

Statistische Auswertung von Vibrationsdaten mit dem linearem gemischten Modell

Bachelorarbeit

zur Erlangung des akademischen Grades
Bachelor of Science (B. Sc.)

vorgelegt dem
Fachbereich Wasser- und Kreislaufwirtschaft
der Hochschule Magdeburg-Stendal

von Frau CHRISTIN SCHRÖDER
geboren am 25. April 1990 in Finsterwalde
MtrNr. 20082636

Gutachter

1. Prof. Dr. JÜRGEN KÖHLER, Hochschule Magdeburg-Stendal
2. Dr. GIZEM FORTA, Bundesanstalt für Arbeitsschutz und Arbeitsmedizin

Magdeburg, 20. August 2012

*Die vorliegende Arbeit entstand während meiner Tätigkeit als studentische
Hilfskraft bei der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin,
in der Gruppe 3.2 Experimentelle Wirkungsforschung.*

Inhaltsverzeichnis

1	Einleitung	2
1.1	Ganzkörpervibrationen unter dem Aspekt des Arbeitsschutzes	3
1.2	Durchführung des Experimentes	3
1.3	Datengrundlage	4
1.4	Problemstellung	6
2	Varianzanalyse	8
2.1	Zweifaktorielle Varianzanalyse	10
2.2	Varianzanalyse mit Messwiederholung	17
2.3	Vergleich beider Analysemethoden	21
3	Lineares gemischtes Modell	23
3.1	Bildung des linearen gemischten Modells	23
3.2	Berechnung der Varianzen im gemischten Modell	26
3.3	Weitere Berechnungsmöglichkeiten für Parameter im linearen gemischten Modell	32
3.4	Weitere Tests im linearen gemischten Modell	36
4	Ergebnisse	39
4.1	Darstellung der Ergebnisse	39
4.2	Vergleich von Mixed Model und Varianzanalyse für wiederholte Messungen	40
4.3	Vergleich von Eichungs-Kurven aus dem Experiment mit Wichtungskurve aus ISO 2631-1	41
5	Zusammenfassung und Ausblick	46
	Literaturverzeichnis	48

1 Einleitung

Im täglichen Leben, ob auf der Arbeit oder in der Freizeit, sind die Menschen häufig Vibrationen ausgesetzt, z.B. bei einer Busfahrt, einer Autofahrt oder dem Halten einer Bohrmaschine. Es ist bekannt, dass Vibrationen ab einer gewissen Expositionsdauer bzw. Stärke zu Schäden führen können. Die Grenzwerte für die Dauer und Stärke von Ganzkörpervibrationen sind unter anderem in der ISO 2631-1 [1] und in der Lärm- und Vibrations-Arbeitschutzverordnung [2] festgelegt. Besonders wichtig sind diese Grenzwerte für Personen, die beruflich Vibrationen ausgesetzt sind, da sie vor möglichen Schäden der Exposition geschützt werden sollen.

Die den Grenzwerten zugrunde liegenden Frequenzbewertungskurven werden momentan von der Arbeitsgruppe 3.2 der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin untersucht. Dazu wurde ein Experiment durchgeführt bei dem Probanden im Labor unterschiedlichen Vibrationen ausgesetzt waren. Vibrationen sind durch unterschiedliche Amplituden, Frequenzen und Richtungen charakterisiert. In dem Experiment wurde die Wirkungsrichtung der Vibrationen konstant gehalten und die Amplituden und Frequenzen verändert. Zur Bewertung der Vibrationen wurden die Probanden gefragt, wie intensiv sie die Vibrationswirkung beurteilen. Im Weiteren wird diese Frage mit dem Begriff Intensitätsfrage bezeichnet. Die Antworten der Probanden stellen die Datengrundlage dieser Arbeit dar.

Diese Arbeit beschäftigt sich mit der Auswertung der Daten aus der Intensitätsfrage und dem Vergleich der im Labor der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin gewonnenen Frequenzbewertungskurven mit denen der ISO 2631-1 dahingehend, ob beide den selben Verlauf haben und sich damit die selbe frequenzbewertete Expositionstärke ergibt.

Im weiteren dieses Kapitels wird genauer auf den Begriff der Vibration, insbesondere der Ganzkörpervibration, eingegangen. Ebenfalls wird die Versuchsdurchführung beschrieben und die Datenstruktur der auszuwertenden Experimentdaten erklärt.

1.1 Ganzkörpervibrationen unter dem Aspekt des Arbeitsschutzes

Ganzkörpervibrationen, auch Ganzkörperschwingungen genannt, werden durch mechanische Schwingungen verursacht. Die mechanischen Schwingungen werden von Maschinen auf den Menschen übertragen und wirken auf den ganzen Körper. Sie entstehen z.B. bei der Benutzung von Baustellenfahrzeugen. Aber auch in normalen PKWs, in Flugzeugen oder Schiffen ist der menschliche Körper Ganzkörpervibrationen ausgesetzt.

Ganzkörperschwingungen können sich auf das Wohlbefinden des Einzelnen auswirken und so die Leistungsfähigkeit verringern. Vibrationen können auch zu gesundheitlichen Schäden führen, beispielsweise zu Schäden an der Lendenwirbelsäule (Berufskrankheit 2110). [3]

Um den Arbeitnehmer vor den schädigenden Auswirkungen von Vibrationen zu schützen wurden Richtlinien entwickelt, die die zulässigen Schwingungsbelastungen, denen ein Mensch ausgesetzt werden darf, regeln. Es gibt beispielsweise die Richtlinie der Vereinigung Deutscher Ingenieure VDI 2057-1 [4] oder die international gültige Norm ISO 2631-1 [1]. Die Belastung, die durch die Schwingung verursacht wird, ist abhängig von der Frequenz, der Amplitude und der Vibrationsrichtung. In den Richtlinien sind Höchstwerte für die Tagesvibrationswerte angegeben, die nicht überschritten werden dürfen. Bei Überschreitungen der maximalen Vibrationsbelastung müssen Maßnahmen zur Verringerung der Exposition getroffen werden. [5]

Die ISO 2631 enthält Kurven, die das Komfortempfinden, das subjektive Empfinden, wie angenehm bzw. unangenehm Schwingungen empfunden werden, in Abhängigkeit von der Frequenz darstellen. Diese Kurven, die das Komfortempfinden in Abhängigkeit von der Frequenz darstellen, werden Wichtungskurven genannt. Sie dienen auch der Beurteilung von Gesundheitsgefährdungen durch Vibrationen. Die Richtigkeit der Verlauf dieser Kurven soll überprüft werden. Dazu führte die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin ein Vibrationsexperiment mit 13 Probanden durch.

1.2 Durchführung des Experimentes

Das Experiment zur Beurteilung des Vibrations-Komfortempfindens wurde in dem Vibrationslabor der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin durchgeführt.

In dem Experiment wurden 13 Versuchspersonen unterschiedlichen Schwingungen ausgesetzt. Während des Versuches saßen die Personen aufrecht auf dem Simulator. Es wurden insgesamt 15 unterschiedliche sinusförmige Schwingungen erzeugt. Dabei wurde zwischen 3 Amplituden ($0.4, 0.8$ und 1.6 ms^{-2}) und 5 ($0.8, 1.6, 3.2, 6.3$ und 12.5 Hz) Frequenzen variiert.

Die Bewertung des Schwingungskomforts durch die Probanden erfolgte mittels einer Linie, die auf die Wand projiziert wurde, die immer kleiner wurde. Die Probanden stoppten die Verkleinerung der Linie durch einen Mausklick. Anschließend wurden sie zu der selben Schwingung nochmals befragt. Diesmal jedoch wurde die Linie größer. Die Probanden stoppten die Bewegung der Linie wieder per Mausklick, wenn die Linie die Größe erreicht hatte, mit der sie die Schwingung bewerten wollten. [6]

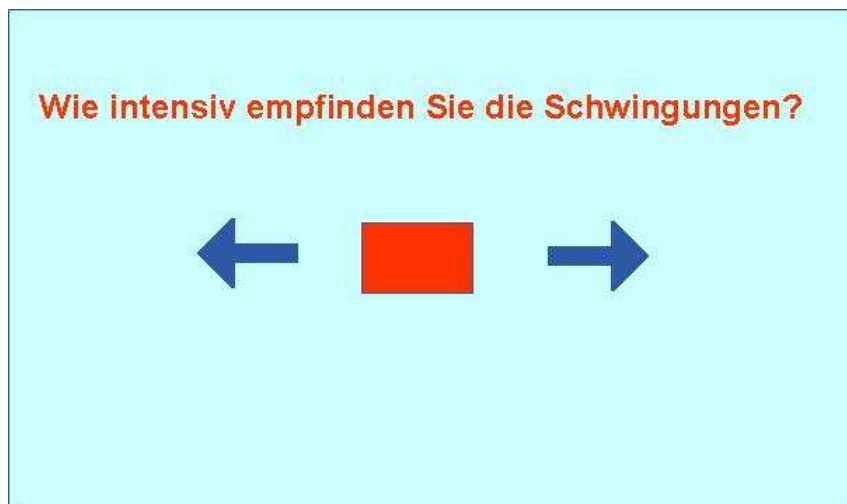


Abbildung 1.1: Frage zum Intensitätsempfinden mit wachsendem Balken

1.3 Datengrundlage

Die auszuwertenden Daten bestehen aus den von Probanden gegebenen Antworten. Zu den Antworten gehören die Linienlängen bei denen die Probanden die Verkleinerung bzw. die Vergrößerung der Linie stoppten. Die Linienlänge wurde als numerischer Wert in der jeweiligen Pixellänge gespeichert. Sie kann theoretisch Werte zwischen 0 und 820 annehmen. In dem Experiment traten Werte von 30 bis 820 auf. Zur Auswertung wird der Mittelwert aus dem Wert, der beim Vergrößern der Linie, und dem Wert, der beim Verkleinern der Linie abgegeben wurde, gebildet.

Zu jeder Linienlänge wurde die jeweilige Probandennummer (von 51 bis 63) ge-

speichert. Außerdem wurde die Frequenz und die Amplitude abgespeichert, die das jeweilige Intensitätsemfinden hervorgerufen haben. Die Frequenzen sind mit Werten von 2001 bis 2005 gekennzeichnet und die Codierung der Amplitude geht von 1001 bis 1003.

Zu jedem Probanden gibt es 15 Messwerte. Insgesamt gibt es also 195 Werte. In Tabelle 1.1 ist ein Auszug der ausgewerteten Daten dargestellt.

Tabelle 1.1: Datenauszug der Intensitätsdaten

Versuchperson	Frequenz	Amplitude	Linienlänge
51	2001	1001	50
51	2002	1001	340
51	2003	1001	180
51	2004	1001	124
⋮	⋮	⋮	⋮
63	2005	1003	400

Um einen ersten Eindruck über die Daten und ihre Verteilung zu bekommen, sind diese grafisch dargestellt. Dazu wurde die gemessene Linienlänge über die Frequenz, über die Amplitude und über die Probandennummer aufgetragen. An den Abbildungen aus 1.3 lässt sich erkennen, dass die Varianz der Daten hoch ist.

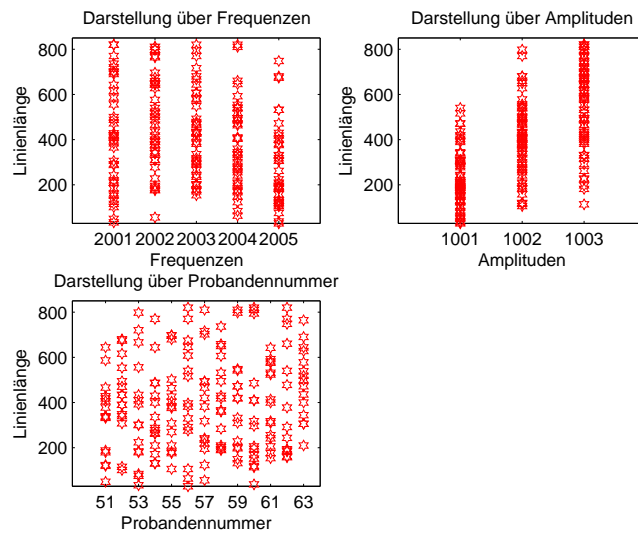


Abbildung 1.2: Darstellung der Ergebnisse der Intensitätsfrage aufgetragen über links oben Frequenz, rechts oben Amplitude, links unten Probandennummer

Abbildung 1.3 zeigt die gemessenen Werte für jede Amplitude in einem einzelnen Diagramm, dabei wurden über die 5 Frequenzen die dazugehörigen Antworten der 13 Probanden aufgetragen. Die Linie verbindet die einzelnen Mittelwerte, die zu jeder Frequenz berechnet wurden. Es wird deutlich, dass der Verlauf der Mittelwerte für jede Amplitude ähnlich ist.

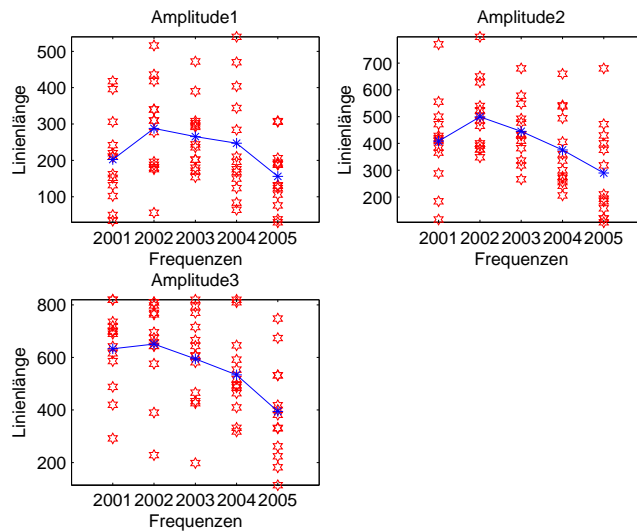


Abbildung 1.3: Grafische Darstellung der Intensitätsdaten getrennt nach den unterschiedlichen Amplitudenstufen. Die Roten Sterne stellen die Rohdaten dar, die blauen Linien sind die Mittelwerte über die jeweiligen Frequenzen.

1.4 Problemstellung

Diese Arbeit soll die Daten auswerten hinsichtlich der Fragestellung, ob die Änderung der Frequenz bzw. der Amplitude einen Einfluss auf das Intensitätsempfinden der Probanden haben. Dazu sollen die erfragten Werte mit zwei statistischen Analysemethoden, der Varianzanalyse und dem linearen gemischten Modell, ausgewertet werden und anschließend die beiden Verfahren verglichen werden.

Desweiteren sollen die im Experiment ermittelten Werte mit denen der ISO-Wichtungskurve verglichen werden um festzustellen, ob die ISO-Kurve einen signifikant anderen Verlauf hat als die Kurve aus den Experimentdaten.

In den folgenden Kapiteln sollen also die Punkte

- Auswertung der Daten mit der Varianzanalyse,

- Auswertung der Daten mit dem linearen gemischten Modell,
- Vergleich der beiden Verfahren und
- Test auf Gleichheit der Kurvenverläufe

behandelt werden.

Das zweite Kapitel beschreibt das Vorgehen bei der Auswertung von Daten mit der Varianzanalyse. Dabei wird sowohl auf die Varianzanalyse als auch auf die Varianzanalyse mit Messwiederholungen eingegangen. Die beiden Methoden werden vorgestellt, die Ergebnisse berechnet und dann verglichen.

Das dritte Kapitel beschäftigt sich mit dem linearen gemischten Modell. Zuerst wird diese Methode beschrieben, dann auf die Modellbildung sowie unterschiedliche Berechnungs- und Testmöglichkeiten eingegangen.

Das vierte Kapitel fasst die Ergebnisse aus den vorherigen Kapiteln zusammen und vergleicht die beiden Methoden miteinander, wobei auf die Vorteile des jeweiligen Verfahrens eingegangen wird. Anschließend wird die ISO-Kurve mit den berechneten Werten aus dem Experiment verglichen.

Das fünfte Kapitel stellt eine kurze Zusammenfassung der Ergebnisse dar und gibt einen kurzen Ausblick auf weitere Problemstellungen, die beim Bearbeiten der oben genannten Punkte auftraten.

In der Arbeit sind wichtige Begriffe **hervorgehoben**. Alle Ergebnisse dieser Arbeit wurden mit Matlab 7.6.0 berechnet und die Abbildungen unter Zuhilfenahme dieses Programmes erstellt.

2 Varianzanalyse

Ziel der Varianzanalyse, auch als ANOVA bezeichnet (*Analysis of Variance*), ist es, den Einfluss der unabhängigen Variablen X_i [$i \in \mathbb{N}$] auf die abhängige Variable Y zu klären und so Unterschiede zwischen den einzelnen unabhängigen Variablen bzw. den unterschiedlichen Ausprägungen der unabhängigen Variablen aufzudecken. Die unabhängigen Variablen nennt man auch Faktoren. Ein Faktor kann in unterschiedlichen Ausprägungen vorliegen (z.B. kein Medikament, Placebo, Medikament). Diese Ausprägungen werden Faktorstufen genannt.

Die Varianzanalyse untersucht also die Abhängigkeit einer abhängigen Variablen von mehreren unabhängigen Variablen. Der Grundgedanke bei der ANOVA ist es, dass sich die gesamte Varianz, die aus allen Daten berechnet werden kann, in zwei Varianzarten unterteilen lässt. Das wäre zum einen die erklärte Varianz. Dabei handelt es sich um die Varianz zwischen den Gruppen. Diese Varianz entsteht dadurch, dass Probanden unterschiedlichen Bedingungen ausgesetzt sind. Es ist zu erwarten, dass zwei Probanden, die der gleichen Gruppe angehören (z.B. gleich behandelt wurden), sich nicht so stark hinsichtlich des Mittelwertes unterscheiden wie zwei Probanden aus unterschiedlichen Gruppen. Zum anderen wäre es die nicht erklärte Abweichung. Jene Varianz wird dadurch verursacht, dass zwischen den Probanden individuelle Unterschiede bestehen und außerdem durch nicht beachtete Faktoren sowie durch zufällige Abweichungen. Diese Varianz wird auch Fehlervarianz genannt. Die Faktoren beeinflussen die abhängige Variable signifikant, wenn durch die Varianz zwischen den Gruppen die Gesamtvarianz besser erklärt wird, als durch die Varianz innerhalb der Gruppen, d.h. wenn die erklärte Varianz höher ist, als die nicht erklärte. Berechnen lässt sich dies über die Prüfgröße. Eine Prüfgröße, auch Testgröße genannt, ist eine Funktion, die unter Annahme der Nullhypothese einer bestimmten Verteilung folgt. Die zu überprüfende Nullhypothese lautet:

$$\mu_1 = \mu_2 = \dots = \mu_i. \quad (2.1)$$

Die Mittelwerte μ_i aller Faktoren sind gleich und es liegt kein signifikanter Unter-

schied zwischen den einzelnen Faktorstufen vor. Es kann also getestet werden, ob von den Unterschieden der Mittelwerte der Gruppen geschlossen werden kann, dass Unterschiede zwischen den Gruppen in der Grundgesamtheit vorliegen. Die Testgröße besteht aus dem Quotienten von erklärter Abweichung und nicht erklärter Abweichung. Sie testet so, ob die Varianz in den Faktorstufen größer ist als zwischen den Faktorstufen. Daher auch der Name des Verfahrens.

Varianzanalysen werden nach Anzahl der unabhängigen Variablen klassifiziert. So wird beispielsweise eine Varianzanalyse mit zwei unabhängigen Variablen zweifaktorielle ANOVA genannt. Wird die Wirkung von mehreren unabhängigen Variablen auf mehrere abhängige Variablen untersucht, so nennt man das MANOVA (*Multiple Analysis of Variance*). Wie die Varianzanalysen zu bezeichnen sind, ist aus Tabelle 2.1 zu entnehmen.

Tabelle 2.1: Klassifizierung der Varianzanalyse

Anzahl abhängiger Variablen	Zahl unabhängiger Variablen	Verfahrensbezeichnung
1	1	Einfaktorielle Varianzanalyse
1	≥ 2	Mehrfaktorielle Varianzanalyse
≥ 2	≥ 1	Mehrdimensionale Varianzanalyse

Zur Überprüfung der Nullhypothese können auch mehrere t-Tests durchgeführt werden, wobei dann alle Vergleiche zwischen den unterschiedlichen Faktorstufen einzeln vorgenommen werden würden. Der Nachteil dieser Vorgehensweise liegt darin, dass es bei der Durchführung mehrerer Tests zu Problemen des multiplen Testens kommen kann. Eines dieser Probleme ist die **Inflation des α -Fehlers**. Bei diesem Problem werden einige Testergebnisse nur zufällig als signifikant angenommen. Werden diese mehreren t-Tests z.B. zum Signifikanzniveau $\alpha = 0.05$ durchgeführt, so werden 5% der Tests zufällig signifikant sein.¹ Die Varianzanalyse stellt eine Erweiterung des t-Testes dar und ist diesem bei mehreren Vergleichen derselben Daten vorzuziehen, um das Problem der α -Inflation und weitere Probleme des multiplen Testens zu vermeiden. Bei der Untersuchung zweier Gruppen auf Gleichheit sind die Ergebnisse des t-Testes und der Varianzanalyse identisch.

Es existieren unterschiedliche Designs der Varianzanalyse, *within-design*, *between-*

¹Dies entspricht der Definition des α -Fehlers: H_0 wird abgelehnt, obwohl H_0 wahr ist.

design und *mixed-design*. Beim *between-design* wird jeder Proband genau einer Gruppe zugeordnet. Beim *within-design* wird jeder Proband in jeder Faktorstufe bzw. jeder Faktorstufenkombination (bei mehrfaktoriellen ANOVAs) getestet. Das *mixed-design* stellt eine Mischung aus den beiden eben genannten dar und ist nur für mehrfaktorielle Varianzanalysen anwendbar. In diesem ANOVA-Design werden einige Faktoren als *within* getestet, andere als *between*. Bei den Daten der Komfortfrage wurden die Probanden mehrfach zu unterschiedlichen Faktorkombinationen befragt. So, dass für die Daten der Komfortfrage sowohl ein *within-design* als auch ein *between-design* anwendbar ist.

Die Formeln von Kapitel 2.1 stammen aus [7] und die Gleichungen aus Kapitel 2.2 aus [8].

2.1 Zweifaktorielle Varianzanalyse

Die vorliegenden gewonnenen Daten aus der Intensitätsfrage beinhalten zwei Faktoren (Frequenz und Amplitude), so dass die einfaktorielle ANOVA nicht genutzt werden kann. Da eine abhängige metrisch skalierte Zielvariable und zwei unabhängige nominalskalierte Faktoren vorliegen, muss eine zweifaktorielle Varianzanalyse durchgeführt werden. Bei der Zielvariablen handelt es sich um die Linienlänge, die von den beiden Faktoren Frequenz und Amplitude abhängt. Der Faktor Frequenz liegt in 5 Ausprägungen vor, die Amplitude in 3 Faktorstufen. Man spricht auch von einem 5×3 faktoriellen Design.

In der zweifaktoriellen Varianzanalyse werden zwei Arten von Effekten unterschieden, sogenannte Haupteffekte und Interaktionseffekte. Der **Haupteffekt** des Faktors X beschreibt den Einfluss des Faktors X auf die abhängigen Variable Y . Neben der Analyse, ob sich einer der beiden Faktoren auf die Zielvariable auswirkt, lässt sich in der zweifaktoriellen ANOVA auch testen, ob sie sich zusammen auf die Zielvariable auswirken. Es lässt sich also herausfinden, ob eine Wechselwirkung zwischen den Einflussgrößen besteht. Ein **Interaktionseffekt** zwischen zwei Faktoren äußert sich darin, dass sich der eine Faktor je nach Ausprägung des anderen Faktors verschieden auf die unabhängige Zielvariable auswirkt.

Vor der eigentlichen Analyse können die Daten in einem **Interaktionsdiagramm** grafisch dargestellt werden. In einem Interaktionsdiagramm werden die Faktorstufenmittelwerte geplottet. Dazu wird auf der Abzisse der Faktor Frequenz aufgetragen und auf der Ordinate die Linienlänge. Anschließend werden die berechneten Mittelwerte für jede Faktorstufenkombination eingezeichnet. 2.1 Für jede Stufe des Faktors

Amplitude ergibt sich so eine eigene Linie. Diese Linien können miteinander verglichen werden und geben Aufschluss über mögliche Wechselwirkungen zwischen den Faktoren. Verlaufen diese drei Linien (da es drei unterschiedliche Amplituden gibt) parallel, heißt das, dass keine Interaktionen zwischen den beiden Faktoren Amplitude und Frequenz auftreten.

Wie in Bild 2.1 ersichtlich, verlaufen die drei Linien fast parallel, d.h. es liegen Effekte der Faktoren Frequenz und Amplitude vor. Die Voraussetzung dafür, dass keine Wechselwirkungen vorliegen, dass die Linien also parallel zueinander verlaufen sind approximativ erfüllt. Deswegen können die Wechselwirkungseffekte näherungsweise als vernachlässigbar angesehen werden.

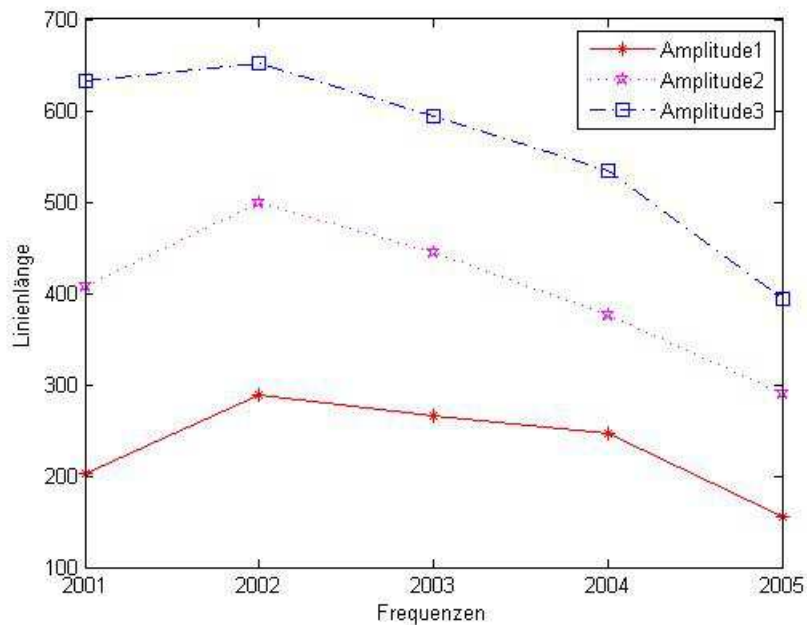


Abbildung 2.1: Interaktionsdiagramm für Intensitätsfrage

Bei der gewählten Darstellung von Interaktionen kann es zu Verfälschungen kommen, da gleichzeitig zu den Wechselwirkungsbeziehungen auch die Haupteffekte berücksichtigt werden. Um eine daraus resultierende falsche Interpretation zu vermeiden, wird die Interaktion anhand der residualen Mittelwerte interpretiert. Die residualen Mittelwerte ergeben sich als Differenz aus den erwarteten Mittelwerten der Interaktionsaktionseffekte und den empirischen Mittelwerten der Interaktionseffekte.

Im Weiteren wird der Faktor der Frequenz mit Faktor A bezeichnet und die Ein-

flussgröße Amplitude mit Faktor B. Faktor A hat $n_i = 5$ Stufen und Faktor B $n_j = 3$ Stufen. Die einzelnen Stufen des Faktors A werden mit i bezeichnet, die von Faktor B mit j . Eine dritte Laufvariable k kennzeichnet die Versuchsperson, so dass $k = 1, \dots, 13$. Die Variable y_{237} enthält die Antwort, die der siebente Proband auf die Intensitätsfrage zu den Einstellungen zweite Frequenz und dritte Amplitude abgegeben hat. Die Daten lassen sich übersichtlich in einer Tabelle, wie in Tabelle 2.1, darstellen, wo die Spalten die Ausprägungen des einen Faktors enthalten und die Zeilen die Faktorstufen des anderen Faktors. Jede mögliche Spalten-Zeilen-Kombination wird Zelle genannt, d.h. alle Werte die bei Frequenz 1 und Amplitude 1 gemessen wurden, stehen in einer Zelle.

Tabelle 2.2: Datenanordnung in einer Tabelle für zweifaktorielle Varianzanalyse

	Frequenz 1	Frequenz 2	...	Frequenz 5	
Amplitude 1	$x_{111}, x_{112},$..., x_{11k}	$x_{211}, x_{212},$..., x_{21k}	...	$x_{511}, x_{512},$..., x_{51k}	\bar{x}_{β_1}
Amplitude 2	$x_{121}, x_{122},$..., x_{12k}	$x_{211}, x_{222},$..., x_{22k}	...	$x_{521}, x_{522},$..., x_{52k}	\bar{x}_{β_2}
Amplitude 3	$x_{131}, x_{132},$..., x_{13k}	$x_{231}, x_{232},$..., x_{23k}	...	$x_{531}, x_{532},$..., x_{53k}	\bar{x}_{β_3}
	\bar{x}_{α_1}	\bar{x}_{α_2}	...	\bar{x}_{α_5}	\bar{x}

Das Modell für die zweifaktorielle Varianzanalyse lautet:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}. \quad (2.2)$$

Y_{ijk} : *Unabhängige Variable*

μ : *Globaler Mittelwert*

α_i : *Haupteffekt des Faktors A*

β_i : *Haupteffekt des Faktors B*

$(\alpha\beta)_{ij}$: *Wechselwirkung der Faktoren*

ϵ : *Residuen*

In dem Modell bezeichnet Y_{ijk} die Zielvariable, also die Antworten der Probanden auf die Intensitätsfrage. μ bezeichnet den globalen Mittelwert, auch als *grand mean*

bezeichnet, und berechnet sich als Mittelwert aus allen gegebenen Antworten.

$$\mu = \frac{1}{n_i n_j} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \mu_{ij} \quad (2.3)$$

α_i und β_i stellen die Haupteffekte dar. α_i beschreibt den Effekt der i -ten Faktorstufe des Faktors A.

$$\alpha_i = \mu_i - \mu \quad (2.4)$$

μ_i ist der Erwartungswert in der i -ten Stufe des Faktors A.

$$\mu_i = \frac{1}{n_j} \sum_{j=1}^{n_j} \mu_{ij} \quad (2.5)$$

β_i beschreibt analog zu α_i die Wirkung der j -ten Faktorstufe des Faktors B. Die Berechnung erfolgt entsprechend zur Berechnung des Effektes α_i . $(\alpha\beta)_{ij}$ beschreibt die Wechselwirkung zwischen den beiden Faktoren.

$$(\alpha\beta)_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_i) \quad (2.6)$$

Zuerst müssen die Schätzer für die Modellparameter berechnet werden. Der Schätzer für den globalen Mittelwert ergibt sich als Mittelwert aus allen Betrachtungen. Die Anzahl der Antworten ist $i \cdot j \cdot k$, da die 13 Probanden eine Frage für jede Frequenz- und jede Amplitudeneinstellung beantwortet haben.

$$\hat{\mu} = \frac{1}{n_i n_j n_k} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} Y_{ijk} = \bar{y} \quad (2.7)$$

Die Effekte der Faktoren A und B werden über die Abweichungen der Mittelwerte der Faktorstufen vom globalen Mittelwert bestimmt.

$$\alpha_i = \bar{y}_i - \hat{\mu} \quad (2.8)$$

$$\hat{y}_i = \frac{1}{n_j n_k} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk} \quad (2.9)$$

$$\beta_j = \bar{y}_j - \hat{\mu} \quad (2.10)$$

$$\hat{y}_j = \frac{1}{n_i n_k} \sum_{i=1}^{n_i} \sum_{k=1}^{n_k} y_{ijk} \quad (2.11)$$

Der Parameter der Wechselwirkungen zwischen den beiden Faktoren lässt sich als die Differenz zwischen dem Gesamtmittelwert und den Mittelwerten für die Faktoren A und B sowie die Mittelwerte je Zelle berechnen.

$$\widehat{(\alpha\beta)}_{ij} = \mu - \bar{y}_{ij} - \bar{y}_i - \bar{y}_j \quad (2.12)$$

$$\bar{y}_{ij} = \frac{1}{n_k} \sum_{k=1}^{n_k} y_{ijk} \quad (2.13)$$

Die Residuen ergeben sich aus:

$$\epsilon_{ijk} = y_{ijk} - (\hat{\mu} + \alpha_i + \beta_j + (\alpha\beta)_{ij}). \quad (2.14)$$

Zur Überprüfung der Einflüsse der einzelnen Effekte muss eine geeignete Prüfgröße gebildet werden. Diese leitet sich aus der Streuungszerlegung, die die Gesamtstreuung in die durch die jeweiligen Effekte erklärte Streuung zerlegt, ab. Die Gesamtstreuung lässt sich in die Streuung der Faktoren A und B, die Streuung der Interaktion und die Reststreuung, zerlegen.

$$SQ_{total} = QS_A + QS_B + QS_{AxB} + QS_r$$

Die Gesamtstreuung berechnet sich wie folgt:

$$SQ_{total} = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} (y_{ijk} - \bar{y})^2. \quad (2.15)$$

Die Streuung der Haupteffekte wird ermittelt, indem das Produkt aus den drei Faktoren Haupteffekt der Stufe i bzw. j , Anzahl der Individuen je Zelle und Anzahl der Stufen des anderen Faktors gebildet wird. Es ergibt sich also ein gewichtetes Produkt der Individuen unter allen Faktorstufenkombinationen, da die Abweichungen der jeweiligen Faktorstufe so multipliziert werden, als ob jeder Messwert dieser Faktorstufe genau die Abweichung des Haupteffektes hervorbringt.

$$SQ_A = n_j \cdot n_k \cdot \sum_{i=1}^{n_i} (\bar{y}_i - \bar{y})^2 \quad (2.16)$$

$$SQ_B = n_i \cdot n_k \cdot \sum_{j=1}^{n_j} (\bar{y}_j - \bar{y})^2, \quad (2.17)$$

die Streuung der Interaktionseffekte:

$$SQ_{A \times B} = n_k \cdot \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\bar{y}_{ij} - \bar{y}_i - \bar{y}_j - \bar{y})^2 = n_k \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\widehat{\alpha\beta})_{ij}^2. \quad (2.18)$$

und die Reststreuung:

$$SQ_r = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} (y_{ijk} - \bar{y}_{ij})^2. \quad (2.19)$$

In der zweifaktoriellen Varianzanalyse gibt es drei Arten von Effekten, so dass drei Hypothesen überprüft werden. Die erste und zweite Hypothese besagen, dass die Faktorstufen der Faktoren A bzw. B sich nicht signifikant unterscheiden, also keine Unterschiede zwischen den verschiedenen Faktorstufen des Faktors vorliegen.

$$H_0^A : \alpha_i = 0 \quad (2.20)$$

$$H_0^B : \beta_i = 0 \quad (2.21)$$

Die dritte Hypothese macht Aussagen zu den Wechselwirkungen, dabei besagt H_0 , dass keine Wechselwirkungen vorliegen.

$$H_0^{A \times B} : (\alpha\beta)_{ij} = 0 \quad (2.22)$$

Die Teststatistiken werden so gebildet, dass die Streuung, die durch die Faktoren bzw. die Wechselwirkungen der Faktoren auftreten in Beziehung zur Reststreuung gesetzt wird. Die Prüfgrößen unterliegen einer F -Verteilung mit den entsprechenden Freiheitsgraden. Die Prüfgrößen, die aus den genannten Streuungen bestehen, die jeweils noch durch die Freiheitsgrade dividiert werden, lauten:

$$F_A = \frac{SQ_A / (n_i - 1)}{SQ_r / n_i n_j (n_k - 1)} \quad (2.23)$$

$$F_B = \frac{SQ_B / (n_j - 1)}{SQ_r / n_i n_j (n_k - 1)} \quad (2.24)$$

$$F_{A \times B} = \frac{SQ_{A \times B} / (n_i - 1)(n_j - 1)}{SQ_r / n_i n_j (n_k - 1)}. \quad (2.25)$$

Zur besseren Übersichtlichkeit lassen sich die jeweiligen Quotienten (Streuung/Freiheitsgrade) noch zu einem mittleren quadratischen Term zusammenfassen, die mit MQ_A , MQ_B ,

MQ_{AxB} und MQ_r bezeichnet werden. So ergibt sich z.B. MQ_A als $MQ_A = SQ_A/(n_i - 1)$. Die Hypothesen in 2.20 und 2.21 werden abgelehnt, wenn:

$$F_A > F_{1-\alpha}[n_i - 1, n_i n_j (n_k - 1)] \quad (2.26)$$

bzw., wenn

$$F_B > F_{1-\alpha}[n - 1, n_i n_j (n_k - 1)]. \quad (2.27)$$

Der Unterschied zwischen den empirischen und theoretischen F-Werten bezüglich der Interaktionen 2.22 ist signifikant, wenn:

$$F_{AxB} > F_{1-\alpha}[(n_i - 1)(n_j - 1), n_i n_j (n_k - 1)]. \quad (2.28)$$

Die Varianzanalyse macht Aussagen darüber, ob zwischen mindestens zwei Faktorstufen ein signifikanter Unterschied besteht. Wenn von Interesse ist, zwischen welchen Faktorstufen signifikante Unterschiede bestehen, müssen weitere Tests für Einzelvergleiche durchgeführt werden, z.B. Scheffé-Test. [8, S. 280]

Dem varianzanalytischen Modell liegen mehrere Modellannahmen zu Grunde, die überprüft werden müssen, wenn die Varianzanalyse eingesetzt werden soll.

- Varianzhomogenität: Die Varianzen der Residuen sollen gleich sein. Überprüfbar ist dies z.B. mit dem Levene-Test.
- Normalverteilung: Die Fehlerkomponenten sind in der Grundgesamtheit normalverteilt. Bei Verletzung dieser Annahme kann statt der Varianzanalyse ein verteilungsfreier Test durchgeführt werden.
- Unabhängigkeit: Die Werte der Stichprobe sind in allen Faktorstufenkombinationen voneinander unabhängig.

Die letzte Voraussetzung, die erfüllt sein muss, um die ANOVA anzuwenden, ist nicht erfüllt, da jede Versuchsperson mehrfach, insgesamt 15-mal, gemessen wurde. Die Werte einer Person sind nicht unabhängig, sondern werden von den persönlichen Voraussetzungen des Menschen bestimmt. Deswegen ist für diese Daten von einer zweifaktoriellen Varianzanalyse Abstand zu nehmen und zur Auswertung der Daten auf die mehrfaktorielle Varianzanalyse für wiederholte Messungen zurückzugreifen.

2.2 Varianzanalyse mit Messwiederholung

Die Varianzanalyse mit Messwiederholung ist eine statistische Auswertemethode für Daten, die untereinander nicht als unabhängig angesehen werden können. Solche abhängigen Daten entstehen z.B. bei Medikamentenstudien, wenn bei dem Teilnehmer vor und nach Medikamentengabe Daten erhoben werden oder wenn in einem Experiment ein Proband mehrmals unter unterschiedlichen Bedingungen befragt wird. Diese Daten können nicht als unabhängig angesehen werden, da die Probanden mehrfach befragt wurden. Auch wenn ein Proband zu jeder Frequenz-Amplituden-Kombination nur einmal befragt wurde, so ist doch zu erwarten, dass jeder Proband dieselbe Kombination unterschiedlich bewerten wird. So wie die Varianzanalyse eine Erweiterung für den t-Test darstellt, so ist die Varianzanalyse mit Messwiederholungen eine Erweiterung für den t-Test für abhängige Stichproben.

Es gibt unterschiedliche Arten von Varianzanalysen mit Messwiederholungen. Wie bereits in Kapitel 2.1 dargestellt werden sie nach Anzahl der unabhängigen Faktoren benannt, d.h. einfaktorielle Varianzanalyse mit Messwiederholung, zweifaktorielle Varianzanalyse mit Messwiederholung usw. Dabei ist zu beachten, dass ein Faktor für die Messwiederholung hinzu kommt. Das bedeutet, dass die einfaktorielle Varianzanalyse nur einen Faktor hat, der die Messwiederholung beinhaltet. Das könnten beispielsweise unterschiedliche Tageszeitpunkte sein (z.B. morgens, mittags, abends). Statt der einfaktoriellen Varianzanalyse mit Messwiederholung lässt sich auch die einfaktorielle Varianzanalyse ohne Messwiederholung durchführen, wenn die Daten des Einzelnen in ipsative Daten umgewandelt werden. Zur Erlangung der ipsativen Daten werden die Daten einer Person am Gesamtmittelwert dieser Person ausgerichtet.

Für die vorliegenden Daten der Intensitätsfrage wäre eine dreifaktorielle Varianzanalyse mit Messwiederholung denkbar. Dabei würden zwei Faktoren, Amplitude und Frequenz, dieselben Faktorstufen haben, wie in Kapitel 2.1 beschrieben, und der dritte Faktor würde die Versuchspersonennummer enthalten, hätte also 13 Faktorstufen. Die beiden Faktoren Frequenz und Amplitude werden als Gruppierungsfaktoren bezeichnet und der Effekt, der durch die Probanden entsteht, als Messwiederholungsfaktor. Es ist zu beachten, dass zwei Gruppierungsfaktoren vorliegen und nur ein Messwiederholungsfaktor, da es ebenfalls eine dreifaktorielle Varianzanalyse mit Messwiederholung auch die Möglichkeit gibt, dass zwei Messwiederholungsfaktoren und ein Gruppierungsfaktor vorliegen. Die Berechnung beider läuft verschieden ab.

Jedoch gehen alle diese Varianzanalysen mit Messwiederholungen davon aus, dass

zu jeder Faktorstufenkombination bzw. zu jedem Faktor eine Stichprobe vorliegt. Dies ist bei den gegebenen Daten nicht der Fall. Es handelt sich um Daten, die eine einmal gezogene Stichprobe darstellen, die dann unter unterschiedlichen Bedingungen getestet wurde. Für einen solchen Fall, wo nur eine Stichprobe unter allen Faktorstufenkombinationen getestet wurde, ist die Varianzanalyse mit kompletter Messwiederholung anzuwenden.

Die Bezeichnung der Faktoren ist in dieser Analyse analog zu Kapitel 2.1. Auch für diesen Fall der Varianzanalyse lassen sich die Daten zur besseren Übersichtlichkeit in einer Tabelle, abgebildet in Tabelle 2.2, darstellen.

Tabelle 2.3: Datenschema komplette Messwiederholung (Auszug)

	Frequenz 1			Frequenz 2			...
Versuchsperson	Amplitude 1	...	Amplitude 3	Amplitude 1	...	Amplitude 3	...
1	x_{111}	...	x_{131}	x_{211}	...	x_{231}	...
2	x_{112}	...	x_{132}	x_{212}	...	x_{232}	...
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

In dieser Analyse wird die Gesamtstreuung ebenfalls zerlegt. Die Gesamtstreuung ist die Summe aus der Streuung zwischen den Versuchspersonen und der Streuung innerhalb der Versuchspersonen. Im weiteren werden die Gesamtstreuung mit SQ_{gesamt} , die Streuung zwischen den Versuchspersonen als SQ_{zwVP} und die Streuung innerhalb der Versuchspersonen mit SQ_{inVP} bezeichnet.

$$SQ_{gesamt} = SQ_{zwVP} + SQ_{inVP} \quad (2.29)$$

Hierbei ist die Streuung innerhalb der Versuchspersonen gegeben als

$$SQ_{inVP} = SQ_A + SQ_B + SQ_{AxB} + SQ_{AxVP} + SQ_{BxVP} + SQ_{AxBxVP}. \quad (2.30)$$

Die Streuungen aus Gleichung 2.30 berechnen sich wie folgt, wobei aus Platzgründen nur die Ergebnisse angegeben sind:

$$SQ_A = \frac{\sum_{i=1}^{n_i} y_i^2}{n_i \cdot n_k} - \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_i \cdot n_j \cdot n_k} = 2.87 \quad (2.31)$$

$$SQ_B = \frac{\sum_{j=1}^{n_j} y_j^2}{n_i \cdot n_k} - \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_i \cdot n_j \cdot n_k} = 2.86 \quad (2.32)$$

$$SQ_{AxB}^2 = \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} y_{ij}^2}{n_k} - \frac{\sum_{j=1}^{n_j} y_j^2}{n_i \cdot n_k} - \frac{\sum_{i=1}^{n_i} y_i^2}{n_j \cdot n_k} - \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_i \cdot n_j \cdot n_k} = -5.68 \quad (2.33)$$

$$SQ_{AxVp}^2 = \frac{\sum_{i=1}^{n_i} \sum_{k=1}^{n_k} y_{ik}^2}{n_j} - \frac{\sum_{i=1}^{n_i} y_i^2}{n_j \cdot n_k} - \frac{\sum_{k=1}^{n_k} y_k^2}{n_j \cdot n_i} + \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_i \cdot n_j \cdot n_k} = -5.35 \quad (2.34)$$

$$SQ_{BxVp}^2 = \frac{\sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{jk}^2}{n_i} - \frac{\sum_{j=1}^{n_j} y_j^2}{n_i \cdot n_k} - \frac{\sum_{k=1}^{n_k} y_k^2}{n_i \cdot n_j} + \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_i \cdot n_j \cdot n_k} = -5.48 \quad (2.35)$$

$$\begin{aligned} SQ_{AxBxVp} &= \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}^2}{n_k} - \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} y_{ij}^2}{n_k} - \frac{\sum_{i=1}^{n_i} \sum_{k=1}^{n_k} y_{ik}^2}{n_j} - \frac{\sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{jk}^2}{n_i} \\ &+ \frac{\sum_{i=1}^{n_i} y_i^2}{n_j \cdot n_j} + \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} y_{ij}^2}{n_k} - \frac{\sum_{i=1}^{n_i} \sum_{k=1}^{n_k} y_{ik}^2}{n_j} - \frac{\sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{jk}^2}{n_i} + \frac{\sum_{i=1}^{n_i} y_i^2}{n_j \cdot n_j} \\ &+ \frac{\sum_{j=1}^{n_j} y_j^2}{n_i \cdot n_k} + \frac{\sum_{k=1}^{n_k} y_k^2}{n_i \cdot n_j} - \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_j \cdot n_j \cdot n_k} = 8.35 \end{aligned} \quad (2.36)$$

$$SQ_{inVP} = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} x_{ijk}^2 - \frac{\sum_{k=1}^{n_k} y_k^2}{n_i \cdot n_j} \quad (2.37)$$

$$SQ_{zwVP} = \frac{\sum_{k=1}^{n_k} y_k^2}{n_i \cdot n_j} - \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_j \cdot n_i \cdot n_k} \quad (2.38)$$

$$SQ_{total} = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} x_{ijk}^2 - \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} y_{ijk}}{n_j \cdot n_i \cdot n_k} \quad (2.39)$$

Wie bereits im vorherigen Kapitel werden die eben berechneten Streuungen noch

durch die Anzahl der Freiheitsgrade df dividiert, bevor aus ihnen die Teststatistiken gebildet werden.

$$\sigma_A = \frac{SQ_A}{df_A}, \sigma_B = \frac{SQ_B}{df_B}, \sigma_{AxB} = \frac{SQ_{AxB}}{df_A \cdot df_B}$$

$$\sigma_{AxVP} = \frac{SQ_{AxVP}}{df_A \cdot df_{VP}}, \sigma_{BxVP} = \frac{SQ_{BxVP}}{df_B \cdot df_{VP}}, \sigma_{AxBxVP} = \frac{SQ_{AxBxVP}}{df_A \cdot df_B \cdot df_{VP}} \quad (2.40)$$

Die Teststatistiken lauten:

$$F_A = \frac{\sigma_A}{\sigma_{AxVP}} \quad (2.41)$$

$$F_B = \frac{\sigma_B}{\sigma_{BxVP}} \quad (2.42)$$

$$F_{AxB} = \frac{\sigma_{AxB}}{\sigma_{AxBxVP}}. \quad (2.43)$$

Ebenfalls wie im vorherigen Kapitel werden 3 Hypothesen überprüft, nämlich ob Faktor A und Faktor B einzeln bzw. die Interaktion der beiden sich signifikant auf die abhängige Variable auswirken. Dazu werden die berechneten empirischen Werte, die nach den Formeln 2.41 bis 2.43 berechnet wurden, mit den theoretischen F-Werten verglichen [9, S.268-272]. Die theoretischen Werte der Fisher-Verteilung sind aus Tabellen ablesbar. Man kann sie sich aber auch mittels der Funktion `finv(1- α , Freiheitsgrade Zähler, Freiheitsgrade Nenner)` in Matlab ausgeben lassen. Sind die empirischen Werte größer als die theoretischen, wird die Nullhypothese verworfen.

Tabelle 2.4: Übersicht empirische und theoretische F-Werte für Intensitätsfrage-Daten

	empirischer F-Wert	theoretischer F-Wert
Faktor A	6.899	2.5652
Faktor B	80.082	3.4028
Interaktion	1.255	2.0363

An dieser Tabelle lässt sich sehr gut erkennen, dass die empirischen Werte der Faktoren A und B größer sind, als die theoretischen, d.h., die Nullhypothese wird abgelehnt und die beiden Faktoren haben einzeln Einfluss auf die Linienlänge. Der Effekt der durch die Interaktion von Faktor A (Frequenz) und Faktor B (Amplitude) ausgeübt wird, ist nicht signifikant.

Die Voraussetzungen für die Varianzanalyse mit Messwiederholungen sind, bis

auf die Voraussetzung der Unabhängigkeit, dieselben wie die für die Varianzanalyse ohne Messwiederholungen (Vgl. Kapitel 2.1). Die Modellannahme der Unabhängigkeit wird durch eine andere Modellannahme ersetzt, die der Sphärizität. Diese wird auch Zirkularitätsannahme genannt. Wenn **Sphärizität** vorliegt, dann sollen alle Varianzen der Differenzen aller möglichen Gruppen-Paare gleich sein. Getestet werden kann diese Annahme mit dem **Mauchly-Test**. Dabei besagt die Nullhypothese des Testes, dass Sphärizität angenommen wird. Ist die Annahme verletzt, kann die Varianzanalyse für messwiederholte Daten trotzdem durchgeführt werden. Jedoch muss eine Berichtigung der F-Werte erfolgen, da sonst bei Verletzung dieser Annahme die Varianzanalyse falsch signifikante Ergebnisse liefern könnte. Die Berichtigung erfolgt, indem eine Korrektur der Freiheitsgrade bei der Berechnung der Teststatistiken bei der Varianzanalyse mit wiederholten Messungen vorgenommen wird. Dabei werden die Freiheitsgrade des Zählers und Nenners mit einem Wert ϵ multipliziert. Dieses Vorgehen ist legitim, da der empirische F-Wert auch dann einer F-Verteilung folgt, wenn Sphärizität nicht angenommen werden kann. Ist die Voraussetzung jedoch verletzt, folgt der empirische F-Wert der F-Verteilung mit anderen Freiheitsgraden. Darum müssen diese nach der beschriebenen Methode angepasst werden.

Es gibt verschiedene Varianten, den Wert ϵ , mit dem der Zähler und Nenner des F-Bruches multipliziert werden, zu bestimmen. Eine der häufigsten angewandten Korrekturen wird Greenhouse-Geiser-Korrektur genannt. Bei dieser Methode zur Bestimmung des ϵ -Wertes wird ϵ nur in Abhängigkeit von der Anzahl der messwiederholten Stufen p ermittelt. ϵ ergibt sich dann als $\epsilon = 1 \cdot (p - 1)^{-1}$ [10]. Der Wert ϵ ist umso geringer, umso stärker die Sphärizitätsannahme verletzt ist. [8]

Bei der Anwendung des Mauchly-Testes ist zu beachten, dass es sich bei dem Test um einen statistischen Test mit geringer Teststärke handelt. Insbesondere bei kleinen Stichproben kommt es häufig zu der Testaussage, dass Sphärizität vorliegt, obwohl dies nicht gegeben ist. Außerdem ist dieser Test anfällig für Verletzungen der Normalverteilungsannahme. Liegen begründete Zweifel für die Sphärizitätsannahme vor, sollten auch dann die Korrekturen angewendet werden, wenn der Test ergeben hat, dass Sphärizität angenommen werden kann. [10]

2.3 Vergleich beider Analysemethoden

Bei der Durchführung der Varianzanalyse ohne Messwiederholung (Kapitel 2.1) ergeben sich für die Daten aus der Intensitätsfrage die folgenden empirischen und theoretischen F-Werte:

Tabelle 2.5: F-Werte im Vergleich für Intensitätsdaten für die Varianzanalyse ohne Messwiederholung und die Varianzanalyse mit Messwiederholung

	Varianzanalyse ohne Messwiederholung			Varianzanalyse mit Messwiederholung		
	empirischer F-Wert	theoretischer F-Wert	Signifikanz	empirischer F-Wert	theoretischer F-Wert	Signifikanz
Frequenz	5	2.441	signifikant	6.899	2.5652	signifikant
Amplitude	0.64	1.8723	nicht signifikant	80.082	3.4028	signifikant
Interaktion	0.72	1.478	nicht signifikant	1.255	2.0363	nicht signifikant

Wie im vorherigen Kapitel ist der Einfluss eines Faktors signifikant, wenn der empirische F-Wert größer ist als der theoretische. Es ist ersichtlich, dass bei der zweifaktoriellen Varianzanalyse die Frequenz einen signifikanten Einfluss auf das Intensitätsempfinden der Versuchspersonen hat. Das Ergebnis der Varianzanalyse mit kompletter Messwiederholung zeigte, dass die beiden Faktoren einzeln einen signifikanten Einfluss haben, die Wechselwirkung jedoch nicht zu einer signifikanten Wechselwirkung der Linienlänge führt. Die beiden Analysen liefern also unterschiedliche Ergebnisse.

Die Ergebnisse der zweifaktoriellen Varianzanalyse ohne Messwiederholung wurden an dieser Stelle dargestellt, um zu zeigen, dass unterschiedliche Ergebnisse möglich sind und deshalb immer die Voraussetzungen der Varianzanalyse überprüft werden sollten, bevor die Auswahl eines Modells getroffen wird. Da eben diese Voraussetzungen für die vorliegenden Daten im Hinblick auf die Unabhängigkeit nicht gegeben waren, wurde die Varianzanalyse für komplette Messwiederholungen als Auswertemethode gewählt. Dabei wurde ersichtlich, dass sowohl die Frequenz als auch die Amplitude Einfluss auf das Intensitätsempfinden der Versuchsteilnehmer hatten. Für die Interaktion der beiden Faktoren ergab sich jedoch kein signifikantes Testresultat.

3 Lineares gemischtes Modell

Lineare Modelle stellen eine Variante von Regressionsmodellen dar. Varianzanalytische Verfahren wiederum sind Spezialfälle des linearen Modells.

Das lineare gemischte Modell ist eine Erweiterung des linearen Modells $Y = X\beta$, wobei Y die abhängige Variable, X die unabhängige Variable und β ein Vektor mit Regressionsparametern ist. Im linearen Modell werden nur feste Effekte betrachtet, d.h., es werden nur solche Effekte untersucht, deren einzelne Stufen vorher bereits bekannt sind. Eine solche Untersuchung ist sinnvoll bei Faktoren deren Stufen alle bekannt sind bzw. deren interessierende Stufen bekannt sind. Bei einer Medikamentenstudie, wo der Einfluss eines Medikaments untersucht werden soll, ist es sinnvoll, den Faktor Medikament als fest anzusehen, da alle Ausprägungen des Faktors bekannt sind, also z.B. Placebo, Medikament A und Medikament B. Anders sähe es jedoch aus, wenn der Probandeneffekt als fester Faktor betrachtet wird. Die Probanden stellen eine zufällig ausgewählte Menge aus einer Grundgesamtheit dar. Würde der Einfluss der Probanden als festen Effekt betrachtet werden, wären die Ergebnisse dieser Studie nicht auf die Grundgesamtheit übertragbar. Deswegen wurde das lineare Modell mit festen Effekten um zufällige Effekte erweitert. Gemischte Modelle ermöglichen es, auch Korrelationen, die durch wiederholte Messungen am selben Subjekt entstehen, einzuschließen.

Es existieren also drei Arten von linearen Modellen. Lineare Modelle, die nur feste Effekte, die nur zufällige Effekte und die gemischte Effekte, d.h. feste und zufällige Effekte werden in einem Modell betrachtet, beinhalten.

Im folgenden soll nur das lineare gemischte Modell betrachtet werden. [11, 12, 13, 14]

3.1 Bildung des linearen gemischten Modells

Feste Effekte sind Faktoren, deren Stufen von Faktoren den interessierenden Variationsbereich der unabhängigen Variable vollständig ausschöpfen.

Zufällige Effekte sind Effekte deren Stufen der Effekte nur eine zufällige Auswahl, also eine Zufallsstichprobe, des interessierenden Bereiches darstellen. [15]

Die abhängigen Faktoren Frequenz und Amplitude werden als feste Faktoren betrachtet, da sie den interessierenden Variationsbereich vollständig ausschöpfen. Der interessierende Bereich ist der, der durch den Versuchsleiter festgelegt wurde. Der messwiederholte Faktor Proband wird als zufälliger Effekt betrachtet, da die ausgewählten Probanden nur eine Zufallsstichprobe aus allen möglichen Probanden darstellen. Die Linienlänge stellt auch hier wieder die abhängige Variable dar. Es wird angenommen, dass die Beziehung zwischen der abhängigen Variable (Linienlänge) und den unabhängigen Variablen (Frequenz, Amplitude und Proband) linear ist. Das Modell des linearen gemischten Modells lautet allgemein:

$$Y = X\beta + Zb + \epsilon \quad (3.1)$$

Y : Vektor der die Werte der abhängigen Variable enthält

X : Designmatrix für die festen Effekte

β : Vektor der den globalen Mittelwert und alle Parameter der festen Effekte enthält

Z : Designmatrix für die zufälligen Effekte

b : Vektor der alle zufälligen Effektvariablen enthält

ϵ : Vektor der alle zufälligen Fehler enthält.

Der Erwartungswert des Vektors der Effektvariablen sei 0, $E(b) = 0$, ebenso wie der Erwartungswert des Vektors der zufälligen Fehler $E(\epsilon) = 0$, d.h., das gebildetete Modell soll die abhängige Variable, möglichst vollständig, erklären. Der Erwartungswert der beobachteten Werte ist das Produkt aus dem Vektor β und der Designmatrix X . Außerdem sind die Varianzen von b und ϵ definiert als

$$\text{var} \begin{bmatrix} Z \\ \epsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}$$

und die Varianz von Y ist definiert als $\text{var}(y) = ZGZ' + R$. Es gelten die Verteilungsannahmen $b_i \sim N(0, D)$ [$i \in \mathbb{N}$] und $\epsilon_i \sim N(0, \Sigma_i)$ [$i \in \mathbb{N}$]. Bei D handelt es sich um eine $q \times q$ Kovarianzmatrix. Σ_i ist eine $n_i \times n_i$ Kovarianzmatrix. Die Vektoren b und ϵ seien voneinander unabhängig. [12, 13]

Designmatrizen sind Matrizen, die alle Werte der unabhängigen Variablen speichern.

Es gibt unterschiedliche Verfahren um Designmatrizen zu erhalten, wie z.B. die Dummycodierung. Eine **dummycodierte Designmatrix** besteht nur aus Nullen und Einsen, wobei die Einsen bedeuten, dass die betrachtete Ausprägung vorliegt (weitere Möglichkeiten zur Codierung siehe [8]).

Bei einem linearen Modell wird von einer

1. Varianzanalyse gesprochen, wenn die Designmatrix der festen Effekte nur aus Nullen und Einsen besteht.
2. Kovarianzanalyse gesprochen, wenn mindestens eine Spalte der Designmatrix X aus Nullen und Einsen besteht und 1. nicht zutrifft.
3. Regressionanalyse gesprochen, wenn weder 1. noch 2. gilt. [16]

Die Designmatrizen für die Vibrationsdaten bestehen nur aus Nullen und Einsen und codieren so die Zugehörigkeit zu der jeweiligen Amplitude bzw. Frequenz.

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 & \dots \\ 1 & 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & 1 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Das heißt, in dem vorliegenden Fall wird das lineare Modell für die Varianzanalyse angewendet.

Zur Auswertung der Vibrationsdaten wird ein Modell benötigt, das zufällige und feste Effekte enthält. Zur Einbeziehung der zufälligen Fehler wird ein Vektor benötigt, der die geschätzten festen Parameter enthält und eine Designmatrix, die die jeweiligen abhängigen Werte den passenden Frequenz und Amplitudenkombinationen zuordnet. Die durch zufälligen Effekte auftretenden Effekte werden in einem Parametervektor gespeichert. Die Formel (3.1) zeigt das lineare gemischte Modell in allgemeiner Form. Für die Vibrationsdaten lautet das Modell:

$$Y_{ij} = a + \alpha_i + \sum_{k=1}^4 b_{Frequ,k} \cdot Ind_{Frequ,k} + \sum_{k=1}^2 b_{Amp,k} \cdot Ind_{Amp,k} + \epsilon_{ij} \quad (3.2)$$

- Y_{ij} : abhängige Variable
 i : indizierte Versuchspersonen
 j : indizierte Beobachtungen innerhalb der Versuchsperson
 α_i : zufällige Abweichung vom festen Mittelwert a
 ϵ_{ij} : Messfehler
 $b_{Frequ,k}$: vier Parameter zur Modellierung der 5 Frequenzstufen
 $Ind_{Frequ,k}$: zugehörige Indikatorvariablen
 $b_{Amp,k}$: zwei Parameter zur Modellierung der 2 Amplituden
 $Ind_{Amp,k}$: zugehörige Indikatorvariablen.

3.2 Berechnung der Varianzen im gemischten Modell

Es gibt unterschiedliche Methoden mittels eines gemischten linearen Modells die Varianzkomponenten, die anschließend getestet werden können, zu bestimmen. Eine Variante ist, bereits fertige Formeln, die aus dem Modell für die Varianzanalyse abgeleitet wurden, zur Berechnung zu benutzen [17, 8]. Bevor geeignete Formeln für die Analyse ausgewählt werden können, ist es notwendig, sich über das zu Grunde liegende Modell Klarheit zu verschaffen. Das bedeutet insbesondere, dass vorher festgelegt werden muss, wie viele Faktoren betrachtet werden und welche dieser Faktoren als zufällig bzw. als fest betrachtet werden sollen. Denn durch die Anwendung des linearen gemischten Modells ändert sich zwar nicht die Berechnung der Quadratsummen, aber die Berechnung der Testgrößen. Die Testgrößen werden je nach Art des Faktors, zufällig oder fest, unterschiedlich berechnet, d.h., soll der Einfluss des festen Faktors Frequenz getestet werden, so berechnet sich die Testgröße, um den Einfluss zu testen, anders als würde man den Einfluss der Probanden testen.

Wie bereits in Kapitel 2.2 erwähnt, werden 3 Faktoren betrachtet. Dabei werden die Faktoren Frequenz und Amplitude als feste Effekte betrachtet und der Faktor Versuchsperson als zufälliger Faktor.

Jeder der Faktoren Frequenz, Amplitude und Versuchsperson stellt einen Haupteffekt dar. Es gibt drei Haupteffekthypothesen. Von denen beziehen sich zwei auf feste Effekte und eine auf den zufälligen Effekt Versuchsperson. Die Hypothesen der beiden festen Haupteffekte lauten, wie aus Kapitel 2 bekannt, dass zwischen den

einzelnen Faktorstufen der jeweiligen Faktoren keine Unterschiede bestehen

$$H_0^{(1)} = \bar{\mu}_{\text{Frequenz1}} = \dots = \bar{\mu}_{\text{Frequenz5}} \quad (3.3)$$

$$H_0^{(2)} = \bar{\mu}_{\text{Amplitude1}} = \dots = \bar{\mu}_{\text{Amplitude3}}. \quad (3.4)$$

Die Hypothese für den zufälligen Effekt lautet, dass sich die Varianz, die durch die Versuchspersonen verursacht wird, im Mittel Null ist

$$H_0^{(3)} : \sigma_{\text{Versuchsperson}}^2 = 0. \quad (3.5)$$

Es können vier unterschiedliche Interaktionen auftreten. Interaktionen werden nach Anzahl der beteiligten Faktoren unterschieden z.B. beziehen Interaktionen erster Ordnung zwei Faktoren ein, Interaktionen zweiter Ordnung beziehen drei Faktoren ein usw. Bei drei der betrachteten Interaktionen handelt es sich um Interaktionen erster Ordnung. Es sind die Interaktionen zwischen Frequenz und Amplitude, Frequenz und Versuchsperson sowie die Wechselbeziehung zwischen Amplitude und Versuchsperson. Die vierte Interaktion beschreibt die wechselseitige Wirkung zwischen Amplitude, Frequenz und Versuchsperson. Alle Interaktionen, die die Wechselbeziehung zu einem zufälligen Faktor darstellen, werden ebenfalls als zufällige Effekte betrachtet, so dass nur solche Interaktionen als feste Effekte behandelt werden, deren Faktoren alle fest sind. Die vier zu überprüfenden Hypothesen für die Interaktionen lauten

$$H_0^{(4)} : \mu_{\text{Amplitude} \times \text{Frequenz}} - \mu_{\text{Amplitude}} - \mu_{\text{Frequenz}} + \mu = 0$$

$$H_0^{(5)} : \sigma_{AC}^2 = 0$$

$$H_0^{(6)} : \sigma_{BC}^2 = 0$$

$$H_0^{(7)} : \sigma_{ABC}^2 = 0.$$

Zur Berechnung der Quadratsummen und F-Summen sollen nun bereits abgeleitete Formeln verwendet werden [17, 8], wobei in dem Werk von [17] der Begriff der Kreuzklassifikation auftritt. **Kreuzklassifikation** ist eine Klassifizierungsmöglichkeit, bei der die betrachtete Menge nach unterschiedlichen Kriterien unterteilt wird. Werden die unterschiedlichen Kriterien zu einer Klassifikation kombiniert, so ergibt sich eine vollständige Kreuzklassifikation [18]. Aus [17] wären demzufolge die Formeln für die dreifache Kreuzklassifikation zu wählen (S. 535 ff.).

In [8] sind die Formeln unter dem Kapitel der dreifaktoriellen Varianzanalyse im Abschnitt Gemischtes Modell zu finden (S. 269 ff.).

Die Quadratsummen, die für jeden Effekt, für jede der vier möglichen Interaktionen und für den Fehleranteil berechnet werden müssen, berechnen sich wie für eine dreifaktorielle Varianzanalyse (Kapitel 2.2 Formel 2.31-2.39). Dabei wird im folgenden dem Faktor A (Frequenz) der Index *Frequ* zugeordnet, dem Faktor B (Amplitude) der Index *Ampl* und dem zufälligen Faktor Versuchsperson der Index *VP*. Ansonsten werden die Bezeichnungen *A*, *B*, *C* für die Faktoren Frequenz, Amplitude und Versuchsperson beibehalten, um die Formeln übersichtlich zu halten. Die Variablen *p*, *q*, *r* geben die Anzahl der Faktorstufen für die Faktoren Frequenz, Amplitude und Versuchsperson an, so dass $p = 5$, $q = 3$ und $r = 13$ sind. Die Anzahl aller Beobachtungen wird mit $n = p \cdot q \cdot r = 3 \cdot 5 \cdot 13 = 195$ angegeben. Die Summe aller gegebenen Antworten, also die Summe aller Werte für die abhängige Variable wird mit *G* bezeichnet. Die Formeln und die berechneten Quadratsummen sind anschließend in Tabelle 3.1 dargestellt.

Tabelle 3.1: Berechnung der Quadratsummen im linearen gemischten Modell

Faktor	Formel zur Berechnung der Quadratsummen	berechnete Quadratsummen
SQ_{Frequ}	$\frac{\sum_{i=1}^p A_i}{n \cdot q \cdot r} - \frac{G^2}{n \cdot p \cdot q \cdot r}$	8.7006e+005
SQ_{Ampl}	$\frac{\sum_{j=1}^q B_j}{n \cdot p \cdot r} - \frac{G^2}{n \cdot p \cdot q \cdot r}$	3.5315e+006
SQ_{VP}	$\frac{\sum_{k=1}^r C_k}{n \cdot q \cdot r} - \frac{G^2}{n \cdot p \cdot q \cdot r}$	3.3390e+005
$SQ_{Frequ \times Ampl}$	$\frac{\sum_{i=1}^p \sum_{j=1}^q AB_{ij}^2}{n \cdot r} - \frac{\sum_{i=1}^p A_i}{n \cdot q \cdot r} - \frac{\sum_{j=1}^q B_j}{n \cdot p \cdot r} - \frac{G^2}{n \cdot p \cdot q \cdot r}$	1.5122e+005
$SQ_{Frequ \times VP}$	$\frac{\sum_{i=1}^p \sum_{k=1}^r AC_{ik}^2}{n \cdot q} - \frac{\sum_{i=1}^p A_i}{n \cdot q \cdot r} - \frac{\sum_{k=1}^r C_k}{n \cdot p \cdot r} - \frac{G^2}{n \cdot p \cdot q \cdot r}$	1.5134e+006
$SQ_{Ampl \times VP}$	$\frac{\sum_{j=1}^q \sum_{k=1}^r BC_{jk}^2}{n \cdot p} - \frac{\sum_{j=1}^q B_j}{n \cdot p \cdot r} - \frac{\sum_{k=1}^r C_k}{n \cdot p \cdot r} - \frac{G^2}{n \cdot p \cdot q \cdot r}$	5.2918e+005
$SQ_{Frequ \times Ampl \times VP}$	$\frac{\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^r ABC_{ijk}^2}{n} - \frac{\sum_{i=1}^p \sum_{j=1}^q AB_{ij}^2}{n \cdot r} - \frac{\sum_{i=1}^p \sum_{k=1}^r AC_{ik}^2}{n \cdot q} - \frac{\sum_{j=1}^q \sum_{k=1}^r BC_{jk}^2}{n \cdot p} + \frac{\sum_{i=1}^p A_i}{n \cdot q \cdot r} + \frac{\sum_{j=1}^q B_j}{n \cdot p \cdot r} + \frac{\sum_{k=1}^r C_k}{n \cdot q \cdot r} - \frac{G^2}{n \cdot p \cdot q \cdot r}$	1.4463e+006
SQ_{Fehler}	$\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^r \sum_{m=1}^n y_{ijkm}^2 - \frac{\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^r ABC_{ijk}^2}{n}$	0
SQ_{Total}	$\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^r \sum_{m=1}^n y_{ijkm}^2 - \frac{G^2}{n \cdot p \cdot q \cdot r}$	8.3755e+006

Um zu prüfen, ob die jeweiligen Abweichungen signifikant sind, müssen jetzt, wie schon im Fall der Varianzanalyse, die mittleren Quadrate gebildet werden. Diese ergeben sich als Quotient aus den Quadratsummen und der Anzahl der entsprechenden Freiheitsgrade. Die Ergebnisse für die Intensitätsfrage können Tabelle 3.2 entnommen werden.

Tabelle 3.2: Berechnung der mittleren Quadrate im linearen gemischten Modell

Faktor	Anzahl der Freiheitsgrade	Ergebnisse mittlere Quadrate
MQ_{Frequ}	$p - 1$	2.1751e+005
MQ	$q - 1$	1.7658e+006
MQ_{VP}	$r - 1$	2.7825e+004
$MQ_{Frequ \times Ampl}$	$(p - 1)(q - 1)$	1.8902e+004
$MQ_{Frequ \times VP}$	$(p - 1)(r - 1)$	3.1528e+004
$MQ_{Ampl \times VP}$	$(q - 1)(r - 1)$	2.2049e+004
$MQ_{Frequ \times Ampl \times VP}$	$(p - 1)(q - 1)(r - 1)$	1.5065e+004
MQ_{Fehler}	$pqr(n - 1)$	nicht definiert
MQ_{Total}	$pqrn - 1$	4.3173e+004

Der Wert für das mittlere Quadrat von MQ_{Fehler} kann nicht berechnet werden, da die Anzahl der Freiheitsgrade $n - 1 = 1 - 1 = 0$ wird. Um das mittlere Quadrat zu erhalten, wird die Quadratsumme durch die Anzahl der Freiheitsgrade dividiert. Da die Division durch 0 jedoch nicht definiert ist, kann das mittlere Quadrat für diese Varianzkomponente nicht berechnet werden.

Anschließend können die Hypothesen überprüft werden, wozu F-Brüche aus den empirischen Daten gebildet werden und mit den entsprechenden Quantilen der F-Verteilung verglichen werden können. Sind die empirischen F-Werte kleiner als die theoretischen F-Werte, so wird die Nullhypothese angenommen.

Die Berechnung der empirischen F-Werte der festen Effekte erfolgt dabei über folgenden Formeln:

$$F_A = MQ_A/MQ_{AC} = 6.90 \quad (3.6)$$

$$F_B = MQ_B/MQ_{BC} = 80.08 \quad (3.7)$$

$$F_{AB} = MQ_{AB}/MQ_{ABC} = 1.25. \quad (3.8)$$

Die entsprechenden theoretischen Werte können aus einer geeigneten Tabelle, die die Werte der F-Verteilung enthält, abgelesen werden oder sind in MATLAB mit

der Funktion `finv(1-alpha, Anzahl der Freiheitsgrade Zähler, Anzahl der Freiheitsgrade Nenner)` abrufbar. Zur besseren Unterscheidbarkeit werden die theoretischen F-Werte nicht mit F, sondern mit tF bezeichnet. Die F-Werte für die eben berechneten empirischen F-Werte lauten:

$$tF_A(0.95, 4, 48) = 2.5652$$

$$tF_B(0.95, 2, 24) = 3.4028$$

$$tF_{AB}(0.95, 96, 195) = 2.0363.$$

Beim Vergleich dieser theoretischen Werte mit den Ergebnissen der Formeln 3.6 bis 3.8, ist ersichtlich, dass die theoretischen Werte der Faktoren A und B unter denen der empirischen für diese Faktoren liegen. D.h., dass für diese beiden Faktoren Frequenz und Amplitude die Hypothese, dass sich die einzelnen Faktorstufen der Faktoren nicht signifikant unterscheiden, abgelehnt werden muss.

Die Wechselwirkungsbeziehung zwischen den beiden festen Faktoren Frequenz und Amplitude ist jedoch nicht signifikant. Dieses Ergebniss entspricht dem der Abbildung 2.1, wonach keine Wechselwirkungen zwischen den Faktoren Amplitude und Frequenz vorliegen.

Anschließend werden die Hypothesen der zufälligen Effekte geprüft. Die Vorgehensweise ist analog der Vorgehensweise bei den festen Effekten, d.h. es werden wieder die F-Werte aus den schon ermittelten mittleren Quadraten gebildet und diese anschließend mit den theoretischen F-Werten verglichen. Die Berechnung der F-Werte aus den mittleren Quadraten erfolgt nach den Formeln [8, S. 270]

$$F_{AC} = MQ_{AC}/MQ_{ABC} = 2.0927 \quad (3.9)$$

$$F_{BC} = MQ_{BC}/MQ_{ABC} = 1.4636 \quad (3.10)$$

$$F_{ABC} = MQ_{ABC}/MQ_{Fehler} = NaN \quad (3.11)$$

Der letzte empirische F-Wert aus Formel 3.11 kann nicht berechnet werden, da kein Wert für den Ausdruck MQ_{Fehler} existiert. Die empirischen Werte können wieder aus entsprechenden Tabellen ermittelt werden. Die von Matlab für tF_{AC} und tF_{BC} ausgegebenen Werte lauten

$$tF_{AC}(0.95, 48, 96) = 1.4889$$

$$tF_{BC}(0.95, 24, 96) = 1.6542.$$

Der zufällige Effekt F_{AC} ist signifikant, da der berechnete Wert größer ist als der abgelesene. Die Überprüfung des Effektes F_{BC} liefert ein nicht signifikantes Ergebnis. Über den Faktor F_{ABC} kann keine Aussage getroffen. Da der Faktor jedoch eine Wechselwirkung beschreibt und bereits in Kapitel 2 grafisch gezeigt wurde [Abbildung 2.1 auf Seite 11], dass die auftretenden Wechselwirkungen als vernachlässigbar anzusehen sind, wäre ein nicht signifikantes Ergebnis zu erwarten.

Zu Beginn dieses Kapitels wurden alle zu überprüfenden Hypothesen genannt. Von den sieben zu testenden Hypothesen wurden sechs bereits mit der Variante, die schon von der Varianzanalyse bekannt ist, überprüft. Die siebente Hypothese $H_0 : \sigma_c^2 = 0$ muss anders überprüft werden, da keines der bisher berechneten mittleren Quadrate dem Erwartungswert, der im Nenner des F-Bruches zur Überprüfung der Hypothese stehen müsste, entspricht. D.h., es kann nicht direkt ein empirischer F-Wert zur Überprüfung der letzten Hypothese gebildet werden. Die Lösung besteht darin, ein mittleres Quadrat zu bilden, das dem eigentlichen Quadrat entspricht. Wird eine Hypothese nicht anhand des eigentlichen Erwartungswertes geprüft, sondern durch eine Kombination verschiedener mittlerer Quadrate, die als Summe den erwarteten Wert ergeben, so wird von einem Quasi-F-Bruch gesprochen. Quasi-F-Brüche stellen eine Möglichkeit dar, die Hypothesen zumindest approximativ zu überprüfen. Approximativ deshalb, weil die Quasi-F-Brüche, nachdem die Freiheitsgrade korrigiert wurden, approximativ F-verteilt sind. **Quasi-F-Brüche** sind also Teststatistiken, die im Nenner oder Zähler statt des direkten Erwartungswertes des zu überprüfenden Erwartungswertes einen Ausdruck enthalten, der den Wert des zu überprüfenden F-Wertes entspricht. Der Quasi-F-Bruch für die zu überprüfende Hypothese lautet

$$F_C = \frac{MQ_C}{MQ_{AC} - MQ_{BC} - MQ_{ABC}} = 0.7225. \quad (3.12)$$

Wie erwähnt müssen die Freiheitsgrade noch korrigiert werden, anschließend kann der theoretische Wert der Student-Verteilung mit dem in Formel 3.12 verglichen werden. Von der Korrektur ist nur der Nennerfreiheitsgrad betroffen, da nur für diesen kein geeignetes mittleres Quadrat zur Verfügung stand und er so ersetzt werden musste, um die Hypothese dennoch überprüfen zu können. Die Korrektur, die für den Nennerfreiheitsgrad notwendig ist, berechnet sich aus der abgerundeten Zahl

aus dem Quotienten von

$$df_{Nenner} = \frac{(MQ_{AC} + MQ_{BC} - MQ_{ABC})^2}{\frac{MQ_{AC}^2}{df_{AC}} + \frac{MQ_{BC}^2}{df_{BC}} + \frac{MQ_{ABC}^2}{df_{ABC}}} = 22.94 \approx 22. \quad (3.13)$$

Die so berechneten Freiheitsgrade werden abgerundet, um die Einhaltung des Signifikanzniveaus α zu gewährleisten. Würde statt abgerundet aufgerundet werden, so könnte das Signifikanzniveau überschritten werden. Der aus den gegebenen Daten berechnete F-Wert $F_C = 0.7225$ kann jetzt mit dem theoretischen $F_{0.95,12,22} = 2.2258$ verglichen werden. Der theoretische Wert ist größer als der empirische, demzufolge kann die Nullhypothese $H_0 : \sigma_C^2 = 0$ nicht abgelehnt werden.

Von den betrachteten sieben Hypothesen sind sechs nach dem Schema der Varianzanalyse überprüfbar. Für eine Hypothese wurde eine neue Möglichkeit zur Überprüfung von Hypothesen eingeführt, die der Quasi-F-Brüche.

Von den drei Hypothesen zu den festen Effekten waren die Hypothesen zu den Haupteffekten von Amplitude und Frequenz signifikant, so dass nicht bewiesen werden konnte, dass sie keinen Einfluss haben. Der feste Wechselwirkungseffekt F_{AB} war nicht signifikant. Dieses Ergebnis war nach Darstellung 2.1 zu erwarten.

Die vier Hypothesen, die die zufälligen Effekte betreffen, konnten nicht vollständig überprüft werden, da es bei der Bildung des F-Wertes von Effekt ABC zu einer Division durch 0 gekommen wäre. Diese ist aber nicht definiert. Die Hypothese, die die Wechselwirkung zwischen Amplitude und Versuchsperson betrifft, liefert ein signifikantes Ergebnis, wohingegen die Hypothese zur Wechselwirkung von Frequenz und Versuchsperson ein nicht signifikantes Ergebnis liefert. Der zufällige Haupteffekt F_C hat keinen signifikanten Einfluss, d.h., die Versuchspersonen haben keinen signifikanten Einfluss auf die gegebene Bewertung der Expositionen.

3.3 Weitere Berechnungsmöglichkeiten für Parameter im linearen gemischten Modell

Neben der eben vorgestellten Vorgehensweise existieren weitere Möglichkeiten um die Parameter im Modell 3.2 zu schätzen [19, 20]. Die drei häufigsten Varianten werden kurz vorgestellt.

Beste lineare erwartungstreue Schätzung und beste lineare unverzerrte Vorhersage

Der beste lineare erwartungstreue Schätzer wird im Englischen als *Best Linear Unbiased Estimator* (**BLUE**) bezeichnet. BLUE schätzt die Parameter für die festen Effekte. Er ist derjenige Schätzer, der unter allen Schätzfunktionen die kleinste Varianz aufweist.

Der beste lineare erwartungstreue Schätzer $\hat{\theta}$ von einem Parameter θ ist derjenige,

1. einen linearen Schätzer darstellt, d.h. die Form $\hat{\theta} = a_1 X_1 + \dots + a_n X_n$ hat,
2. der erwartungstreu ist, d.h. es soll gelten $(E(\hat{\theta}) = \theta)$ oder $Bias(\hat{\theta}, \theta) = 0$
3. und der die kleinste Varianz unter allen linearen erwartungstreuen Schätzern hat, d.h., er ist effizienter und es für jeden beliebigen anderen Schätzer $\hat{\theta}_{bel}$ gilt $\hat{\theta} < \hat{\theta}_{bel}$. [21, 22]

Die beste lineare unverzerrte Vorhersage wird in der englischen Literatur *Best Linear Unbiased Prediction* (**BLUP**) genannt. Während BLUE diejenige Schätzfunktion mit der kleinsten Varianz ist, ist BLUP der im Durchschnitt minimierte Schätzfehler.

Der beste lineare unverzerrte Schätzer ist derjenige, der

1. linear ist, d.h., er ist in dem Sinne linear, als dass er eine lineare Funktion der y-Werte ist ,
2. erwartungstreu ist, d.h., es soll gelten $(E(\hat{\theta}) = \theta)$ oder $Bias(\hat{\theta}, \theta) = 0$
3. der "beste" Schätzer ist, d.h., er soll den kleinsten quadratischen Fehler von allen linearen unverzerrten Schätzern haben. [23]

Methode der kleinsten Quadrate

Der Vektor der Parameter der festen Effekte lässt sich ebenfalls durch die Methode der kleinsten Quadrate schätzen. Der Ansatz ist hierbei derselbe wie bei allgemeinen Regressionsmodellen. D.h., die Summe der quadrierten Residuen werden minimiert, um den Parametervektor β zu schätzen. Die Methode sei hier am Beispiel der Schätzers für feste Effekte dargestellt.

$$RSS : \sum_{i=1}^n (y_i - x_i' \beta_j)^2 \rightarrow \min_{\beta_1, \dots, \beta_k} !$$

Zur Minimierung der Funktion müssen die partiellen Ableitungen nach β gebildet werden. Dabei wird für die erste Ableitung die Nullstelle errechnet und für den Wert der Nullstelle mittels der zweiten Ableitung überprüft, ob der Funktionswert größer 0 ist.

$$\begin{aligned} f'(\beta) &= -2X'y + 2X'X\beta \\ 0 &= -2X'y + 2X'X\beta \\ 2X'y &= 2X'X\beta \\ f''(\beta) &= 2X'X \end{aligned}$$

Die eindeutige Lösung dieser Gleichungen erhält man mit

$$\hat{\beta} = (X'X)^{-1}X'y. \quad (3.14)$$

Der Schätzer aus 3.14 erfüllt die Bedingungen 1 bis 3 des besten linearen erwartungstreuen Schätzers.

Sei jetzt die Varianz von y bekannt und mit V bezeichnet und nicht-singulär, dann muss die Funktion $(y - X\beta)'V^{-1}(y - X\beta)$ minimiert werden und es ergibt sich der Schätzer

$$\hat{\beta} = (X'V^{-1}X)^{-1}X^TV^{-1}y.$$

[13]

Maximum-Likelihood-Schätzung

Während die ersten beiden vorgestellten Methoden keine Angaben über eine Verteilung der Beobachtung benötigen, sondern nur Voraussetzungen über die Beobachtungsfehler, muss bei der Maximum-Likelihood-Schätzung die Dichte der Beobachtungen in Abhängigkeit von den unbekanntem zu schätzenden Parametern gegeben sein. [12]

Die gemeinsame Dichte für die gegebenen Antworten auf die Intensitätsfrage y_1, \dots, y_n ist gegeben als

$$f(y_1, \dots, y_n | \theta).$$

Sie hängt von dem unbekanntem Parameter θ ab. Dieser soll mit der Maximum-Likelihood-Methode geschätzt werden. Die Likelihoodfunktion, die die Wahrschein-

lichkeit des unbekanntes Parameters θ beschreibt, lautet

$$L(\theta) = f(y_1, \dots, y_n | \theta).$$

Der Schätzer $\hat{\theta}$ für θ soll jetzt so gewählt werden, dass er die Likelihoodfunktion maximiert. Wenn die Likelihoodfunktion differenzierbar nach dem zu schätzenden Parameter θ ist, dann kann das Maximum bestimmt werden, indem die Likelihoodfunktion differenziert und Null gesetzt wird und aus dem sich ergebenden Gleichungssystem durch Auflösen die Lösung für θ bestimmt wird. Jedoch wird üblicherweise nicht die Likelihoodfunktion zum differenzieren, Nullsetzen und berechnen genutzt, sondern die log-Likelihoodfunktion. Diese ergibt sich durch Logarithmieren der Likelihoodfunktion. Diese Verfahrensweise ist legitim, da es sich bei dem Logarithmus um eine streng monotone Funktion handelt und so die Lage der Extremwerte unverändert bleibt.

Die Schätzer ergeben sich als

$$\beta = (X' \hat{V}^{-1} X)^{-1} X' \hat{V}^{-1} y$$

und

$$u = G \hat{U}' \hat{V}^{-1} (y - X\beta).$$

Hendersons Mixed Model Equations

Es ist möglich, die Schätzer BLUE und BLUP gleichzeitig zu bestimmen. Unter der Annahme, dass die Varianzparameter fest sind, lassen sich die beiden Schätzer als Lösung eines System von linearen Gleichungen bestimmen. Diese Vorgehensweise wurde 1959 von Charles Roy Henderson beschrieben und das Ergebnis wurde als Mixed Model Equations bekannt. [24]

$$X' R^{-1} X \hat{\beta} + X' R^{-1} Z \hat{u} = X' R^{-1} y$$

$$Z' R^{-1} X \beta + (Z' R^{-1} Z + D^{-1}) u = Z' R^{-1} y$$

Oder kompakter in Matrixschreibweise

$$\begin{bmatrix} X' R^{-1} X & X' R^{-1} Z \\ Z' R^{-1} X & R^{-1} Z + D^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X' R^{-1} y \\ Z' R^{-1} y \end{bmatrix}$$

3.4 Weitere Tests im linearen gemischten Modell

Die eben berechneten Parameter können genutzt werden, um zu testen, ob einzelne Faktorstufen bzw. ein Faktor sich mit all seinen Faktorstufen auf das Intensitätsempfinden der Probanden auswirkt.

Test auf Einfluss einzelner Faktorstufen

Die geschätzten Parameter sind im Vektor β zusammengefasst. Um zu testen, ob eine Komponente des Vektors Einfluss ausübt, wird die j -te Komponente des Vektors β getestet. Dieser Test stellt eine Variante des t-Testes dar.

Die Nullhypothese lautet $H_0 : \beta_j = d_j$. d_j ist ein beliebiger Wert gegen den getestet wird, $d_j \in \mathbb{R}$. Wird $d_j = 0$ gewählt, so kann getestet werden, ob das j -te Element des Parametervektors Einfluss auf die abhängige Variable hat.

Die Teststatistik lautet

$$t_j = \frac{\hat{\beta}_j - d_j}{\hat{\sigma}_j} \stackrel{A}{\sim} N(1 - 0.5\alpha) \quad (3.15)$$

$\hat{\beta}_j$ ist der Vektor mit den geschätzten Parametern. Die Schätzung der Parameter kann mit einer der Methoden aus Kapitel 3.3 erfolgen. $\hat{\sigma}_j$ ist das $[j, j]$ -te Element der Kovarianzmatrix von $\hat{\beta}$.

Die Teststatistik ist nur in Spezialfällen unter Annahme der Nullhypothese t -verteilt. Im Allgemeinen wird davon ausgegangen, dass die Teststatistik bei Annahme der Normalverteilung für große Stichproben approximativ normalverteilt ist. Die Nullhypothese wird abgelehnt, wenn der berechnete Wert größer ist als der theoretische. [12, S. 267]

Test auf Einfluss eines Faktors

Häufig ist ein globaler Test, also ein Test der untersucht, ob ein Faktor mit all seinen Faktorstufen einen Einfluss auf die abhängige Variable hat, von Interesse. Um das zu testen, kann die Methode nach Scheffé genutzt oder ein Wald-Test durchgeführt werden.

Mit der **Scheffé-Methode**, auch Scheffé-Test genannt, kann der Einfluss eines festen Faktors in einem Regressionsmodell getestet werden. Beim Test nach Scheffé

werden alle Mittelwerte aller Faktorstufen eines Faktors miteinander paarweise verglichen um festzustellen, ob signifikante Unterschiede zwischen den Faktorstufen eines Faktors vorliegen.

Dazu werden Konfidenzintervalle zu der Nullhypothese $\varphi = 0$, d.h. es liegt kein Unterschied zwischen den einzelnen Faktorstufen vor, gebildet. Um die Nullhypothese zu testen, müssen lineare Kontraste gebildet werden. Lineare Kontraste sind Funktionen von den Mittelwerten der einzelnen Faktorstufen $i = 1, \dots, k$ $k \in \mathbb{N}$ und den zugehörigen Koeffizienten

$$\varphi = \sum_{i=1}^k c_i \mu_i,$$

wobei für die Koeffizienten gelten soll, dass die Summe der Koeffizienten gleich Null ist

$$0 = \sum_{i=1}^k c_i.$$

Als Schätzer für μ_i wird der Mittelwert der jeweiligen Faktorstufe \bar{y}_i angenommen

$$\hat{\mu}_i = \bar{y}_i$$

n_i ist die Anzahl der betrachteten Werte. In dem Experiment für die Intensitätsdaten würde i von 1 bis 5 reichen, wenn der Faktor Frequenz mit all seinen Faktorstufen betrachtet wird. Die Mittelwerte \bar{y}_i sind die Mittelwerte der jeweiligen Faktorstufe, also z.B. der Mittelwert der Faktorstufe Frequenz 1. Anschließend können die Konfidenzintervalle gebildet werden

$$\varphi - S \cdot s(\varphi) \leq \varphi \leq \varphi + S \cdot s(\varphi)$$

Dabei ist $s(\varphi) = \sqrt{\hat{\sigma}_e^2 \sum_{i=1}^k \frac{c_i^2}{n_i}}$ und $S = \sqrt{(k-1)F_{k-1, n-k, 1-\alpha}}$. Die Nullhypothese wird verworfen, wenn der Wert $\varphi = 0$, gegen den getestet wird, außerhalb des Intervalls

$$[\varphi - S \cdot s(\varphi), \varphi + S \cdot s(\varphi)]$$

liegt. [25, S. 63-64]

Eine Fehlerquelle bei der Methode nach Scheffé ist das Festlegen der Kontraste, da dadurch eine Beziehung zwischen den Faktorstufen unterstellt wird. So bedeutet z.B. der Kontrast $[0, -1, 2, -1]'$, dass der Einfluss der dritten Stufe doppelt so hoch ist wie der der zweiten und vierten Stufe. Deshalb sollte sich die Festlegung an Beziehungen

orientieren, die aus der Praxis bekannt sind. Um eine mögliche Kombination für die Kontraste der Daten der Intensitätsfrage festzulegen, können Grafiken genutzt werden, die die Daten wiedergeben. Um die Kontraste für einen Vergleich, ob der Faktor Frequenz insgesamt einen Einfluss auf das Intensitätsempfinden der Probanden hat, zu ermitteln, kann z.B. die Abbildung 1.3 auf Seite 6 genutzt werden. Es ist ersichtlich, dass die Probanden sich bei niedrigsten und höchsten Frequenzen am wohlsten fühlten. D.h. diesen Faktoren wäre ein geringer Wert im Kontrast zuzuordnen. Eine mögliche Kombination um die Kontraste festzulegen wäre $c = [-2, 3, 2, 1, -4]'$.

Eine andere Methode um den Einfluss eines Faktors zu testen, ist der **Wald-Test**.

Die Nullhypothese des Testes lautet $H_0 : C\beta = d$. Bei Ablehnung der Nullhypothese wird die Differenz $C\beta - d$ klein. Testen lässt sich diese Hypothese mit der Prüfgröße

$$W = (C\hat{\beta} - d)'(C'A_{11}C)^{-1}(C\hat{\beta} - d). \quad (3.16)$$

$(C'A_{11}C)^{-1}$ sei nicht nicht-singulär. C ist eine $[m, r]$ Matrix. Die Matrix enthält die zu untersuchenden Linearkombinationen. Der Vektor d hat die Größe $[r, 1]$. A_{11} ist die Varianz der geschätzten Parameter $Var(\hat{\beta})$.

Unter Annahme der Nullhypothese ist die Testgröße χ^2 -verteilt mit $f_c = rg(C)$. Die Nullhypothese wird angenommen, wenn der theoretische Wert größer ist als der berechnete.

Es lässt sich zeigen, dass der in Kapitel 3.4 aufgeführte Test auf einen Parameter als Spezialfall des Wald-Testes betrachtet werden kann. Um einen einzelnen Parameter auf Signifikanz zu überprüfen, wird $d = 0$ und $C = [0, \dots, 1, \dots, 0]$ gewählt. Wird das in die Teststatistik (3.16) eingesetzt, ergibt sich $W = (\hat{\beta}_r \frac{1}{\sigma_j} \hat{\beta}_r)$. Das wird zur Teststatistik 3.2 und ist unter Annahme der Nullhypothese normalverteilt. [12, S. 267]

4 Ergebnisse

In diesem Kapitel sind die Ergebnisse aus den bisherigen Kapitel dargestellt. Anschließend werden die beiden Methode, Varianzanalyse und lineares gemischtes Modell, miteinander verglichen. Eine der in der Einleitung beschriebenen Aufgaben ist der Vergleich zwischen der ISO-Linie und den aus den Experimentdaten gewonnenen Kurve. Die Frequenzbewertungskurven aus der ISO und Experiment werden in Kapitel 4.3 miteinander verglichen.

4.1 Darstellung der Ergebnisse

Die Ergebnisse der Kapitel 2 und 3 sind in der folgenden Tabelle zusammenfassend dargestellt.

Tabelle 4.1: Ergebnisse der Varianzanalyse und Ergebnisse des linearen gemischten Modells

	Varianzanalyse mit Messwiederholung			Lineares gemischtes Modell		
	empirischer F-Wert	theoretischer F-Wert	Signifikanz	empirischer F-Wert	theoretischer F-Wert	Signifikanz
Frequenz	6.899	2.5652	signifikant	6.9	2.5652	signifikant
Amplitude	80.082	3.4028	signifikant	80.08	3.4028	signifikant
Interaktion Frequ, Ampl	1.255	2.0363	nicht signifikant	1.25	2.0363	nicht signifikant
Interaktion Frequ, Prob	-	-	-	2.0927	1.4889	signifikant
Interaktion Ampl, Prob	-	-	-	1.4636	1.6542	nicht signifikant
Proband	-	-	-	0.7225	2.2258	nicht signifikant

Es ist erkennbar, dass die Varianzanalyse nur Ergebnisse für feste Effekte liefert, wohingegen das lineare gemischte Modell sowohl Ergebnisse für feste als auch für zufällige Effekte liefert.

Die Ergebnisse für die festen Effekte der beiden Methoden können miteinander verglichen werden. Es ist ersichtlich, dass die Berechnungen der beiden Methoden dieselben Ergebnisse liefern und so die Ergebnisse übereinstimmen. Beide liefern als Ergebnis, dass die Faktoren Amplitude und Frequenz das Intensitätsempfinden der Probanden signifikant beeinflussen.

Der Einfluss der zufälligen Effekte kann nur mit dem linearen gemischten Modell betrachtet werden. Die Nullhypothese $H_0 : \sigma_c^2 = 0$ muss abgelehnt werden, da der berechnete Wert größer ist als der theoretische F-Wert. Die Interaktionen zwischen Amplitude und Proband bzw. zwischen Frequenz und Proband sind nicht signifikant.

Es ist also festzustellen, dass das Intensitätsempfinden der Probanden signifikant von der Größe der Amplitude und der Höhe der Frequenz beeinflusst wird.

4.2 Vergleich von Mixed Model und Varianzanalyse für wiederholte Messungen

In Kapitel 2 und Kapitel 3 wurden die beiden Methoden vorgestellt. Bei der Anwendung von unterschiedlichen Verfahren zur Auswertung stellen sich die Fragen, ob beide dieselben Ergebnisse liefern, ob sie beide dieselben Voraussetzungen haben und worin die Vorteile der einzelnen Verfahren liegen. Diese Fragen sollen in diesem Unterkapitel beantwortet werden.

Aus Tabelle 4.1 ist ersichtlich, dass beide Methoden dieselben Ergebnisse für die Signifikanztests auf feste Effekte liefern. Dies war bei genauerer Betrachtung der Formeln aus Kapitel 3.2 zu erwarten, da die Formeln sich zu denen der Varianzanalyse für wiederholte Messungen vereinfachen, wenn $n = 1$ gewählt wird. Mit dem linearen gemischten Modell war es jedoch möglich, Aussagen über den Einfluss der zufälligen Effekte, die Probanden, zu treffen. Dies ist bei der Anwendung der Varianzanalyse für wiederholte Messungen nicht möglich.

Die Voraussetzungen für die Varianzanalyse mit wiederholten Messungen sind Normalverteilung, Sphärizität und Varianzhomogenität. Insbesondere die Voraussetzung der Sphärizität führte dabei zu Problemen, da von dem Mauchly-Test auf Sphärizität bekannt ist, dass er zu falschen Ergebnissen führen kann und deshalb bei Zweifeln am Vorliegen der Sphärizität immer die korrigierte Testgröße benutzt werden sollte.

Die Voraussetzungen des linearen gemischten Modells verlangen im wesentlichen nur, dass das gefundene bzw. benutzte Modell den Verlauf der Daten erklären sollen. Das Modell soll also die wirkliche Verteilung der Daten darstellen.

Ein weitere Vorteil des linearen Modells ist, dass es auch mit Daten umgehen kann, die nicht in allen Zellen gleichbesetzt sind, es also fehlende Werte in der Messreihe gibt. Es existieren zwar Ansätze, wie auch mit solchen Daten eine Varianzanalyse durchführbar ist, allerdings liefern diese unterschiedliche Ergebnisse [8, S. 249]. Bei dem linearen gemischten Modell können zur Auswertung von unbalancierten Versuchsplänen die Henderson Methoden I-III genutzt werden [24, S. 168-231].

Das lineare gemischte Modell kann für Varianzanalyse, Kovarianzanalyse und Regressionen genutzt werden und ist so vielseitiger einsetzbar, als die Varianzanalyse. Daher sind die vorliegenden Variablen in ihrer Struktur auch nicht festgelegt, so müssen die unabhängigen Variablen nicht nur kategorial, sondern können auch metrisch sein (dann würde eine Kovarianzanalyse vorliegen).

Die beiden Verfahren unterscheiden sich auch hinsichtlich ihres Haupteinsatzgebietes. Lineare gemischte Modell werden insbesondere genutzt, wenn hierarchische oder longitudinale, bzw. messwiederholte, Daten auszuwerten sind. Um mit einer Varianzanalyse messwiederholte Daten auszuwerten ist ein besonderes ANOVA-Design erforderlich. Varianzanalyse werden größtenteils eingesetzt, wenn der Einfluss mehrerer kategorialer Variablen auf eine oder mehrere abhängige metrische Variablen untersucht werden soll.

Allerdings ist der Test auf die Signifikanz eines Faktors komplizierter als in der Varianzanalyse, z.B. mussten bei der Methode nach Scheffè zusätzliche Annahmen getroffen werden, um die Kontraste formulieren zu können. Andererseits können im linearen gemischten Modell über die geschätzten Parameter und die Datenmatrix sowohl Tests auf einzelne Faktorstufen als auch auf den Einfluss eines Faktors, mit all seinen Faktorstufen, durchgeführt werden.

4.3 Vergleich von Eichtungs-Kurven aus dem Experiment mit Wichtungs-Kurve aus ISO 2631-1

Um die ISO-Kurve mit den Daten des Experimentes vergleichen zu können, müssen die Daten zuerst aufbereitet werden. Nachdem die gemessenen Werte normalisiert, logarithmisiert und mit -1 multipliziert wurden, sind sie mit den ISO-Kurven vergleichbar. Die Kurven der ISO-Wichtungskurven zu dem Komfortempfinden, wie

komfortabel sich ein Proband bei einer gewissen Amplitude-Frequenz-Einstellung fühlt, und die Kurven aus den Experimentdaten sind vergleichbar, da die Kurven zur Intensivitätsfrage an der x-Achse gespiegelt den selben Verlauf haben, wie die Kurven zum Komfortempfinden. In Bild 4.1 sind die errechneten Kurven (die oberen) und die ISO-Kurve (die untere) eingezeichnet.

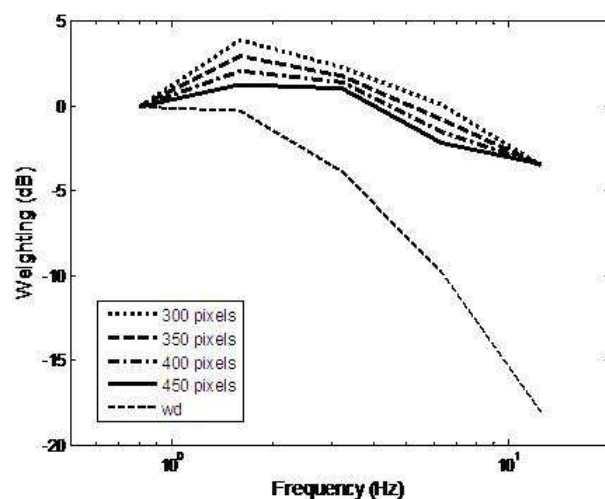


Abbildung 4.1: Darstellung der aus den Experimentdaten ermittelten Wichtungskurve und der ISO-Wichtungskurve
Quelle: [6]

Es soll der Verlauf der Grafen hinsichtlich der Fragestellung verglichen werden, ob sie einen signifikant anderen Kurvenverlauf haben. Es handelt sich um zwei unverbundene Stichproben, da die zu testenden Daten in unterschiedlichen Experimenten mit unterschiedlichen Probanden ermittelt wurden. Da nicht davon ausgegangen werden kann, dass die Daten einer Normalverteilung folgen, wird zur Auswertung ein nichtparametrischer Test benutzt.

Es wird der **U-Test**, auch Wilcoxon-Mann-Whitney-Test genannt, genutzt [9]. Die Voraussetzungen des Testes sind

- die Unabhängigkeit der zu vergleichenden Stichproben,
- eine näherungsweise gleiche Verteilungsform der Stichproben, d.h. insbesondere Varianzhomogenität $\sigma_{BAuA} = \sigma_{ISO}$,
- eine stetige Verteilung der Stichproben

- und dass die Anzahl der Messungen pro Stichprobe größer gleich 4 sein soll.

Die Unabhängigkeit ist erfüllt so lange eine Kurve aus dem Experiment mit der ISO-Kurve verglichen wird, da die Werte in unterschiedlichen Experimenten, die an verschiedenen Probanden durchgeführt wurden, gewonnen wurden. Die Verteilung ist kontinuierlich. Es existieren keine Unstetigkeitsstellen. Pro Stichprobe sollen 5 Werte verglichen werden, so dass $n_{BAuA}, n_{ISO} \geq 4$ gilt. Mit dem Test wird geprüft ob jeweils die ISO-Kurve und eine der Kurven aus den Experimentdaten der selben Verteilung folgen. Unter Annahme der Nullhypothese ist die Verteilung gleich. Es ist zu beachten, dass bei stark abweichenden Verteilungen die Effizienz des Testes sinkt. Aus der Abbildung 4.1 ist eher zu erwarten, dass die Gleichheit der Varianzen nicht angenommen werden kann. Die ISO-Kurve zeigt einen viel größeren Abfall der Werte als die Werte des BAuA-Experiments.

Bei dem U-Test handelt es sich um einen Rangsummentest. Es werden also zuerst alle Werte der Stichproben ihrer Größe nach sortiert und anschließend die Ränge vergeben, so dass der kleinste Wert den Rang 1 bekommt. Existieren mehrere gleichgroße Werte, so bekommen die gleichgroßen Werte als Rang den arithmetischen Mittelwert der als nächstes zu vergebenden Ränge. Das Vorgehen sei kurz an folgendem Beispiel in Tabelle 4.3 demonstriert.

Tabelle 4.2: Beispiel zur Rangvergabe

Stichprobe 1	Rang	Stichprobe 2	Rang
-0.0700	8.5	-0.0700	8.5
0.9793	10	-0.2800	6
-0.1603	7	-3.8500	5
-4.0813	4	-9.8100	2
-7.3161	3	-15.6000	1

Anschließend werden die Summen $r_i [i \in \mathbb{N}]$ der gebildeten Ränge für jede Stichprobe einzeln gebildet. So würde sich für das Beispiel aus Tabelle 4.3 für die erste Stichprobe die Summe 32.5 und für die Zweite 22.5 ergeben. Mit dem Test wird überprüft, ob die Wahrscheinlichkeit das ein beliebig ausgewähltes Element zu einer der beiden Stichproben gehört $\frac{1}{2}$ ist.

Der Test wird zweiseitig durchgeführt, d.h. es ist nur von Interesse, ob die beiden Verteilungen sich unterscheiden. Die Nullhypothese lautet

$$H_0 : F_1 = F_2$$

und die Alternativhypothese

$$H_0 : F_1 \neq F_2.$$

Die Prüfgröße ist

$$U = \min\{U_1, U_2\},$$

dabei sind

$$U_1 = r_1^* - \frac{n_1(n_1 + 1)}{2}$$

$$U_2 = r_2^* - \frac{n_2(n_2 + 1)}{2}$$

Der sich so ergebene empirische Wert für die Testgröße wird mit dem theoretischen Wert $U_{n_1, n_2, \alpha/2}$ verglichen. Die Nullhypothese ist anzunehmen, wenn der berechnete Wert kleiner ist als der theoretische Wert.

$$H_0 \text{ annehmen} : U \geq U_{n_1, n_2, \alpha/2}$$

$$H_0 \text{ ablehnen} : U < U_{n_1, n_2, \alpha/2}$$

Die Werte für die theoretischen U-Werte können entsprechenden Tabellen entnommen werden, z.B. [9, S. 275-279].

Es wurden die ISO-Wichtungskurve mit denen aus den Experiment gewonnenen Kurven verglichen.

Tabelle 4.3: Vergleich der ISO-Wichtungskurve mit den Wichtungskurven, die sich aus dem Experiment ergaben

	Vergleich ISO Kurve und Wichtungskurve bei Liniennlänge 100	Vergleich ISO Kurve und Wichtungskurve bei Liniennlänge 450
$U_{n_1, n_2, \alpha/2}$	2.0	2.0
U	3.5	4.5
Testentscheid	Nullhypothese annehmen	Nullhypothese annehmen

Beim Vergleich der Daten wurde die Nullhypothese, dass die jeweilige Kurve und die ISO-Kurve denselben Verlauf haben, angenommen.

Dieses Ergebnis ist bei Betrachtung von Grafik 4.1 überraschend. Eine mögliche Ursache für die Annahme der Nullhypothese könnte die geringe Stichprobengröße sein.

Eine Alternative, um zwei Stichproben miteinander zu vergleichen, ist der Cramér-von-Mises-Test. Der Zweistichprobenfall des Testes vergleicht die Verteilung der beiden betrachteten Stichproben. [26, 27]

5 Zusammenfassung und Ausblick

Diese Arbeit sollte die Daten der Intensitätsfrage dahingehend auswerten, ob der Faktor Frequenz bzw. der Faktor Amplitude sich signifikant auf das Intensitätsempfinden der Probanden auswirkt. Um diese Frage zu beantworten, wurden die Daten sowohl mit der Varianzanalyse für wiederholte Messungen als auch mit dem linearen gemischten Modell ausgewertet. Beide Verfahren lieferten dieselben Ergebnisse. Es zeigte sich, dass sowohl die Frequenz als auch die Amplitude einen signifikanten Einfluss auf das Wohlbefinden der Probanden hatten. Um zu testen, welche Faktorstufen sich signifikant auf die abhängige Variable auswirken kann der Bonferroni-Test oder der Tukey-Test genutzt werden. Um die einzelnen Faktorstufen auf Signifikanz zu testen, kann der Test in Kapitel 3.4 genutzt werden.

Anschließend wurde aus dem im Experiment gewonnenen Daten Wichtungskurven gebildet und der Verlauf dieser Kurven mit dem Verlauf der Kurve aus der ISO 2631 verglichen. Der Vergleich wurde durch einen U-Test realisiert. Der Test lieferte das Ergebnis, dass sich die ISO-Kurve nicht signifikant von denen aus den Daten gebildeten Kurven unterscheidet.

Bei der Auswertung der Daten mit den unterschiedlichen Analysemethoden zeigten sich die Vor- und Nachteile dieser Verfahren. Das lineare gemischte Modell ist anpassbarer als die Varianzanalyse. Es ist möglich, mit dieser Methode auch die zufälligen Effekte zu testen. Die Varianzanalyse erfordert zwar ebenfalls, dass vor Beginn der Analyse die zu untersuchenden Faktoren festgelegt werden und eine Vereinbarung darüber getroffen wird, wie sie zu behandeln sind, z.B. als messwiederholte Faktoren oder als Gruppierungsfaktoren, jedoch ist es nicht notwendig, ein Modell aufzustellen, da ein solches bereits durch die Benutzung der Varianzanalyse gegeben ist. Nachteilig an der Varianzanalyse ist jedoch, dass die Voraussetzungen, wann die Methode anwendbar ist, strenger sind als bei dem linearen gemischten Modell.

Bei dem Experiment wurden zwar die Frequenzen und die Amplituden vorgegeben, jedoch konnten diese von dem Simulator nicht immer in einem noch tolerierbaren Rahmen umgesetzt werden. Deshalb wäre es denkbar, die Analyse zu wiederholen und die Daten, von denen bekannt ist, dass der Simulator die zu erzeugenden Frequenzen

nicht erzeugt hat, wegzulassen und so eine Analyse für unbalancierte Daten durchzuführen. Ebenfalls wäre es möglich, statt der kategorialen Variablen für die Faktoren Frequenz und Amplitude metrische Variablen einzuführen. Diese würden dann statt der Werte 2001, ..., 2005 und 1001, ..., 1003 die tatsächlich gemessenen Frequenzen und Amplituden enthalten. Wird bei nur einem der beiden Faktoren die kategoriale Variable durch eine metrische ersetzt, kann eine Kovarianzanalyse durchgeführt werden. Werden beide Faktoren ersetzt, so kann eine Regressionsanalyse durchgeführt werden.

Desweiteren könnte noch getestet werden, ob die Probanden mit fortschreitender Zeit des Experimentes die Schwingungen intensiver fühlen. Um eine solche Fragestellung beantworten zu können, wurde am Anfang und am Ende jedes Experimentes der Proband zur selben Frequenz und Amplitude befragt, wie intensiv er die Schwingung empfindet. Durch eine solche Fragestellung könnte geklärt werden, ob die Probanden mit der Zeit Schwingungen intensiver fühlen.

Das Ziel des Experimentes war es zu klären, ob die Wichtungskurven der ISO-Kurven das tatsächliche Komfortempfinden der Exponierten wiedergeben oder ob die ISO-Kurven angepasst werden müssen. Dazu wurden aus den Daten des Experimentes eigene Wichtungskurven erstellt und diese mit denen der ISO verglichen. Es ist wichtig, dass die Wichtungskurven der ISO das tatsächliche Komfortempfinden anzeigen, da diese Kurven als Grundlage für die Ermittlung von Grenzwerten, wie lange ein Arbeiter einer Vibration mit einer gewissen Stärke ausgesetzt sein darf, um nicht zu Gesundheitsschäden beim Arbeiter zu führen. Der durchgeführte Test zeigte zwar, dass sich die ISO-Wichtungskurve und die Kurven aus dem Experiment sich nicht signifikant unterscheiden, allerdings kann dies auch auf die geringe Stichprobengröße zurückzuführen sein. Bei einem erneuten Experiment sollten die Probanden zu mehr Frequenzen als in diesem Experiment befragt werden, so dass mehr Werte zur Auswertung vorhanden sind.

Literaturverzeichnis

- [1] 2631-1, ISO: *Mechanical vibration and shock - Evaluation of human exposure to whole-body vibration - Part 1: General requirements*. 1997
- [2] *Verordnung zum Schutz der Beschäftigten vor Gefährdungen durch Lärm und Vibrationen (Lärm- und Vibrations-Arbeitsschutzverordnung - LärmVibrations-ArbSchV) vom 6. März 2007 (BGBl. I S. 261), die zuletzt durch Artikel 3 der Verordnung vom 19. Juli 2010 (BGBl. I S. 960) geändert worden ist*
- [3] *Berufskrankheiten-Verordnung vom 31. Oktober 1997 (BGBl. I S. 2623), die zuletzt durch Artikel 1 der Verordnung vom 11. Juni 2009 (BGBl. I S. 1273) geändert worden ist*
- [4] 2057-1, VDI: *Einwirkung mechanischer Schwingungen auf den Menschen - Ganzkörper-Schwingungen*. Düsseldorf, 2002
- [5] SCHNAUBER, H. ; TREIER, C.: *Zur Beurteilung von Vibrationen am Arbeitsplatz*. www.ruhr-uni-bochum.de/rubin/rbin2_95/rubin6.htm. Version: 5. August 2012
- [6] FORTA, N. ; SCHUST, M. ; LÖWIS, P. von ; KAISER, H. ; KREISEL, A.: Equivalent intensity contours for fore-and-aft whole-body vibration. In: *United Kingdom Conference on Human Response to Vibration 46* (2011), S. 141–154
- [7] BACKHAUS, K. ; ERICHSON, B. ; PLINKE, W. ; WEIBER, R.: *Multivariate Analysemethoden: Eine Anwendungsorientierte Einführung*. Springer, 2005
- [8] BORTZ, J. ; SCHUSTER, C.: *Statistik für Human- und Sozialwissenschaftler*. Springer, 2010
- [9] WERNER, J.: *Biomathematik und medizinische Statistik*. Urban & Schwarzenberg, 1992

- [10] RASCH, B. ; FRIESE, M. ; HOFMANN, W.J. ; NAUMANN, E.: *Quantitative Methoden 2. Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Springer, 2006
- [11] DUEMBGEN, L.: *Biometrie*. Vieweg+Teubner Verlag, 2009
- [12] FAHRMEIR, L. ; KNEIB, T. ; LANG, S.: *Regression: Modelle, Methoden Und Anwendungen*. Springer, 2009
- [13] RIEDL, Regina: *Lineare gemischte Modelle für longitudinale Daten*. <http://portal.tugraz.at/portal/page/portal/Files/i5060/files/students/theses/riedl08.pdf>. Version: 31. Juli 2012
- [14] TOUTENBURG, H.: *Lineare Modelle: Theorie Und Anwendungen*. Physica-Verlag, 2002
- [15] KOHN, W.: *Statistik: Datenanalyse und Wahrscheinlichkeitsrechnung*. Springer, 2004
- [16] SCHACH, S. ; SCHÄFER, T.: *Regressions- und Varianzanalyse: e. Einf.* Springer, 1978
- [17] RASCH, D.: *Einführung in die mathematische statistik*. VEB Deutscher Verlag der Wissenschaften, 1978 (Bd. 2)
- [18] *Kreuzklassifikation*. <http://www.christianlehmann.eu/ling/epistemology/concepts/index.html?http://www.christianlehmann.eu/ling/epistemology/concepts/klassifikation.php>. Version: 14. August 2012
- [19] WITKOVSKÝ, Viktor: *MATLAB algorithm mixed.m for solving Henderson's mixed model equations*. <http://aiolos.um.savba.sk/~viktor/PDF/Vw2001f.pdf>. Version: 31. Juli 2012
- [20] MCLEAN, Robert A. ; SANDERS, William L. ; STROUP, Walter W.: A Unified Approach to Mixed Linear Models. In: *American Statistician* 45 (1991), S. 54–64
- [21] *Estimable Functions And Gauss-Markov Theorem*. <http://courses.washington.edu/b533/lect10.pdf>. Version: 8. August 2012
- [22] WANG, G.C.S. ; JAIN, C.L.: *Regression Analysis: Modeling & Forecasting*. Graecyway Pub., 2003

- [23] *Best Linear Unbiased Prediction (BLUP) of Random Effects in the Normal Linear Mixed Effects Model*. <http://www.public.iastate.edu/~dnett/S511/26BLUP.pdf>. Version: 15. August 2012
- [24] SEARLE, S.R. ; CASELLA, G. ; MCCULLOCH, C.E.: *Variance components*. Wiley, 1992
- [25] PRUSCHA, H.: *Statistisches Methodenbuch: Verfahren, Fallstudien, Programm-codes*. Springer, 2005
- [26] SACHS, L. ; HEDDERICH, J.: *Angewandte Statistik*. Springer, 2006
- [27] BORTZ, J. ; LIENERT, G.A. ; BOEHNKE, K.: *Verteilungsfreie Methoden in Der Biostatistik*. Springer, 2010
- [28] FAHRMEIR, L. ; PIGEOT, I. ; KÜNSTLER, R. ; TUTZ, G.: *Statistik: Der Weg Zur Datenanalyse*. Springer, 2009

Tabellenverzeichnis

1.1	Datenauszug der Intensitätsdaten	5
2.1	Klassifizierung der Varianzanalyse	9
2.2	Datenanordnung in einer Tabelle für zweifaktorielle Varianzanalyse	12
2.3	Datenschema komplette Messwiederholung (Auszug)	18
2.4	Übersicht empirische und theoretische F-Werte für Intensitätsfrage-Daten	20
2.5	F-Werte im Vergleich für Intensitätsdaten für die Varianzanalyse ohne Messwiederholung und die Varianzanalyse mit Messwiederholung	22
3.1	Berechnung der Quadratsummen im linearen gemischten Modell	28
3.2	Berechnung der mittleren Quadrate im linearen gemischten Modell	29
4.1	Ergebnisse der Varianzanalyse und Ergebnisse des linearen gemischten Modells	39
4.2	Beispiel zur Rangvergabe	43
4.3	Vergleich der ISO-Wichtungskurve mit den Wichtungskurven, die sich aus dem Experiment ergaben	44

Abbildungsverzeichnis

1.1	Frage zum Intensitätsempfinden mit wachsendem Balken	4
1.2	Darstellung der Ergebnisse der Intensitätsfrage aufgetragen über links oben Frequenz, rechts oben Amplitude, links unten Probandennummer	5
1.3	Grafische Darstellung der Intensitätsdaten getrennt nach den unterschiedlichen Amplitudenstufen. Die Roten Sterne stellen die Rohdaten dar, die blauen Linien sind die Mittelwerte über die jeweiligen Frequenzen.	6
2.1	Interaktionsdiagramm für Intensitätsfrage	11
4.1	Darstellung der aus den Experimentdaten ermittelten Wichtungskurve und der ISO-Wichtungskurve Quelle: [6]	42

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Magdeburg, der 28.08.2012

Anhang

Die benutzten Programme befinden sich auf der beiliegenden CD.

- `Plots_Einleitung.m`
enthält die Programme um die Grafiken der Einleitung zu erstellen.
- `VA_mwh.m`
berechnet die Varianzanalyse für wiederholte Messungen aus Kapitel 2.2.
- `GeMo.m`
berechnet die Ergebnisse des linearen gemischten Modells nach Kapitel 3.2.
- `U_Test.m`
enthält den U-Test aus Kapitel 4.3 zum Vergleich zweier Stichproben.