixschreibweise formuliert.

Das erste Beispiel ist das Modell mit zufälligem Intercept. Angenommen, es werden Beobachtungen auf zehn verschiedenen Bergen gemacht, die Anzahl der Gruppen  $M=10$ . Falls innerhalb der Berge 6 verschiedene Areale abgesteckt werden, gilt  $n_i = n = 6$ , und es handelt sich um ein balanciertes Design. Die entsprechenden Matrizen  $Z_i$  und  $X_i$  lauten

$$Z_i = X_i = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{für } i = 1, \dots, 10.$$

Die Dimension des Parametervektors  $\beta$  ist  $1 \times 1$ . Es wird ein Gesamt-Intercept geschätzt. Die Dimension von  $b_i$  ist ebenfalls  $1 \times 1$ , für jeden Berg wird ein zufälliger Effekt geschätzt. Die Varianzkovarianzmatrix  $\Psi$  reduziert sich zu einem Element,  $\sigma_B^2$ . Das Beispiel stellt die einfachste Form eines gemischten Modells dar.

In dem Beispiel mit den Wachstumskurven für Städte handelt es sich um ein Modell mit zufälliger Steigung. Angenommen es werden 12 Städte betrachtet, die Anzahl der Gruppen  $M = 12$ . Falls die Einwohnerzahlen zum Beispiel in der Zeitperiode 2001 bis 2010 herangezogen werden, und Daten für alle Zeitpunkte zur Verfügung stehen, handelt es sich um ein balanciertes Design mit  $n_i = n = 10$ . Die entsprechenden Matrizen  $X_i$  und  $Z_i$  lauten

$$Z_i = X_i = \begin{bmatrix} 1 & 2001 \\ 1 & 2002 \\ \vdots & \vdots \\ 1 & 2010 \end{bmatrix} \quad \text{für } i = 1, \dots, 12.$$

Es wird eine zweite Spalte für die Kovariable der Zeit in die Matrizen aufgenommen. Die Dimension des festen Parametervektors  $\beta$  und die des Vektors der zufälligen Effekte  $b_i$  ist  $2 \times 1$ , somit ergibt sich  $p = q = 2$ . Für die Varianzkovarianzmatrix  $\Psi$  sind zwei Parametrisierungen möglich,

$$\Psi_1 = \begin{bmatrix} \sigma_B^2 & \sigma_{SB} \\ \sigma_{SB} & \sigma_S^2 \end{bmatrix}$$

mit dem Kovarianzparameter  $\sigma_{SB}$  und

$$\Psi_2 = \begin{bmatrix} \sigma_B^2 & 0 \\ 0 & \sigma_S^2 \end{bmatrix}$$

ohne Kovarianzparameter. Die Wahl von  $\Psi$  hängt davon ab, ob der zusätzliche Parameter eine signifikante Verbesserung der Aussagekraft des Modells mit sich bringt.

Das in der Einführung gebrachte Beispiel mit den Medikamenten, welche in verschiedenen Krankenhäusern verabreicht werden, könnte folgendermaßen aussehen. Angenommen, es sollen drei verschiedene Medikamente getestet werden, und jeder Patient bekommt ein Medikament verabreicht, dessen Wirkung den Wert der Responsevariable ausmacht. Die Krankenhäuser stellen die Gruppierungsvariable dar, sie werden mit  $i = 1, \dots, M$  nummeriert. Der Einfachheit halber wird angenommen, dass pro Krankenhaus jeweils 9 Patienten getestet werden und jedes Medikament an drei Personen verab-

reicht wird. Dadurch handelt es sich um ein balanciertes Design. Die entsprechenden Designmatrizen haben folgende Gestalt,

$$Z_i = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{9 \times 1}$$

und die Matrix der festen Effekte in Form der Treatmentcodierung ist

$$X_i = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Die Zufallseffekte  $b_i$  sind eindimensional und messen die zufällige Schwankung zwischen den Krankenhäusern in Form der Varianzkomponente  $\sigma_B^2$ , welche gleichzeitig die reduzierte Varianzkovarianzmatrix  $\Psi$  darstellt.

Der Vektor der festen Effekte  $\beta$  ist  $3 \times 1$ , das erste Element steht für den Intercept des Modells, das zweite misst den Effekt des zweiten Medikaments in Bezug zum Mittelwert, und das dritte Element misst den Effekt des dritten Medikaments. Die Effekte können auf deren Signifikanz geprüft werden wie das für ein gewöhnliches lineares Modell der Fall ist. Der Vorteil des gemischten Modells liegt darin, dass der zufällige Effekt der Krankenhäuser heraus gerechnet ist, und somit ein Vergleich der Medikamente möglich ist. Wenn die Gruppierung durch die Krankenhäuser ignoriert wird, kann das zu anderen Ergebnissen führen.

## 2.4 Das mehrstufige gemischte Modell

Bis jetzt wird eine Gruppierungsvariable zugelassen, welche die Beobachtungen in  $M$  Gruppen unterteilt. Es gibt Situationen, wo die einzelnen Gruppen wiederum eine Gruppierung aufweisen, und eine zweite oder mehrere Stufen der Unterteilung vorhanden sind. Solche Modelle werden mehrstufige oder verschachtelte Modelle genannt. Im Fall einer zweistufigen Gruppierung wird das Modell auf folgende Weise formuliert,

$$y_{ij} = X_{ij}\beta + Z_{i,j}b_i + Z_{ij}b_{ij} + \epsilon_{ij}.$$

Die Laufvariable  $i = 1, \dots, M$  bezeichnet die  $i$ -te Gruppe der ersten Stufe. Die Laufvariable  $j = 1, \dots, M_i$  bezeichnet die  $j$ -te Gruppe innerhalb der  $i$ -ten Gruppe, die Gruppierung auf der zweiten Stufe. Die Beobachtungsanzahl der Gruppen auf der ersten Stufe wird mit  $n_i$ , die auf der zweiten Ebene mit  $n_{ij}$  bezeichnet. Die Gesamtanzahl der Beobachtungen  $N = \sum_{i=1}^M n_i = \sum_{i=1}^M \sum_{j=1}^{M_i} n_{ij}$ . Der Responsevektor  $y_{ij}$  ist ein  $n_{ij} \times 1$  Vektor.

Die Designmatrix  $X_{ij}$  der festen Effekte hat Dimension  $n_{ij} \times k$ . Der zugehörige Parametervektor  $\beta$  ist über alle Gruppen invariant und hat die Form  $k \times 1$ .

Die Regressormatrix der zufälligen Effekte  $X_{i,j}$  der ersten Stufe hat Dimension  $n_{ij} \times q_1$ . Der Beistrich bezeichnet den  $j$ -ten Abschnitt der Matrix  $X_i$ . Da das Modell auf der innersten Stufe formuliert ist, muss die Matrix für jede Gruppe der ersten Stufe in  $j$  Teilabschnitte zerlegt werden. Der Parametervektor der zufälligen Effekte der ersten Stufe,  $b_i$ , ist von der Form  $q_1 \times 1$ .

$X_{ij}$  ist die Regressormatrix der zweiten Stufe und hat die Dimension  $n_{ij} \times q_2$ . Es ist kein Beistrich zwischen den Subskripts, da es auf dieser Ebene theoretisch für jedes  $i$  und  $j$  eine eigene Matrix geben kann. Der Parametervektor der zufälligen Effekte  $b_{ij}$  beinhaltet die zufälligen Effekte auf der zweiten Stufe und ist von der Form  $q_2 \times 1$ . Die Dimension des Fehlervektors  $\epsilon_{ij}$  ist  $n_{ij} \times 1$ .

Die Eigenschaften der Vektoren  $b_i$  und  $b_{ij}$  sind

$$E(b_i) = 0 \quad \text{und} \quad E(b_{ij}) = 0 \quad \text{für alle } i = 1, \dots, M, j = 1, \dots, M_i.$$

Die Varianzkovarianzmatrizen werden analog zu den Stufen der zufälligen Effekte mit  $\Psi_1$  und  $\Psi_2$  bezeichnet und sind symmetrisch und positiv definit. Die Dimensionen der Varianzkovarianzmatrizen sind  $q_1 \times q_1$  bzw.  $q_2 \times q_2$ .

Die zufälligen Effekte  $b_i$  auf der ersten Stufe sind unabhängig für alle  $i = 1, \dots, M$ , die zufälligen Effekte auf der zweiten Stufe  $b_{ij}$  sind unabhängig für alle  $i = 1, \dots, M$  und  $j = 1, \dots, M_i$ . Weiters sind die  $b_i$  von den  $b_{ij}$  unabhängig für alle  $i = 1, \dots, M$  und  $j = 1, \dots, M_i$ .

Die Eigenschaft der Fehlervektoren  $\epsilon_{ij}$  ist

$$E(\epsilon_{ij}) = 0, \quad VC(\epsilon_{ij}) = \sigma^2 I \quad \text{für alle } i, j.$$

Sie bestehen aus voneinander unabhängigen Elementen und sind unabhängig für alle  $i = 1, \dots, M$  und  $j = 1, \dots, M_i$ . Weiters sind die  $\epsilon_{ij}$  unabhängig von  $b_i$  sowie von  $b_{ij}$  für alle  $i$  und  $j$ . Es bestehen keine Abhängigkeiten zwischen den zufälligen Effekten für verschiedene Stufen und den Fehlertermen.

Das Modell kann für beliebig viele Verschachtelungen formuliert werden, für den Fall von drei ineinander verschachtelten Ebenen lautet das Modell

$$y_{ijk} = X_{ijk}\beta + Z_{i,jk}b_i + Z_{ij,k}b_{ij} + Z_{ijk}b_{ijk} + \epsilon_{ijk}.$$

Es gibt eine zusätzliche Laufvariable  $k = 1, \dots, M_{ij}$ , welche die Gruppe der dritten Stufe innerhalb der  $ij$ -ten Gruppe der zweiten Stufe bezeichnet. Die Beobachtungszahl der  $ijk$ -ten Gruppe ist  $n_{ijk}$ . Die Laufvariable  $i$  geht wie vorhin von  $1, \dots, M$  und  $j = 1, \dots, M_i$ . Es gibt drei zufällige Vektoren  $b_i$ ,  $b_{ij}$  und  $b_{ijk}$ , welche als die zufälligen Effekte auf den jeweiligen Stufen definiert sind. Die Regressormatrizen  $Z_{i,jk}$ ,  $Z_{ij,k}$  und  $Z_{ijk}$  unterscheiden sich in ihrer Bezeichnung nur durch die Position des Beistrichs, die Variablen vor dem Beistrich bezeichnen die Gruppierungsstufe, die Variablen nach dem Beistrich die Unterteilung der Matrix. Die Anzahl der Zeilen der Matrizen ist  $n_{ijk}$ , die der Spalten entsprechen der Länge des jeweiligen Vektors der zufälligen Effekte,  $q_1$ ,  $q_2$  und  $q_3$ . Die Eigenschaften der Zufallseffekte und deren Unabhängigkeit untereinander lassen sich aus dem zweistufigen Modell erweitern. In einem Modell der Form sind drei Varianzkovarianzmatrizen für die Vektoren der zufälligen Effekte nötig.

Das in der Einleitung angeführte Beispiel der Schulen entspricht einem einfachen Modell mit verschachtelter Gruppenstruktur. Die Schulen stellen die erste Gruppierungsstufe dar, sie werden mit  $i = 1, \dots, M$  nummeriert. Auf der zweiten Ebene stellen die einzelnen Klassen innerhalb der Schulen die Gruppierungsvariable dar. Da jede Klasse zu einer bestimmten Schule gehört, ist das Modell ineinander verschachtelt. Die Klassen innerhalb der  $i$ -ten Schule werden mit  $j = 1, \dots, M_i$  nummeriert.  $M_i$  bezeichnet die Anzahl der Klassen in der  $i$ -ten Schule. Innerhalb der Klassen sind die einzelnen Beobachtungen zu finden, die Testergebnisse der Schüler. Die Anzahl der Beobachtungen pro Klasse ist  $n_{ij}$ . Falls es sich bei dem Beispiel um ein balanciertes Design handelt, zum Beispiel mit 10 Beobachtungen pro Klasse, sehen die entsprechenden Matrizen folgendermaßen aus,

$$X_{ij} = Z_{i,j} = Z_{ij} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{10 \times 1},$$

die zufälligen Effekte  $b_i$  und  $b_{ij}$  sind jeweils Skalare, und deren Kovarianzmatrizen  $\Psi_1$  und  $\Psi_2$  reduzieren sich zu jeweils einem Parameter,  $\sigma_1^2$  und  $\sigma_2^2$ . Das ist ein Beispiel für ein zufälliges Intercept Modell mit zwei Stufen.

### 3 Schätzen der Parameter

In einem linearen gemischten Modell müssen die festen Effekte  $\beta$ , die Parameter welche die Varianzkovarianzmatrix  $\Psi$  bestimmen und die Fehlervarianz  $\sigma^2$  geschätzt werden. Es gibt verschiedene Schätzmethoden, um die Parameter zu schätzen. Eine der ersten verwendeten Methoden, um Varianzkomponenten zu schätzen, ist die ANOVA-Methode. Der Schätzer ist im Kontext der Varianzanalyse entwickelt worden. Eine weitere Schätzmethode ist die Maximum-Likelihood-Schätzung. Das ist eine der am häufigsten verwendeten Methoden, um Parameter zu schätzen, da sie bei größeren Stichprobenumfängen sehr gute Eigenschaften besitzt. Aus dieser leitet sich die restringierte Maximum-Likelihood-Schätzung her, welche im Zusammenhang mit gemischten Modellen einen wichtigen Stellenwert hat. Eine weitere Methode ist die Minimum-Norm-Schätzung (MINQUE), welche in einigen Softwarepaketen implementiert ist.

#### 3.1 ANOVA-Methode

Die als erstes entwickelte Methode, um Varianzkomponenten zu schätzen, ist die ANOVA-Methode. Angenommen, es soll in einer einfachen Varianzanalyse mit balancierten Gruppen die Varianzkomponente geschätzt werden, also die Streuung zwischen den Gruppen. Das Modell hat folgende Gestalt,

$$y_{ij} = \mu + b_i + \epsilon_{ij} \quad \text{für } i = 1, \dots, M, \quad j = 1, \dots, n.$$

Es gibt  $M$  Gruppen mit jeweils  $n$  Beobachtungen.  $\mu$  beschreibt den Gesamt-Intercept, und  $y_{ij}$  ist die Responsevariable. Die Koeffizienten  $b_i$  werden als zufällige Effekte angesehen. Die Basis für eine Varianzanalyse ist die Zerlegung der Gesamtstreuung um den Mittelwert in eine Streuung innerhalb und zwischen den Gruppen. Das wird berechnet durch

$$\begin{aligned} TSS &= \sum_{i=1}^M \sum_{j=1}^n (y_{ij} - \bar{y})^2, \\ ESS &= \sum_{i=1}^M n(\bar{y}_i - \bar{y})^2, \\ RSS &= \sum_{i=1}^M \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2, \end{aligned}$$

wobei  $\bar{y}$  der Gesamtmittelwert und  $\bar{y}_i$  der Mittelwert der  $i$ -ten Gruppe ist.  $TSS$  bezeichnet die Quadratsumme um den Gesamtmittelwert,  $ESS$  die



Quadratsumme zwischen den Gruppen und  $RSS$  die Quadratsumme der Residuen. Daraus lassen sich die mittleren Quadratsummen zwischen ( $MSE$ ) und innerhalb ( $MSR$ ) der Gruppen berechnen,

$$MSR = RSS/(M(n-1))$$

sowie

$$MSE = ESS/(n-1).$$

$MSR$  ist ein unverzerrter Schätzer für  $\sigma^2$ . Es kann gezeigt werden, dass die Varianzkomponente  $\sigma_B^2$  durch den Schätzer

$$\hat{\sigma}_B^2 = \left( \frac{ESS}{M-1} - \hat{\sigma}^2 \right) / n = \frac{MSE - MSR}{n}$$

unverzerrt geschätzt werden kann. Der Schätzer hat einen Nachteil, im Falle von  $MSR > MSE$  nimmt er negative Werte an. Für die Schätzung einer Varianz ist das ein unbefriedigendes Resultat, Varianzen sind niemals negativ. Es gibt eine Reihe von Methoden, um die negativen Schätzungen zu verhindern, dadurch werden die einfachen Formeln aber komplizierter, und die Eigenschaft der Unverzerrtheit ist nicht mehr gegeben. Eine Alternative ist, einen negativen Wert für  $\hat{\sigma}_B^2$  als ein Indiz für das Nichtvorhandensein einer Varianzkomponente zu deuten.

Die ANOVA-Methode kann auch bei unbalancierten Designs oder mehrfachen Varianzanalysen mit Interaktionseffekten angewendet werden, die Möglichkeit einer negativen Schätzung kann aber nicht vermieden werden. Deshalb hat sich die einfache Schätzmethode, deren Vorteil eine explizite Darstellung der Schätzer ist, in der Praxis nicht durchgesetzt, und es wird ein allgemeines Schätzprinzip verwendet, das der Maximum-Likelihood-Schätzung.

### 3.2 Maximum-Likelihood-Methode

Die Maximum-Likelihood(ML)-Methode ist ein allgemeines Schätzprinzip, welches für die Schätzung unbekannter Parameter eines Modells herangezogen wird. Die Schätzer haben gute Eigenschaften wie asymptotische Erwartungstreue und Effizienz, das heißt, die Varianz erreicht asymptotisch die Cramer-Rao-Schranke. Bei großen Stichprobenumfängen ist die ML-Methode sehr gut geeignet, um Schätzer zu bestimmen. Ein Nachteil ist, dass die Schätzer nicht unverzerrt sind, was bei kleinen Stichprobenumfängen zu berücksichtigen ist. Eine Nebenbedingung ist, dass den Daten eine Verteilungseigenschaft zugewiesen werden muss. Das ist bei der ANOVA-Methode nicht der Fall.

Beim linearen Modell ist die Verteilung meistens die Normalverteilung,

$$y \sim N(X\beta, I\sigma^2).$$

Die ML-Methode führt zu den bekannten ML-Schätzern, welche in expliziter Form angegeben werden können. Im linearen gemischten Modell wird ebenfalls mit der Normalverteilung gearbeitet, da sonst die Komplexität des Modells weiter erhöht wird. Die Eigenschaft der Reproduzierbarkeit ist notwendig, da Zufallsvariablen auf mehreren Ebenen vorhanden sind. Die erste Annahme betrifft den Vektor der zufälligen Effekte  $b_i$ ,

$$b_i \sim N(0, \Psi) \quad \text{für } i = 1, \dots, M.$$

Wie schon erwähnt, sind die  $b_i$  unabhängig für verschiedene  $i$ . Die zweite Annahme betrifft die Fehlervektoren  $\epsilon_i$ , deren Verteilung ebenfalls die Normalverteilung ist,

$$\epsilon_i \sim N(0, I\sigma^2) \quad \text{für } i = 1, \dots, M.$$

Die  $\epsilon_i$  sind nicht nur für verschiedene  $i$ , sondern auch innerhalb des Vektors unabhängig, was an der diagonalen Struktur der Kovarianzmatrix zu erkennen ist. Weiters sind die  $\epsilon_i$  von den  $b_i$  unabhängig.

Aus den Annahmen lässt sich eine Verteilungseigenschaft für den Responsevektor  $y_i$  herleiten,

$$y_i \sim N(X_i\beta, \sigma^2 I + Z_i\Psi Z_i^T).$$

Das entspricht einer multivariate Normalverteilung der Dimension  $n_i$ . Die Varianzkovarianzmatrix  $\Psi$  kann auf verschiedene Weisen parametrisiert werden. Wenn alle Kovarianzen geschätzt werden, ergeben sich  $q(q+1)/2$  Parameter, wenn nur die Diagonalelemente geschätzt, die Kovarianzen auf Null gesetzt werden, sind das  $q$  Elemente. Diese werden im Parametervektor  $\theta$  zusammengefasst.

Die Dichtefunktion des Responsevektors  $y_i$  lautet

$$f_{y_i}(y; \beta, \sigma^2, \theta) = \frac{1}{(2\pi)^{n_i/2} |\sigma^2 I + Z_i\Psi Z_i^T|^{1/2}} * \exp\left\{-\frac{1}{2}(y - X_i\beta)^T (\sigma^2 I + Z_i\Psi Z_i^T)^{-1} (y - X_i\beta)\right\}.$$

Die Likelihoodfunktion ist definiert als

$$L(\beta, \sigma^2, \theta | y_i) = f_{y_i}(y; \beta, \sigma^2, \theta),$$

die Dichtefunktion von  $y_i$ , wobei das Argument nicht  $y$ , sondern die gesuchten Parameter  $\beta$ ,  $\sigma^2$  und  $\theta$  sind. Die Likelihoodfunktion wird auf den Vektor der beobachteten  $y_i$ s bedingt.

Die einzelnen Gruppen sind voneinander unabhängig, daraus folgt, dass die  $y_i$ s für verschiedene  $i$  unabhängig sind. Somit kann die Dichtefunktion für den gesamten Responsevektor  $y = [y_1^T, y_2^T, \dots, y_M^T]^T$  definiert werden,

$$f_y = \prod_{i=1}^M f_{y_i}.$$

Die entsprechende Likelihoodfunktion lautet

$$L(\beta, \sigma^2, \theta|y) = \prod_{i=1}^M \frac{1}{(2\pi)^{n_i/2} |\sigma^2 I + Z_i \Psi Z_i^T|^{1/2}} * \exp \left\{ -\frac{1}{2} (y_i - X_i \beta)^T (\sigma^2 I + Z_i \Psi Z_i^T)^{-1} (y_i - X_i \beta) \right\}.$$

Durch die monotone Transformation des Logarithmierens kann die Likelihoodfunktion durch Summen ausgedrückt und leichter optimiert werden. Die resultierende Funktion wird Log-Likelihood-Funktion genannt und mit einem  $l$  bezeichnet,

$$l(\beta, \sigma^2, \theta|y) = -\frac{1}{2} \sum_{i=1}^M n_i \log(2\pi) - \frac{1}{2} \sum_{i=1}^M \log |\sigma^2 I + Z_i \Psi Z_i^T| - \frac{1}{2} \sum_{i=1}^M (y_i - X_i \beta)^T (\sigma^2 I + Z_i \Psi Z_i^T)^{-1} (y_i - X_i \beta)$$

Die Funktion ist in den Parametern  $\beta, \sigma^2, \theta$  zu optimieren, um die ML-Schätzer zu bestimmen. Es handelt sich um ein komplexes Optimierungsproblem, da der Parametervektor  $\theta$  teilweise positiv und die Matrix  $\Psi$  positiv-definit ist. Weiters stellt sich heraus, dass die Form der Likelihood computational schwierig zu optimieren ist, daher gibt es eine andere äquivalente Form, welche sich besser eignet. Zunächst wird die Varianzkovarianzmatrix  $\Psi$  in Form der relativen Präzisionsfaktoren  $\Delta$  ausgedrückt,

$$\frac{\Psi^{-1}}{1/\sigma^2} = \Delta^T \Delta.$$

Ein solches  $\Delta$  existiert, falls  $\Psi$  positiv-definit ist. Im Folgenden steht der Vektor  $\theta$  für die Parameter, welche die Matrix  $\Delta$  bestimmen. Die Likelihood

wird hergeleitet, indem die Dichte der Daten, unter Bedingung der zufälligen Effekte, integriert wird.

$$L(\beta, \theta, \sigma^2 | y) = \prod_{i=1}^M f(y_i; \beta, \theta, \sigma^2) \prod_{i=1}^M \int f(y_i | b_i, \beta, \sigma^2) f(b_i | \theta, \sigma^2) db_i.$$

Hier wird zusätzlich verwendet, dass  $b_i$  und  $\epsilon_i$  unabhängig sind. Die  $b_i$ s sind Zufallsvektoren, falls  $\theta$  und  $\sigma^2$  bekannt sind, kann deren Dichte  $f(b_i | \theta, \sigma^2)$  bestimmt werden. Es handelt sich um eine multivariate Normalverteilung. Die bedingte Dichte der Daten, gegeben die zufälligen Effekte,  $f(y_i | b_i, \beta, \sigma^2)$ , ist ebenfalls eine multivariate Normalverteilung mit der Form

$$f(y_i | b_i, \beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n_i/2}} \exp \left\{ - \frac{\| y_i - X_i\beta - Z_i b_i \|^2}{2\sigma^2} \right\},$$

wobei  $\| \cdot \|$  die euklidische Norm bezeichnet. Die Dichte von  $b_i$  ist

$$\begin{aligned} f(b_i | \theta, \sigma^2) &= \frac{1}{(2\pi)^{q/2} \sqrt{|\Psi|}} \exp \left\{ - \frac{b_i^T \Psi^{-1} b_i}{2} \right\} \\ &= \frac{1}{(2\pi\sigma^2)^{q/2} \text{abs}|\Delta|^{-1}} \exp \left\{ - \frac{\| \Delta b_i \|^2}{2\sigma^2} \right\}, \end{aligned}$$

wobei  $|\Delta|$  die Determinante bezeichnet. Die Likelihoodfunktion kann durch

$$L(\beta, \theta, \sigma^2 | y) = \prod_{i=1}^M \frac{\text{abs}|\Delta|}{(2\pi\sigma^2)^{n_i/2}} \int \frac{\exp \left\{ - (\| y_i - X_i\beta - Z_i b_i \|^2 + \| \Delta b_i \|^2) / 2\sigma^2 \right\}}{(2\pi\sigma^2)^{q/2}} db_i$$

formuliert werden. Mithilfe des Pseudo-Daten-Ansatzes kann die Likelihood anders ausgedrückt werden. Hierfür sind folgende Erweiterungen des Vektors  $y_i$  und der Matrizen  $X_i$  und  $Z_i$  nötig,

$$\tilde{y}_i = \begin{bmatrix} y_i \\ 0 \end{bmatrix}, \tilde{X}_i = \begin{bmatrix} X_i \\ 0 \end{bmatrix}, \tilde{Z}_i = \begin{bmatrix} Z_i \\ \Delta \end{bmatrix}.$$

Die Dimension von  $\tilde{y}_i$  ist  $(n_i + q) \times 1$ ,  $\tilde{X}_i$  ist eine  $(n_i + q) \times k$  und  $\tilde{Z}_i$  eine  $(n_i + q) \times q$  Matrix. Die Likelihood kann jetzt geschrieben werden als

$$L(\beta, \theta, \sigma^2 | y) = \prod_{i=1}^M \frac{\text{abs}|\Delta|}{(2\pi\sigma^2)^{n_i/2}} \int \frac{\exp \left\{ - \| \tilde{y}_i - \tilde{X}_i\beta - \tilde{Z}_i b_i \|^2 / 2\sigma^2 \right\}}{(2\pi\sigma^2)^{q/2}} db_i.$$

Der Name Pseudo-Daten-Ansatz kommt daher, dass der Responsevektor  $y_i$  um zusätzliche Pseudodaten, Nullen, erweitert wird. Im Exponenten steht eine aus der Theorie der linearen Modelle bekannte Form, eine Residuenquadratsumme. Diese kann durch den OLS-Schätzer

$$\hat{b}_i = (\tilde{Z}_i^T \tilde{Z}_i)^{-1} \tilde{Z}_i^T (\tilde{y}_i - \tilde{X}_i \beta)$$

minimiert werden, welcher von dem unbekannten  $\Delta$  und dem  $\beta$  abhängt. Weiters kann die Residuenquadratsumme auf folgende Weise umgeschrieben werden,

$$\begin{aligned} \|\tilde{y}_i - \tilde{X}_i \beta - \tilde{Z}_i b_i\|^2 &= \|\tilde{y}_i - \tilde{X}_i \beta - \tilde{Z}_i \hat{b}_i\|^2 + \|\tilde{Z}_i(b_i - \hat{b}_i)\|^2 \\ &= \|\tilde{y}_i - \tilde{X}_i \beta - \tilde{Z}_i \hat{b}_i\|^2 + (b_i - \hat{b}_i)^T \tilde{Z}_i^T \tilde{Z}_i (b_i - \hat{b}_i). \end{aligned}$$

Der Exponent im Integral der Likelihoodfunktion kann auf folgende Weise aufgeteilt werden,

$$\exp \left\{ - \frac{\|\tilde{y}_i - \tilde{X}_i \beta - \tilde{Z}_i \hat{b}_i\|^2}{2\sigma^2} \right\} \exp \left\{ - \frac{(b_i - \hat{b}_i)^T \tilde{Z}_i^T \tilde{Z}_i (b_i - \hat{b}_i)}{2\sigma^2} \right\}.$$

Der erste Term hängt nicht von  $b_i$  ab, er kann aus dem Integral herausgezogen werden. Damit hat das Integral folgende Form,

$$\int \frac{\exp \left\{ - (b_i - \hat{b}_i)^T \tilde{Z}_i^T \tilde{Z}_i (b_i - \hat{b}_i) / 2\sigma^2 \right\}}{(2\pi\sigma^2)^{q/2}} db_i.$$

Die Terme im Integral gleichen bis auf eine Normierungskonstante einer multivariaten Normalverteilung mit Erwartungswert  $\hat{b}_i$  und Varianzkovarianzmatrix  $(\tilde{Z}_i^T \tilde{Z}_i)^{-1}$ . Die Normierungskonstante ist genau

$$\frac{1}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}.$$

Somit kann das gesamte Integral der Likelihoodfunktion geschrieben werden als

$$\begin{aligned} \int \frac{\exp \left\{ - \|\tilde{y}_i - \tilde{X}_i \beta - \tilde{Z}_i b_i\|^2 / 2\sigma^2 \right\}}{(2\pi\sigma^2)^{q/2}} db_i &= \\ &= \frac{\exp \left\{ - \|\tilde{y}_i - \tilde{X}_i \beta - \tilde{Z}_i \hat{b}_i\|^2 / 2\sigma^2 \right\}}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}. \end{aligned}$$

Die Likelihoodfunktion hat nun folgende Gestalt,

$$L(\beta, \theta, \sigma^2 | y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left\{ -\frac{\sum_{i=1}^M \|\tilde{y}_i - \tilde{X}_i\beta - \tilde{Z}_i\hat{b}_i\|^2}{2\sigma^2} \right\} \prod_{i=1}^M \frac{\text{abs}|\Delta|}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}.$$

Die Funktion ist in den Parametern  $\beta$ ,  $\theta$  und  $\sigma^2$  zu maximieren. Die Likelihoodfunktion kann vereinfacht werden, indem die einzelnen Parameter als Funktionen der restlichen Parameter ausgedrückt werden. Für diese werden die ML-Schätzer bestimmt und für die wahren Parameter in die Likelihood eingesetzt. Die daraus resultierende Funktion wird profilierte Likelihoodfunktion genannt. In dem Fall kann die Likelihood so profiliert werden, dass sie nur eine Funktion von  $\theta$  ist. Dazu werden die Schätzer  $\hat{\beta}(\theta)$  und  $\hat{\sigma}^2(\theta)$  benötigt. Aus der Formel ist zu erkennen, dass es sich bei den ersten zwei Termen um die Likelihood eines linearen Regressionsmodells handelt, für welches die Schätzer für  $\beta$  und  $\sigma^2$  bekannt sind.

Der ML-Schätzer  $\hat{\beta}$  hängt von  $\hat{b}_i$  ab, diese wiederum sind eine Funktion von  $\beta$ . Daher müssen die Parameter simultan geschätzt werden, was zu folgendem Minimierungsproblem führt,

$$(\hat{b}_1^T, \dots, \hat{b}_M^T, \hat{\beta}^T)^T = \arg \min_{b_1, \dots, b_M, \beta} \|y_e - X_e(b_1^T, \dots, b_M^T, \beta^T)^T\|^2,$$

wobei

$$X_e = \begin{bmatrix} Z_1 & 0 & \cdots & 0 & X_1 \\ \Delta & 0 & \cdots & 0 & 0 \\ 0 & Z_2 & \cdots & 0 & X_2 \\ 0 & \Delta & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & Z_M & X_M \\ 0 & 0 & \cdots & \Delta & 0 \end{bmatrix} \quad \text{und} \quad y_e = \begin{bmatrix} y_1 \\ 0 \\ y_2 \\ 0 \\ \vdots \\ y_M \\ 0 \end{bmatrix}$$

die Designmatrix und den Responsevektor darstellen.  $X_e$  hängt nur vom Parametervektor  $\theta$  ab. Die Lösung des Minimierungsproblems ist

$$(\hat{b}_1^T, \dots, \hat{b}_M^T, \hat{\beta}^T)^T = (X_e^T X_e)^{-1} X_e^T y_e.$$

Die Matrix  $X_e$  hat den Nachteil, dass sie von großer Dimension ist, daher wird aus computationalen Gründen nicht direkt mit ihr gearbeitet.

Der bedingte Schätzer für  $\sigma^2$  ist ebenfalls aus der Theorie des linearen Modells zu entnehmen, mit der Form

$$\hat{\sigma}^2(\theta) = \frac{\|y_e - X_e(\hat{b}_1^T, \dots, \hat{b}_M^T, \hat{\beta}^T)^T\|^2}{N}.$$

Zu beachten ist, dass beim Maximum-Likelihood-Schätzer durch  $N$  dividiert wird, und nicht durch  $N - k$ . Die bedingten Schätzer werden in die Likelihoodfunktion eingesetzt, die profilierte Likelihoodfunktion für  $\theta$  ist gegeben durch

$$L(\theta) = L(\hat{\beta}(\theta), \theta, \hat{\sigma}^2(\theta)) = \frac{\exp(-N/2)}{(2\pi\hat{\sigma}^2(\theta))^{N/2}} \prod_{i=1}^M \frac{\text{abs}|\Delta|}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}.$$

Der Zähler des ersten Terms vereinfacht sich zu  $\exp(-N/2)$ , da

$$-\sum_{i=1}^M \|\tilde{y}_i - \tilde{X}_i\beta - \tilde{Z}_i\hat{b}_i\|^2 = -\|y_e - X_e(\hat{b}_1^T, \dots, \hat{b}_M^T, \beta^T)^T\|^2$$

gesetzt werden kann.

Bei der Form der Likelihood kommen die zufälligen Effekte  $\hat{b}_i$  und die festen Effekte  $\hat{\beta}$  nicht mehr in direkter Form vor, sondern nur in der Residuenquadratsumme, welche  $\hat{\sigma}^2(\theta)$  bestimmen. Für die Bestimmung der Residuenquadratsumme wird die Regressormatrix in orthogonale Dreiecksmatrizen zerlegt, welche eine schnelle Berechnung erlauben. Mittels Optimierungsverfahren wird die Likelihoodfunktion maximiert, und  $\hat{\theta}_{ML}$  bestimmt. Auf die Optimierungsalgorithmen wird später eingegangen, zunächst wird die restringierte ML-Schätzung näher diskutiert.

### 3.3 Restringierte ML-Schätzung

Der Nachteil des gewöhnlichen Maximum-Likelihood-Schätzers ist, dass er nicht unverzerrt ist. So ist etwa für unabhängige normalverteilte Zufallsvariablen  $X_1, X_2, \dots, X_n$  der ML-Schätzer für die Varianz  $\hat{\sigma}_{ML}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ , und nicht der unverzerrte Schätzer  $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ . Das Problem gibt es bei gemischten Modellen in ähnlicher Weise, die Freiheitsgrade  $k$  für die Anzahl der festen Effekte  $\beta$  werden nicht berücksichtigt.

Die restringierte Maximum-Likelihood(REML)-Schätzung, oder auch residuale Maximum-Likelihood-Schätzung, ist eine Möglichkeit, den Bias in den Griff zu bekommen. Der Hauptgedanke ist, den Teil der Likelihood zu maximieren, welcher invariant gegenüber den festen Effekten  $\beta$  ist. Es gibt mehrere Ansätze, welche alle zum selben Ergebnis führen. Im Folgenden wird der bayesianische Ansatz skizziert.

Angenommen,  $\beta$  ist ein Zufallsvektor und folgt einer Gleichverteilung auf  $R^k$ . Das entspricht einer nichtinformativen Aprioriverteilung. Dann kann  $\beta$  aus der Likelihood heraus integriert werden,

$$L_R(\theta, \sigma^2|y) = \int_{R^k} L(\beta, \theta, \sigma^2|y) d\beta,$$

wobei  $L_R$  die restringierte Likelihoodfunktion ist. Die daraus resultierende Formel hat folgende Gestalt,

$$L_R(\theta, \sigma^2 | y) = \frac{(2\pi\sigma^2)^{-(N-k)/2}}{|\sum_{i=1}^M X_i^T V_i^{-1} X_i|^{1/2}} \prod_{i=1}^M \frac{1}{|V_i|^{1/2}} * \\ \exp \left\{ -\frac{1}{2\sigma^2} (y_i - X_i \hat{\beta})^T V_i^{-1} (y_i - X_i \hat{\beta}) \right\},$$

wobei  $V_i = I + Z_i \Psi Z_i^T / \sigma^2$  die skalierte Varianzkovarianzmatrix ist, und  $\hat{\beta}$  der gewichtete Kleinst-Quadrate-Schätzer mit der Form

$$\hat{\beta} = \left[ \sum_{i=1}^M X_i^T V_i^{-1} X_i \right]^{-1} \left[ \sum_{i=1}^M X_i^T V_i^{-1} y_i \right].$$

Beim ersten Term steht  $N - k$  im Exponenten, woran zu erkennen ist, dass die Freiheitsgrade in die Formel mit einfließen. Im Gegensatz zur normalen Likelihood ist der Term  $|\sum_{i=1}^M X_i^T V_i^{-1} X_i|^{-1/2}$  hinzugekommen. Die restringierte Log-Likelihood ist

$$l_R(\theta, \sigma^2 | y) = -\frac{N-k}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log \left| \sum_{i=1}^M X_i^T V_i^{-1} X_i \right| \\ - \frac{1}{2} \sum_{i=1}^M \log |V_i| - \frac{1}{2\sigma^2} \sum_{i=1}^M (y_i - X_i \hat{\beta})^T V_i^{-1} (y_i - X_i \hat{\beta}).$$

Wie bei der ML-Schätzung gilt, dass die Formel für die Maximierung der Funktion computational nicht geeignet ist. Eine äquivalente Form sowie die profilierte Likelihoodfunktion für  $\theta$  können auf ähnliche Weise wie im vorigen Abschnitt hergeleitet werden. Diese sind für die Optimierung besser geeignet.

Aus dem REML-Schätzer für  $\theta$  lässt sich eine Schätzung für die Kovarianzmatrix der  $b_i$ s,  $\hat{\Psi}$ , erstellen. Daraus wird der gewichtete Kleinst-Quadrate-Schätzer  $\hat{\beta}$  für die festen Effekte berechnet.

Eine Eigenschaft der REML-Schätzer ist, dass sie bei balancierten Daten dieselben Ergebnisse liefern wie die ANOVA-Schätzer, falls diese positiv sind. In dem Fall sind sie erwartungstreu. Das ist ein Vorteil gegenüber den ML-Schätzern, was der Grund dafür ist, dass sie in der Praxis häufiger verwendet werden.

Ein weiterer Unterschied zwischen ML- und REML-Schätzern ist, dass die Wahl der Kontraste für die fixen Effekte die ML-Schätzung nicht, die REML-Schätzung aber schon beeinflusst. Eine Veränderung der Designmatrix hat eine Veränderung des Terms  $|\sum_{i=1}^M X_i^T V_i^{-1} X_i|^{-1/2}$  zur Folge, welcher die restringierte Likelihoodfunktion verändert. Eine Konsequenz daraus ist, dass



Modelle mit verschiedener Struktur für die festen Effekte nicht anhand der restringierten Likelihood miteinander verglichen werden können. In dem Fall können Likelihood-Quotienten-Tests nicht durchgeführt werden.

### 3.4 Optimierungsalgorithmen

Es gibt einige Spezialfälle von linearen gemischten Modellen, wo explizite Formeln für die ML- bzw. REML-Schätzer verfügbar sind. Dazu gehört das balancierte Modell mit der Eigenschaft, dass

$$Z = X_i = Z_i \quad \text{für } i = 1, \dots, M.$$

Die Regressormatrizen für die festen und die zufälligen Parameter sind identisch. Ein Beispiel dafür ist das Modell der Wachstumskurven von Städten aus der Einleitung.

Eine weitere Form des Modells mit einer expliziten Lösung für die ML-Schätzer ist das zufällige Intercept-Modell, ebenfalls für balancierte Gruppen.

Im allgemeinen Fall sind die Lösungen für die ML- und REML-Schätzer nicht in expliziter Form vorhanden, und die Likelihoodfunktion muss mit numerischen Optimierungsverfahren maximiert werden. In der Praxis wird die Log-Likelihoodfunktion maximiert, da sie numerisch stabiler ist.

Bei linearen gemischten Modellen wird die profilierte Log-Likelihoodfunktion für  $\theta$  maximiert. Die festen Effekte  $\beta$  und die Fehlervarianz  $\sigma^2$  können daraus rückgerechnet werden.

Um eine Funktion analytisch zu maximieren, werden im Allgemeinen die ersten Ableitungen Null gesetzt und das daraus resultierende Gleichungssystem gelöst. Bei linearen gemischten Modellen handelt es sich um komplizierte nichtlineare Gleichungssysteme, wobei zusätzlich der Parameterraum beschränkt sein kann. Falls das Maximum am Rand eines Parameterraumes auftritt, ist es eher unwahrscheinlich, dass die erste Ableitung Null ist. Weiters kann es sich bei den Lösungen auch um Minima oder Sattelpunkte handeln. Daher wird nicht mit den ersten Ableitungen, sondern direkt mit der Log-Likelihoodfunktion gearbeitet.

Das Maximieren der Funktion ist eine schwierige Aufgabe, da sie komplizierte nichtlineare Terme enthält, und bei Vorhandensein von Kovarianzen zwischen den zufälligen Effekten weitere Bedingungen wie Positiv-Definitheit der Varianzkovarianzmatrix in den Iterationsschritten vorliegen.

Allgemein gilt, dass iterative Verfahren, welche die erste und zweite Ableitung der Funktion verwenden, am besten geeignet sind, um Maxima zu finden. Daneben gibt es ein Verfahren, welches einen statistischen Hintergrund hat, den EM-Algorithmus.

Von der Grundstruktur sind alle Algorithmen ähnlich, ausgehend von einem Startpunkt, bestehen sie aus einer Regel, welche den nächsten Iterationspunkt bestimmt. Dazu gehören die Richtung des Schrittes und seine Länge. In Form einer Gleichung hat das folgende Gestalt,

$$\theta_{s+1} = \theta_s + \lambda_s \delta_s, \quad s = 0, 1 \dots,$$

wobei

$$\delta_s = H_s^{-1} g_s,$$

und  $g_s$  beschreibt den Gradienten der Log-Likelihoodfunktion berechnet an der Stelle  $\theta_s$ , also  $\partial l / \partial \theta|_{\theta_s}$ . Die Schrittweite wird durch den Skalar  $\lambda_s$  bestimmt. Die Matrix  $H$  muss positiv definit sein, um eine weitere Verbesserung der Zielfunktion in der nächsten Iteration zu gewährleisten.

Einer der am häufigsten verwendeten Algorithmen ist der Newton-Raphson-Algorithmus. Bei quadratischen Funktionen benötigt der Algorithmus nur eine Iteration, um das Maximum zu erreichen, da die zweite Ableitung verwendet wird. Deshalb besteht die Matrix  $H$  aus der negativen zweiten Ableitung der Log-Likelihood. Bei mehrdimensionalem  $\theta$  ist das die negative Hesse-Matrix ( $-\partial^2 l / \partial \theta^2$ ). Weiters wird eine Schrittlänge  $\lambda_s$  benötigt. Diese wird mit einem Line-Search-Verfahren in jeder Iteration so bestimmt, dass sich der Funktionswert der Likelihood maximal verbessert. Ein Vorteil des Verfahrens ist, dass es in der Nähe des Maximums ein quadratisches Konvergenzverhalten aufweist, sich relativ rasch dem Maximum annähert. Andererseits kann es vorkommen, dass die Matrix  $H$  während der Iteration nicht positiv-definit bleibt, was am Rand des Parameterraumes oder weit entfernt des Maximums eintreffen kann. In solchen Fällen muss der Algorithmus abgebrochen werden, da das Maximum nicht erreicht wird. Die Berechnung der Hesse-Matrix in jeder Iteration ist computational recht aufwendig, es gibt Verfahren, welche einfachere Approximationen der Matrix verwenden, diese werden Quasi-Newton-Algorithmen genannt.

Eine häufig verwendete Methode ist das Fisher-Scoring. Für die Matrix  $H$  wird der Erwartungswert der negativen Hesse-Matrix verwendet. Die Matrix wird auch Informationsmatrix genannt. Falls das Modell korrekt spezifiziert ist, bleibt die Matrix über die gesamte Iteration positiv-definit. Das ist der Vorteil gegenüber dem Newton-Raphson-Algorithmus. Weiters ist das Fisher-Scoring robuster gegenüber Ausreißern, da die erwartete und nicht die beobachtete Information verwendet wird. Die durch den Algorithmus mitgelieferte asymptotische Schätzung der Varianzkovarianzmatrix ist etwas besser als bei NR-Verfahren.

Das dritte Verfahren, das im Bereich der Maximierung von Likelihoodfunktionen entwickelt wurde, ist der Expectation-Maximization-Algorithmus

(EM-Algorithmus). Im Zusammenhang mit linearen gemischten Modellen werden die zufälligen Effekte  $b_i$ ,  $i = 1, \dots, M$ , als nichtbeobachtbare Daten angesehen. Bei jedem Iterationsschritt wird der aktuelle Parametervektor  $\theta_s$  verwendet, um die Verteilung von  $b_i|y$  zu bestimmen und den bedingten Erwartungswert der Log-Likelihood für einen Punkt  $\theta$  zu berechnen. Der Schritt wird *Expectation Step* genannt. Dann wird die Log-Likelihood maximiert, um den neuen Wert  $\theta_{s+1}$  zu bestimmen (*Maximization Step*). Der Funktionswert der Likelihood wird automatisch vergrößert, deshalb wird die Schrittweite  $\lambda_s = 1$  gesetzt.

Der Nachteil des Algorithmus ist, dass er sich in der Nähe des Maximums nur langsam fortbewegt und dass eine Schätzung der Varianzkovarianzmatrix der Parameter nicht automatisch mitgeliefert wird.

Bei der praktischen Optimierung werden hybride Verfahren verwendet. Die ersten Schritte werden mit dem EM-Algorithmus iteriert, sodass der Prozess in die Nähe des Maximums gebracht wird, nach einer gewissen Anzahl an Schritten wird zum Newton-Raphson-Verfahren übergegangen, welches ein schnelleres Konvergenzverhalten in der Nähe des Maximums aufweist.

Die Startwerte für die Iterationen können aus den Daten selbst geschätzt werden. Mittels Kleinst-Quadrate-Schätzung kann eine erste Lösung für die festen Effekte  $\beta$  berechnet werden. Die Korrelation für die Daten innerhalb der Gruppen kann vernachlässigt werden. Aus der Theorie folgt, dass der Kleinst-Quadrate-Schätzer trotzdem erwartungstreu ist. Aus den Schätzern können erste Werte für die Varianzkomponenten berechnet werden. Ebenfalls können die ANOVA-Schätzer als Startpunkte herangezogen werden.

Die Frage, wann der aktuelle Parametervektor der Iteration zum ML- oder REML-Schätzer erklärt wird, kann mittels Abbruchkriterien beantwortet werden. Es gibt drei verschiedene Kriterien, anhand derer die Entscheidung zu treffen ist. Entweder wenn die Funktionswerte der Likelihoodfunktion, oder die Werte des Parametervektors sich nur geringfügig verändern, oder die Werte des Gradienten nahe bei Null sind. Eines haben die Kriterien und die drei Algorithmen gemeinsam, keines kann das Erreichen eines globalen Maximums garantieren.

## 4 Inferenz im linearen gemischten Modell

In dem Kapitel wird das Testen verschiedener Hypothesen und die Vorhersage von festen und zufälligen Effekten im linearen gemischten Modell behandelt. Zuerst wird auf Verteilungseigenschaften der ML- bzw. REML-Schätzer eingegangen.

## 4.1 Verteilung der ML- bzw. REML-Schätzer

Das Testen von Hypothesen bezüglich der Parameter eines linearen gemischten Modells beruht auf approximativen Verteilungseigenschaften der ML- bzw. REML-Schätzer. Unter bestimmten Annahmen, welche in der Praxis meistens erfüllt sind, sind die ML- bzw. REML-Schätzer asymptotisch normalverteilt. Deren Varianzkovarianzmatrix ist die Inverse der Informationsmatrix, welche aus der Log-Likelihoodfunktion  $l(\beta, \theta, \sigma^2)$  berechnet wird.

Im Folgenden werden die Eigenschaften des ML-Schätzers näher diskutiert. Ein wichtiges Resultat ist, dass die festen Parameter  $\beta$  von den Varianzkomponenten unabhängig sind, die Informationsmatrix in zwei Blöcke aufgeteilt werden kann,

$$E(\partial^2 l / \partial \beta \partial \theta) = 0,$$

$$E(\partial^2 l / \partial \beta \partial \sigma^2) = 0.$$

Ein Beweis dafür kann in Demidenko (2004) S. 121-122 nachgelesen werden. Daraus resultiert, dass die ML-Schätzer für  $\beta$  und  $\theta$  sowie  $\beta$  und  $\sigma^2$  asymptotisch unkorreliert sind, die Varianzkovarianzmatrizen lassen sich getrennt voneinander anschreiben,

$$VC(\hat{\beta}_{ML}) = \left[ \sum_{i=1}^M X_i^T (I\sigma^2 + Z_i \Psi Z_i^T)^{-1} X_i \right]^{-1}.$$

Die asymptotische Verteilung der Varianzkomponenten ist

$$\begin{bmatrix} \hat{\theta}_{ML} \\ \hat{\sigma}_{ML}^2 \end{bmatrix} \sim N \left( \begin{bmatrix} \theta \\ \sigma^2 \end{bmatrix}, I^{-1}(\theta, \sigma^2) \right),$$

und die empirische Informationsmatrix  $I$  ist definiert als

$$I(\theta, \sigma^2) = - \begin{bmatrix} \frac{\partial^2 l}{\partial \theta \partial \theta^T} & \frac{\partial^2 l}{\partial \theta \partial \sigma^2} \\ \frac{\partial^2 l}{\partial \sigma^2 \partial \theta} & \frac{\partial^2 l}{\partial \sigma^2 \partial (\sigma^2)^2} \end{bmatrix}.$$

Hier wird die für  $\theta$  und  $\sigma^2$  profilierte Log-Likelihood  $l$  verwendet. Für die praktische Durchführung werden die jeweiligen ML-Schätzer für  $\theta$  und  $\sigma^2$  eingesetzt. Für die REML-Schätzer gelten approximativ dieselben Resultate, statt der Log-Likelihood  $l$  wird die restringierte Log-Likelihood  $l_R$  für die Berechnung herangezogen. Die Verteilungseigenschaften können für das Testen von Hypothesen und die Konstruktion von Konfidenzintervallen verwendet werden.

## 4.2 Testen der festen Effekte

Um Hypothesen bezüglich der festen Effekte  $\beta$  zu testen, wird dieselbe Prozedur angewendet wie bei linearen Modellen, da eine Schätzung für die Varianzkovarianzmatrix  $I\sigma^2 + Z_i\Psi Z_i^T$  zur Verfügung steht. Das Modell hat folgende Gestalt

$$y_i = X_i\beta + \eta_i \quad i = 1, \dots, M,$$

wobei  $\eta_i$  der Vektor der Fehlerterme mit Erwartungswert Null und Varianzkovarianzmatrix  $I\sigma^2 + Z_i\Psi Z_i^T$  ist,

$$\eta_i \sim N(0, I\sigma^2 + Z_i\Psi Z_i^T).$$

In die Varianzkovarianzmatrix werden entweder die ML- oder REML-Schätzer  $\hat{\sigma}^2$  und  $\hat{\Psi}$  eingesetzt. Mit dem F-Test können Hypothesen der Form  $H_0 : R\beta = r$  auf gewöhnliche Weise getestet werden. Die Freiheitsgrade im Zähler sind gleich der Anzahl der Zeilen von  $R$ . Der einzige Unterschied zur Testprozedur im linearen Modell ist die Anzahl der Freiheitsgrade im Nenner. Diese ist abhängig davon, auf welcher Gruppierungsstufe der Effekt getestet wird. Wenn sich die Ausprägungen eines Faktors innerhalb einer Gruppe verändern, wird der Faktor innerhalb der Gruppierungsstufe getestet. Wenn sich die Ausprägungen innerhalb der Gruppen nicht verändern, wird der Faktor außerhalb der Gruppierungsstufe getestet.

Der Unterschied zwischen Faktoren eines Modells wird im Folgenden anhand eines Beispiels erläutert. Angenommen, ein neues Präparat soll an Patienten getestet werden. Dafür werden jeweils vor und nach der Einnahme die Werte der Responsevariable gemessen, pro Patient sind zwei Messungen vorhanden. Da an der Studie Frauen und Männer teilnehmen, soll der Geschlechtereffekt zusätzlich quantifiziert werden. Die Gruppen werden durch die einzelnen Patienten definiert. Der Geschlechtereffekt ist außerhalb der Gruppierungsvariable, da sich das Geschlecht nicht innerhalb einer Gruppe verändern kann. Andererseits ist der Behandlungseffekt des Präparates innerhalb der Gruppierung, da sich die Ausprägung innerhalb eines Patienten verändert.

Die Anzahl der Freiheitsgrade im Nenner hängt von der Gruppierungsstufe des zu testenden Faktors ab. Ein Faktor wird auf dem  $i$ -ten Level getestet, falls er innerhalb der  $i - 1$ -ten und außerhalb der  $i$ -ten Gruppierungsstufe liegt. Ist ein Faktor innerhalb aller Gruppierungsstufen, wird er auf dem  $Q + 1$ -ten Level getestet.  $Q$  ist die Anzahl der Gruppierungsstufen im Modell.

Der Interceptterm wird, obwohl er außerhalb jeder Gruppierung ist, nicht auf dem ersten, sondern auf dem  $Q + 1$ -ten Level getestet.

Sei  $m_i$  die Anzahl der Gruppen der  $i$ -ten Gruppierungsstufe mit der Konvention, dass  $m_0 = 1$  falls ein Interceptterm vorhanden ist und  $m_0 = 0$  falls

keiner vorhanden ist. Weiters sei  $m_{Q+1} = N$  die Stichprobenanzahl und  $p_i$  die Freiheitsgrade des zu testenden Faktors auf dem  $i$ -ten Level.  $p_i$  stellt die Anzahl der linearen Hypothesen dar. Bei einem Faktor mit drei Ausprägungen wäre  $p_i$  gleich zwei, da für diesen Faktor zwei lineare Hypothesen zu testen sind.

Die Freiheitsgrade im Nenner werden definiert durch

$$denDF_i = m_i - (m_{i-1} + p_i), \quad i = 1, \dots, Q + 1.$$

Bei einem balancierten Datensatz stimmt die Definition mit der Formel der Freiheitsgrade im linearen Modell überein, im allgemeinen Fall ist das eine Approximation.

*Profilierte Likelihoodintervalle.* Eine weitere Möglichkeit ist die Verwendung von profilierten Likelihoodintervallen, welche verwendet werden können, um einzelne Komponenten des  $\hat{\beta}$  Vektors zu testen. Z.B. soll die  $j$ -te Komponente von  $\beta$  getestet werden. Zuerst wird der ML-Schätzer für den gesamten Parametervektor berechnet,  $\hat{\beta}_{ML}$ , und es gilt  $l_{max} = l(\hat{\beta}_{ML})$ . Nun wird die  $j$ -te Komponente aus dem Vektor entfernt,  $\tilde{\beta} = [\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_k]^T$ , welcher die Dimension  $(k - 1) \times 1$  hat. Jetzt handelt es sich um ein anderes Modell,

$$\tilde{y}_i = \tilde{X}_i \tilde{\beta} + Z_i b_i + \epsilon_i \quad i = 1, \dots, M,$$

wobei  $\tilde{y}_i = y_i - \beta_j x_{i,j}$ , und  $x_{i,j}$  ist die  $j$ -te Spalte von  $X_i$ . Bei der  $n_i \times (k - 1)$  Regressormatrix  $\tilde{X}_i$  ist die  $j$ -te Spalte entfernt. Für das Modell wird der ML-Schätzer  $\tilde{\beta}_{ML}$  berechnet und die  $j$ -te Komponente fixiert. Es gilt  $\tilde{l}_{max} = \tilde{l}(\tilde{\beta}_{ML}; \beta_j)$ , das Maximum der Likelihood ist eine Funktion von  $\beta_j$ . Ein  $(1 - \alpha)100\%$ -Konfidenzintervall ist gegeben durch die Lösungen  $x$ , der Gleichung

$$\tilde{l}(x) = l_{max} - \frac{1}{2} z_{1-\alpha/2}^2,$$

und  $z_{1-\alpha/2}$  ist das  $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung. Wenn die Anzahl der Parameter relativ groß ist, empfiehlt es sich, die  $t$ -Verteilung für die Berechnung des Intervalls heranzuziehen,

$$\tilde{l}(x) = l_{max} - \frac{1}{2} q_{1-\alpha/2, N-k}^2,$$

wobei  $q_{1-\alpha/2, N-k}^2$  das  $(1 - \alpha/2)$ -Quantil der  $t$ -Verteilung mit  $N - k$  Freiheitsgraden ist. Der Vorteil gegenüber der ersten Form ist, dass das die exakten Konfidenzintervalle bei Annahme eines linearen Modells und bei bekannter Varianzkovarianzmatrix ist. Allgemein gilt, dass die profilierten Likelihoodintervalle eine etwas bessere Performance aufweisen als die auf der Informationsmatrix beruhenden Waldschen Konfidenzintervalle.

### 4.3 Testen der Varianzkomponenten

Das Testen der Varianzkomponenten in einem linearen gemischten Modell stellt ein schwierigeres Problem dar als das Testen der festen Effekte. Der Grund dafür ist, dass die Hypothese  $H_0 : \sigma^2 = 0$  bzw.  $H_1 : \sigma^2 > 0$  am Rand des Parameterraumes von  $\sigma^2$  liegt. Das Problem gibt es bei Parametern welche in ganz  $R$  definiert sind nicht.

Bei der einfachen Varianzanalyse mit balancierten Gruppen stellt sich heraus, dass die Signifikanz der Varianzkomponente  $\sigma_B^2$  mit dem gewöhnlichen  $F$ -Test beantwortet werden kann. Ähnliche Resultate gibt es für unbalancierte oder mehrfaktorielle Designs, die Teststatistiken sowie deren Verteilungen werden jedoch zunehmend komplizierter und müssen für jedes Modell speziell hergeleitet werden.

Mit dem Likelihood-Quotienten-Test ist es möglich, zwei verschachtelte Modelle miteinander zu vergleichen. Der Parameterraum des ersten Modells muss eine Untermenge des Parameterraumes des zweiten Modells sein. Ist die Teststatistik signifikant, ist das komplexe Modell dem einfachen vorzuziehen. Der Likelihood-Quotienten-Test ist wie folgt definiert. Sei  $L_1$  das Maximum der Likelihoodfunktion des einfachen und  $L_2$  das Maximum des komplexen Modells. Dann folgt, dass  $L_2 > L_1$  und  $\log L_2 > \log L_1$ . Die Teststatistik ist

$$2 \log(L_2/L_1) = 2[\log(L_2) - \log(L_1)].$$

Sei  $k_1$  die Anzahl der Parameter des einfachen und  $k_2$  die Anzahl der Parameter des komplexen Modells, dann ist die Teststatistik unter der Nullhypothese  $\chi^2$  verteilt mit  $k_2 - k_1$  Freiheitsgraden. Die Nullhypothese bedeutet, dass das komplexe Modell keine Verbesserung in der Modellanpassung gegenüber dem einfachen Modell zur Folge hat. Die Tests können mit Maximum-Likelihood und restringierter Maximum-Likelihood Anpassung durchgeführt werden. Bei REML-Schätzungen können nur die Varianzkomponenten verglichen werden, die Spezifikation der festen Effekte muss in beiden Modellen übereinstimmen.

Da die Varianzkomponenten im einfachen Modell auf Null gesetzt werden, dem Rand des Parameterraumes, gelten die asymptotischen Eigenschaften des Likelihoodquotienten nur mehr bedingt. Im speziellen werden die P-Werte zu groß sein, was zu konservativen Testergebnissen führt. Das Vorhandensein einer Varianzkomponente wird zu häufig abgelehnt. Es gibt Adjustierungen der  $\chi^2$ -Verteilung, welche aber nicht allgemein anwendbar sind.

Eine Alternative stellen Computersimulationen dar, womit die Verteilung der Teststatistik unter der Nullhypothese simuliert werden kann. In jeder Iteration werden Daten des Responsevektors  $y$  unter dem einfachen Modell generiert. Dann werden die beiden zu vergleichenden Modelle angepasst, die resultierenden Werte der Likelihoodfunktion der ML- oder REML-Schätzer

berechnet und der Likelihoodquotient bestimmt. Somit ist für jede Iteration ein simulierter Wert der Teststatistik unter der Nullhypothese vorhanden. Der Vektor entspricht einer simulierten Verteilung der Teststatistik, anhand derer ein P-Wert für das aktuelle Testproblem angegeben werden kann. Je nach Anzahl der Simulationen, kann der Standardfehler des P-Wertes verkleinert werden. Die Resultate sind verlässlicher, als die auf den theoretischen Quantilen der  $\chi^2$ -Verteilung beruhenden Ergebnisse.

Informationskriterien bieten eine weitere Möglichkeit, verschiedene Modelle zu vergleichen. Dazu gehört das

$$AIC = -2l_{max} + 2n_{par}$$

und das

$$BIC = -2l_{max} + n_{par} \log(N),$$

wobei  $n_{par}$  die Anzahl der Parameter im Modell ist. Bei den Definitionen gilt, je kleiner der Wert des Kriteriums, desto besser ist die Anpassung. Für die REML-Schätzungen wird die Log-Likelihood  $l_R$  verwendet und  $\log(N)$  durch  $\log(N - k)$  ersetzt.

#### 4.4 Konfidenzintervalle für Varianzkomponenten

Mit der ML- oder REML-Methode können Schätzer für den Parametervektor  $\theta$  berechnet werden. Die Elemente des Vektors bestimmen die Varianz-kovarianzmatrix der zufälligen Effekte  $b_i$ . Im allgemeinen Fall ist das eine symmetrische  $q \times q$  Matrix mit den Varianzen der einzelnen Effekte in der Hauptdiagonale und den Kovarianzen oder Korrelationen in den Nebendiagonalen.

Angenommen, es soll ein Konfidenzintervall für eine Varianzkomponente  $\sigma^2$  des Vektors  $\theta$  konstruiert werden. Die Schätzung ist approximativ  $N(\sigma^2, Var(\hat{\sigma}^2))$ -verteilt. Durch das Logarithmieren des Parameters kann der Parameterraum auf  $(-\infty, \infty)$  erweitert werden. Mithilfe der Deltamethode wird die Varianz des transformierten Schätzers bestimmt, welcher approximativ folgende Verteilung hat,

$$N(\log(\sigma^2), \sigma^{-4}Var(\hat{\sigma}^2)).$$

In der Praxis kann der Parameter  $\sigma^2$  durch die ML- oder REML-Schätzung ersetzt werden. Ein  $(1 - \alpha)100\%$ -Konfidenzintervall ist gegeben durch

$$\log(\hat{\sigma}^2) \pm z_{1-\alpha/2} \hat{\sigma}^{-2} \sqrt{Var(\hat{\sigma}^2)}.$$



Die originale Skalierung des Intervalls wird rückgerechnet durch,

$$\exp \left( \log(\hat{\sigma}^2) \pm z_{1-\alpha/2} \hat{\sigma}^{-2} \sqrt{\text{Var}(\hat{\sigma}^2)} \right).$$

Für Konfidenzintervalle von Korrelationsparametern  $\rho$ , welche auf  $[-1, 1]$  definiert sind, ist die allgemeine Logit-Transformation  $\log[(1 + \rho)/(1 - \rho)]$  angebracht. Das Konfidenzintervall hat folgende Gestalt,

$$\tanh \left( \arctan(\hat{\rho}) \pm z_{1-\alpha/2} (1 - \hat{\rho}^2)^{-1} \sqrt{\text{Var}(\hat{\rho})} \right).$$

Die Konfidenzintervalle basieren auf der beobachteten Informationsmatrix und gelten daher nur approximativ.

## 4.5 Vorhersage im gemischten Modell

Obwohl die zufälligen Effekte  $b_i$  keine Modellparameter in herkömmlicher Weise sind, können ihre Werte im Nachhinein geschätzt werden. Die Schätzer können als der bedingte Erwartungswert der  $b_i$ s, gegeben die Responsevariable  $y_i$ , berechnet werden. Die Herleitung hat einen starken Zusammenhang mit bayesianischer Statistik. Die Schätzungen für  $b_i$  können wie folgt berechnet werden,

$$\hat{b}_i = \Psi Z_i^T (I\sigma^2 + Z_i \Psi Z_i^T)^{-1} (y - X_i \beta) \quad i = 1, \dots, M,$$

in der Praxis werden für  $\beta$  der gewichtete Kleinst-Quadrate-Schätzer  $\hat{\beta}(\theta)$  und für  $\Psi$  sowie  $\sigma^2$  die jeweiligen ML- oder REML-Schätzer eingesetzt. Die  $\hat{b}_i$ s werden BLUPs (*Best Linear Unbiased Predictor*) genannt.

Die Vorhersage von Werten der Responsevariable für gegebene Beobachtungen ist eine der wichtigsten Aufgaben eines statistischen Modells. Es können für die im Datensatz vorhandenen Beobachtungen Vorhersagen gemacht werden, diese werden gefittete Werte genannt, oder eine Vorhersage für neue Kovariableninformation ist von Interesse. Gefittete Werte sind im Zusammenhang mit der Modellprüfung von großer Bedeutung, da die Eigenschaften der Fehlerterme verifiziert werden können.

Für gemischte Modelle können Vorhersagen auf den verschiedenen Stufen des Modells getroffen werden. Die erste Stufe des Modells entspricht der Gesamtpopulation, hier ist die Information für alle Gruppen gleich. Sei  $x_h$  ein Kovariablenvektor für die festen Effekte, dann ist der Erwartungswert der Responsevariable  $y_h$  gegeben durch

$$E(y_h) = x_h^T \beta.$$

Wenn Vorhersagen auf einer tieferen Stufe  $k$  des Modells getroffen werden sollen, sind die bedingten Erwartungswerte der Responsevariablen, gegeben die zufälligen Effekte der Stufen  $\leq k$ , zu berechnen. Sei  $z_h(i)$  ein Kovariablenvektor, welcher Informationen der  $i$ -ten Gruppe der ersten Verschachtelung beinhaltet. Der Erwartungswert der Vorhersage ist

$$E(y_h|b_i) = x_h^T \beta + z_h(i)^T b_i.$$

Für Vorhersagen auf der zweiten Stufe wird zusätzliche Kovariableninformation benötigt. Sei  $z_h(i, j)$  ein Kovariablenvektor für die  $i$ -te Gruppe der ersten und  $j$ -te Gruppe der zweiten Verschachtelung, dann ist der Erwartungswert der Responsevariablen

$$E(y_h|b_i, b_{ij}) = x_h^T \beta + z_h(i)^T b_i + z_h(i, j)^T b_{ij}.$$

Das Schema lässt sich auf beliebig viele Stufen erweitern.

In der Praxis können die Vorhersagen durch die Verwendung der bedingten Schätzer  $\hat{\beta}(\theta)$  und den BLUPs  $\hat{b}_i(\theta)$  berechnet werden. Die besten linearen unverzerrten Prädiktoren der oben angeführten Erwartungswerte lauten

$$\begin{aligned}\hat{y}_h &= x_h^T \hat{\beta}(\theta), \\ \hat{y}_h(i) &= x_h^T \hat{\beta}(\theta) + z_h(i)^T \hat{b}_i(\theta), \\ \hat{y}_h(i, j) &= x_h^T \hat{\beta}(\theta) + z_h(i)^T \hat{b}_i(\theta) + z_h(i, j)^T \hat{b}_{ij}(\theta).\end{aligned}$$

Für den unbekannten Parametervektor  $\theta$  wird der ML- oder REML-Schätzer eingesetzt.

Die folgenden Kapitel beschäftigen sich mit der Anwendung gemischter Modelle anhand eines speziellen Datensatzes, welcher aus einer Untersuchung bezüglich der Haltung von Milchziegen stammt. Die Daten wurden an der Veterinärmedizinischen Universität aufgenommen und dankenswerterweise von Frau Prof. Waiblinger bereitgestellt. Zuerst werden deskriptive Analysen durchgeführt, dann werden mit den Statistikpaketen R, SPSS und SAS lineare gemischte Modelle angepasst und deren Ergebnisse verglichen.

## 5 Der Datensatz

Das Thema der Untersuchung ist der Vergleich von behornten und unbehornten Milchziegen in Bezug auf deren Sozialverhalten. Grundaktivität, Verletzungen und Stress der Tiere sind die interessierenden Merkmale. Die Anzahl der Stresshormone kann durch den Kortisolgehalt, welcher die Responsevariable darstellt, in den Milch- bzw. Kotproben der Tiere gemessen werden.

In dem Betrieb wurden je 79 behorn- und unbehorn- Tiere getrennt voneinander gehalten. Für jedes Tier wurde ein Index erstellt, welcher die sozialen Interaktionen und die Aktivität misst. Dann wurden die zehn rangtiefsten, zehn rangmittlersten und zehn ranghöchsten Tiere der unbehorn- und behorn- Gruppe ausgewählt. Die 60 Tiere sind in dem Datensatz vorhanden, wobei die Tiere 1-30 die Gruppe der behorn- und die Tiere 31-60 die Gruppe der unbehorn- Tiere darstellen. An 8 verschiedenen Tagen wurden Milch- bzw. Kotproben entnommen. Diese sind der 19. und 20. Mai, 3. und 4. Juni, 1. und 2. Juli sowie 13. und 14. Juli des Jahres 2008. An den Tagen stehen Messungen des Kortisolgehalts zu Verfügung. Für 43 Tiere sind an allen Zeitpunkten Messungen vorhanden, für 13 Tiere fehlt jeweils eine Messung, und für 4 Tiere fehlen zwei Messungen. Im Datensatz ist für die Messung des 38-ten Tieres am 2. Juli ein Kortisolgehalt von 3786,53 eingetragen. Bei dem Wert wird es sich wahrscheinlich um einen Mess- oder Tippfehler handeln, daher geht er ebenfalls als fehlender Wert in die Analyse ein. Die Stichprobenanzahl im Datensatz ist somit 459.

Die zur Verfügung stehenden Faktoren sind die Behornung mit den Ausprägungen 0 für „unbehorn-“ und 1 für „behorn-“, der Probezeitpunkt mit 8 verschiedenen Ausprägungen entsprechend dem Datum der Probeentnahme und dem Faktor Zeitpunkt mit vier Ausprägungen, wobei die jeweils aufeinander folgenden Tage zusammengefasst werden.

Die einzige in dem Datensatz vorhandene Kovariable ist der Dominanzindex, welcher für die Auswahl der Tiere herangezogen wurde und Werte im Intervall  $[0; 1]$  annimmt.

Eine fachspezifische Fragestellung ist die Untersuchung des Faktors Be-

hornung, der Vergleich zwischen der behornten und unbehornten Gruppe in Bezug auf das Stressverhalten ist maßgebend. Die restlichen Faktoren und die Kovariable dienen dazu, die Streuung in den Daten zu verringern, um schärfere Aussagen über die Signifikanz der Behornung machen zu können.

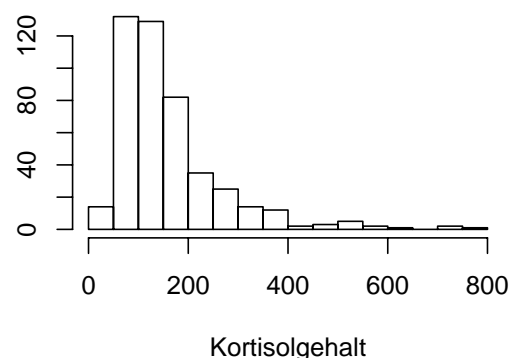
Die Analysen werden mit den Softwarepaketen R (Version 2.8.1), SPSS Statistics (Version 17.0.1) sowie SAS (Version 9.2) durchgeführt. Im Softwarepaket R wird das Paket `nlme` (Version 3.1-96), in SPSS und SAS die Prozedur MIXED zum Anpassen linearer gemischter Modelle verwendet.

## 6 Deskriptive Analyse

Die Responsevariable ist der an 8 Zeitpunkten gemessene Kortisolgehalt der Tiere. Es sind 459 Messungen vorhanden. Die erste Tabelle zeigt einige deskriptive Statistiken der Variable, unterhalb ist ein Histogramm abgebildet. Zu beachten ist, dass es sich bei den Daten um wiederholte Messungen handelt, Korrelationsstrukturen für die Messungen innerhalb eines Tieres sind vorhanden.

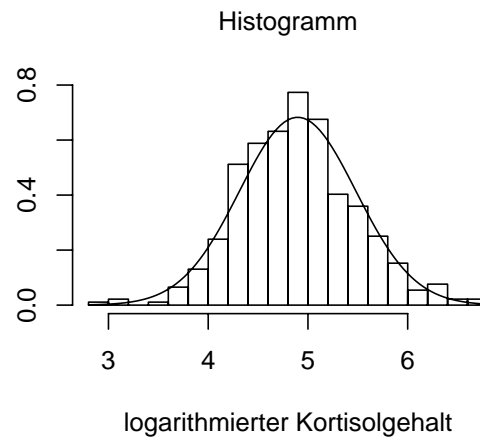
Mittelwert	159,9
Median	129,5
Minimum	19,5
Maximum	757,5
Standardabweichung	108,5

Histogramm



Die Verteilung der Variable ist nicht symmetrisch, einige extreme Werte befinden sich am rechten Rand. Das ist auch daran zu erkennen, dass der Mittel-

wert um einiges höher liegt als der Median. Durch die monotone Transformation des Logarithmus kann die Verteilung in eine symmetrische Form gebracht werden. Zusätzlich kann die Annahme einer Normalverteilung gemacht werden, was an der in folgendem Histogramm eingezeichneten Dichtekurve zu sehen ist.

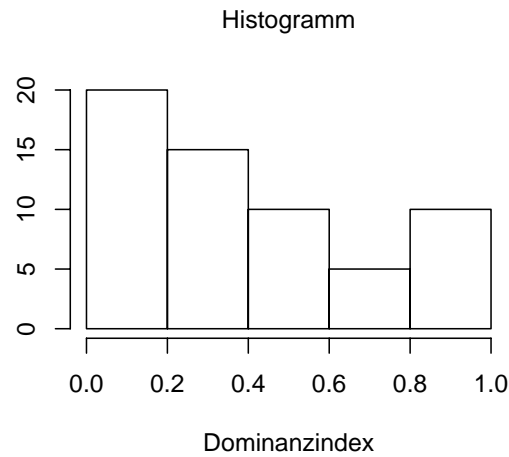


In folgender Tabelle sind einige deskriptive Statistiken der transformierten Responsevariable eingetragen.

Mittelwert	4,897
Median	4,863
Minimum	2,969
Maximum	6,630
Standardabweichung	0,584

Der Dominanzindex ist eine Maßzahl für die Interaktionen und die Aktivität jedes Tieres. Die Variable ist invariant innerhalb der Tiere, für jedes Tier ist ein Wert im Intervall  $[0; 1]$  erstellt worden. Folgende Tabelle zeigt deskriptive Statistiken der Variable, unterhalb ist ein Histogramm abgebildet.

Mittelwert	0,41
Median	0,35
Minimum	0
Maximum	1
Standardabweichung	0,313



Da aus der Gruppe mit 79 Tieren die rangniedrigsten, rangmittlersten und ranghöchsten Tiere ausgewählt wurden, kann für die Gruppe der unbehornen und behornen Tiere eine Bandbreite des Dominanzindex angegeben werden.

Behornung	Tier	Dominanzindex
0	1-10	0-0,154
0	11-20	0,333-0,555
0	21-30	0,8-1
1	1-10	0-0,125
1	11-20	0,308-0,364
1	21-30	0,545-0,813

An den Bandbreiten ist zu erkennen, dass die unbehornen Tiere einen etwas höheren Dominanzindex aufweisen.

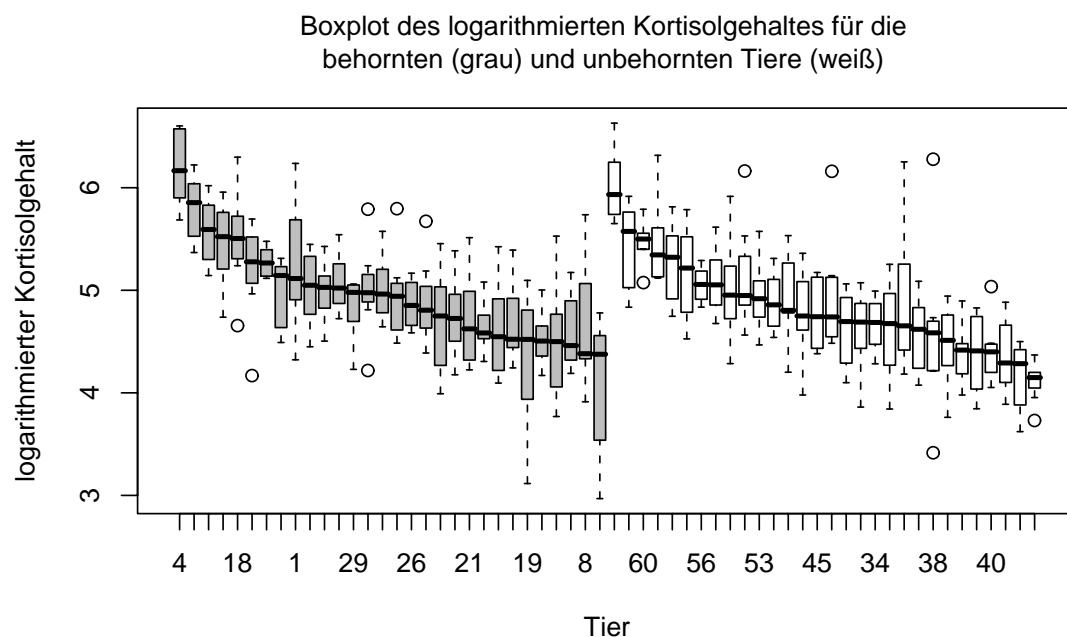
Im Datensatz befinden sich Werte von je 30 behornen und unbehornen Tieren. Eine Berechnung des Mittelwertes der Responsevariable für die einzelnen Gruppen über alle Messungen der Tiere wird nicht angegeben, da die Messungen für ein Tier korreliert sind. Aussagekräftiger ist eine Mittelung der Messungen für jedes Tier, darauf basierend können deskriptive Statistiken für die behornen und unbehornen Tiere gerechnet werden. Mit der R-Funktion `gsummary`, welche in dem Paket `nlme` zur Verfügung steht, können solche Datenmanipulationen auf einfache Weise durchgeführt werden.

Behornung	0	1
Mittelwert	150,9	169,3
Median	141	146,9

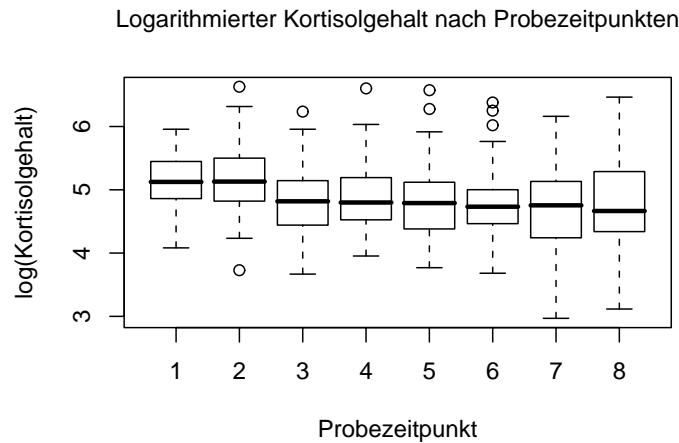
Die Statistiken basieren auf den jeweils 30 Mittelwerten für jedes Tier, zu erkennen ist, dass die Werte des Kortisolgehalts für die behornnte Gruppe etwas höher liegen. Der logarithmierte Kortisolgehalt ist eine für die Analyse geeignetere Größe, da die Verteilung symmetrischer ist. In der folgenden Tabelle sind die Statistiken für die Transformation angegeben.

Behornung	0	1
Mittelwert	4,838	4,95
Median	4,834	4,907

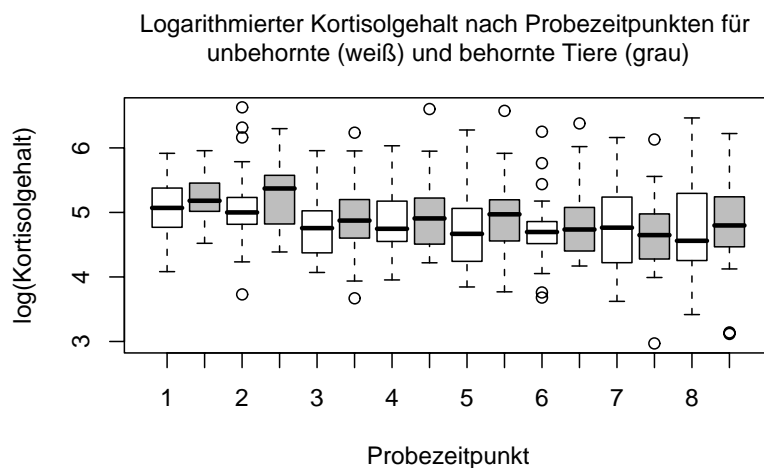
Eine visuelle Vergleichsmöglichkeit zwischen den unbehornten und behornnten Tieren bietet folgender Boxplot, welcher auf dem logarithmierten Kortisolgehalt basiert. Ein Unterschied zwischen den Gruppen ist schwer auszumachen.



Die Variable Probezeitpunkt besteht aus 8 Ausprägungen, welche die Zeitpunkte der Probeentnahme durchnummerieren. Sie ist nicht als Kovariable verwendbar, da die Abstände zwischen den Ausprägungen nicht den tatsächlichen Zeitspannen entsprechen. Zwischen den Probezeitpunkten 1 und 2 liegt gerade ein Tag, zwischen den Probezeitpunkten 2 und 3 liegen 14 Tage. Dadurch ist die Variable als Faktor zu verstehen und kann mittels Boxplots graphisch dargestellt werden.



Es ist zu erkennen, dass die logarithmierten Werte des Kortisolgehalts für die Probezeitpunkte 1 und 2 am höchsten liegen, die letzten zwei Probezeitpunkte weisen das niedrigste Niveau auf. Das lässt sich auch an den Mittelwerten erkennen, für die Probezeitpunkte 1 und 2 sind die Mittelwerte 5,14 bzw. 5,17, für den 7-ten und 8-ten Probezeitpunkt 4,73 bzw. 4,77. Wechselwirkungen zwischen dem Faktor Behornung und dem Faktor Probezeitpunkt können in den folgenden Boxplots untersucht werden.



Die Gruppe der behornen Tiere liegt in den meisten Fällen etwas höher, gerade der 7-te Probezeitpunkt macht eine Ausnahme. Das Ausmaß der Unterschiede zwischen der behornen und unbehornen Gruppe ist für die Probezeitpunkte ähnlich, es sind keine Wechselwirkungen vorhanden.

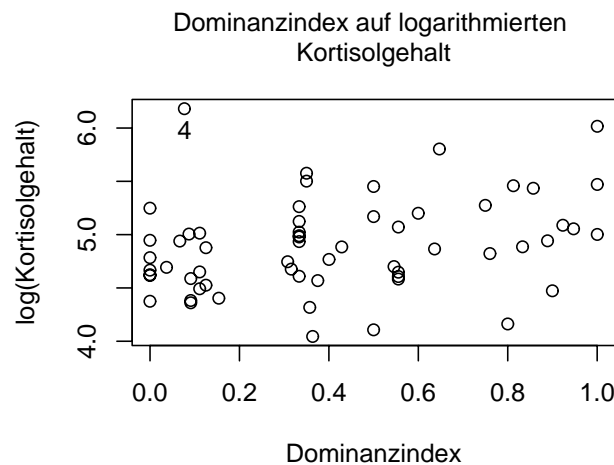


Der paarweise Zusammenhang zwischen den Probezeitpunkten für aufeinander folgende Tage ist offensichtlich. Die Variable Zeitpunkt fasst die aufeinander folgenden Tage zusammen und besteht aus vier Ausprägungen, welche mit 7,8,9,10 nummeriert sind. In der folgenden Tabelle sind einige deskriptive Merkmale der Variable zusammengefasst.

Zeitpunkt	7	8	9	10
Mittelwert	5,16	4,88	4,8	4,75
Median	5,13	4,81	4,77	4,72

Vom ersten bis zum letzten Zeitpunkt ist ein durchgehender Abfall im logarithmierten Kortisolgehalt zu erkennen.

In der folgenden Graphik ist eine Gegenüberstellung des Dominanzindex und des logarithmierten Kortisolgehalts zu sehen. Für den logarithmierten Kortisolgehalt ist für jedes Tier der Wert eingetragen, welcher durch die Mittelung über die Probezeitpunkte berechnet wird. Bis auf das Tier 4 ist ein leicht positiver Zusammenhang zu erkennen, welcher sich durch eine Korrelation von 0,29 mit einem signifikanten P-Wert von 0,027 quantifizieren lässt.



Die deskriptiven Statistiken und Graphiken dienen dazu, einen ersten Überblick über die Struktur der Daten zu bekommen. In den folgenden Abschnitten werden statistische Modelle angepasst, um genaue Schätzungen der Effekte zu berechnen und Hypothesen zu testen.

Mit folgendem R-Code können die Statistiken und die Graphiken der deskriptiven Analyse erstellt werden. Die Variablen sind in einem `data.frame` mit dem Namen `Kortisol` gespeichert.

```

> attach(Kortisol)

## Deskriptive Statistiken
> summary(Kortisolgehalt)
> sd(Kortisolgehalt)

## Histogramm Kortisolgehalt
> hist(Kortisolgehalt,20)

## Histogramm LogKortisolgehalt mit Normalverteilungskurve
> hist(LogKortisolgehalt,20,prob=T)
> curve(dnorm(x,mean(LogKortisolgehalt),sd(LogKortisolgehalt))
,add=T)

## Histogramm Dominanzindex
> hist(Dominanzindex,5)

## Boxplot Logarithmierter Kortisolgehalt für die
## behornten (grau) und unbehornten Tiere (weiß)
> Median_behornt=tapply(Kortisol[Kortisol$Behornung==1,]
  $LogKortisolgehalt,Kortisol[Kortisol$Behornung==1,]$Tier,
  median)
> Median_unbehornt=tapply(Kortisol[Kortisol$Behornung==0,]
  $LogKortisolgehalt,Kortisol[Kortisol$Behornung==0,]$Tier,
  median)
> Reihenfolge_behornt=order((1:30)[order(Median_behornt,
  decreasing=T)])
> Reihenfolge_unbehornt=order((31:60)[order(Median_unbehornt,
  decreasing=T)])
> par(font.main=1,cex.main=1)
> boxplot(LogKortisolgehalt~factor(Tier),data=Kortisol,
  col=rep(c("grey","white"),each=30),at=c(Reihenfolge_behornt,
  (Reihenfolge_unbehornt+30)),ylab="logarithmierter Kortisol
  gehalt",xlab="Tier",main="Boxplot des logarithmierten
  Kortisolgehalts für die behornten (grau) und unbehornten
  Tiere (weiß)")

## Boxplot Logarithmierter Kortisolgehalt nach Probezeitpunkten
> boxplot(LogKortisolgehalt~factor(Probezeitpunkt),data=Kortisol,
  main="Logarithmierter Kortisolgehalt nach Probezeitpunkten",
  xlab="Probezeitpunkt",ylab="log(Kortisolgehalt)")

```

```
## Boxplot Logarithmierter Kortisolgehalt nach Probezeit-
## punkten für unbehornte (weiß) und behornte Tiere (grau)
> boxplot(LogKortisolgehalt~factor(Behornung)+factor
  (Probezeitpunkt),data=Kortisol,main="Logarithmierter
  Kortisolgehalt nach Probezeitpunkten für unbehornte
  (weiß) und behornte Tiere (grau)",xlab="Probezeitpunkt",
  ylab="log(Kortisolgehalt)",col=rep(c("white","grey"),8),
  names=c("1","","2","","3","","4","","5","","6","","7","","
  "8",""))

## Scatterplot Dominanzindex auf logarithmierten Kortisolgehalt
> x=gsummary(Kortisol,mean,group=Tier)$Dominanzindex
> y=gsummary(Kortisol,mean,group=Tier)$LogKortisolgehalt
> plot(x,y,col=1,main="Dominanzindex auf logarithmierten
  Kortisolgehalt",xlab="Dominanzindex",ylab="log(Kortisolgehalt)")
> text(x[4],y[4],labels=gsummary(Kortisol,mean,group=Tier)
  $Tier[4],pos=1)
```

## 7 Zufälliges Intercept-Modell

Zu Beginn wird das zufällige Intercept-Modell angepasst. Es stellt, wie schon im Theorieteil beschrieben, das einfachste lineare gemischte Modell dar, da es nur einen festen Effekt, den Intercept, beinhaltet. Die Tiere stellen die Gruppierungsvariable dar, deren Schwankung um den Mittelwert in Form einer Varianzkomponente geschätzt wird. Das Modell hat folgende Form,

$$\log(Kortisolgehalt)_{ij} = \beta + b_i + \epsilon_{ij}.$$

$\log(Kortisolgehalt)_{ij}$  beschreibt den  $j$ -ten logarithmierten Kortisolgehalt des  $i$ -ten Tieres. Für den Datensatz gilt  $i = 1, \dots, 60$ , und  $j = 1, \dots, n_i$  nummeriert die Beobachtungen für das  $i$ -te Tier,  $n_i$  ist in dem Fall 6, 7 oder 8.  $\beta$  ist der feste Interceptparameter, die  $b_i$ s sind die zufälligen Effekte, welche normalverteilt sind mit  $N(0, \sigma_f^2)$ .  $\epsilon_{ij}$  sind die Fehlerterme mit  $N(0, \sigma^2)$ . In dem Modell wird auf den logarithmierten Kortisolgehalt regressiert, da das zu besseren empirischen Verteilungseigenschaften der zufälligen Effekte sowie der Residuen führt. Aus Sicht der ursprünglichen Skalierung der Responsevariable handelt es sich um ein log-lineares Modell. Die Schätzungen können auf folgende Weise zurück transformiert werden. Sei  $y$  der Kortisolgehalt, dann lautet das Modell

$$\log(y_{ij}) = \beta + b_i + \epsilon_{ij}.$$

Durch exponentieren kann die Responsevariable in die ursprüngliche Lage gebracht werden,

$$y_{ij} = \exp(\beta + b_i + \epsilon_{ij}).$$

Der Erwartungswert kann wie folgt berechnet werden,

$$\begin{aligned} E(y_{ij}) &= E(\exp(\beta) \exp(b_i + \epsilon_{ij})) \\ &= \exp(\beta) E(\exp(b_i + \epsilon_{ij})) \\ &= \dots \end{aligned}$$

$b_i$  und  $\epsilon_{ij}$  sind unabhängige normalverteilte Größen, deren Summe nach dem Reproduktionssatz wieder normalverteilt ist. Der Erwartungswert ist Null, und die Varianz der Summe ist  $\sigma_I^2 + \sigma^2$ . Falls  $X$  einer Normalverteilung folgt, wird  $\exp(X)$  als log-normalverteilte Zufallsgröße bezeichnet, deren Erwartungswert  $\exp(\mu + \sigma^2/2)$  ist. Zusammengesetzt ergibt das

$$\begin{aligned} \dots &= \exp(\beta) \exp((\sigma_I^2 + \sigma^2)/2) \\ &= \exp(\beta + (\sigma_I^2 + \sigma^2)/2). \end{aligned}$$

Die Berechnung der Varianz der  $y_{ij}$  verläuft auf ähnliche Weise,

$$\begin{aligned} \text{Var}(y_{ij}) &= E(y_{ij}^2) - E(y_{ij})^2 \\ &= \exp(2\beta) E(\exp(2(b_i + \epsilon_{ij}))) - E(y_{ij})^2 \\ &= \exp(2\beta) \exp(2(\sigma_I^2 + \sigma^2)) - E(y_{ij})^2 \\ &= \exp(2(\beta + \sigma_I^2 + \sigma^2)) - \exp(2(\beta + \sigma_I^2 + \sigma^2)) \end{aligned}$$

Für die Berechnung werden die Parameterschätzungen eingesetzt.

## 7.1 Analyse in R

Die im Statistikpaket R zur Verfügung stehende Funktion zum Anpassen linearer gemischter Modelle ist `lme`, welche in dem Paket `nlme` zur Verfügung steht. Für das zufällige Intercept-Modell wird folgender Befehl verwendet,

```
lme(log(Kortisolgehalt)~1,data=Kortisol,random=~1|Tier).
```

`Kortisol` bezeichnet den `data.frame`, in welchem die Variablen gespeichert sind. Zuerst werden die festen Effekte spezifiziert, dann der Datensatz und zuletzt wird die Struktur der zufälligen Effekte und die Gruppierungsvariable angegeben. Es ist nicht notwendig die Variable `log(Kortisolgehalt)` vorher zu berechnen, das kann in der Modellformel direkt eingegeben werden. Der erste R-Output hat folgende Gestalt,

```

Linear mixed-effects model fit by REML
  Data: Kortisol
  Log-restricted-likelihood: -323.1352
  Fixed: log(Kortisolgehalt) ~ 1
(Intercept)
  4.894722
Random effects:
  Formula: ~1 | Tier
          (Intercept)  Residual
StdDev:    0.4083588  0.4256231
Number of Observations: 459
Number of Groups: 60

```

Das Modell wird mit der restringierten Maximum-Likelihood-Methode (REML) angepasst. Der geschätzte Interceptterm  $\hat{\beta} = 4,8947$ . Obwohl der Wert bis auf die dritte Kommastelle mit dem einfachen Mittelwert des logarithmierten Kortisolgehalts identisch ist, ist er der „korrekt“ berechnete Mittelwert, da die Korrelationen innerhalb der Tiere mit einbezogen sind. Die Ähnlichkeit zum einfachen Mittelwert ist darauf zurückzuführen, dass die Anzahl der Beobachtungen für die einzelnen Tiere in etwa gleich groß ist, zusätzlich ein großer Stichprobenumfang zur Verfügung steht. Dadurch ist der einfach gerechnete Mittelwert sehr stabil.

Die weiteren zwei Parameter des Modells sind die Varianzkomponenten  $\sigma_I^2$  und  $\sigma^2$ . Die Schätzungen werden von R in Form der Standardabweichungen angegeben, für die Varianzen lauten sie  $\hat{\sigma}_I^2 = 0,167$  und  $\hat{\sigma}^2 = 0,181$ .

Daraus lässt sich eine Schätzung für die Intra-Klassen-Korrelation berechnen mit einem Wert von  $\hat{\sigma}_I^2 / (\hat{\sigma}_I^2 + \hat{\sigma}^2) = 0,48$ .

Der Mittelwert, in die ursprüngliche Skalierung rückgerechnet, ist 158,96. Das ist im Vergleich zum einfach gerechneten Mittelwert mit 159,91 kein großer Unterschied. Die aus dem Modell errechnete Standardabweichung des Kortisolgehalts ist 102,542 im Vergleich zu einem einfach gerechneten Wert mit 108,5.

Die Eigenschaft der approximativen Normalverteilung für Maximum-Likelihood-Schätzer wird verwendet, um Konfidenzintervalle zu konstruieren. Die R Funktion `intervals` berechnet die auf der Inversen der empirischen Informationsmatrix beruhenden Intervalle für die festen Effekte und die Varianzkomponenten.

```

Approximate 95% confidence intervals
Fixed effects: lower      est.    upper
(Intercept) 4.783924 4.894722 5.00552
Random Effects:

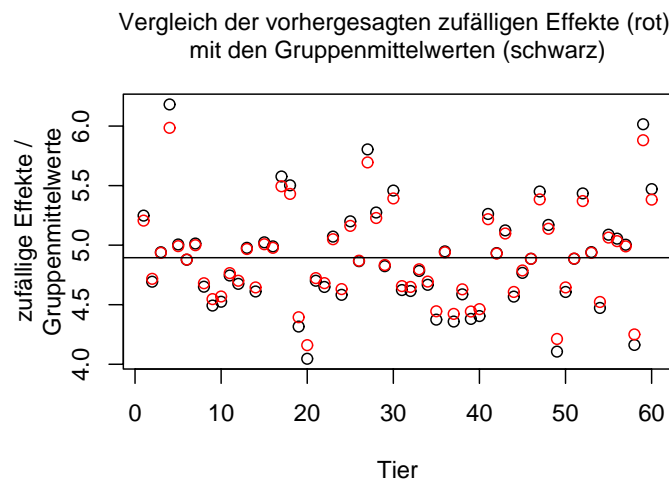
```

```

Level: Tier          lower      est.      upper
sd((Intercept)) 0.3318964 0.4083588 0.5024367
Within-group standard error:
      lower      est.      upper
0.3970768 0.4256231 0.4562216

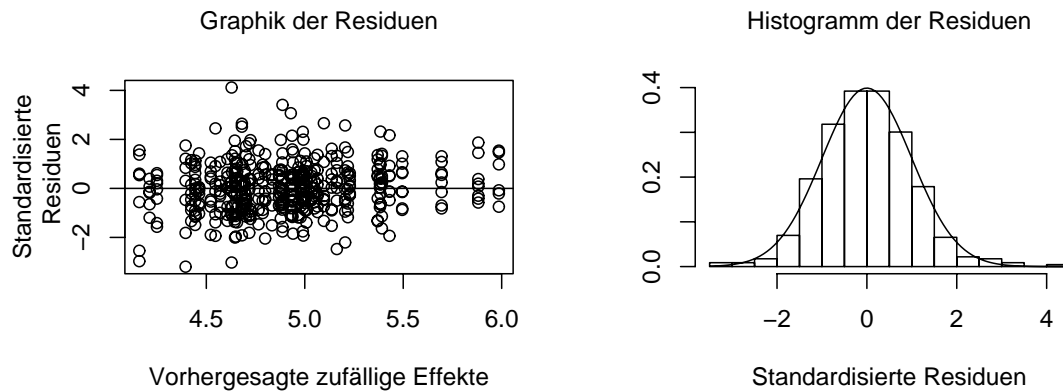
```

Durch den Befehl `ranef` können die vorhergesagten zufälligen Effekte des Modells abgerufen werden. Falls das Modell verschachtelte zufällige Effekte aufweist, muss in dem Befehl das Level der gewünschten Vorhersagen angegeben werden. Im zufälligen Intercept-Modell gibt es nur eine Stufe, welche durch die Gruppierungsvariable Tier definiert ist. Folgende Graphik zeigt die Werte der vorhergesagten zufälligen Effekte für jedes Tier, zusätzlich sind die Gruppenmittelwerte eingetragen.



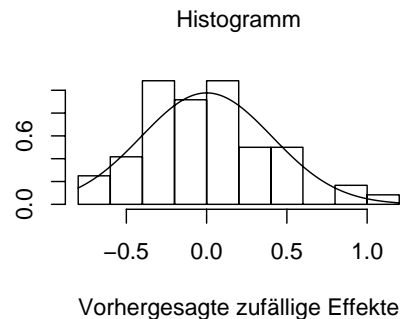
Zu erkennen ist, dass die Gruppenmittelwerte etwas weiter vom Gesamtmittelwert entfernt sind als die vorhergesagten zufälligen Effekte. Vor allem für die vom Mittelwert weit entfernten Punkte ist das der Fall. Die Vorhersagen tendieren eher zum Zentrum, da in die Berechnung das gesamte Kollektiv der Subjekte mit einfließt. Die Tiere 4 und 59 weisen die nach oben hin extremsten Werte des Kortisolgehalts auf.

Die Residuenanalyse spielt eine wichtige Rolle, um die Modellannahmen zu überprüfen, im speziellen die Verteilungsannahmen der Residuen. In folgender Graphik sind diagnostische Plots abgebildet.



Die Residuen weisen konstante Varianz über die vorhergesagten Werte auf. Die empirische Verteilung der Residuen wird durch die Normalverteilung perfekt beschrieben.

Die Verteilung der zufälligen Effekte wird ebenfalls als normalverteilt angenommen. Auch diese Annahme lässt sich in einem Histogramm überprüfen, da im Datensatz eine große Anzahl von Subjekten vorhanden ist.



Die zufälligen Effekte können als normalverteilt angenommen werden. Die Varianz der Normalverteilungsdichte in der Graphik ist der Schätzer für die Varianzkomponente  $\hat{\sigma}_I^2 = 0,167$ .

Eine weitere Möglichkeit der Modellierung eines Datensatzes dieser Art ist durch die Spezifikation der Varianzkovarianzmatrix der Residuen. Diese kann in Blöcke unterteilt sein, welche die Korrelationsstruktur der Messungen für ein Tier berücksichtigen. Für so eine Analyse ist es notwendig, eine Ordnungsstruktur der Messungen innerhalb eines Tieres anzugeben, welche als wiederholte Messungen bezeichnet werden. In dem Datensatz werden die Messungen

durch den Faktor Probezeitpunkt zeitlich geordnet. Die Zeitspannen zwischen den Messungen variieren, dadurch kann der Faktor Probezeitpunkt nicht als kontinuierliche Kovariable angesehen werden.

Die einfachste Art der Parametrisierung der Varianzkovarianzmatrix ist ein  $\sigma^2$  in der Hauptdiagonale, welches invariant ist bezüglich der Probezeitpunkte und ein Korrelationsparameter  $\rho$ , welcher die Abhängigkeit der Messungen für ein Tier misst. Die Form wird „compound symmetry“ genannt. Die folgende Matrix stellt einen Block dar,

$$\Psi = \begin{bmatrix} \sigma^2 & \rho & \cdots & \rho \\ \rho & \sigma^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & \sigma^2 \end{bmatrix}_{8 \times 8} .$$

Die Matrix kann äquivalent parametrisiert werden durch

$$\Psi = \begin{bmatrix} \sigma^2 + \sigma_I^2 & \sigma_I & \cdots & \sigma_I \\ \sigma_I & \sigma^2 + \sigma_I^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_I \\ \sigma_I & \cdots & \sigma_I & \sigma^2 + \sigma_I^2 \end{bmatrix}_{8 \times 8} .$$

Es gibt weitere Möglichkeiten der Parametrisierung der Blöcke wie z.B. ein autoregressiver Prozess erster Art. Hier werden Korrelationsparameter für die jeweils aufeinander folgenden Zeitpunkte geschätzt. Aufgrund der unterschiedlichen Zeitspannen zwischen den Messungen ist eine Modellierung dieser Art inhaltlich nicht sinnvoll.

In R können Modelle mit heterogenen Varianzstrukturen mit der Funktion `glS` angepasst werden, welche sich ebenfalls im `nlme`-Paket befindet. Hier können keine zufälligen Effekte angegeben werden. Durch die Option `correlation` wird die Form der Varianzkovarianzmatrix angegeben. Das zufällige Intercept-Modell kann auf folgende Weise gerechnet werden,

```
glS(log(Kortisolgehalt)~1,data=Kortisol,correlation=+
    corCompSymm(0.5,form=~1|Tier,fixed=F))
```

Es werden die festen Effekte und der Datensatz angegeben, die Form der Varianzkovarianzmatrix ist „compound symmetry“.

Generalized least squares fit by REML

Model: LogKortisolgehalt ~ 1

Data: Kortisol



```

Log-restricted-likelihood: -323.1352
Coefficients:
(Intercept)
  4.894722
Correlation Structure: Compound symmetry
Formula: ~1 | Tier
Parameter estimate(s):
  Rho
0.4793078
Degrees of freedom: 459 total; 458 residual
Residual standard error: 0.5898406

```

Der Wert der restringierten Log-Likelihood und die Schätzung für den Intercept sind identisch mit dem zuvor gerechneten Modell. Der Unterschied besteht in der Parametrisierung der Varianzkovarianzparameter. Die Residuenvarianz wird mit 0,348 geschätzt, im Modell zuvor kann die Größe durch  $\hat{\sigma}^2 + \hat{\sigma}_I^2$  berechnet werden. Die Korrelation von 0,479 wird im anderen Modell durch  $\hat{\sigma}_I^2 / (\hat{\sigma}_I^2 + \hat{\sigma}^2)$  berechnet.

Weiters besteht die Möglichkeit für jeden Probezeitpunkt eigene Varianz- und Kovarianzparameter zu schätzen. Für den Datensatz sind das 36 Parameter. Die Form der Varianzkovarianzmatrix ist unstrukturiert,

$$\Psi = \begin{bmatrix} \sigma_1^2 & \sigma_{21} & \cdots & \sigma_{81} \\ \sigma_{21} & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_{87} \\ \sigma_{81} & \cdots & \sigma_{87} & \sigma_8^2 \end{bmatrix}_{8 \times 8}.$$

In R kann das Modell mit folgender Syntax gerechnet werden,

```

gls(log(Kortisolgehalt)~1,data=Kortisol,correlation=+
    corSymm(form=~Probezeitpunkt|Tier,fixed=F),+
    weights = varIdent(form = ~ 1|Probezeitpunkt))

```

Als fester Effekt wird wiederum nur der Intercept angegeben. Die Korrelationsstruktur ist jetzt unstrukturiert symmetrisch, zwischen jeden Probezeitpunkten wird ein Korrelationsparameter geschätzt. Durch die Option `weights` wird die Form der Varianzen der Residuen angegeben, für jeden Probezeitpunkt wird eine eigene Varianz geschätzt. Folgender Output wird ausgegeben,

```

Coefficients:
(Intercept)

```

```

5.022023
Correlation:
  1      2      3      4      5      6      7
2 0.631
3 0.506 0.352
4 0.589 0.363 0.805
5 0.294 0.051 0.529 0.565
6 0.318 0.298 0.417 0.620 0.637
7 0.490 0.282 0.631 0.614 0.630 0.602
8 0.467 0.240 0.778 0.656 0.558 0.443 0.762
Variance function:
Parameter estimates:
  1      2      3      4      5      6      7      8
1.000 1.248 1.243 1.250 1.433 1.280 1.590 1.628
Residual standard error: 0.4481657

```

Es werden 36 Varianzkovarianzparameter geschätzt, die Kovarianzen werden in Form der Korrelationen angegeben. Die Residuenvarianz wird für den ersten Probezeitpunkt geschätzt, die restlichen können durch Skalierung mit dem Wert berechnet werden. Die Varianz der Werte des ersten Probezeitpunktes ist 0,2, die des zweiten ist  $(0,448 * 1,248)^2 = 0,313$ . Die Kovarianz des ersten und zweiten Probezeitpunktes  $\sigma_{12}$  kann durch  $0,631 * 0,448 * (0,448 * 1,248) = 0,158$  berechnet werden. Zu erkennen ist, dass die Varianzen der 7. und 8. Probezeitpunkte am höchsten sind.

Mit der Funktion `intervals` können die Konfidenzintervalle der Parameter betrachtet werden. Die Varianzparameter sind durchgehend signifikant, da ihre Konfidenzintervalle deutlich von Null entfernt sind.

Bei den Korrelationsparametern gibt es zwei nichtsignifikante Parameter. Diese betreffen die Korrelation des 5. mit dem 2. Probezeitpunkt mit einem Wert von 0,051 sowie die des 8. und des 2. Probezeitpunktes mit einem Wert von 0,240. Bereits an den Werten ist zu erkennen, dass die Korrelationen sehr niedrig sind. Das hat zu bedeuten, dass bei Tieren, welche bei Probezeitpunkt 5 z.B. einen hohen Kortisolgehalt aufweisen, keine Aussage über den Kortisolgehalt bei Probezeitpunkt 2 möglich ist. Ähnlich verhält es sich mit Probezeitpunkt 8 und 2. Andererseits ist ein hoher Wert des Kortisolgehalts bei Probezeitpunkt 3 ein hohes Indiz für einen hohen Wert bei Probezeitpunkt 4, da der Korrelationsparameter 0,805 beträgt. Insgesamt wird ersichtlich, dass die Messungen des 2. Probezeitpunktes strukturelle Abweichungen vom Gesamtschema aufweisen, da auch die restlichen Korrelationsparameter nicht all zu hoch sind. Die Tiere weisen beim 2. Probezeitpunkt andere Werte des Kortisolgehalts auf, als im Vergleich zu den restlich Probezeitpunkten.

## 7.2 Analyse in SPSS

In SPSS kann das zufällige Intercept-Modell mit folgender Syntax gerechnet werden,

```
MIXED LogKortisolgehalt BY Tier  
/FIXED= | SSTYPE(3)  
/METHOD=REML  
/PRINT=SOLUTION TESTCOV  
/RANDOM=Tier | COVTYPE(VC).
```

`LogKortisolgehalt` bezeichnet die bereits logarithmierte Variable Kortisolgehalt. Zuerst wird die Prozedur `MIXED` angegeben, dann folgt die Responsevariable. Durch das `BY` Statement werden die Faktoren bestimmt, in dem Fall nur die Gruppierungsvariable `Tier`. Durch den `/FIXED`-Befehl werden die festen Effekte spezifiziert, keine Angabe bedeutet, dass nur der Intercept-term in dem Modell inkludiert ist. Der `/PRINT`-Befehl dient der Ausgabe der Schätzungen und der Konfidenzintervalle. Im `/RANDOM`-Befehl wird die Struktur der zufälligen Effekte angegeben. Hier kann weiters die Form der Varianzkovarianzmatrix der zufälligen Effekte spezifiziert werden. `VC` bedeutet Varianzkomponenten. Es gibt eine für den `/RANDOM`-Befehl äquivalente Syntax,

```
/RANDOM=INTERCEPT | SUBJECT(Tier) COVTYPE(VC).
```

wobei die Gruppierungsvariable nach dem Strich angegeben wird. Das `/METHOD`-Statement definiert die Schätzmethode. Es sind `ML` für Maximum-Likelihood und `REML` für restringierte Maximum-Likelihood-Methode möglich.

Das Ergebnis wird in Form eines Ausgabefensters angezeigt, wobei zu Beginn Informationen über die Modelldimension und die Anzahl der im Modell geschätzten Parameter angegeben werden. Weiters wird eine Tabelle mit dem Maximalwert der logarithmierten Likelihoodfunktion und Informationskriterien wie das AIC und BIC angegeben. Zwei weniger bekannte Informationskriterien nach Hurvich und Tsai (AICC oder IC) und das Bozdogan-Kriterium (CAIC) sind ebenfalls aufgelistet. Diese werden wie folgt definiert,

$$AICC = -2l_{max} + \frac{2n_{par}N}{N - n_{par} - 1},$$

$$CAIC = -2l_{max} + n_{par}(\log(N) + 1).$$

Das AICC ist das für endliche Stichproben korrigierte AIC, das CAIC wird als konsistentes AIC bezeichnet. Beim BIC wird die Anzahl der Parameter

stärker „bestraft“ als beim AIC, dadurch werden beim BIC eher kleinere Modelle vorgezogen, beim AIC Modelle mit mehr Parametern. Die Berechnung der Kriterien für die REML-Schätzungen ist in den Statistikpaketen teilweise unterschiedlich. So wird in R die Anzahl der Parameter  $n_{par}$  durch die Anzahl der festen Parameter plus die Anzahl der Varianzkomponenten definiert, in SPSS und in SAS ist  $n_{par}$  nur die Anzahl der Varianzkomponenten. Allgemein gilt jedenfalls, je kleiner das Kriterium, desto besser ist das Modell. Folgende Tabelle zeigt die Informationskriterien für die REML-Schätzung des zufälligen Intercept-Modells.

Informationskriterien <sup>a</sup>	
Eingeschränkte -2 Log Likelihood	646,270
Akaike- Informationskriterium (AIC)	650,270
Hurvich und Tsai (IC)	650,297
Bozdogan-Kriterium (CAIC)	660,524
Bayes-Kriterium von Schwarz (BIC)	658,524

Der Wert der Likelihoodfunktion ist identisch mit dem von R berechneten Wert. In folgender Tabelle ist die Schätzung und der Test des festen Parameters angegeben,

Schätzungen fester Parameter <sup>a</sup>					
Parameter	Schätzung	Standardfehler	Freiheitsgrade	T-Statistik	Signifikanz
Konstanter Term	4,894726	,055884	59,481	87,588	,000

a. Abhängige Variable: LogResponse.

Die Schätzung des Intercept ist identisch mit dem Ergebnis in R. Die Ergebnisse für die Varianzkomponenten und Waldsche Signifikanztests sind in folgender Tabelle eingetragen,

Schätzungen von Kovarianzparametern <sup>a</sup>						
Parameter	Schätzung	Std.-Fehler	Wald Z	Sig.	Konfidenzintervall 95%	
					Untergrenze	Obergrenze
Residuum	,181155	,012834	14,116	,000	,157670	,208138
Tier Varianz	,166757	,035281	4,727	,000	,110153	,252449

a. Abhängige Variable: LogKortisolgehalt.

Die Werte entsprechen den Schätzern in R, die Konfidenzintervalle sind ebenfalls identisch. Die Testergebnisse sprechen für eine Signifikanz der Varianzkomponenten. Das ist auch an den Intervallen zu sehen, welche um einiges vom Nullpunkt entfernt sind.

Das zufällige Intercept-Modell kann in SPSS auch mit dem **repeated**-Befehl angepasst werden. Hier werden die Messungen für ein Tier als wiederholte Messungen betrachtet, wobei der Faktor Probezeitpunkt die Ordnung der Messungen angibt. Für die Varianzkovarianzmatrix können eine Vielzahl von Parametrisierungen angegeben werden, diese reichen von der einfachen „compound symmetry“ bis hin zur unstrukturierten Form. Das folgende Modell ist mit dem zufälligen Intercept-Modell identisch,

```
MIXED LogKortisolgehalt BY Probezeitpunkt Tier
/REPEATED=Probezeitpunkt | sub(Tier) covtype(cs)
```

Der Faktor Probezeitpunkt wird als Ordnungsstruktur der Messungen benötigt, die Blockstruktur der Matrix wird durch ein Tier bestimmt. Die Kovarianzstruktur ist „compound symmetry“. Folgende Schätzungen der Varianzkovarianzparameter werden ausgegeben,

**Schätzungen von Kovarianzparametern<sup>a</sup>**

Parameter		Schätzung	Std.-Fehler	Wald Z	Sig.
Meßwiederholungen	CS, diagonaler Offset	,181155	,012834	14,116	,000
	CS, Kovarianz	,166757	,035281	4,727	,000

Es werden dieselben Schätzungen ausgegeben wie bei dem Modell mit zufälligen Effekten. Der diagonale Offset entspricht der Residuenvarianz, und die Kovarianz entspricht der Varianz für den Faktor Tier.

Folgendes Modell beinhaltet eine unstrukturierte Form der Varianzkovarianzmatrix,

```
MIXED LogKortisolgehalt BY Probezeitpunkt Tier
/REPEATED=Probezeitpunkt | sub(Tier) covtype(un)
```

In dem Modell werden 37 Parameter geschätzt, davon 36 Varianzkovarianzparameter und ein Intercept. Die Ergebnisse sind mit dem in R gerechneten Modell identisch, der Wert der restringierten  $-2\text{Log-Likelihood}$  ist 511,843. Der Unterschied besteht in der Parametrisierung, in SPSS werden die Kovarianzen, in R die Korrelationen geschätzt. Zur Veranschaulichung werden die ersten Elemente der Matrix angegeben,

Schätzungen von Kovarianzparametern<sup>a</sup>

Parameter		Schätzung	Std.-Fehler	Wald Z	Sig.
Meßwiederholungen	UN (1,1)	,200851	,042034	4,778	,000
	UN (2,1)	,158277	,044441	3,562	,000
	UN (2,2)	,312911	,062395	5,015	,000
	UN (3,1)	,126405	,036880	3,427	,001
	UN (3,2)	,109827	,043369	2,532	,011
	UN (3,3)	,310374	,062300	4,982	,000

Diese sind mit den in R geschätzten Parametern identisch. Bis auf die Kovarianzen für Probezeitpunkt 2 und 5 sowie 2 und 8 sind alle Parameter signifikant.

Der Unterschied in der Parametrisierung zwischen „compound symmetry“ und „unstrukturiert“ lässt sich anhand der Modellinformationskriterien verdeutlichen. Beim ersten Modell wird ein AIC von 650 und ein BIC von 658 ausgegeben, beim zweiten, komplexeren Modell sind die Werte 583 für das AIC und 732 für das BIC. Die große Anzahl der Parameter führt zu einer deutlichen Verbesserung des AIC, aber zu einer Verschlechterung des BIC.

### 7.3 Analyse in SAS

In SAS können, ähnlich wie in SPSS, die Befehle in einem eigenen Fenster eingegeben werden. Mit den Prozeduren PROC IMPORT bzw. PROC DATA STEP können die Daten eingelesen und in die entsprechende Form gebracht werden. Die Variable für den logarithmierten Kortisolgehalt wird vor der Analyse berechnet und als LogKortisolgehalt gespeichert. Mit der PROC MIXED werden lineare gemischte Modelle angepasst. Das zufällige Intercept-Modell wird mit folgender Syntax gerechnet,

```
proc mixed data=WORK.Kortisol covtest;
  class Tier;
  model LogKortisolgehalt = /s;
  random Tier;
run;
```

Der Name des Datensatzes ist `Kortisol`, welcher im WORK-Verzeichnis gespeichert ist. Die zusätzlich angegebene Option `covtest` erzeugt Tests der Varianzkomponenten im Ausgabefenster. Für das Modell gibt es einen Faktor, `Tier`, welcher nach dem `Class`-Statement angegeben wird. Nach dem `Model`-Statement werden die festen Parameter des Modells spezifiziert, die Leerstelle bedeutet, dass nur der Interceptterm geschätzt wird. Im `Random`-

Statement wird die Gruppierungsvariable angegeben. Ähnlich wie in SPSS gibt es auch hier eine alternative Syntax,

```
random Intercept / Subject=Tier type=un;
```

Im Unterschied zu den festen Effekten muss der Interceptterm explizit angegeben werden. Die Option `type` bestimmt die Form der Varianzkovarianzmatrix der zufälligen Effekte. In dem Fall wird nur eine Varianzkomponente geschätzt, daher ist der Eintrag redundant. Als Grundeinstellung wird die restringierte Maximum-Likelihood-Schätzung (REML) angewendet. Weitere Möglichkeiten sind `ML` und `MIVQUE0`. Bei der Methode wird die quadrierte Varianz minimiert, was zu unverzerrten Schätzern für die Varianzkomponenten führt.

Die Ergebnisse der Analyse werden in einem Ausgabefenster präsentiert. Es gibt mehrere Dateiformate, die zur Verfügung stehen, als Grundeinstellung wird eine RTF-Datei erzeugt. Zuerst werden Modell- und Klassifizierungsausprägungsinformationen angegeben. Der Iterationsverlauf für die Berechnung der Schätzer hat folgende Gestalt,

#### Iterationsverlauf

Iteration	Evaluations	-2 Res Log Like	Kriterium
0	1	813.66863078	
1	2	646.27248257	0.00002101
2	1	646.27040696	0.00000001

Bereits nach zwei Iterationen ist das Konvergenzkriterium erfüllt. Die Schätzungen der Varianzkomponenten werden auf folgende Weise angegeben,

#### Covariance Parameter Estimates

Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr > Z
Intercept	Tier	0.1667	0.03527	4.73	<.0001
Residual		0.1812	0.01283	14.12	<.0001

Die Tabelle ist mit der von SPSS ausgegebenen identisch. Die Berechnung der Teststatistiken folgt auf dieselbe Weise.

Einen Unterschied zwischen SAS und den anderen Paketen gibt es bei der Berechnung der Konfidenzintervalle für die Varianzkomponenten. Als Grundeinstellung wird die Satterthwaite-Approximation verwendet, welche folgende Form hat,

$$\frac{\nu \hat{\sigma}^2}{\chi_{\nu, 1-\alpha/2}^2} \leq \sigma^2 \leq \frac{\nu \hat{\sigma}^2}{\chi_{\nu, \alpha/2}^2}$$

wobei  $\nu = 2Z^2$ , und  $Z$  ist die Waldsche Teststatistik  $\hat{\sigma}^2/se(\hat{\sigma}^2)$ . Im Nenner stehen die Quantile einer  $\chi^2$ -Verteilung mit  $\nu$  Freiheitsgraden. Folgende Intervalle werden ausgegeben,

	Untere	Schätzwert	Obere
Intercept Tier	0.1146	0.1667	0.2650
Residual	0.1584	0.1812	0.2092

Durch die Option `nobound` können die Waldschen Konfidenzintervalle ausgegeben werden. Diese werden durch die approximative Normalverteilungseigenschaft der Statistik  $\hat{\sigma}^2/se(\hat{\sigma}^2)$  konstruiert und liegen daher symmetrisch um den geschätzten Parameterwert. Die Werte für das zufällige Intercept-Modell sind

	Untere	Schätzwert	Obere
Intercept Tier	0.0976	0.1667	0.2359
Residual	0.1584	0.1812	0.2092

Das Konfidenzintervall für  $\sigma_I^2$  liegt symmetrisch um den Parameterwert.

In SAS gibt es die Möglichkeit, die Struktur der Varianzkovarianzmatrix durch das `repeated`-Statement anzugeben. Das zufällige Intercept-Modell kann mit folgender Syntax gerechnet werden, wobei die Varianzkovarianzmatrix in Form der „compound symmetry“ parametrisiert ist,

```
proc mixed data=WORK.Kortisol covtest;
  class Tier Probezeitpunkt;
  model LogKortisolgehalt = /s;
  repeated Probezeitpunkt /sub=Tier type=cs;
run;
```

Wie an den folgenden Ergebnissen zu sehen ist, handelt es sich um äquivalente Schätzungen der Parameter,

Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr Z
CS	Tier	0.1667	0.03527	4.73	<.0001
Residual		0.1812	0.01283	14.12	<.0001

Für das nächste Modell wird eine unstrukturierte Varianzkovarianzmatrix angegeben, indem die Option `type` adjustiert wird,

```
repeated Probezeitpunkt /sub=Tier type=un;
```

Die Ergebnisse stimmen mit den Schätzungen in R und SPSS überein, folgende Tabelle beinhaltet die ersten 6 der insgesamt 36 geschätzten Varianzkovarianzparameter.



Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr Z
UN(1,1)	Tier	0.2009	0.04203	4.78	<.0001
UN(2,1)	Tier	0.1583	0.04444	3.56	0.0004
UN(2,2)	Tier	0.3129	0.06239	5.02	<.0001
UN(3,1)	Tier	0.1264	0.03688	3.43	0.0006
UN(3,2)	Tier	0.1098	0.04337	2.53	0.0113
UN(3,3)	Tier	0.3104	0.06230	4.98	<.0001

## 8 Marginale Effekte

Der nächste Schritt in der Modellierung sind marginale Modelle. Sie dienen dazu, den eigenständigen Erklärungswert der unabhängigen Variablen zu untersuchen. Die Komplexität der Modelle ist nicht allzu hoch, sodass inhaltliche Interpretationen möglich sind.

Eine wichtige Fragestellung in der Analyse ist die Bedeutung des Faktors Behornung. Im Folgenden wird ein Modell gerechnet, welches den marginalen Erklärungswert der Variable quantifiziert. Der Faktor Tier wird, wie bereits im zufälligen Intercept-Modell, als Gruppierungsvariable verwendet und mit zufälligen Effekten modelliert. Die Responsevariable Kortisolgehalt wird einer Log-Transformation unterzogen. Das Modell sieht folgendermaßen aus,

$$\log(Kortisolgehalt)_{ij} = \beta_0 + \beta_1 Behornung_j + b_i + \epsilon_{ij}.$$

Die Laufvariable  $i = 1, \dots, 60$  nummeriert die Tiere für die behornte und die unbehornte Gruppe. Falls das  $j$ -te Tier behornt ist, wird  $Behornung_j = 1$  gesetzt, andernfalls ist  $Behornung_j = 0$ . Die zufälligen Effekte werden mit  $b_i$  bezeichnet. Theoretisch ist es möglich, zwei getrennte Varianzkomponenten zu schätzen, welche jeweils auf den 30 Ausprägungen der behornten und unbehornten Tiere basieren. In dem Modell wird davon ausgegangen, dass die Streuung der Tiere für die behornte und unbehornte Gruppe identisch ist.

$$b_i \sim N(0, \sigma_I^2), \quad \epsilon_{ij} \sim N(0, \sigma^2).$$

$\beta_0$  stellt den Intercept für die unbehornten Tiere dar,  $\beta_1$  die Differenz zwischen der unbehornten und der behornten Gruppe. Der Unterschied zum zufälligen Intercept-Modell liegt darin, dass hier zwei Interceptterme geschätzt werden, jeweils für die Gruppe der behornten und unbehornten Tiere. Die Laufvariable  $j = 1, \dots, n_i$  zählt die  $j$ -te Beobachtung innerhalb der  $i$ -ten Gruppe.

## 8.1 Analyse in R

Das Modell wird in R mit folgendem Befehl gerechnet,

```
> lme(log(Kortisolgehalt)~factor(Behornung),data=Kortisol,+  
+ random=~1|Tier)
```

Der Faktor Behornung wird als fester Effekt angegeben. Folgende Schätzungen werden ausgegeben,

```
Fixed: log(Kortisolgehalt) ~ factor(Behornung)  
      (Intercept) factor(Behornung)1  
      4.8382672      0.1128778  
Random effects:  
Formula: ~1 | Tier  
      (Intercept) Residual  
StdDev:  0.4083249 0.4256261
```

Der geschätzte Intercept für die unbehornete Gruppe ist  $\hat{\beta}_1 = 4,838$ , die behornete Gruppe hat einen Wert von  $\hat{\beta}_1 + \hat{\beta}_2 = 4,951$ . Die Varianzkomponenten sind  $\hat{\sigma}_I^2 = 0,1667$  und  $\hat{\sigma}^2 = 0,181$ . Die Werte sind mit den Schätzungen des zufälligen Intercept-Modells nahezu identisch. Daran ist zu erkennen, dass der Faktor Behornung keine zusätzliche Erklärung der Streuung mit sich bringt. Mit dem `anova`-Befehl kann die Hypothese getestet werden. Folgender Output wird ausgegeben,

	numDF	denDF	F-value	p-value
(Intercept)	1	399	7543.772	<.0001
factor(Behornung)	1	58	1.003	0.3208

Der P-Wert ist weit von der Signifikanzgrenze entfernt, was für ein beibehalten der Nullhypothese spricht. Die Freiheitsgrade im Nenner werden wie folgt berechnet.  $Q$  ist die Anzahl der Gruppierungsstufen, in dem Fall 1. Der Interceptterm wird auf der innersten Stufe getestet, die Formel für die Freiheitsgrade ist somit  $m_2 - (m_1 + p_2) = 459 - (60 + 0)$ . Der Faktor Behornung wird auf dem ersten Level getestet, da er außerhalb der Gruppierungsvariable liegt,  $m_1 = 60$ ,  $m_0 = 1$ , und  $p_1 = 1$  ist die Summe der Freiheitsgrade der Terme, welche auf diesem Level getestet werden. Daraus ergibt sich eine Anzahl von 58 Freiheitsgraden im Nenner für den Faktor Behornung.

Die Variable Dominanzindex ist invariant innerhalb eines Tieres. Dadurch wird sie außerhalb des Gruppierungsfaktors getestet. Das marginale Modell hat folgende Form,

$$\log(Kortisolgehalt)_{ij} = \beta_1 + \beta_2 Dominanzindex_i + b_i + \epsilon_{ij}.$$

Die Variable *Dominanzindex<sub>i</sub>* ist der Dominanzindex des *i*-ten Tieres, die restlichen Parameter und die Indizierung sind identisch mit dem vorhergehenden Modell. Der R Befehl zum Ausführen des Modells ist folgender,

```
> lme(log(Kortisolgehalt)~Dominanzindex,data=Kortisol,+
+ random=~1|Tier)
```

Das Modell untersucht den eigenständigen Erklärungswert der Variable Dominanzindex. Folgende Parameterschätzungen werden von R berechnet,

```
(Intercept) Dominanzindex
  4.7304861      0.4025991
Random effects:
Formula: ~1 | Tier
      (Intercept)  Residual
StdDev:   0.3918591  0.4256751
```

Der Interceptterm ist ähnlich zu interpretieren wie bei einer einfachen Regressionsanalyse. Es ist der Vertikalabstand der Regressionslinie beim Wert Dominanzindex gleich Null. Der Parameter für den Dominanzindex ist die Steigung der Geraden, bei einer Veränderung des Dominanzindex um 1 verändert sich der logarithmierte Kortisolgehalt um 0,4 Einheiten. Die Signifikanz der Kovariablen ist in folgender Tabelle eingetragen,

	numDF	denDF	F-value	p-value
Dominanzindex	1	58	5.291	0.0251

Der Dominanzindex hat auf dem 5%-Niveau einen knapp signifikanten Einfluss auf die Responsevariable.

Die Schätzung der Varianzkomponente  $\hat{\sigma}_I^2 = 0,154$ , der Wert ist etwas niedriger als beim zufälligen Intercept-Modell. Die Schätzung für die Residuenvarianz  $\sigma^2$  ist hingegen nahezu identisch geblieben. Das ist darauf zurückzuführen, dass der Dominanzindex außerhalb der Gruppierungsvariable liegt, der Term kann nur zur Verkleinerung der Varianz zwischen den Tieren führen. Die Varianz innerhalb der Tiere wird dadurch nicht beeinflusst.

Bis jetzt wird die Variable Dominanzindex als kontinuierliche Variable aufgefasst. Durch das Diskretisieren ist es möglich, den nichtlinearen Einfluss der Variable zu untersuchen. Dazu wird die Variable in drei Gruppen aufgeteilt, welche mit 1,2,3 kodiert werden. Die erste Gruppe entspricht den 20 Tieren, welche in der behornten und unbehornten Gruppe die jeweils zehn

rangniedrigsten Werte des Dominanzindex aufweisen. Die zweite Gruppe umfasst die 20 rangmittleren und die dritte Gruppe die ranghöchsten Tiere. Das marginale Modell hat folgende Form,

$$\log(Kortisolgehalt)_{ij} = \beta_1 + \beta_2 Dominanzindex_2 + \beta_3 Dominanzindex_3 + b_i + \epsilon_{ij}.$$

Es werden insgesamt drei feste Parameter geschätzt, *Dominanzindex<sub>2</sub>* und *Dominanzindex<sub>3</sub>* stellen Indikatorfunktionen für die jeweiligen Gruppen dar, die erste Gruppe bildet die Referenzkategorie. In R wird das Modell auf folgende Weise gerechnet,

```
> lme(log(Kortisolgehalt)~factor(DominanzindexDiskret),+
  data=Kortisol,random=~1|Tier)
```

*DominanzindexDiskret* stellt die diskrete Variable Dominanzindex dar. Folgender R-Output wird ausgegeben,

```
Linear mixed-effects model fit by REML
  Data: Kortisol
Log-restricted-likelihood: -323.2392
Fixed: log(Kortisolgehalt) ~ factor(DominanzindexDiskret)
      (Intercept) factor(DominanzindexDiskret)2
              4.7663904                      0.1029524
factor(DominanzindexDiskret)3
              0.2821139
Random effects:
Formula: ~1 | Tier
      (Intercept) Residual
StdDev:   0.3985529 0.4256716
```

Die erste Gruppe besitzt die niedrigsten Werte des logarithmierten Kortisolgehalts, mit einem Mittelwert von 4,77 , bei den rangmittleren Tieren ist der Mittelwert um 0,103 , bei den ranghöchsten Tieren um 0,28 größer. Durch die *anova*-Funktion kann die Signifikanz des Faktors berechnet werden,

	numDF	denDF	F-value	p-value
(Intercept)	1	399	7869.254	<.0001
factor(DominanzindexDiskret)	2	57	2.231	0.1167

Diese ist bei der diskreten Variable Dominanzindex nicht gegeben.

Im Folgenden wird der Einfluss des Probezeitpunktes analysiert. Der Faktor befindet sich innerhalb der Gruppierungsvariable, das Modell hat dadurch

eine etwas andere Gestalt. Es gibt acht Ausprägungen, welche durch die Laufvariable  $k = 1, \dots, 8$  nummeriert werden. Jede Beobachtung der Responsevariable kann durch das Tier und den Probezeitpunkt eindeutig indiziert werden, es gibt für jede Faktorkombination eine Messung. Zusätzlich kommt es vor, dass für ein Tier nicht alle 8 Messungen der Probezeitpunkte vorhanden sind.

$$\log(Kortisolgehalt)_{ik} = \beta_k \text{Probezeitpunkt}_k + b_i + \epsilon_{ik}.$$

Die Regressormatrix der festen Effekte wird in die Form einer Treatmentkodierung für den Faktor Probezeitpunkt gebracht. Die erste Ausprägung stellt die Referenzkategorie dar. Die Fehlerterme  $\epsilon_{ik}$  messen den Unterschied zwischen den vorhergesagten und den tatsächlichen Werten für den logarithmierten Kortisolgehalt. Das Modell wird in R auf ähnliche Weise wie zuvor gerechnet. Folgender Output wird ausgegeben,

```
Fixed: log(Kortisolgehalt) ~ factor(Probezeitpunkt)
      (Intercept) Probezeitpunkt2 Probezeitpunkt3
      5.13807613    0.04012447    -0.29771881
Probezeitpunkt4 Probezeitpunkt5 Probezeitpunkt6
      -0.23628207    -0.34932642    -0.35577806
Probezeitpunkt7 Probezeitpunkt8
      -0.38789677    -0.35402644
Random effects:
Formula: ~1 | Tier
      (Intercept) Residual
StdDev:    0.4122480 0.3940707
```

Der logarithmierte Kortisolgehalt hat für den ersten Probezeitpunkt einen Mittelwert von 5,138, der Probezeitpunkt zwei liegt etwas höher. Die restlichen Probezeitpunkte weisen einen geringeren Wert der Responsevariable auf. Die geschätzten Varianzkomponenten sind  $\hat{\sigma}_T^2 = 0,1699$  und  $\hat{\sigma}^2 = 0,155$ . Im Unterschied zum zufälligen Intercept-Modell ist ein starker Rückgang der Residuenvarianz zu bemerken, die Varianz zwischen den Tieren ist konstant geblieben. Der Faktor Probezeitpunkt liegt innerhalb der Gruppierungsvariable, daher verringert er die Streuung der Residuen. Mit der `anova`-Funktion kann die Signifikanz des Faktors verifiziert werden.

Weiters ist im Datensatz der Faktor Zeitpunkt angegeben, welcher die jeweils aufeinander folgenden Probezeitpunkte zusammenfasst. Dadurch gibt es 4 Ausprägungen, welche durch die Laufvariable  $k = 1, \dots, 4$  nummeriert werden. Das marginale Modell für die Zeitpunkte hat die Form,

$$\log(Kortisolgehalt)_{ijk} = \beta_k \text{Zeitpunkt}_k + b_i + \epsilon_{ijk}.$$

Hier wird eine neue Indexvariable  $j = 1, 2$  verwendet, um die Beobachtung innerhalb des Tieres und des Zeitpunktes zu bezeichnen. Die Ausgabe in R hat folgende Gestalt,

```
Fixed: log(Kortisolgehalt) ~ factor(Zeitpunkt)
(Intercept)  Zeitpunkt8  Zeitpunkt9  Zeitpunkt10
  5.1579476   -0.2871910   -0.3724618   -0.3908116
Random effects:
Formula: ~1 | Tier
(Intercept)  Residual
StdDev:      0.4124127 0.3926711
```

Der Zeitpunkt 7 hat einen Mittelwert von 5,158, die restlichen drei Zeitpunkte liegen etwas niedriger. In folgender Tabelle sind die Mittelwerte der Responsevariable für die vier Zeitpunkte eingetragen. Die Transformation der Werte in die ursprüngliche Lage erfolgt wie im zufälligen Intercept-Modell, für den Zeitpunkt 7 ist das etwa  $\exp(\hat{\beta}_1 + (\hat{\sigma}_I^2 + \hat{\sigma}^2)/2)$ .

	log(Kortisolgehalt)	Kortisolgehalt
Zeitpunkt 7	5,16	204,4
Zeitpunkt 8	4,87	153,38
Zeitpunkt 9	4,79	140,84
Zeitpunkt 10	4,78	138,28

Die Schätzungen der Varianzkomponenten sind mit dem vorigen Modell nahezu identisch. Somit kann die Streuung der Residuen durch das Zeitpunkt-Modell genauso gut erklärt werden wie durch das Probezeitpunkt-Modell, welches jedoch um vier feste Parameter mehr beinhaltet.

Die zwei Modelle lassen sich durch Modellinformationskriterien miteinander vergleichen. Dazu müssen die Modelle mit der Maximum-Likelihood-Methode angepasst werden. Die REML-Methode lässt keine Modellvergleiche zu, welche in der Spezifikation der festen Effekte variieren. In R kann die Methode durch das zusätzliche Argument `method="ML"` geändert werden.

Für das Probezeitpunkt-Modell sind die ML-Schätzungen der festen Parameter nahezu identisch, die Varianzkomponenten weisen kleinere Differenzen zu den REML-Schätzungen auf. Die Werte sind  $\hat{\sigma}_{I,ML}^2 = 0,167$  und  $\hat{\sigma}_{ML}^2 = 0,1526$ . Diese sind etwas kleiner als die REML-Schätzungen, da die Anzahl der festen Parameter  $k$  im Nenner nicht berücksichtigt wird.

Das gleiche Resultat liefert das mit der ML-Methode geschätzte Zeitpunkt-Modell.

Nun lassen sich die Modelle durch die Modellinformationskriterien vergleichen, welche in R mit der `summary`-Funktion abgerufen werden können. Folgende Tabelle zeigt deren Werte,

Modell	Probezeitpunkt	Zeitpunkt
AIC	593,75	586,96
BIC	635,05	611,74

Beide Informationskriterien sprechen für das Modell mit dem Faktor Zeitpunkt. Daher wird in den weiteren Modellen mit diesem Faktor gearbeitet.

## 8.2 Analyse in SPSS

Für das marginale Modell des Faktors Behornung werden folgende Adjustierungen im Syntaxbefehl vorgenommen,

```
MIXED LogKortisolgehalt BY Tier Behornung
/FIXED=Behornung | SSTYPE(3)
```

Jetzt wird der Faktor Behornung in das Modell hinein genommen und dessen Signifikanz geprüft.

Schätzungen fester Parameter <sup>b</sup>					
Parameter	Schätzung	Standardfehler	Freiheitsgrade	T-Statistik	Signifikanz
Konstanter Term	4,951145	,079687	57,454	62,132	,000
[Behornung=0]	-,112878	,112710	57,486	-1,001	,321
[Behornung=1]	0 <sup>a</sup>	0	.	.	.

Die Schätzungen der festen Effekte sind identisch, die behornete Gruppe wird als Referenzkategorie verwendet.

Die Anzahl der Freiheitsgrade für die Signifikanztests wird in SPSS auf eine andere Weise berechnet. Es wird die Satterthwaite-Approximation (1964) zur Bestimmung der Freiheitsgrade im Nenner vorgenommen. Angenommen, es soll eine Hypothese über einen oder mehrere Parameter des Vektors  $\beta$  der festen Effekte getestet werden. Diese kann in folgender Form angeschrieben werden,

$$H_0 : R\beta = 0,$$

wobei  $R$  eine Matrix mit schätzbaren Linearkombinationen des Vektors  $\beta$  darstellt. Die Matrix  $R$  ist von der Dimension  $p \times k$  und  $\text{rank}(R) = p$ . Sei  $C = VC(\hat{\beta})$ , welche auf Seite 35 definiert ist, und  $\hat{C}$  die geschätzte Varianzkovarianzmatrix. Es handelt sich dabei um symmetrische Matrizen mit Dimension  $k \times k$ . Die verwendeten Schätzungen für  $\hat{C}$  sind  $\hat{\theta}$  und  $\hat{\sigma}^2$ . Mit folgender F-Statistik kann die Hypothese geprüft werden,

$$F = \frac{\hat{\beta}^T R^T (R \hat{C} R^T)^{-1} R \hat{\beta}}{p},$$

die Anzahl der Freiheitsgrade im Zähler ist  $p$ . Um die Anzahl der Freiheitsgrade im Nenner zu berechnen, wird die Matrix  $R\hat{C}R^T$  in Form der Eigenwerte und Eigenvektoren dargestellt,

$$R\hat{C}R^T = P^T D P.$$

$P$  ist die orthogonale Matrix der Eigenvektoren und  $D$  ist die Diagonalmatrix der Eigenwerte. Beide Matrizen sind von der Dimension  $p \times p$ . Sei weiters  $l_m$  die  $m$ -te Reihe der Matrix  $PR$ , welche von der Dimension  $p \times k$  ist. Sei nun

$$v_m = \frac{2(D_m)^2}{g_m^T A g_m},$$

wobei  $(D_m)$  das  $m$ -te Diagonalelement von  $D$  bezeichnet,  $g_m$  ist der Gradient der Funktion  $f(\theta, \sigma^2) = l_m C l_m^T$  ausgewertet im Punkt  $[\hat{\theta}^T, \hat{\sigma}^2]^T$ , und  $A$  stellt die asymptotische Varianzkovarianzmatrix von  $\hat{\theta}$  bzw.  $\hat{\sigma}^2$  dar, die Inverse der beobachteten Informationsmatrix. Weiters ist

$$E = \sum_{m=1}^p \frac{v_m}{v_m - 2} I(v_m > 2),$$

wobei durch die Indikatorfunktion die Terme  $v_m \leq 2$  ausgeschlossen werden. Die Freiheitsgrade  $v$  im Nenner werden dann berechnet durch

$$v = \frac{2E}{E - p},$$

unter der Bedingung  $E > p$ . Falls das nicht zutrifft werden die Freiheitsgrade im Nenner gleich null gesetzt.

Für den Faktor Behornung führt die Berechnung der Freiheitsgrade im Nenner durch die Satterthwaite-Approximation in SPSS (57,468) zu keiner substanziellen Veränderung des P-Wertes gegenüber der in R verwendeten Methode.

Die Schätzungen der Varianzkomponenten sind in folgender Tabelle gelistet, die Ergebnisse sind mit R identisch.

Schätzungen von Kovarianzparametern<sup>a</sup>

Parameter	Schätzung	Std.-Fehler	Wald Z	Sig.	Konfidenzintervall 95%	
					Untergrenze	Obergrenze
Residuum	,181158	,012834	14,116	,000	,157672	,208142
Tier Varianz	,166729	,035584	4,685	,000	,109735	,253326

Die Parameterschätzer des marginalen Modells für den Dominanzindex ergeben dieselben Resultate.



**Schätzungen fester Parameter<sup>a</sup>**

Parameter	Schätzung	Standardfehler	Freiheitsgrade	T-Statistik	Signifikanz
Konstanter Term	4,730311	,089808	57,323	52,672	,000
Dominanzindex	,402573	,174999	57,357	2,300	,025

**Schätzungen von Kovarianzparametern<sup>a</sup>**

Parameter	Schätzung	Std.-Fehler	Wald Z	Sig.	Konfidenzintervall 95%	
					Untergrenze	Obergrenze
Residuum	,181199	,012840	14,112	,000	,157703	,208196
Tier Varianz	,153552	,033182	4,628	,000	,100533	,234530

Für das Modell mit diskreter Variable Dominanzindex werden folgende Schätzungen für die festen Effekte ausgegeben,

**Schätzungen fester Parameter<sup>b</sup>**

Parameter	Schätzung	Standardfehler	Freiheitsgrade	T-Statistik	Signifikanz
Konstanter Term	5,048504	,095578	56,334	52,821	,000
[Dominanzindex Diskret=1]	-,282114	,135160	56,323	-2,087	,041
[Dominanzindex Diskret=2]	-,179161	,135160	56,323	-1,326	,190

Die Referenzkategorie ist im Gegensatz zu R die dritte Gruppe. Für das Modell gilt, dass in SPSS und R identische Werte ausgegeben werden.

Für das Modell mit dem Faktor Probezeitpunkt kommen in SPSS und R dieselben Schätzungen heraus, der Wert der Likelihood ist, wie auch in den zuvor gerechneten Modellen, ebenfalls identisch. Analoges gilt für das Modell mit dem Faktor Zeitpunkt.

Für die ML-Schätzungen der Modelle sind die Ergebnisse in R und SPSS identisch. Zusätzlich werden zwei weitere Modellinformationskriterien von SPSS ausgegeben. Der Vergleich zwischen dem Probezeitpunkt-Modell und dem Zeitpunkt-Modell kann auch nach den zwei Kriterien vorgenommen werden.

Modell	Probezeitpunkt	Zeitpunkt
Hurvich/Tsai	594,25	587,15
Bozdogan	645,05	617,74

Das Modell mit dem Faktor Zeitpunkt erweist sich auch nach den Kriterien als das Effektivere.

### 8.3 Analyse in SAS

Die Syntaxbefehle zum Anpassen der marginalen Modelle in SAS sind für das `class`- und das `model`-Statement zu verändern. Es wird jeweils der zu untersuchende Faktor eingetragen.

```
class Tier Behornung;
model LogKortisolgehalt = Behornung /s;
```

Die Schätzungen für die Varianzkomponenten des Modells für den Faktor Behornung ergeben folgende Werte,

Covariance Parameter Estimates					
Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr > Z
Intercept	Tier	0.1667	0.03557	4.69	<.0001
Residual		0.1812	0.01283	14.12	<.0001

Die Schätzungen sind identisch mit R und SPSS. Ein Unterschied zwischen den Programmen ist das Hypothesentesten für die festen Effekte. In SAS gibt es fünf Methoden für die Berechnung der Freiheitsgrade im Nenner.

Die Grundeinstellung ist die Methode **containment**. Die Vorgangsweise zur Bestimmung der Freiheitsgrade im Nenner ist folgende. Zuerst wird der zu testende Faktor bestimmt. Nun werden die zufälligen Effekte des Modells herangezogen, welche syntaktisch innerhalb des Faktors liegen. Die zufälligen Effekte sind innerhalb des Faktors verschachtelt. Im Beispieldatensatz **Kortisol** liegt der zufällige Effekt der Tiere innerhalb des Faktors Behornung, die Tiere sind innerhalb des Faktors Behornung verschachtelt. Dann wird der Rang des Teiles der Matrix  $Z$  bestimmt, welcher die Dummykodierung für den zufälligen Effekt beinhaltet. Die Anzahl der Freiheitsgrade im Nenner ist der kleinste Rang der in Betracht gezogenen zufälligen Effekte. Falls sich kein zufälliger Effekt innerhalb des Faktors befindet, werden die Freiheitsgrade gleich  $N - rank(XZ)$  gesetzt. Grundsätzlich ist die Methode der in R verwendeten sehr ähnlich. Einen Unterschied gibt es bei der Handhabung kontinuierlicher Variablen. In SAS ist es nicht möglich, diese in Form einer verschachtelten Struktur in der Syntax anzugeben, daher werden die Freiheitsgrade im Nenner gleich  $N - rank(XZ)$  gesetzt. Weiters ist zu bemerken, dass in SAS die verschachtelte Struktur der zufälligen Effekte im **random**-Statement anzugeben ist, da sie nicht automatisch erkannt wird.

Für das obige Modell wird für den festen Effekt Behornung folgende Anzahl der Freiheitsgrade im Nenner ausgegeben.

Type 3 Tests of Fixed Effects				
	Num	Den		
Effekt	DF	DF	F Value	Pr > F
Behornung	1	399	1.00	0.3172

Die Anzahl der Freiheitsgrade im Nenner ist gleich  $N - rank(XZ) = 459 - 60 = 399$ . Die Matrix  $X$  besteht aus 2, die Matrix  $Z$  aus 58 linear unabhängigen Spalten.

Falls im `random`-Befehl explizit die verschachtelte Struktur `Tier(Behor-`  
`nung)` angegeben wird, werden die Freiheitsgrade anders berechnet,

#### Type 3 Tests of Fixed Effects

	Num	Den		
Effect	DF	DF	F Value	Pr > F
Behornung	1	58	1.00	0.3207

Hier wird von SAS erkannt, dass sich der zufällige Effekt `Tier` innerhalb des Faktors `Behornung` befindet, es wird der Rang der  $Z$ -Matrix als Anzahl der Freiheitsgrade herangezogen. Da jeweils ein Tier für die behornnte und unbehornnte Gruppe die Referenzkategorie darstellen, ist die Anzahl  $60 - 2 = 58$ .

Eine weitere, in SAS zur Verfügung gestellte Methode zur Berechnung der Freiheitsgrade im Nenner ist `residual`. Hier wird die Anzahl gleich  $N - \text{rank}(X)$  gesetzt. Der Nachteil der Methode ist, dass die Abhängigkeiten der wiederholten Messungen nicht in die Anzahl der Freiheitsgrade mit einfließen, sie ist daher nicht empfehlenswert.

Weiters wird die Option `betwithin` für die Berechnung angeboten. Diese ist die Grundeinstellung für Modelle mit dem `repeated`-Statement. Die durch die Methode `residual` berechnete Anzahl der Freiheitsgrade wird in zwei Teile aufgespalten. Ein Teil entspricht den Freiheitsgraden, welche innerhalb des zufälligen Effektes liegen, der andere Teil entspricht den Freiheitsgraden, welche außerhalb des Effektes liegen. Je nachdem, ob sich der zu testende Faktor innerhalb oder außerhalb des zufälligen Effektes befindet, werden die Freiheitsgrade bestimmt.

Die Methode nach Satterthwaite (1946) ist wie auch in SPSS optional möglich. Die Berechnung ist auf Seite 70 beschrieben. Die Ergebnisse der Berechnungen für die Modelle des Datensatzes stimmen in SPSS und SAS überein. Die Methode führt zu guten Ergebnissen, für kleine Stichprobenumfänge ist sie weniger geeignet. Die in R verwendete Methode ist der Satterthwaite-Methode sehr ähnlich.

Ein Nachteil der Satterthwaite-Methode ist, dass für die Berechnung eine Schätzung der Varianzkovarianzmatrix der festen Effekte  $VC(\hat{\beta})$  benötigt wird. Das Einsetzen der geschätzten Parameter  $\hat{\theta}$  bzw.  $\hat{\sigma}^2$  führt jedoch zu einer verzerrten Schätzung, da die Parameter als bekannt voraus gesetzt werden. In Littell et al. (2006), Seite 188, wird dazu folgendes bemerkt,

*„Kackar and Harville (1984) showed that using estimated covariance parameters in this way results in test statistics that are biased upward and standard errors that are biased downward,...“*

Kenward und Roger (1997) entwickelten Korrekturen für die Standardfehler und die Teststatistiken. Diese können durch die Option `kenwardroger`

in SAS berechnet werden. In der Methode nach Satterthwaite wird die korrigierte Varianzkovarianzmatrix für die Berechnung der Anzahl der Nennerfreiheitsgrade herangezogen. Die Adjustierung wird vor allem bei kleinen Stichprobenumfängen sowie unbalancierten Datensätzen empfohlen (Littell et al. (2006)).

Für den Beispieldatensatz `Kortisol` ist in den gerechneten Modellen kein Unterschied zwischen den Methoden nach Satterthwaite und Kenward-Roger zu bemerken. Das ist auf die große Stichprobenanzahl und das nahezu balancierte Design zurückzuführen. Daher wird in den folgenden in SAS angepassten Modellen die Satterthwaite-Methode zur Berechnung der Nennerfreiheitsgrade verwendet.

Für das Modell mit der Kovariable Dominanzindex ist das Ergebnis für die festen Parameter folgendes,

Lösung für feste Effekte					
Effekt	Schätzwert	Std.fehler	DF	t-Wert	Pr >  t
Intercept	4.7305	0.08975	57.3	52.71	<.0001
Dominanzindex	0.4026	0.1750	57.4	2.30	0.0251

Die Schätzwerte stimmen mit den Ergebnissen aus R und SPSS überein, die Schätzungen für die Varianzkomponenten sind in allen drei Paketen identisch.

Für das Modell mit diskretem Dominanzindex werden in SAS folgende Schätzungen für die festen Effekte ausgegeben,

Solution for Fixed Effects						
Effect		Estimate	Std.Error	DF	t-Value	Pr> t
Intercept		5.0485	0.09558	56.3	52.82	<.0001
DomIndDiskret	1	-0.2821	0.1352	56.3	-2.09	0.0414
DomIndDiskret	2	-0.1792	0.1352	56.3	-1.33	0.1903

Der P-Wert für den Faktor Dominanzindex ist 0,1168. Die Schätzungen der Parameter sind mit R und SPSS identisch.

Ähnliches gilt für die Modelle mit den Faktoren Zeitpunkt und Probezeitpunkt. Die Ergebnisse für die REML- und ML-Schätzungen sind in SAS, R und SPSS identisch.

## 9 Verschachtelte zufällige Effekte

Der Faktor Zeitpunkt liegt innerhalb der Gruppierungsvariable, die Ausprägungen ändern sich innerhalb eines Tieres. Für 43 Tiere gibt es zwei Messungen für alle vier Zeitpunkte, bei 13 Tieren fehlt eine Messung. Bei 4 Tieren fehlen zwei Messungen, welche genau einen Zeitpunkt betreffen. In folgender Tabelle sind diese aufgelistet.

Tier	4	24	45	60
Zeitpunkt	10	8	8	7

Für die restlichen Tiere kann ein zusätzlicher zufälliger Effekt für den Faktor Zeitpunkt berechnet werden. Der Effekt ist innerhalb der Tiere verschachtelt. In einem Modell dieser Art gibt es keine festen Effekte für die einzelnen Zeitpunkte, es wird eine Varianzkomponente  $\sigma_Z^2$  für den Faktor Zeitpunkt geschätzt. Die Hypothese, ob der Faktor Zeitpunkt einen signifikanten Einfluss auf die Responsevariable hat, kann getestet werden. Das ist der Fall wenn  $\sigma_Z^2 > 0$ . Das Modell lässt keine Vergleiche zwischen speziellen Zeitpunkten zu, da sie als zufällige Effekte modelliert werden. Für den Faktor Probezeitpunkt ist kein verschachteltes Modell möglich, da jede Beobachtung durch das Tier und den Probezeitpunkt eindeutig bestimmt ist, und somit keine Varianzkomponente geschätzt werden kann.

Das Modell wird wie folgt definiert,

$$\log(Kortisolgehalt)_{ijk} = \beta + b_i + b_{ij} + \epsilon_{ijk}.$$

$$b_i \sim N(0, \sigma_I^2) \quad b_{ij} \sim N(0, \sigma_Z^2) \quad \epsilon_{ijk} \sim N(0, \sigma^2)$$

Die Laufvariable  $i = 1, \dots, 60$  nummeriert die Tiere, die Laufvariable  $j = 1, \dots, 4$  bezeichnet den  $j$ -ten Zeitpunkt und die Variable  $k = 1, 2$  bezeichnet die Messung innerhalb des Zeitpunktes bzw. Tieres. Für vier Tiere gibt es Zeitpunkte, wo keine Messung zur Verfügung steht, hier gibt es keinen zufälligen Effekt für den Zeitpunkt und kein Residuum.  $b_i$  sind die zufälligen Effekte auf der Stufe der Tiere,  $b_{ij}$  die zufälligen Effekte auf der Stufe der Zeitpunkte innerhalb der Tiere. Die Varianzkomponenten sind  $\sigma_I^2$  und  $\sigma_Z^2$ .

## 9.1 Analyse in R

In R lässt sich ein Modell dieser Art mit dem Syntaxbefehl

```
lme(log(Kortisolgehalt)~1,data=Kortisol,+
    +random=~1|Tier/Zeitpunkt)
```

rechnen. Der einzige feste Effekt ist der Interceptterm. Im **random**-Befehl wird die Struktur des verschachtelten Modells mit der Syntax **Tier/Zeitpunkt** eingegeben.

```
Fixed: log(Kortisolgehalt) ~ 1
(Intercept)
      4.895996
Random effects:
```

```

Formula: ~1 | Tier
          (Intercept)
StdDev:   0.3893783
Formula: ~1 | Zeitpunkt %in% Tier
          (Intercept) Residual
StdDev:   0.3000138 0.3253174

```

In dem Modell gibt es  $(60 * 4 - 4) = 236$  Ausprägungen für den Faktor Zeitpunkt innerhalb des Tieres. Für den Interceptterm wird nahezu die gleiche Schätzung ausgegeben wie beim zufälligen Intercept-Modell. Die Schätzungen der Varianzkomponenten und die 95%-Konfidenzintervalle sind in folgender Tabelle zusammengefasst,

	untere Schranke	Schätzung	obere Schranke
$\sigma_I^2$	0,096	0,152	0,239
$\sigma_Z^2$	0,063	0,09	0,129
$\sigma^2$	0,088	0,106	0,128

Der Großteil der Streuung wird durch den Faktor Tier erklärt. Durch den Faktor Zeitpunkt wird in etwa soviel Varianz erklärt wie durch die Residuen. Die Summe der drei Komponenten ist 0,347. Beim zufälligen Intercept-Modell sind die Varianzkomponenten  $\hat{\sigma}_I^2 = 0,167$  und  $\hat{\sigma}^2 = 0,181$ , was in Summe ebenfalls 0,348 ergibt. Daran ist zu erkennen, dass die einzelnen Varianzkomponenten auf verschiedenen Stufen einen additiven Beitrag zur Gesamtstreuung beitragen.

Die Signifikanz der Varianzkomponente  $\hat{\sigma}_Z^2$  kann durch einen Likelihood-Quotienten-Test verifiziert werden. Das Modell mit der zusätzlichen Varianzkomponente ist das Alternativmodell, das Nullmodell entspricht dem Modell mit zufälligem Intercept. Da die Spezifikation der festen Effekte in beiden Modellen die gleiche ist, kann die REML-Schätzung für den Test herangezogen werden. Die Log-Likelihood für das Nullmodell ist  $-323,135$  bei 3 Freiheitsgraden, die des Alternativmodells  $-300,998$  bei 4 Freiheitsgraden. Bereits hier ist zu erkennen, dass die Log-Likelihood des Alternativmodells einen weit größeren Wert hat. Die Teststatistik  $Q = 44,274$ , sie folgt einer  $\chi^2$ -Verteilung mit einem Freiheitsgrad. Die Wahrscheinlichkeit eine solche oder noch größere Teststatistik zu erhalten, ist kleiner 0,0001. Somit ist die Varianzkomponente hoch signifikant, der Faktor Zeitpunkt eine wichtige Einflussgröße. In R kann der Test mit folgendem Statement gerechnet werden,

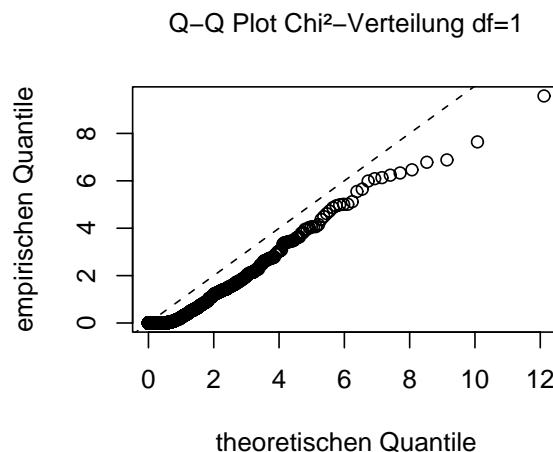
```
anova(lme.alt,lme.null)
```

wobei `lme.alt` das Alternativmodell und `lme.null` das Nullmodell darstellt.

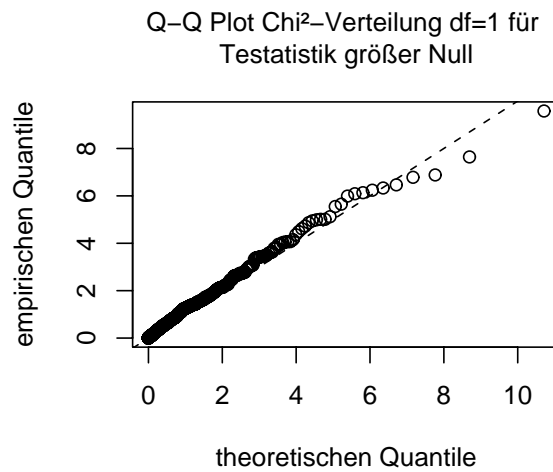
Der Test basiert auf dem theoretischen Resultat, dass die Likelihood-Quotienten-Statistik einer  $\chi^2$ -Verteilung folgt. Da die Varianzkomponenten am Rand des Parameterraumes getestet werden, ist die tatsächliche Verteilung der Statistik eine andere. Die Verteilung kann mit einer Simulation untersucht und mit der  $\chi^2$ -Verteilung verglichen werden. Mit der R-Funktion `simulate` können simulierte Responsevektoren erzeugt werden, welche nach dem Nullmodell verteilt sind. Den Responsevektoren werden die zwei Modelle angepasst, und die Likelihood-Quotienten-Statistik errechnet. Daraus resultiert eine empirische Verteilung der Likelihood-Quotienten-Statistik. Mit folgendem Befehl werden 1000 Zufallsvariablen der Verteilung erzeugt,

```
simulate.lme(lme.null,m2=lme.alt,nsim=1000)
```

Die simulierte und die  $\chi^2$ -Verteilung können durch einen Quantile-Quantile-Plot verglichen werden,



Es ist zu erkennen, dass die empirischen Quantile kleiner sind als die theoretischen. Dadurch sind die theoretischen P-Werte zu groß und die Entscheidungen zu konservativ. Es werden signifikant von Null verschiedene Varianzen nicht als solche ausgegeben. Weiters ist zu bemerken, dass etwa die Hälfte der simulierten Teststatistiken nahezu Null sind. Die restlichen folgen annähernd einer  $\chi^2$ -Verteilung wie in folgender Graphik zu sehen ist.



Die tatsächliche Verteilung der Likelihood-Quotienten-Teststatistik kann als eine Mischverteilung aus einer zu 50% bestehenden Masse auf Null und einer  $\chi^2$ -Verteilung mit einem Freiheitsgrad betrachtet werden.

## 9.2 Analyse in SPSS

In SPSS wird für die Analyse eines verschachtelten Modells der `random-`Befehl abgeändert. Es gibt drei äquivalente Möglichkeiten die Struktur der zufälligen Effekte anzugeben,

```
/RANDOM=Intercept Zeitpunkt | SUBJECT(Tier) COVTYPE(VC).
/RANDOM=Tier Tier*Zeitpunkt | COVTYPE(VC).
/RANDOM=Tier Tier(Zeitpunkt) | COVTYPE(VC).
```

Der feste Effekt besteht jeweils nur aus dem Interceptterm. Folgende Schätzungen der Varianzkomponenten werden ausgegeben,

**Schätzungen von Kovarianzparametern<sup>a</sup>**

Parameter		Schätzung	Std.-Fehler	Wald Z	Sig.
Residuum		,105831	,010113	10,465	,000
Tier	Varianz	,151616	,035221	4,305	,000
Zeitpunkt * Tier	Varianz	,090008	,016682	5,396	,000

Die Waldschen Tests ergeben klare Signifikanz für die drei Parameter. Im Folgenden sind die Konfidenzintervalle aufgelistet,



Parameter		Konfidenzintervall 95%	
		Untergrenze	Obergrenze
Residuum		,087756	,127631
Tier	Varianz	,096163	,239046
Zeitpunkt * Tier	Varianz	,062593	,129431

Die Werte entsprechen, bis auf Rundungsfehler, den Werten aus R.

### 9.3 Analyse in SAS

Für die Anpassung des verschachtelten Modells in SAS ist das `random`-Statement zu adjustieren. Wie in SPSS gibt es mehrere äquivalente Möglichkeiten der Syntax,

```
random Tier Tier*Zeitpunkt;
random intercept Zeitpunkt / subject=Tier;
```

Folgende Schätzungen werden für das Modell ausgegeben,

Covariance Parameter Estimates					
Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr > Z
Intercept	Tier	0.1516	0.03522	4.30	<.0001
Zeitpunkt	Tier	0.0900	0.01668	5.40	<.0001
Residual		0.1058	0.01011	10.46	<.0001

Die Werte sind mit SPSS und R identisch.

## 10 Zufällige Wechselwirkungen

Wie bereits im Theorieteil vorgestellt, gibt es bei gemischten Modellen die Möglichkeit, Wechselwirkungen als zufällige Größen zu modellieren. Insbesondere sind diese zwischen dem Faktor Tier und dem Faktor Zeitpunkt von Interesse. Werden Wechselwirkungen als feste Effekte geschätzt, gibt es in den meisten Fällen eine beträchtliche Anzahl an zusätzlichen Parametern, das Schätzen von zufälligen Wechselwirkungen benötigt nur eine Varianzkomponente. Die Effekte können nicht als feste Größen modelliert werden, da der Faktor Tier als zufälliger Faktor geschätzt wird. Aussagen über die speziellen Wechselwirkungseffekte können keine gemacht werden.

Im Gegensatz zum Modell aus dem vorigen Abschnitt werden feste Effekte für den Faktor Zeitpunkt mit geschätzt. Das Modell mit zufälligen Wechselwirkungen hat folgende Gestalt,

$$\log(Kortisolgehalt)_{ijk} = \beta_j \text{Zeitpunkt}_j + b_i + b_{ij} + \epsilon_{ijk}.$$

$$b_i \sim N(0, \sigma_I^2) \quad b_{ij} \sim N(0, \sigma_W^2) \quad \epsilon_{ijk} \sim N(0, \sigma^2)$$

Die Laufvariable  $j = 1, \dots, 4$  nummeriert die Anzahl der festen Effekte. Diese sind durch die Treatmentkodierung definiert,  $\beta_1$  steht für den Mittelwert des 7-ten Zeitpunktes,  $\beta_2$  für den Unterschied zu Zeitpunkt 8 usw. Durch  $i = 1, \dots, 60$  werden die Tiere indiziert, und  $k = 1, 2$  nummeriert die Beobachtung innerhalb eines Tieres und eines Zeitpunktes. In 13 Fällen gibt es nur ein  $k=1$ , da nur eine Messung für den Zeitpunkt vorhanden ist. Es gibt 236 Ausprägungen der Zufallsvariable  $b_{ij}$ , welche die Wechselwirkung zwischen dem festen Effekt für den Zeitpunkt  $j$  und dem zufälligen Effekt für das  $i$ -te Tier messen. Die Varianzkomponente wird mit  $\sigma_W^2$  bezeichnet. Mit den Hypothesen  $\sigma_W^2 = 0$  und  $\sigma_W^2 > 0$  kann das Vorhandensein von Wechselwirkungen getestet werden.

## 10.1 Analyse in R

Das Modell mit zufälligen Wechselwirkungen für den Faktor Zeitpunkt wird in R mit folgendem Befehl angepasst,

```
lme(log(Kortisolgehalt)~factor(Zeitpunkt),
+data=Kortisol,random=~1|Tier/Zeitpunkt)
```

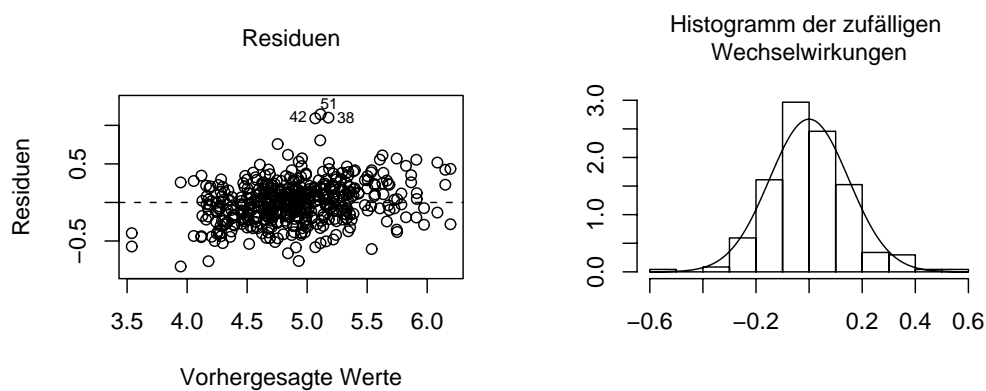
Der einzige Unterschied zum verschachtelten Modell ist die Hereinnahme des Zeitpunktes als festen Effekt.

```
Fixed: log(Kortisolgehalt) ~ factor(Zeitpunkt)
(Intercept)  Zeitpunkt8  Zeitpunkt9  Zeitpunkt10
   5.1565704  -0.2860011  -0.3629341  -0.3920700
Random effects:
Formula: ~1 | Tier
          (Intercept)
StdDev:    0.4003866
Formula: ~1 | Zeitpunkt %in% Tier
          (Intercept)  Residual
StdDev:    0.2387518  0.3272982
```

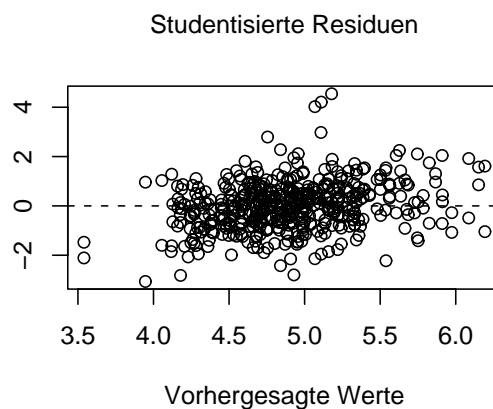
In der Ausgabe ist zu erkennen, dass vier feste Effekte und drei Varianzkomponenten geschätzt werden. Die Varianzkomponente für den Faktor Tier und die Residuenvarianz ist gegenüber dem verschachtelten Modell nahezu unverändert geblieben. Die Schätzung für die Varianzkomponente der Wechselwirkungen ist  $\hat{\sigma}_W^2 = 0,057$ . Das Konfidenzintervall für den Parameter ist  $[0,036; 0,091]$ . Das Intervall ist deutlich vom Nullpunkt entfernt, die Varianzkomponente scheint eine wichtige Einflussgröße zu sein. Mit der

`anova`-Funktion kann die Hypothese verifiziert werden. Das Nullmodell beinhaltet nur die Varianzkomponente für den Faktor Tier. Der P-Wert der Likelihood-Quotienten-Statistik ist kleiner 0,0001.

In der folgenden Graphik werden die Modellannahmen überprüft. Die Residuen links, zeigen keine Unregelmäßigkeiten entlang der vorhergesagten Werte. Für drei Beobachtungen werden sehr hohe Residuen angezeigt, welche die Tiere 38, 42 und 51 betreffen. Die Zeitpunkte der Messungen sind 9, 10 und 9. Auf der rechten Seite ist die Verteilung der 236 zufälligen Wechselwirkungen abgebildet, die Annahme der Normalverteilung ist gerechtfertigt.

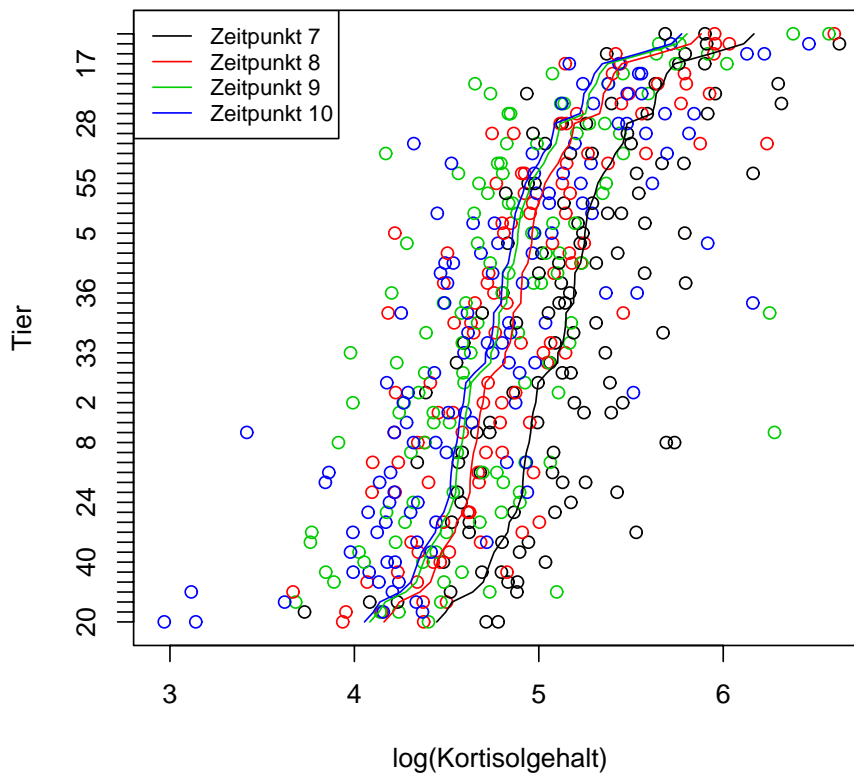


In folgender Graphik sind die studentisierten Residuen abgebildet. Diese berücksichtigen den unterschiedlichen Standardfehler der Vorhersage und sind dadurch besser untereinander vergleichbar. In R und SPSS wird eine Berechnung der Größen nicht automatisch mitgeliefert. In SAS können diese durch das zusätzliche Argument `residuals` im `model`-Statement ausgegeben werden.



In der nächsten Graphik sind die Beobachtungen für jedes Tier und die festen Effekte des Faktors Zeitpunkt eingetragen. Die Tiere sind nach dem zufälligen Intercept aufsteigend sortiert. Die Linien für die festen Effekte verlaufen parallel, der horizontale Abstand ist der Unterschied in den Zeitpunkten. Die Wechselwirkungen sind daran zu erkennen, dass sich der Wert für den festen Effekt für ein Tier teilweise nicht direkt bei den Beobachtungen befindet. Für das Tier Nummer 20 etwa ist der feste Effekt für die Zeitpunkte 10 und 7 weit von den tatsächlichen Beobachtungen entfernt. Der Abstand wird durch die zufälligen Wechselwirkungen modelliert.

Unterschiede im logarithmierten Kortisolgehalt für die vier Zeitpunkte  
Tiere sind nach zufälligem Intercept aufsteigend sortiert



Die Graphik wird mit folgendem R-Code generiert,

```
## R-Code für Graphik
## lme.1 ist gespeichertes Modell
> reihenfolge=as.numeric(rownames(ranef(lme.1,level=1)))
      [order(ranef(lme.1,level=1)[,1])]
> plot(1,1,"n",xlim=c(2.95,6.63),ylim=c(1,60),xlab="log
```

```

(Kortisolgehalt)",ylab="Tier",main="Unterschiede im
logarithmierten Kortisolgehalt für die vier Zeitpunkte
Tiere sind nach zufälligem Intercept aufsteigend sortiert",
font.main=1,cex.main=1,axes=F)
> axis(1)
> axis(2,1:60,labels=reihenfolge,cex=0.8)
> box()
> legend("topleft",c("Zeitpunkt 7","Zeitpunkt 8","Zeitpunkt
9","Zeitpunkt 10"),lty=1,col=1:4,cex=0.8)
> for (i in 1:60){
  > LR=Kortisol[Kortisol$Tier==reihenfolge[i],]$LogKortiso
    lgehalt
  > ZP=Kortisol[Kortisol$Tier==reihenfolge[i],]$Zeitpunkt
  > l=length(LR)
  > points(LR, rep(i,l),col=ZP-6)
> }
> L= sort(ranef(lme.1,level=1)[,1])+lme.1$coef$fixed[1]
> lines(L,1:60,col=1)
> L= sort(ranef(lme.1,level=1)[,1])+lme.1$coef$fixed[1]
  +lme.1$coef$fixed[2]
> lines(L,1:60,col=2)
> L= sort(ranef(lme.1,level=1)[,1])+lme.1$coef$fixed[1]
  +lme.1$coef$fixed[3]
> lines(L,1:60,col=3)
> L= sort(ranef(lme.1,level=1)[,1])+lme.1$coef$fixed[1]
  +lme.1$coef$fixed[4]
> lines(L,1:60,col=4)

```

Anhand der Modellinformationskriterien ist zu erkennen, dass das Modell dem verschachtelten und dem Modell ohne zufällige Wechselwirkungen vorzuziehen ist. Ein Vergleich ist bei den mit ML-Schätzung angepassten Modellen möglich. In folgender Tabelle sind die drei Modelle gegenübergestellt,

Modell	AIC	BIC
verschachtelt	606,1	622,6
ohne zufällige Wechselwirkung	587,0	611,7
mit zufälliger Wechselwirkung	567,5	596,4

## 10.2 Analyse in SPSS

In SPSS wird das Modell mit folgendem Befehl gerechnet,

```

/FIXED=Zeitpunkt | SSTYPE(3)
/RANDOM=Intercept Zeitpunkt | SUBJECT(Tier) COVTYPE(VC).

```

Für den random-Befehl sind, wie vorhin, drei verschiedene Syntaxvarianten möglich. Die Option SSTYPE(3) bedeutet, dass die Tests der festen Effekte mit der Quadratsummenzerlegung des dritten Typs durchgeführt werden. Der Vorteil des Testverfahrens ist, dass nur der eigenständige Beitrag des Effektes ausschlaggebend ist, die restlichen Effekte sind heraus gerechnet.

**Schätzungen von Kovarianzparametern<sup>a</sup>**

Parameter	Schätzung	Std.-Fehler	Wald Z	Sig.
Residuum	,107124	,010327	10,374	,000
Tier Varianz	,160309	,035164	4,559	,000
Zeitpunkt * Tier Varianz	,057002	,013756	4,144	,000

Die Schätzungen für die Varianzkomponenten sind identisch mit R. Dasselbe gilt für die Schätzungen der festen Effekte, in SPSS ist die Referenzkategorie Zeitpunkt 10.

**Schätzungen fester Parameter<sup>b</sup>**

Parameter	Schätzung	Standardfehler	Freiheitsgrade	T-Statistik	Signifikanz
Konstanter Term	4,764500	,067652	114,970	70,427	,000
[Zeitpunkt=7]	,392070	,061825	161,685	6,342	,000
[Zeitpunkt=8]	,106069	,062366	163,540	1,701	,091
[Zeitpunkt=9]	,029136	,061489	161,462	,474	,636
[Zeitpunkt=10]	0 <sup>a</sup>	0	.	.	.

## 10.3 Analyse in SAS

Folgende Statements werden in SAS eingegeben, um das Modell mit zufälligen Wechselwirkungen für den Faktor Zeitpunkt zu rechnen,

```

class Tier Zeitpunkt;
model LogKortisolgehalt=Zeitpunkt /s DDFM=sat;
random Tier Zeitpunkt*Tier;

```

Die Ausgabe der Schätzungen für die Varianzkomponenten ist folgende,

### Covariance Parameter Estimates

Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr > Z
Intercept	Tier	0.1603	0.03516	4.56	<.0001
Zeitpunkt	Tier	0.0570	0.01376	4.14	<.0001
Residual		0.1071	0.01033	10.37	<.0001

Die Schätzungen sind in den Softwarepaketen identisch. Für den Test des Faktors Zeitpunkt werden in SAS durch die Satterthwaite-Methode 163 Nennerfreiheitsgrade berechnet, in SPSS sind das 162,869 und in R werden 173 Freiheitsgrade verwendet. Die Unterschiede in der Teststatistik sind minimal, das Testergebnis ist eine hohe Signifikanz für den Faktor Zeitpunkt.

## 11 Modellselektion

Im folgenden Kapitel werden Modelle angepasst, welche aus mehreren unabhängigen Variablen bestehen. Ziel der Modelle ist es, die Responsevariable so gut als möglich zu erklären und Aussagen über den Einfluss der Behornung zu treffen. Die Modelle beinhalten die Haupteffekte, wie die Faktoren Behornung und Zeitpunkt sowie die Kovariable Dominanzindex. Zusätzlich wird das Vorhandensein von festen Wechselwirkungstermen geprüft. Der Faktor Tier und Wechselwirkungen für den Faktor Zeitpunkt werden wie vorhin durch zufällige Effekte modelliert. Die Responsevariable Kortisolgehalt wird einer Log-Transformation unterzogen, um die Verteilungseigenschaften der Modelle zu gewährleisten.

### 11.1 Haupteffektemodell

Das folgende Modell umfasst die Hauptfaktoren ohne Wechselwirkungen, der gemeinsame Erklärungswert der Variablen wird untersucht. Der Faktor Tier wird als zufälliger Effekt modelliert.

$$\log(Kortisolgehalt)_{ijk} = \beta_0 + \beta_1 Behornung_i + \beta_{2,j} Zeitpunkt_j + \beta_3 Dominanzindex_i + b_i + \epsilon_{ijk}$$

$$b_i \sim N(0, \sigma_I^2) \quad \epsilon_{ijk} \sim N(0, \sigma^2)$$

Für die Modellformulierung werden drei Laufvariablen benötigt, um die richtige Kombination der erklärenden Variablen für die jeweilige Ausprägung der Responsevariable zu bestimmen.  $\beta_0$  stellt den Interceptterm dar,  $\beta_1$  misst den Unterschied zwischen den behornten und unbehornten Tieren.  $Behornung_i = 0$ , falls das  $i$ -te Tier nicht behornt ist, und 1 falls es behornt ist. Für den Faktor Zeitpunkt gibt es drei Spalten,  $j = 1, 2, 3$ , der erste Zeitpunkt stellt die Referenzkategorie dar. Der Dominanzindex wird durch das  $i$ -te Tier eindeutig bestimmt, da es eine kontinuierliche Variable ist, wird ein Parameter benötigt. Für das Residuum gibt es zwei Möglichkeiten  $k = 1, 2$ , da es zwei Ausprägungen für jeden Zeitpunkt gibt.

## 11.2 Analyse in R

In R wird das Modell mit folgendem Befehl ausgeführt,

```
lme(log(Kortisolgehalt)~factor(Zeitpunkt)+Dominanzindex+  
+factor(Behornung),Kortisol,random=~1|Tier)
```

Der Faktor Behornung wird als letzter Effekt in der Modellgleichung angegeben, da beim Testen der festen Effekte in R ein hierarchisches Verfahren angewendet wird. So kann der eigenständige Erklärungswert der Behornung getestet werden. Die Teststatistik ist äquivalent mit der Quadratsummenzerlegung des dritten Typs in SPSS und SAS. Die `anova`-Funktion liefert folgende Testergebnisse,

	numDF	denDF	F-value	p-value
(Intercept)	1	396	8270.286	<.0001
factor(Zeitpunkt)	3	396	24.277	<.0001
Dominanzindex	1	57	5.569	0.0217
factor(Behornung)	1	57	1.937	0.1694

Der Faktor Behornung wird als nicht signifikante Größe ausgegeben, der P-Wert hat sich im Vergleich zum marginalen Modell um einiges in Richtung Signifikanzgrenze verändert. Die eigenständige Bedeutung der restlichen zwei Einflussgrößen kann durch das Verändern der Modellgleichung im R-Befehl überprüft werden. Für den Dominanzindex wird ein P-Wert von 0,013 ausgegeben und der Faktor Zeitpunkt ist hoch signifikant.

Die Schätzungen der Varianzkomponenten sind nahezu identisch mit  $\hat{\sigma}_I^2 = 0,153$  und  $\hat{\sigma}^2 = 0,154$ . Gegenüber den zuvor gerechneten Modellen sind die Werte nur geringfügig kleiner geworden.

Im marginalen Modell für den Dominanzindex hat sich herausgestellt, dass die diskrete Variable Dominanzindex keine signifikante Einflussgröße ist. In folgendem Modell werden alle Haupteffekte und die Variable Dominanzindex als diskrete Variable mit drei Gruppen entsprechend den rangniedrigsten, rangmittleren und ranghöchsten Tieren inkludiert. Das wird mit folgendem R-Befehl ausgeführt,

```
lme(log(Kortisolgehalt)~factor(Behornung)+factor(Zeitpunkt)+  
factor(DominanzindexDiskret),Kortisol,random=~1|Tier)
```

Folgende Schätzungen werden ausgegeben,

```
Linear mixed-effects model fit by REML  
Intercept Behornung1 Zeitpunkt8 Zeitpunkt9 Zeitpunkt10  
4.9774620 0.1042525 -0.2864791 -0.3724200 -0.3906299
```



```

DominanzindexDiskret2 DominanzindexDiskret3
              0.1005612              0.2838538
Random effects:
Formula: ~1 | Tier
          (Intercept) Residual
StdDev:    0.4028718 0.3927127

```

Die festen Effekte für den Dominanzindex sind in etwa in der Größenordnung wie im marginalen Modell. Durch die `anova`-Funktion wird ein P-Wert von 0,1129 für die diskrete Variable Dominanzindex ausgegeben. Weiters ist zu bemerken, dass die Signifikanz der Behornung mit 0,3488 um einiges schwächer ist als im zuvor gerechneten Modell. Bei den Modellinformationskriterien ist eine starke Präferenz für das Modell mit kontinuierlichem Dominanzindex zu beobachten.

Im vorigen Abschnitt hat sich herausgestellt, dass zufällige Wechselwirkungen für den Faktor Zeitpunkt zu einer Verbesserung der Modellanpassung führen, diese werden in folgendem Modell inkludiert, die Variable Dominanzindex geht als kontinuierliche Variable in das Modell ein,

```

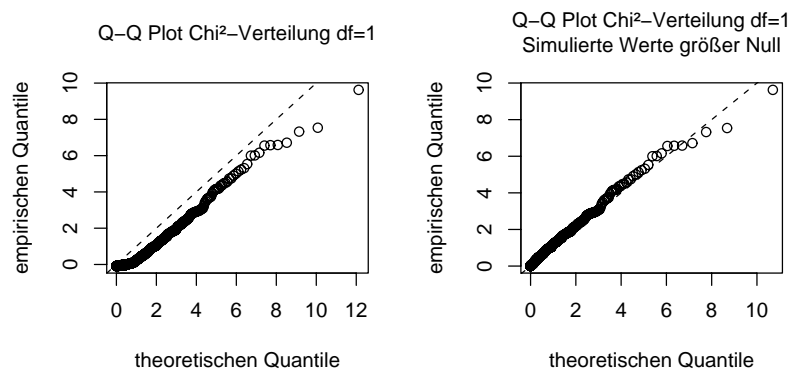
lme(log(Kortisolgehalt)~factor(Zeitpunkt)+Dominanzindex+
    +factor(Behornung),Kortisol,random=~1|Tier/Zeitpunkt)

```

Ein Drittel der Residuenvarianz kann durch die zufälligen Wechselwirkungen erklärt werden, die Schätzungen sind  $\hat{\sigma}^2 = 0,107$  und  $\hat{\sigma}_W^2 = 0,057$ . Mit der `anova`-Funktion wird das Modell ohne, und das Modell mit inkludierter zufälliger Wechselwirkung verglichen,

	Model	df	AIC	BIC	logLik	L.Ratio	p-value
	<code>lme.null</code>	1	8	604.8098	637.7369	-294.4049	
	<code>lme.alt</code>	2	9	584.5595	621.6025	-283.2797	22.25027 <.0001

An den Modellinformationskriterien ist zu erkennen, dass das Alternativmodell dem Nullmodell vorzuziehen ist. Der Likelihood-Quotienten-Test zeigt ein hoch signifikantes Ergebnis an. Mit einer Simulation basierend auf 1000 Durchläufen kann, wie im verschachtelten Modell, die Verteilung der Teststatistik verifiziert werden. Auf der linken Seite ist ein Q-Q-Plot der simulierten Verteilung abgebildet, etwa die Hälfte der Werte sind nahezu Null. In der rechten Graphik sind die Werte nicht inkludiert, die Anpassung mit der  $\chi^2$ -Verteilung ist perfekt. Die Verteilung der Teststatistik ist, wie im verschachtelten Modell, eine Mischverteilung aus einer Punktmasse auf Null und einer  $\chi^2$ -Verteilung mit einem Freiheitsgrad.



### 11.3 Analyse in SPSS

In SPSS wird das Haupteffektemodell mit folgendem Befehl erzeugt,

```
MIXED LogKortisolgehalt BY Zeitpunkt Tier Behornung
      With Dominanzindex
/FIXED=Zeitpunkt Behornung Dominanzindex | SSTYPE(3)
/RANDOM Intercept | sub(Tier) COVTYPE(vc).
```

Der Wert der Likelihoodfunktion an der Stelle der Parameterschätzungen ist 588,8 und somit identisch mit dem Wert aus R. Auch die Parameterschätzungen der festen und zufälligen Effekte gleichen den in R berechneten Werten. An den Type III Tests der festen Effekte ist zu erkennen, dass der Faktor Behornung keine signifikante Einflussgröße ist.

Tests auf feste Effekte, Typ III<sup>a</sup>

Quelle	Zähler-Freiheitsgrade	Nenner-Freiheitsgrade	F-Wert	Signifikanz
Konstanter Term	1	56,359	2759,224	,000
Behornung	1	56,356	1,937	,169
Dominanzindex	1	56,385	6,571	,013
Zeitpunkt	3	396,359	24,264	,000

Für das Haupteffektemodell mit diskretem Dominanzindex werden folgende Tests der festen Effekte ausgegeben,

Tests auf feste Effekte, Typ III<sup>a</sup>

Quelle	Zähler-Freiheitsgrade	Nenner-Freiheitsgrade	F-Wert	Signifikanz
Konstanter Term	1	55,390	7874,308	,000
DominanzindexDiskret	2	55,386	2,269	,113
Behornung	1	55,397	,893	,349
Zeitpunkt	3	396,352	24,251	,000

Der Wert der Log-Likelihoodfunktion mit 593,621 stimmt für das Modell mit dem in R berechneten Wert überein. Somit handelt es sich um identische Schätzungen für die Parameter.

Für die Modellierung mit zufälligen Wechselwirkungen muss der `Random`-Befehl adaptiert werden,

```
/RANDOM Intercept Zeitpunkt | sub(Tier) COVTYPE(vc).
```

## 11.4 Analyse in SAS

Ähnlich wie in SPSS ist die Syntax für das Haupteffektmodell zu adjustieren,

```
class Tier Zeitpunkt Behornung;
model LogKortisolgehalt=Zeitpunkt Dominanzindex Behornung /s
      DDFM=satterthwaite;
random Intercept / sub=Tier;
```

Folgende, in R und SPSS identische Schätzungen, werden für die Varianzkomponenten ausgegeben,

Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr > Z
Intercept	Tier	0.1535	0.03278	4.68	<.0001
Residual		0.1542	0.01097	14.06	<.0001

Die Tests der festen Effekte ergeben einen P-Wert von 0,1694 für den Faktor Behornung. Die Berechnung der Freiheitsgrade über die Satterthwaite-Methode ist identisch zu SPSS, und damit stimmen die P-Werte überein. In R werden 57 Freiheitsgrade im Nenner verwendet.

Effekt	Zähler Freiheitsgrade	Nenner Freiheitsgrade	F-Statistik	Pr > F
Zeitpunkt	3	396	24.26	<.0001
Behornung	1	56.4	1.94	0.1694
Dominanzindex	1	56.4	6.57	0.0131

Für die Modellierung mit der diskreten Variable Dominanzindex werden folgende Tests für die festen Effekte ausgegeben,

	Num	Den		
Effect	DF	DF	F Value	Pr > F
DomIndDiskret	2	55.4	2.27	0.1130
Zeitpunkt	3	396	24.25	<.0001
Behornung	1	55.4	0.89	0.3488

Die Methode zur Berechnung der Freiheitsgrade im Nenner ist **satterthwaite**. Der diskrete Dominanzindex hat keine signifikante Wirkung. Die Parameterschätzungen für das Modell sind in allen drei Paketen identisch.

Das Haupteffektemodell mit zufälligen Wechselwirkungen für den Faktor Zeitpunkt wird mit folgendem Befehl erzeugt,

```
random Intercept Zeitpunkt/ sub=Tier;
```

Die Konfidenzintervalle werden in SAS über die Satterthwaite-Approximation (siehe S.62) berechnet. Hier eine Gegenüberstellung der ausgegebenen Intervalle für die drei Varianzkomponenten,

Kov.Parm	Subjekt	Untere	Schätzwert	Obere
Intercept	Tier	0.09673	0.1444	0.2387
Zeitpunkt	Tier	0.03718	0.0568	0.0974
Residual		0.08957	0.1073	0.1309

Die folgenden Konfidenzintervalle werden von SPSS und R ausgegeben, die Intervallgrenzen liegen niedriger,

	Untere	Schätzwert	Obere
Intercept	0,0924	0,1444	0,2256
Zeitpunkt	0,0353	0,0568	0,0913
Residuen	0,0888	0,1073	0,1296

## 11.5 Modell mit Wechselwirkungstermen

Abweichungen der Linearität in den Haupteffekten können mit Wechselwirkungen modelliert werden. Falls der Zeitpunkteffekt für die behornete Gruppe eine andere Wirkung hat als für die Unbehornte, sind Wechselwirkungen vorhanden. Genauso können Wechselwirkungen mit der Variable Dominanzindex interpretiert werden. Falls der Effekt der Behornung sich für hohe Werte des Dominanzindex anders verhält als für niedrige Werte, sind auch hier Wechselwirkungen vorhanden. Das Modell mit drei zusätzlichen Wechselwirkungstermen hat 13 feste Parameter und zwei Varianzkomponenten, die dreifachen Wechselwirkungen sind nicht inkludiert.

## 11.6 Analyse in R

In R wird das Modell auf folgende Weise gerechnet,

```
lme(log(Kortisolgehalt)~factor(Zeitpunkt)*Dominanzindex*  
  factor(Behornung)-factor(Zeitpunkt):factor(Behornung):  
  Dominanzindex,data=Kortisol,random=~1|Tier)
```

Die Testergebnisse für die festen Effekte sind folgende,

	numDF	denDF	F-value	p-value
(Intercept)	1	390	8351.280	<.0001
factor(Zeitpunkt)	3	390	24.886	<.0001
Dominanzindex	1	56	5.623	0.0212
factor(Behornung)	1	56	1.956	0.1674
factor(Zeitpunkt):Dominanzindex	3	390	5.021	0.0020
factor(Zeitpunkt):factor(Behornung)	3	390	0.611	0.6080
Dominanzindex:factor(Behornung)	1	56	0.600	0.4420

Da in R ein hierarchisches Testverfahren durchgeführt wird, ist der Faktor Behornung zuletzt aufgelistet. Der P-Wert hat sich gegenüber dem Haupteffektmodell kaum verändert. Die einzige signifikante Wechselwirkung ist zwischen dem Zeitpunkt und dem Dominanzindex vorhanden.

Für das Modell mit zufälligen Wechselwirkungen für den Faktor Zeitpunkt, diese betreffen die Gruppierungsvariable Tier, werden ähnliche Resultate ausgegeben. Die Modellanpassung ist etwas besser, das AIC ist 598,3 gegenüber einem Wert von 616,2 für das vorhergehende Modell.

Aufgrund der Testergebnisse empfiehlt es sich, die Wechselwirkungen für den Faktor Behornung nicht zu inkludieren, da sie keine signifikante Größe sind. Das folgende Modell umfasst daher nur die Wechselwirkung Zeitpunkt-Dominanzindex, die zufälligen Wechselwirkungen für den Faktor Zeitpunkt werden aufgrund der besseren Anpassung inkludiert.

```
lme(log(Kortisolgehalt)~factor(Zeitpunkt)*Dominanzindex+  
  factor(Behornung),data=Kortisol,random=~1|Tier/Zeitpunkt)
```

Für die Testergebnisse werden folgende Werte ausgegeben,

	numDF	denDF	F-value	p-value
(Intercept)	1	223	8424.206	<.0001
factor(Zeitpunkt)	3	170	17.580	<.0001
Dominanzindex	1	57	5.527	0.0222
factor(Behornung)	1	57	1.886	0.1750
factor(Zeitpunkt):Dominanzindex	3	170	3.802	0.0113

Der P-Wert des Faktors Behornung ist gleich geblieben, alle restlichen Parameter haben einen signifikanten Einfluss. Die Parameterschätzungen sind in folgender Ausgabe ersichtlich,

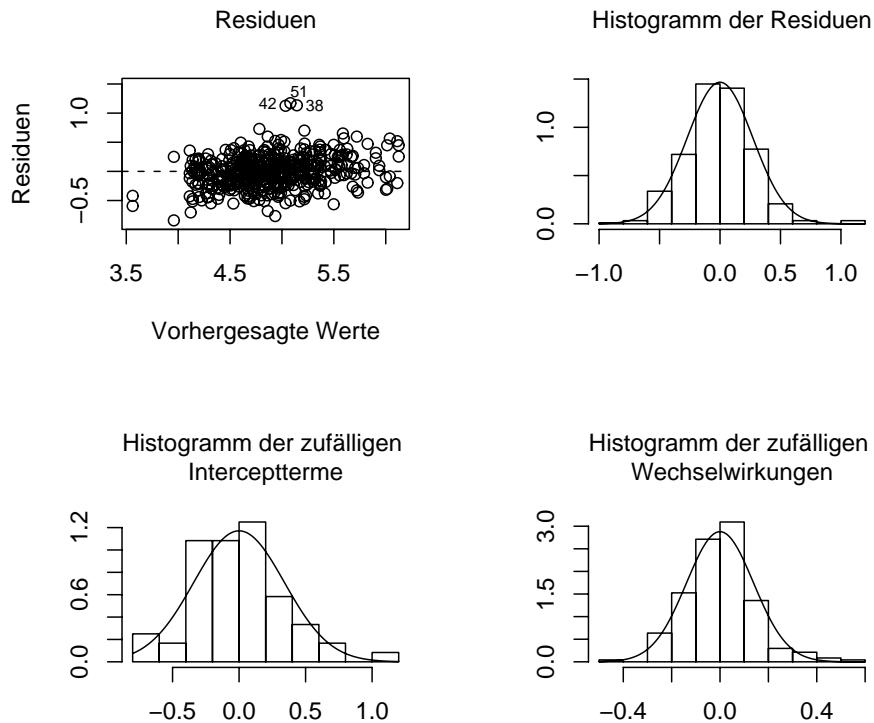
(Intercept)	Zeitpunkt8	Zeitpunkt9	Zeitpunkt10
5.0193643	-0.4249071	-0.4341489	-0.6521457
Dominanzindex	Behornung1	Zeitpunkt8:Dominanzindex	
0.1489337	0.1482779		0.3481315
Zeitpunkt9:Dominanzindex	Zeitpunkt10:Dominanzindex		
	0.1793187		0.6367413

Der Interceptterm ist der Mittelwert der Zeitpunktkategorie 7 und hat daher auch den höchsten Wert aller Zeitpunkte. Der Parameter Dominanzindex stellt die Steigung der Responsevariable für den Zeitpunkt 7 dar. Diese ist mit 0,149 weit geringer als die Steigung aus dem marginalen Modell mit 0,4. Der Wert für die Interaktion des Zeitpunktes 8 mit dem Dominanzindex ist die zusätzliche Steigung, insgesamt macht das einen Wert von 0,497. Für den Zeitpunkt 9 ist die Gesamtsteigung 0,328, und für den letzten Zeitpunkt 0,786. Das bedeutet, dass für den letzten Zeitpunkt die Kovariable Dominanzindex die größte Aussagekraft in Bezug auf die Responsevariable hat, für den Zeitpunkt 7 hingegen hat der Dominanzindex kaum einen Erklärungswert. Das lässt sich auch an den einzelnen Tests für die Parameterschätzer erkennen, welche mit der `summary`-Funktion verfügbar sind,

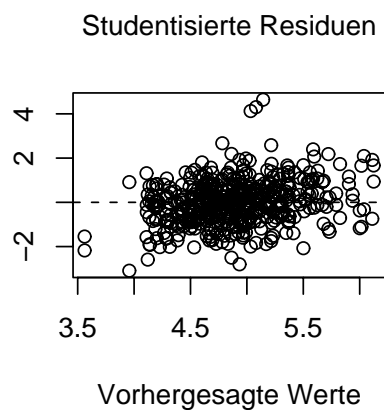
	Value	Std.Error	DF	t-value	p-value
(Intercept)	5.019	0.126	223	39.70	0.0000
Zeitpunkt8	-0.424	0.100	170	-4.22	0.0000
Zeitpunkt9	-0.434	0.099	170	-4.35	0.0000
Zeitpunkt10	-0.652	0.100	170	-6.48	0.0000
Dominanzindex	0.148	0.213	57	0.69	0.4893
Behornung1	0.148	0.108	57	1.36	0.1762
Zeitpunkt8:Dominanzindex	0.348	0.198	170	1.75	0.0812
Zeitpunkt9:Dominanzindex	0.179	0.196	170	0.91	0.3627
Zeitpunkt10:Dominanzindex	0.636	0.197	170	3.22	0.0015

Die Wechselwirkung des letzten Zeitpunktes mit dem Dominanzindex hat den geringsten P-Wert.

Die Schätzungen für die Varianzkomponenten des Modells sind  $\hat{\sigma}_I^2 = 0,143$ ,  $\hat{\sigma}_W^2 = 0,053$  und  $\hat{\sigma}^2 = 0,107$ . Die Residuenvarianz, welche die Streuungskomponente ist, die nicht durch das Modell erklärt werden kann, hat sich im Unterschied zum zufälligen Intercept-Modell fast halbiert. Im Folgenden sind diagnostische Graphiken der Residuen und der zufälligen Effekte abgebildet,



Die Residuen weisen eine homogene Varianz über die vorhergesagten Werte auf, die Normalverteilungseigenschaft kann für alle drei Ebenen der Streuung akzeptiert werden. In folgendem Plot sind die studentisierten Residuen aufgetragen, diese weisen jedoch keine großen Unterschiede zu den normalen Residuen auf.



In dem Modell sind alle Parameter bis auf den Faktor Behornung signifikante Einflussgrößen. Im Folgenden wird der Parameter aus dem Modell entfernt, um etwaige Verbesserungen der Modellinformationskriterien zu untersuchen. Folgender R Befehl führt das Modell aus,

```
lme(log(Kortisolgehalt)~factor(Zeitpunkt)*Dominanzindex,
    data=Kortisol,random=~1|Tier/Zeitpunkt)
```

Die Ergebnisse zeigen, dass sich nur mehr signifikante Parameter im Modell befinden,

	numDF	denDF	F-value	p-value
(Intercept)	1	223	8295.727	<.0001
factor(Zeitpunkt)	3	170	17.569	<.0001
Dominanzindex	1	58	5.441	0.0232
factor(Zeitpunkt):Dominanzindex	3	170	3.800	0.0114

Da sich die Modelle in der Spezifikation der festen Terme unterscheiden, ist ein Vergleich nur mit Maximum-Likelihood-Schätzungen möglich. In folgender Tabelle sind einige Modelle und deren Modellinformationskriterien eingetragen.

Modelle (ML-Schätzungen)	AIC	BIC
Zufälliges Intercept-Modell	648,35	660,73
Marginales Modell Zeitpunkt	567,53	596,44
Marginales Modell Behornung	649,32	665,83
Marginales Modell Dominanzindex	645,11	661,63
Haupteffekte	564,31	601,47
Haupteffekte + WW	564,90	630,97
Haupteffekte + WW für Zpt. und Domind.	558,80	608,35
Zeitpunkt + Domind. + WW	558,75	604,17

Beim zufälligen Intercept-Modell und den marginalen Modellen für den Faktor Behornung und Dominanzindex gibt es eine Varianzkomponente für die Gruppierungsvariable Tier. Alle anderen Modelle sind mit zufälligen Effekten für die Variable Tier und zufälligen Wechselwirkungen für den Faktor Zeitpunkt angepasst worden. An den Werten ist zu erkennen, dass das BIC Modelle mit weniger Parametern bevorzugt. Das marginale Modell für den Zeitpunkt mit 7 Parametern hat den geringsten Wert. Die letzten beiden Modelle werden vom AIC bevorzugt, der Unterschied besteht nur aus dem festen Effekt für den Faktor Behornung. Die Werte des AIC sind nahezu identisch, der Faktor Behornung hat keinen zusätzlichen Erklärungswert.



Eine weitere Möglichkeit, die Signifikanz des Faktors Behornung zu überprüfen, ist mit dem Likelihood-Quotienten-Test. Das Nullmodell beinhaltet den Faktor Behornung nicht, beim Alternativmodell ist er inkludiert. Die Modelle werden mit der ML-Methode angepasst. Folgenden Output liefert die `anova`-Funktion,

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
	<code>lme.null</code>	1	11	558.75	604.17	-268.37		
	<code>lme.alt</code>	2	12	558.80	608.35	-267.40	1 vs 2	1.946676 0.1629

Der P-Wert spricht für ein beibehalten der Nullhypothese.

Wie bereits beim zufälligen Intercept-Modell vorgestellt, gibt es den Modellierungsansatz der wiederholten Messungen. Es ist möglich, die Varianzkovarianzmatrix der Residuen entsprechend der Tiere in Blöcke aufzuteilen und auf diese Weise Varianzkovarianzparameter zu schätzen. Die Kombination von zufälligen Effekten und wiederholten Messungen ist prinzipiell möglich, bei dem speziellen Datensatz aber nicht sinnvoll, da die Parameter teilweise dasselbe schätzen. Im Folgenden wird das Haupteffektemodell mit Wechselwirkungen für den Faktor Zeitpunkt und Dominanzindex im Sinne einer Analyse für wiederholte Messungen angepasst. Demnach wird die Varianzkovarianzmatrix spezifiziert. Die Form „compound symmetry“ entspricht exakt dem Modell für einen zufälligen Intercept für die Tiere, somit wird die Form nicht spezifiziert. Einige Parametrisierungen basieren auf äquivalenten zeitlichen Abständen der Messungen, wie zum Beispiel AR(1)- oder ARMA-Prozesse. Diese sind für den Datensatz nicht gegeben. Die Form der Varianzkovarianzmatrix in dem Modell ist „unstrukturiert“.

```
gls(log(Kortisolgehalt)~factor(Zeitpunkt)*Dominanzindex+
    factor(Behornung),data=Kortisol, correlation=+
    corSymm(form=~Probezeitpunkt|Tier,fixed=F),+
    weights = varIdent(form = ~ 1|Probezeitpunkt))
```

In dem Modell werden 45 Parameter geschätzt, davon 36 Varianzkovarianzparameter. Die Signifikanz der festen Parameter wird durch die `anova`-Funktion abgerufen,

	numDF	F-value	p-value
(Intercept)	1	11970.806	<.0001
factor(Zeitpunkt)	3	22.780	<.0001
Dominanzindex	1	3.644	0.0569
factor(Behornung)	1	6.758	0.0096
factor(Zeitpunkt):Dominanzindex	3	4.026	0.0076

Der Faktor Behornung wird als signifikanter Parameter angezeigt. Bis auf die Korrelation des 2. mit dem 5. Probezeitpunkt sind alle Korrelationsparameter signifikant, die Tabelle wird nicht angezeigt. Der Vergleich mit den zuvor gerechneten Modellen ist durch die ML-Schätzungen möglich. Für das Modell wird ein AIC von 538 ausgegeben, der Wert ist um einiges besser als bei den Modellen mit zufälligen Effekten. Das BIC ist mit 724 weit schlechter, was auf die Anzahl der Parameter in dem Modell zurückzuführen ist.

## 11.7 Analyse in SPSS

Das Haupteffektemodell mit zweifachen Wechselwirkungstermen wird in SPSS auf folgende Weise eingegeben,

```
/FIXED=Behornung Dominanzindex Zeitpunkt Zeitpunkt*
Behornung Zeitpunkt*Dominanzindex
Behornung*Dominanzindex | SSTYPE(3)
```

Der Wert der restringierten -2Log-Likelihood ist 586,209 und somit identisch mit R. Folgende mit R identische Varianzkomponenten werden ausgegeben,

**Schätzungen von Kovarianzparametern<sup>a</sup>**

Parameter	Schätzung	Std.-Fehler	Wald Z	Sig.
Residuum	,150458	,010782	13,955	,000
Konstanter Term [Subjekt = Tier] Varianz	,152338	,032723	4,655	,000

Für das Modell mit den Haupteffekten, der Wechselwirkung für den Faktor Zeitpunkt und Dominanzindex sowie zufälligen Wechselwirkungen für den Faktor Zeitpunkt wird folgende Syntax eingegeben,

```
/FIXED=Behornung Dominanzindex Zeitpunkt
Zeitpunkt*Dominanzindex | SSTYPE(3)
/RANDOM=intercept Zeitpunkt | sub(Tier) COVTYPE(VC).
```

Folgende Testergebnisse für die festen Effekte werden ausgegeben,

**Tests auf feste Effekte, Typ III<sup>a</sup>**

Quelle	Zähler-Freiheitsgrade	Nenner-Freiheitsgrade	F-Wert	Signifikanz
Konstanter Term	1	56,165	2815,188	,000
Behornung	1	56,170	1,875	,176
Dominanzindex	1	56,246	6,372	,014
Zeitpunkt	3	160,566	14,723	,000
Zeitpunkt * Dominanzindex	3	159,706	3,802	,011

Die Varianzkomponenten des Modells lauten,

**Schätzungen von Kovarianzparametern<sup>a</sup>**

Parameter	Schätzung	Std.-Fehler	Wald Z	Sig.
Residuum	,107247	,010344	10,368	,000
Konstanter Term [Subjekt = Tier] Varianz	,143192	,032376	4,423	,000
Zeitpunkt [Subjekt = Tier] Varianz	,052110	,013363	3,900	,000

Die Ergebnisse sind identisch mit R. In der folgenden Tabelle sind 95%-Konfidenzintervalle der Varianzkomponenten aufgelistet,

	Untere	Schätzwert	Obere
$\hat{\sigma}_I^2$	0,092	0,143	0,223
$\hat{\sigma}_W^2$	0,032	0,052	0,086
$\hat{\sigma}^2$	0,089	0,107	0,130

Die größte Spannweite besitzt das Intervall für die Varianzkomponente  $\hat{\sigma}_I^2$ . Das ist darauf zurückzuführen, dass die Anzahl der Ausprägungen mit 60 weit weniger ist, als für die zufälligen Wechselwirkungen mit 236 oder für die Residuen mit 459 Ausprägungen, hier kann der Schätzwert präziser angegeben werden.

Das Modell mit den Haupteffekten sowie der Wechselwirkung für den Faktor Zeitpunkt und Dominanzindex wird nun mit der Analyse für wiederholte Messungen angepasst. In SPSS kann das mit folgender Syntax eingegeben werden,

```
MIXED LogKortisolgehalt BY Zeitpunkt Probezeitpunkt
      Tier Behornung With Dominanzindex
/FIXED=Zeitpunkt Dominanzindex Behornung
      Zeitpunkt*Dominanzindex|sstype(3)
/REPEATED=Probezeitpunkt | sub(Tier) covtype(un)
```

Durch den Befehl `covtype(un)` wird die Form einer unstrukturierten Varianzkovarianzmatrix spezifiziert. Die Ordnungsstruktur der Messungen wird durch den Probezeitpunkt angegeben, die Blockstruktur durch die Tiere. Der Wert der restringierten  $-2\text{Log-Likelihood}$  ist 476,124 und somit identisch mit dem Wert in R. Die festen Effekte und die Varianzkovarianzparameter werden in SPSS und R identisch geschätzt, die Parametrisierung ist eine andere. Für die festen Effekte werden folgende Testergebnisse ausgegeben,

Tests auf feste Effekte, Typ III<sup>a</sup>

Quelle	Zähler-Freiheitsgrade	Nenner-Freiheitsgrade	F-Wert	Signifikanz
Konstanter Term	1	55,380	2830,554	,000
Zeitpunkt	3	54,753	13,130	,000
Dominanzindex	1	55,653	11,842	,001
Behornung	1	54,762	6,822	,012
Zeitpunkt * Dominanzindex	3	55,156	4,025	,012

Der Faktor Behornung sowie alle anderen Terme sind signifikant, in folgender Tabelle sind Informationskriterien für die Maximum-Likelihood-Schätzung angegeben,

Informationskriterien<sup>a</sup>

-2 Log-Likelihood	448,251
Akaike- Informationskriterium (AIC)	538,251
Hurvich und Tsai (IC)	548,275
Bozdogan-Kriterium (CAIC)	769,058
Bayes-Kriterium von Schwarz (BIC)	724,058

Zu erkennen sind sehr gute Werte für AIC und IC, die hohen Werte für BIC und CAIC sind ein Anzeichen für die große Anzahl an Parametern in dem Modell.

## 11.8 Analyse in SAS

Das Haupteffektemodell mit Wechselwirkungen wird in SAS auf folgende Weise generiert,

```
class Tier Zeitpunkt Behornung;
model LogKortisolgehalt=Zeitpunkt Dominanzindex Behornung
      Zeitpunkt*Dominanzindex    Zeitpunkt*Behornung
      Dominanzindex*Behornung /s DDFM=sat;
random Intercept / sub=Tier;
```

Bei dem Modell gibt es zwei Varianzkomponenten, folgende mit SAS und R identische Schätzungen werden ausgegeben,

Kov.Parm	Subjekt	Schätzwert	Standardfehler	Z-Wert	Pr > Z
Intercept	Tier	0.1523	0.03272	4.66	<.0001
Residual		0.1505	0.01078	13.95	<.0001

Für die festen Effekte des Modells werden in allen drei Softwarepaketen dieselben Schätzer ausgegeben.

Nun zu dem Modell, welches nur einen Wechselwirkungseffekt für den Faktor Zeitpunkt und den Dominanzindex sowie zufällige Wechselwirkungen für den Faktor Zeitpunkt innerhalb der Tiere beinhaltet. Bei der Syntax sind folgende Statements zu verändern,

```
class Tier Zeitpunkt Behornung;
model LogKortisolgehalt=Zeitpunkt|Dominanzindex Behornung /s
      DDFM=satterthwaite;
random Intercept Zeitpunkt / sub=Tier;
```

Folgende 95%-Konfidenzintervalle für die Varianzkomponenten werden von SAS ausgegeben,

Kov.Parm	Subjekt	Untere	Schätzwert	Obere
Intercept	Tier	0.09614	0.1432	0.2359
Zeitpunkt	Tier	0.03337	0.0521	0.0927
Residual		0.08954	0.1072	0.1308

Die Schätzungen stimmen mit SAS und R überein, die Intervallgrenzen sind etwas unterschiedlich. Weiters werden die Tests auf feste Parameter ausgegeben,

Effekt	Zähler Freiheitsgrade	Nenner Freiheitsgrade	F-Statistik	Pr > F
Zeitpunkt	3	161	14.72	<.0001
Behornung	1	56.2	1.88	0.1763
Dominanzindex	1	56.2	6.37	0.0144
Dominanzindex*Zeitpunkt	3	160	3.80	0.0114

Die Freiheitsgrade im Nenner werden durch die Satterthwaite-Methode berechnet und sind identisch mit SPSS. In R werden für den Faktor Zeitpunkt und den Wechselwirkungsterm 170 Freiheitsgrade im Nenner verwendet. Das unterscheidet sich etwas von der Anzahl in SAS und SPSS mit ca. 160, die P-Werte sind davon aber nicht substantiell betroffen.

Das Modell kann auch über den Ansatz der wiederholten Messungen geschätzt werden. Hierfür wird die Varianzkovarianzmatrix der Residuen

entsprechend der Versuchseinheiten in Blöcke gegliedert. Der Faktor Probezeitpunkt stellt die Ordnungsstruktur der Messungen dar. Es gibt einige Möglichkeiten der Parametrisierung, wie zum Beispiel die einfache Form „compound symmetry“. Die Ergebnisse einer solchen Analyse sind aber identisch mit einem zufälligen Effekt für den Faktor Tier. Im Folgenden wird eine unstrukturierte Varianzkovarianzmatrix verwendet, welche mit folgender Syntax eingegeben wird,

```
proc mixed data=Kortisol method=REML covtest cl;
  class Tier Probezeitpunkt Zeitpunkt Behornung;
  model LogKortisolgehalt=Zeitpunkt|Dominanzindex Behornung /s;
  repeated Probezeitpunkt / sub=Tier type=un;
run;
```

Der Wert der  $-2\text{Log-Likelihood}$  ist 476,124, das Modell wird auf die gleiche Weise geschätzt wie in R und SPSS. Die Testergebnisse der festen Effekte bringen leicht unterschiedliche Ergebnisse, bei dieser Art der Analyse ist der Faktor Behornung signifikant.

Effekt	Zähler Freiheitsgrade	Nenner Freiheitsgrade	F-Statistik	Pr > F
Behornung	1	57	6.82	0.0115
Dominanzindex	1	57	11.84	0.0011
Zeitpunkt	3	57	13.13	<.0001
Dominanzin*Zeitpunkt	3	57	4.03	0.0115

Die Parameterschätzungen der Varianzkovarianzparameter sind in den Softwarepaketen identisch, der 5. und der 2. Probezeitpunkt weisen keine signifikante Kovarianz auf.

Folgende Werte für die Informationskriterien werden von SAS für die ML-Schätzungen berechnet, AIC=538, BIC=632. In dem Modell werden 36 Varianzkovarianzparameter geschätzt. Zum Vergleich dazu, das vorher gerechnete Modell mit den Haupteffekten, der Wechselwirkung *Zeitpunkt \* Dominanzindex* und den zufälligen Wechselwirkungen für den Faktor Zeitpunkt hat folgende Werte der Informationskriterien für die ML-Schätzungen, AIC=558, BIC=583. Hier werden 3 Varianzparameter geschätzt. Es bleibt die Frage offen, welches Modell „besser“ geeignet ist, die Signifikanz des Faktors Behornung sicherzustellen bzw. welches Informationskriterium für die Wahl des „besten“ Modells herangezogen wird.

Zusammenfassend kann gesagt werden, dass die Schätzungen der festen Parameter und der Varianzkomponenten in allen drei Softwarepaketen

identisch sind. Unterschiede gibt es bei den Testprozeduren und bei der Berechnung der Konfidenzintervalle für Varianzkomponenten. In R und SPSS werden die Konfidenzintervalle identisch berechnet, in SAS wird die Satterthwaite-Methode verwendet. Die in SPSS verwendete Methode zur Berechnung der Freiheitsgrade im Nenner für die Testprozeduren der festen Effekte ist die Satterthwaite-Approximation, welche in SAS ebenfalls angeboten wird. Die in R verwendete Berechnung bringt ähnliche Resultate. In SAS sind weitere Optionen für die Berechnung der Größen vorhanden. Die Adjustierung nach Kenward-Roger ist bei kleinen oder unbalancierten Designs empfehlenswert.

## Literaturverzeichnis

Pinheiro José C. und Bates Douglas M. (2000). *Mixed-Effects Models in S and S-PLUS*, Springer-Verlag, ISBN 0-387-98957-9

Die Modellformulierung für das einstufige und das mehrstufige Modell sind in dem Kapitel 2.1.1 bzw. 2.1.2 zu finden. Die Herleitung der Likelihoodfunktion wird in Kapitel 2.2.1 beschrieben, die der restringierten Likelihoodfunktion in Abschnitt 2.2.5. Die verschiedenen Optimierungsalgorithmen werden in Kapitel 2.2.8 erläutert. Approximative Verteilungseigenschaften und Inferenzmethoden werden in den Abschnitten 2.3-2.4 angegeben. Das Kapitel 2.5 beschäftigt sich mit den Vorhersagen im linearen gemischten Modell.

Searle Shayle R., Casella George und McCulloch Charles E. (1992). *Variance Components*, Wiley, ISBN 0-471-62162-5

Die Herleitung der ANOVA-Schätzer ist in dem Kapitel 3.5 beschrieben. In den Abschnitten 3.6-3.7 werden die Maximum-Likelihood- und die restringierte Maximum-Likelihood-Methode für das Modell der einfachen Varianzanalyse erläutert. In Kapitel 6 wird eine generelle Abhandlung über die ML- bzw. REML-Schätzung gegeben, Kapitel 7 beschäftigt sich mit der Vorhersage zufälliger Effekte. Optimierungsalgorithmen für die Likelihoodfunktion sind in den Abschnitten 8.2-8.3 zu finden.

Demidenko Eugene (2004). *Mixed Models, Theory and Applications*, Wiley, ISBN 0-471-60161-6

Die Modellformulierung ist in Abschnitt 2.2.1 zu finden, die Herleitung der Likelihoodfunktion in Abschnitt 2.2.2. In Kapitel 2.2.4 werden die profilierten Likelihoodfunktionen angegeben, das darauf folgende Kapitel 2.2.5 beschäftigt sich mit der REML-Schätzung. Abschnitt 2.8 beschäftigt sich mit der generellen Formulierung von Maximierungsalgorithmen, in Abschnitt 2.10 wird der Newton-Raphson-Algorithmus vorgestellt. In Kapitel 2.11 wird das Fisher-Scoring, in Kapitel 2.12 der EM-Algorithmus erläutert. Die zwei folgenden Kapitel behandeln die Startpunkte und die Optimierung für die REML-Schätzung. In Abschnitt 3.3 wird die Informationsmatrix der Schätzer angegeben, in Abschnitt 3.4 werden die profilierten Likelihoodintervalle vorgestellt. In Kapitel 3.5 wird das Testen der zufälligen Effekte behandelt, in Kapitel 3.7 wird die Vorhersage besprochen.

Davidian Marie, Giltinan David M. (1995). *Nonlinear Models for Repeated*



*Measurement Data* ,Chapman & Hall, ISBN 0-412-98431-9

In Kapitel 3.2.1 wird das lineare gemischte Modell formuliert, die Abschnitte 3.3.1 und 3.3.2 beschäftigen sich mit Inferenzmethoden für die zufälligen Effekte. In Kapitel 3.4.1 wird der EM-Algorithmus, in Kapitel 3.4.2 der Newton-Raphson und das Fisher-Scoring vorgestellt.

Faraway, Julian J. (2010). *Extending the Linear Model with R* , Chapman & Hall, ISBN 978-1-58488-424-8

In Abschnitt 8.1 wird das Modell mit zufälligen Effekten eingeführt, im folgenden Kapitel die Inferenzmethoden erläutert. Im Abschnitt 8.3 wird die Vorhersage der zufälligen Effekte behandelt, und Kapitel 8.6 beschäftigt sich mit verschachtelten Modellen.

Venables W.N., Ripley B.D. (2002). *Modern Applied Statistics with S* , Springer-Verlag, ISBN 978-0-387-95457-8

Die Abschnitte 10.1 und 10.2 beschäftigen sich mit der Anwendung linearer gemischter Modelle

Johnston J. (1984). *Econometric Methods*, McGraw-Hill, ISBN 0-07-032685-1

Littell, Milliken, Stroup, Wolfinger, Schabenberger (2006). *SAS for Mixed Models* , Second Edition. Cary, NC: SAS Institute Inc.

Weitere verwendete Quellen sind die von SAS und SPSS zur Verfügung gestellten Dokumentations- und Hilfsseiten.

## **Zusammenfassung**

Das Thema der Arbeit sind lineare gemischte Modelle, welche in vielen Bereichen der Biostatistik angewendet werden. Wiederholte Messungen an Subjekten sowie gruppierte Datenstrukturen können mit dieser Art von Modellen analysiert werden. Die Korrelationen der Messungen werden berücksichtigt und in Form eigener Varianzkomponenten geschätzt. Im ersten Teil der Arbeit werden die theoretischen Grundlagen des Modells beschrieben, der zweite Teil beschäftigt sich mit der Analyse eines realen Datensatzes.

Im ersten Kapitel wird anhand von einführenden Beispielen die Anwendung linearer gemischter Modelle erläutert. Im folgenden Abschnitt wird das Modell in Matrixschreibweise formuliert, dessen Annahmen und Eigenschaften besprochen. Das Thema des dritten Kapitels sind die zur Verfügung stehenden Schätzmethoden, hauptsächlich wird die Maximum-Likelihood-Schätzung und die restringierte Maximum-Likelihood-Schätzung behandelt. Das darauf folgende Kapitel beschäftigt sich mit den Methoden der Inferenz im linearen gemischten Modell, dem Testen von Hypothesen für die festen sowie zufälligen Effekte und der Konstruktion von Konfidenzintervallen. Zusätzlich wird anhand von Simulationen die Verteilung der Likelihood-Quotienten-Teststatistik untersucht.

Im weiteren Verlauf der Arbeit werden anhand eines Datensatzes lineare gemischte Modelle angepasst. Zu Beginn werden deskriptive Analysen durchgeführt, um die zur Verfügung stehenden Einflussvariablen zu untersuchen. Das zufällige Intercept-Modell bildet den Startpunkt der Modellierung, später werden komplexe Modelle, welche Haupteffekte und Wechselwirkungsterme beinhalten angepasst. Die Technik der Modellierung von zufälligen Effekten wird in eigenen Kapiteln angesprochen, diese sind das verschachtelte Modell und zufällige Wechselwirkungen.

Der Vergleich der Softwarepakete R, SPSS und SAS führt zu dem Ergebnis, dass die ausgegebenen Schätzungen der Parameter übereinstimmen, lediglich die Testprozeduren und die Konfidenzintervalle werden geringfügig anders berechnet, in SAS sind optionale Berechnungsvarianten möglich.

## Abstract

The issue of this paper is linear mixed-effects-models, which are able to handle grouped data structures as well as repeated measures. Correlation structures between measurements on the same subject are taken into account. These kinds of data arise in many areas of biostatistics. The first chapters provide a theoretical basis on the subject, in the second part the model is applied on a real dataset.

The first chapter contains introductory examples of mixed-effects-models, followed by the general model formulation in the second chapter. The third chapter is about the estimation of the unknown parameters the main methods available are Maximum-Likelihood and restricted Maximum-Likelihood estimation. The following chapter covers the hypothesis testing for the fixed and random effects and the construction of confidence intervals. Simulation studies are made to investigate the distribution of the Likelihood-Ratio-Teststatistic.

The following chapters deal with the applied analysis of linear mixed-effects-models on a dataset about milk goats. At first a descriptive analysis is given to provide an overview of the data structure, and to investigate the explanatory variables. Mixed-effects-models are fitted to the data, starting with simple random-intercept-models up to complex models which contain the main effects as well as interaction terms. The technique of modelling random effects is illustrated in own chapters which are about nested classification factors and random interaction effects.

The comparison of the results of the statistical software R, SPSS und SAS shows, that the values of the estimated parameters are equal. There are little differences in the hypothesis-testing-procedure and in the construction of confidence intervals, SAS provides optional methods to compute these quantities.

# Lebenslauf

## Persönliche Daten:

Name: Raimund Möser  
Geburtsdatum: 4. Juni 1981  
Geburtsort: Wien  
Staatsbürgerschaft: Österreich  
e-mail: raimundmoeser@hotmail.com

## Ausbildung:

1999 AHS-Matura BRG 9  
2001-2003 Studium der Geographie an der Universität Wien  
2004-2010 Studium der Statistik an der Universität Wien  
2009 Bakkalaureus der Sozial- und Wirtschaftswissenschaften, Statistik