

Statistische Analysen zu Training und Leistung im Skisprung

Masterarbeit (SS 2009)

Institut für Statistik

Ludwig-Maximilians-Universität München

Valentin Wimmer

Betreuer:

Prof. Dr. Ludwig Fahrmeir (LMU)

Prof. Dr. Ansgar Schwirtz (TUM)

Nora Fenske (LMU)

Patricia Pyrka (TUM)

3. November 2009

Zusammenfassung

In dieser Arbeit werden die Diagnostikdaten von Athleten der Sportarten Skisprung und Nordische Kombination ausgewertet. Im Rahmen des Projekts „Leistungsdiagnostik Ski Nordisch“ werden seit dem Jahr 2000 mit den deutschen A-, B-, C- und D/C-Kaderathleten Sprungkrafttests und isometrische und dynamische Bein Streckkrafttests am Diagnostikgerät IsoMed2000 durchgeführt. An den Untersuchungsorten Freiburg, München und Oberhof wurden so über 1000 Beobachtungen von 99 Skispringern und 92 Nordischen Kombinierern gesammelt und in einem Datensatz zusammengefasst. Für einen Teil der Athleten liegen zusätzlich Ergebnisse von Übungssprüngen von der Schanze in Oberstdorf vor. In der Schanzendiagnostik wurden unter anderem Sprungweite, Absprunggenauigkeit und vertikale Absprunggeschwindigkeit erfasst.

In einem ersten Schritt werden die erhobenen Daten zur (Labor-) Leistungsdiagnostik ausgewertet und dabei insbesondere die Zusammenhänge der Variablen betrachtet. Mit einer Faktorenanalyse wird ein Faktorenmodell aufgestellt, das die beobachteten Variablen durch eine kleinere Menge an latenten Variablen beschreibt. Es lassen sich drei Faktoren konstruieren, wobei der erste die allgemeine Sprungkraftfähigkeit und relative Maximalkraft, der zweite die allgemeine Maximalkraft und Schnellkraft und der dritte das Kraftanstiegsverhalten zusammenfasst. Mit Bootstrap-Verfahren werden Konfidenzintervalle für die geschätzten Parameter des Faktorenmodells konstruiert, welche bestätigen, dass die Ergebnisse nicht von einzelnen Beobachtungen oder Athleten abhängen.

Im zweiten Schritt wird die Schanzendiagnostik beschrieben. Hier wurden im Rahmen verschiedener Lehrgänge Übungssprünge der Athleten von der Schattenbergschanze (Hillsize 137) in Oberstdorf dokumentiert. Es wird beschrieben, wie anhand der verschiedenen Sprünge ein Datensatz konstruiert wird, der für jeden Athleten eine Beobachtung der interessierenden Größen zu jedem Lehrgang enthält. Eine besondere Bedeutung hat hier die Sprungweite von der Schanze, da sie die Leistung eines Athleten unter Wettkampfbedingungen zeigt.

Im dritten Schritt werden die Ergebnisse der Schanzendiagnostik und Leistungsdiagnostik zu einem gemeinsamen Datensatz mit 107 Beobachtungen bei 46 Athleten zusammengefasst. Ziel ist die Identifikation von Parametern der Leistungsdiagnostik, die einen Einfluss auf die Sprungweite haben. Zur Modellierung des Problems werden Mischansätze aus Regression und Faktorenanalyse, Latente-Variablen-Modelle und strukturierte additive Regression mit Bayesianischer Regularisierung verwendet. Es zeigt sich, dass insbesondere die Variablen zum Squat Jump aus der Leistungsdiagnostik einen positiven Zusammenhang mit der Sprungweite und der vertikalen Absprunggeschwindigkeit vom Schanzentisch haben. Ein solcher Zusammenhang zeigt sich bei den Variablen zur Beinkraftdiagnostik nicht.

Vorwort

Diese Arbeit wurde in Kooperation des Instituts für Statistik der Ludwig-Maximilians-Universität München und dem Fachgebiet Biomechanik im Sport der Technischen Universität München geschrieben. Das Thema dieser Arbeit kann für verschiedene Leser von Interesse sein. Der Schwerpunkt dieser Arbeit liegt auf der Beschreibung und Anwendung statistischer Methodik für einen Datensatz. An den Stellen der Arbeit, wo Vorwissen benötigt wird, wird auf die entsprechende Literatur zu dem Themen verwiesen. Wichtige Methoden werden in einem Exkurs vorgestellt. Außerdem enthält der Anhang E eine Übersicht und Beschreibung wichtiger Grundlagenliteratur.

In Anhang A finden sich Abkürzungen, die im Rahmen dieser Arbeit verwendet wurden. Im Anhang B.1 sind Hinweise zu der Anwendung des statistischen Programmpakets R. Am Ende jedes Kapitels werden die wichtigsten Ergebnisse des Kapitels diskutiert und kurz zusammengefasst.

Diese Arbeit wurde mit \LaTeX erstellt. Alle Berechnungen und Grafiken wurden mit den statistischen Programmpaketen R (R Development Core Team, 2009) und `BayesX` (Brezger u. a., 2009) erzeugt und mit `Sweave` (Leisch, 2002) in den \LaTeX -Code eingebunden. Die Daten von Skispringern und Nordischen Kombinierern, die in dieser Arbeit analysiert werden, sind Gegenstand von weiteren Untersuchungen und von großem Interesse für Wissenschaftler, Trainer und Athleten. Aus Datenschutzgründen werden in dieser Arbeit keine Namen von Athleten genannt und der Arbeit keine CD mit Daten angefügt.

Mein Dank für die Unterstützung bei dieser Arbeit gilt in alphabetischer Reihenfolge:

- Prof. Dr. Ludwig Fahrmeir für die hervorragende Betreuung und die vielen hilfreichen methodischen Anmerkungen und Vorschläge.
- Nora Fenske für die hervorragende Betreuung meiner Arbeit verbunden mit vielen inhaltlichen Vorschlägen und Diskussionen. Außerdem danke ich für die Hinweise zur strukturierten Gliederung der Arbeit.
- Sebastian Kaiser für die Beantwortung praktischer Fragen zur Faktorenanalyse.
- Susanne Konrath für die Beantwortung zahlreicher Fragen zur Bayesianischen Regularisierung.
- Patricia Pyrka danke ich für Diskussionen und Anmerkungen zu meiner Arbeit. Außerdem bedanke ich mich dafür, dass ich am Institutskolloquium am 09.06.2009 teilnehmen konnte und so zahlreiche neue Ideen bekommen habe und außerdem danke ich für die Teilnahme an einem Termin zur Leistungsdiagnostik, der mir einen tieferen Einblick in die Entstehung der Daten geliefert hat.
- Katharina Schüller für die Vermittlung des Themas und die Teilnahme am Kolloquium.
- Prof. Dr. Ansgar Schwirtz für die hervorragende fachliche Betreuung und zahlreichen inhaltlichen Anregungen und Diskussionen.
- Christian Seiler für die Bereitstellung des R-Codes zu Latente-Variablen-Modellen.
- Außerdem danke ich meinen Korrekturlesern Julia Meyer, Sebastian Meyer, Magdalene Wimmer, Raphael Wimmer und Janna-Levke Wösthoff.

Inhaltsverzeichnis

1. Skisprung und Nordische Kombination	1
1.1. Inhaltliche Einführung	1
1.1.1. Historische Entwicklung und Regeln	1
1.1.2. Wissenschaftliche Aspekte der Leistungsdiagnostik beim Skispringen	2
1.2. Statistische Analyse von Daten aus der Leistungssportforschung	5
1.3. Fragestellungen dieser Arbeit	6
2. Projekt dezentrale Leistungsdiagnostik Ski Nordisch	8
2.1. Beschreibung des Projekts	8
2.2. Anzahl der Messungen im Zeitverlauf	9
2.3. Erhobene Parameter und Beschreibung der zugehörigen Variablen	10
2.3.1. Persönliche Daten	12
2.3.2. Sprungkraftdiagnostik	12
2.3.3. Beinkraftdiagnostik	13
2.3.4. Abgeleitete Variablen	15
2.4. Analyse fehlender Werte	16
2.5. Explorative Analysen zur Leistungsdiagnostik	18
2.5.1. Univariate Verteilungen	18
2.5.2. Entwicklung über die Zeit	20
2.5.3. Altersspezifische Leistungsentwicklung	21
2.5.4. Profilbild einzelner Athleten	22
2.5.5. Vergleich von Skispringern und Nordischen Kombinierern unter Berücksichtigung des longitudinalen Datendesigns	24
2.5.6. Korrelationen der Variablen	28
2.6. Diskussion	28
3. Faktorenmodelle für die Daten zur Leistungsdiagnostik	31
3.1. Statistische Methodik: Faktorenanalyse	31
3.1.1. Faktorenanalyse bei Longitudinaldaten	31
3.1.2. Notation	34
3.1.3. Ausgangslage und Zielsetzung	35
3.1.4. Modell der Faktorenanalyse	36
3.1.5. Modellannahmen	38
3.1.6. Schätzung der Ladungsmatrix	39
3.1.7. Rotation der Faktoren	42
3.1.8. Interpretation der Ergebnisse	43
3.1.9. Schätzung der Faktorenwerte	43
3.1.10. Quantifizierung der Unsicherheit der Ergebnisse (Inferenzstatistik)	44
3.2. Auswertung und Ergebnisse	46
3.2.1. Motivation zur Verwendung einer Faktorenanalyse in der Leistungsdiagnostik	47
3.2.2. Vorbereitung der Analyse	47

3.2.3.	Festlegung der Anzahl der Faktoren	48
3.2.4.	Schätzung der Ladungsmatrix	48
3.2.5.	Darstellung der Faktorenwerte	52
3.2.6.	Unsicherheitsabschätzungen	53
3.2.7.	Schätzung von Faktoren zweiter Ordnung	56
3.3.	Diskussion	58
4.	Auswertung der Schanzenparameter	60
4.1.	Datenlage und Datenaufbereitung	60
4.1.1.	Erster Schritt: Einlesen der Rohdaten	60
4.1.2.	Zweiter Schritt: Aggregation der Rohdaten	60
4.2.	Explorative Analysen	62
4.3.	Diskussion	64
5.	Gemeinsame Auswertung der Leistungsdiagnostik und Schanzendiagnostik	67
5.1.	Datenvorbereitung	67
5.2.	Deskription	68
5.3.	Analysen und Ergebnisse	69
5.3.1.	Faktoren- und Regressionsanalyse	70
5.3.2.	Latente-Variablen-Modell	71
5.3.3.	Strukturierte additive Regressionsmodelle	77
5.4.	Diskussion	86
6.	Diskussion und Ableitungen für die weitere Forschungspraxis	88
6.1.	Leistungsdiagnostik	88
6.2.	Verknüpfung der Leistungsdiagnostik mit der Schanzendiagnostik	89
7.	Zusammenfassung	91
	Anhang	93
	A. Abkürzungsverzeichnis	94
	B. Hinweise zur statistischen Software	95
B.1.	Vorliegende Funktionen aus dem Basis-R-Paket oder Erweiterungen	95
B.2.	Eigene Funktionen	96
profilPlot	96
bootFA	97
plotKILoadings	98
	C. Ergebnisse der Regressionsmodelle mit Ridge-Regularisierung	100
	D. Samplingpfade	102
	E. Hinweise zur Literatur	105
	Literaturverzeichnis	106

1. Skisprung und Nordische Kombination

In dieser Arbeit werden Leistungen von Kaderathleten aus dem Bereich des Skisprungs und der Nordischen Kombination mit statistischen Methoden untersucht. Dieses Kapitel bietet in Abschnitt 1.1 eine inhaltliche Einführung zum Thema Skisprung und eine Beleuchtung der wissenschaftlichen Aspekte im Skisprung. Anschließend werden die zentralen Fragestellungen dieser Arbeit formuliert.

1.1. Inhaltliche Einführung

1.1.1. Historische Entwicklung und Regeln

Die Sportart Skisprung ist in Norwegen entstanden, wo 1879 der erste Wettkampf ausgetragen wurde. Seit 1924 ist die Sportart Teil der olympischen Winterspiele. Neben Norwegen zählen heute unter anderem Deutschland, Österreich und Finnland zu den großen Skisprungnationen. Der letzte große internationale Erfolg eines deutschen Springers ist der Sieg bei der Vierschanzentournee 2001/02 von Sven Hannawald. (FIS, 2009a) Die Wettkämpfe werden live und in Fernsehübertragungen weltweit von vielen Menschen verfolgt, wobei nur wenige diese Sportart selbst aktiv betreiben.

Beim Skispringen fährt der Athlet in hockender Haltung in einer vorbereiteten Spur den Anlauf der Sprungschanze hinab. Erreicht er den sogenannten Schanzentisch, so springt er bei einer Geschwindigkeit von etwa 90 km/h ab. Nun richtet er sich mit einem kräftigen Sprung auf, führt die Skier zum Körper und breitet sie gleichzeitig zu einem V aus. Beim Einleiten der Landung bringt er die Skier in eine parallele Stellung und schließt den Sprung optimalerweise mit einem sogenannten „Telemark“ ab. In Wettkämpfen wird der Sprung zu gleichen Teilen aus der erzielten Weite und den von Kampfrichtern vergebenen Noten für Flughaltung und Landung bewertet (FIS, 2009c).

Bei der Bewertung der Weite spielt es eine Rolle, von welcher Schanze der Sprung ausgeführt wurde. Maßgeblich ist hier der *Konstruktionspunkt* (K-Punkt) der Schanze. Dieser gibt den Punkt an, an dem das Gefälle des Aufsprunghangs flacher wird. Durch die Angabe des K-Punktes werden die Sprünge von unterschiedlichen Schanzen vergleichbar. So werden bei einer Sprungweite die genau dem K-Punkt der Schanze entspricht 60 Punkte vergeben. Bei einer größeren Weite werden Punkte addiert, andernfalls abgezogen. Seit dem Jahr 2004 werden die Schanzen anhand der *Hillsize* in Kategorien eingeteilt. Die Hillsize gibt die Strecke zwischen der Kante des Schanzentisches und dem Hillsize-Punkt der Schanze an. Der Hillsize-Punkt liegt an der Stelle, an dem der Auslaufhang eine Steigung von 32° hat (vgl. dazu auch Abbildung 1.1). Die Hillsize für kleine Schanzen ist kleiner oder gleich 49 m, bei mittleren Schanzen 50-84 m, bei Normalschanzen 85-109 m, bei Großschanzen größer als 110 m und bei Flugschanzen größer als 185m. (FIS, 2009d).

Neben der Einzeldisziplin Skisprung ist das Skispringen auch Teil der Nordischen Kombination. In diesem Mehrkampf wird nach aktuell gültigem Reglement zuerst ein Sprungdurchgang durchgeführt. Die erzielten Sprungpunkte werden nach der *Gundersen-Methode* in Zeitrückstände für den anschließenden Skilanglauf umgerechnet. Auch die Nordische Kombination ist seit 1924 Teil der olympischen Winterspiele, wobei erst seit 1948 zuerst gesprungen und dann gelaufen wird. Auch bei der Nordischen Kombination ist Deutschland eine erfolgreiche Nation und bei internationalen Wettkämpfen der letzten Jahre konnten deutsche Athleten viele Erfolge verbuchen. Die Sprünge werden bei der Nordischen

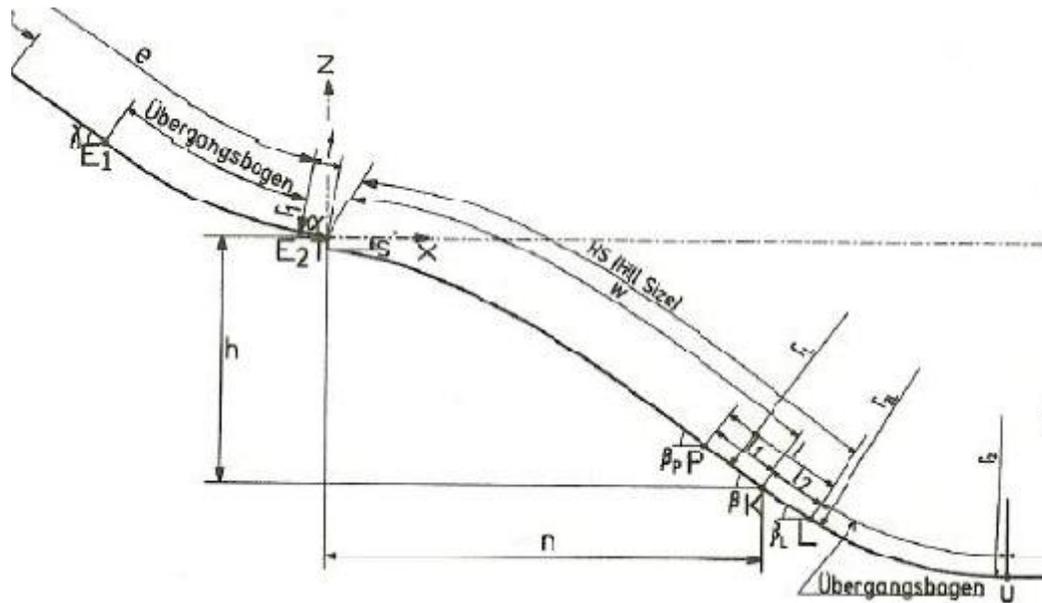


Abbildung 1.1.: Schematische Darstellung einer Skisprungschanze (übernommen aus FIS (2009d)). Mit „K“ ist der Konstruktionspunkt und mit „T“ die Tischkante der Schanze gekennzeichnet.

Kombination von Normal- oder Großschanzen ausgeführt (FIS, 2009b).

Im deutschen Skiverband gehören die Athleten, je nach ihren bisherigen Erfolgen, Alter und Leistungsvermögen zu verschiedenen Kadern. Dabei werden in den A- und B-Kadern die aktuellen Spitzenathleten und im C-Kader die jüngeren Athleten bundesweit betreut. Die D/C-Kader sind regionale Kader. Außerdem werden die Athleten unterschiedlichen Lehrganggruppen zugeordnet, je nachdem wo sie trainieren. Jeder Kader verfügt über eine umfangreiche Trainergruppe zur sportlichen und wissenschaftlichen Betreuung (DSV, 2009).

1.1.2. Wissenschaftliche Aspekte der Leistungsdiagnostik beim Skispringen

Bereits seit über 60 Jahren ist der Skisprung Gegenstand wissenschaftlichen Interesses. Die Forschung wird dadurch erleichtert, dass beim Skispringen – mit Ausnahme von Windeinflüssen – im Vergleich zu anderen Sportarten relativ konstante Bedingungen vorliegen (Neumaier u. a., 1997). Dies macht die Sprünge von einer Schanze untereinander leichter vergleichbar. Erste Untersuchungen beschäftigten sich mit dem physikalisch-mechanischen Bereich des Skisprungs. Themen sind hier die Ermittlung der optimalen Flugposition unter aerodynamischen Gesichtspunkten oder die Technik des richtigen Absprungsverhaltens vom Schanzentisch. Beim Verlassen der Schanze sind Richtung und Betrag des Geschwindigkeitsvektors wichtige Einflussgrößen für die Sprungweite.

Grundlage für Spitzenleistungen im Skisprung ist ein jahrelanges und umfangreiches Training. Besonders der Sprungkraft der Athleten wird eine besondere Bedeutung beigemessen. Daher wird intensiv an der Entwicklung und Verbesserung von Trainingsmethoden gearbeitet. Wichtig ist hier besonders ein sportartspezifisches Training. Beim Skisprung springt der Athlet aus einer Kauerstellung ab, die vorwiegend *konzentrische* Muskelarbeit erfordert (Müller u. Wachter, 1989). Ein Überblick über die verschiedenen Arbeitsweisen der Muskulatur findet sich in Tabelle 1.1.

Entscheidend sind auch die Kraftfähigkeiten eines Athleten. Mit *Kraft* wird dabei die „Fähigkeit des Nerv-Muskelsystems, durch Muskeltätigkeit Widerstände zu überwinden (konzentrische Kontraktion),

Kontraktionsform	isometrisch	(dynamisch-) konzentrisch	(dynamisch-) exzentrisch
Beschreibung	Die intramuskuläre Spannung ändert sich und die Muskellänge verändert sich nicht	Die intramuskuläre Spannung ändert sich und die Muskeln verkürzen sich	Die intramuskuläre Spannung ändert sich und die Muskeln verlängern sich
Beispiel	Gewicht hochhalten	Bizeps beim Klimmzug (Aufwärtsbewegung)	Bizeps beim Klimmzug (Herablassen)

Tabelle 1.1.: Arbeitsweisen der Muskulatur – Kontraktionsformen (Nigg u. Herzog, 1994)

ihnen entgegenzuwirken (exzentrische Kontraktion) bzw. sie zu halten (isometrische Kontraktion)” (vgl. Wolf, 1997, S.52), bezeichnet. Dabei wird zusätzlich noch zwischen Maximal- und Schnellkraft unterschieden.

Als *Maximalkraft* wird die maximale Kraft bezeichnet, die ein Sportler gegen einen Widerstand aufbringen kann. Hier unterscheidet man wieder zwischen isometrischer, (dynamisch-) konzentrisch und (dynamisch-) exzentrischer Maximalkraft. In der Theorie sind unter dynamisch-exzentrischen Muskelaktionsbedingungen die stärksten Kräfte zu erwarten, gefolgt von einer isometrischen Muskelaktion. Die geringsten Kräfte können unter dynamisch-konzentrischen Bedingungen entwickelt werden.

Die Differenz zwischen exzentrischer und isometrischer Maximalkraft wird als *Kraftdefizit* bezeichnet (Schwartz u. a., 2004). Dieser Parameter kann in der Trainingssteuerung der Belastungsdosierung berücksichtigt werden. „Ein hohes Kraftdefizit deutet auf die Möglichkeit hin, etwa durch forciertes Training der intramuskulären Koordination vorhandene Kraftreserven mobilisieren zu können, ohne dabei den Muskelquerschnitt und damit das Körpergewicht zu erhöhen” (Bruhn u. a., 2002). Neben der Maximalkraft ist auch die *Schnellkraftfähigkeit* der Muskulatur eines Athleten entscheidend. Damit wird die Fähigkeit bezeichnet, innerhalb von einer kurzen Zeit eine möglichst große Kraft aufzubauen. Zum Anforderungsprofil eines Skispringers gehört es, in der Absprunghase eine hohe vertikale Geschwindigkeit zu erreichen, wofür die Schnellkraft des Athleten ein wesentlicher Faktor ist (Wolf, 1997).

Begleitend zum Training der Kaderathleten werden mit einer *Leistungsdiagnostik* die Trainingserfolge verfolgt und dokumentiert. Nach Martin (1980) erfassen leistungsdiagnostische Verfahren

„[...] durch standardisierte Tests, Beobachtungen und/oder Untersuchungen den Trainingszustand und dessen Faktoren mit dem Ziel, den Grad der Ausprägung des Trainingszustandes und die Wirkung des Trainings auf seine Faktoren und Komponenten festzustellen. Durch die dabei gewonnenen Informationen wird der Trainingsprozess unmittelbar gesteuert.”

In Starischka (1981) wird die Leistungsdiagnostik in den Kontext der Trainingssteuerung eingebunden. Anhand der Leistungsdiagnostik können eventuelle Schwächen des Athleten identifiziert werden und die Trainingsinhalte entsprechend angepasst werden. Da insbesondere Kraft gut trainierbar ist, ist die Kraftdiagnostik wichtig bei der Trainingssteuerung (Letzelter u. a., 1990). Wird die Leistungsdiagnostik über einen längeren Zeitraum hinweg unter standardisierten Bedingungen für eine Gruppe von Athleten durchgeführt, so können sowohl die individuelle Entwicklung der Athleten, als auch Vergleiche zwischen den Athleten betrachtet werden. Auch Verletzungen können im Rahmen der Leistungsdiagnostik identifiziert werden.

Die Leistungsdiagnostik beruht auf dem Messen bestimmter Parameter mit Testverfahren und Messgeräten. Letzelter u. a. (1990) diskutieren die Ergebnisse der Kraftdiagnostik vor dem Hintergrund von

Messfehlern. Damit sind weniger Fehler an den Messgeräten gemeint, als vielmehr der Rückschluss von Kraftwerten auf die dahinterstehende Fähigkeit, also die inhaltliche Validität des Tests. Einerseits müssen Verbesserungen bei einem Krafttest nicht zwingend auf einen Kraftzuwachs, sondern können lediglich auf eine wirkungsvolle Technik bei dieser Übung zurückzuführen sein. Andererseits können manche Ergebnisse auch Ausreißer nach oben oder unten sein. Die Messwerte sind somit nur *Schätzer* des richtigen Niveaus. In Letzelter u. a. (1990) wird jedoch angemerkt, dass bei der Auswertung von Kraft-Zeit-Kurven, wie sie auch bei der hier untersuchten Leistungsdiagnostik (vgl. Kapitel 2) zum Einsatz kommen, eine sehr hohe Validität vorliegt.

Neben der Kraftdiagnostik ist die Sprungkraftdiagnostik ein zentraler Bestandteil der Leistungsdiagnostik von Skispringern. Hier kann man zwischen Sprüngen ohne und mit einleitender Gegenbewegung unterscheiden. Bei ersteren werden geringere Sprunghöhen erzielt, da die Beinmuskulatur nur konzentrisch arbeitet während bei zweiteren die Muskulatur einen *Dehnungs-Verkürzungszyklus* (DVZ) durchläuft. Hier wird Energie aus der Gegenbewegung gespeichert und kann anschließend in der Streckphase genutzt werden. Die Sprungformen stellen unterschiedliche Voraussetzungen an die Athleten und so sollte eine differenzierte Diagnostik aus mehreren Sprungformen bestehen (Frick u. a., 1991). Das Standardsprungkraftprogramm setzt sich aus den folgenden Sprungtests zusammen (vgl. Frick u. a. (1991) und Bruhn u. a. (2002)):

Squat Jump (SJ) Ausgangshaltung für diesen Sprung ist die Kauerstellung (90 Grad Knie- und Hüftwinkel). Um eine Gegenbewegung der Arme auszuschließen, bleiben diese während des gesamten Sprungs in die Hüfte gestützt. Der Athlet versucht nun, rein mit seiner Beinkraft möglichst hoch zu springen. Dieser Sprung überprüft die Sprungkraftfähigkeit bei einer rein konzentrischen Absprungbewegung.

Counter-Movement Jump (CMJ) Dieser Sprung wird aus dem aufrechten Stand ausgeführt. Im Anschluss an eine schnelle Ausholbewegung der Beine (bis zu einem Kniewinkel von ca. 90 Grad) holt der Athlet Schwung und springt in die Höhe. Hier durchläuft die Beinmuskulatur einen Dehnungs-Verkürzungszyklus. Diese Sprungform misst die Fähigkeit eines Athleten, die kinetische Energie der exzentrischen Ausholbewegung zu einem großen Teil in eine konzentrische Bewegung zu übertragen.

Drop Jump (DJ) Die Ausgangshaltung hier entspricht dem des CMJ, nur dass hier der Athlet auf einem Kasten mit festgelegter Höhe (Fallhöhe 24 oder 32 cm) steht. Der Sportler springt von dem Kasten und springt nach einem kurzen Bodenkontakt (möglichst < 200 Millisekunden, vgl. Bruhn u. a. (2002)) maximal in die Höhe. Diese Sprungform gibt Auskunft über das reaktive Kraftverhalten in einem schnell ablaufenden Dehnungs-Verkürzungszyklus.

Bei allen Sprungformen ist primär die Sprunghöhe von Interesse, daneben jedoch auch noch Parameter wie die Absprunggeschwindigkeit oder Bodenkontaktzeit. Die Sprunghöhe wird durch die Körperschwerpunkterhöhung zwischen Verlassen des Bodens und dem Erreichen des Scheitels der Flugphase ermittelt. Dazu können verschiedene Messverfahren wie Videoanalyse oder Kraftmessplatten eingesetzt werden. Frick u. a. (1991) beschreiben und vergleichen die verschiedenen Messmethoden.

Die Ergebnisse der Leistungsdiagnostik können auf unterschiedliche Art und Weise von Trainern und Athleten verwendet werden. Im Rahmen der Diagnostik können Stärken und Schwächen der Athleten identifiziert und das Trainingsprogramm hierauf ausgerichtet werden. Nach Vdovičenko (1980) gibt es hier sowohl den Ansatz, die Priorität auf unterentwickelte Fähigkeiten zu legen, als auch das weitere Training der Stärken des Athleten zu verfolgen. Trainingserfolge lassen sich dann in einer Längsschnittuntersuchung verfolgen. Für einzelne Athleten kann auch der Zusammenhang zwischen Maximalkraftparametern und Sprunghöhen ermittelt werden. Laut Bruhn u. a. (2002) bringt die Nähe

dieses Zusammenhangs die Fähigkeit des Springers zum Ausdruck, sein vorhandenes Kraftpotenzial in die Absprungbewegung einzusetzen. Ist die Korrelation hoch, so kann eine bessere Sprunghöhe durch Steigerung der Kraft erzielt werden (vgl. auch Pyrka u. a. (2008)).

Neben der individuellen Leistungsanalyse der Athleten kann die Diagnostik auch zum Vergleich zwischen Athleten herangezogen werden. Voraussetzung hierfür ist eine Leistungsdiagnostik unter standardisierten Voraussetzungen, wie sie in dem hier untersuchten Projekt vorliegt (vgl. Kapitel 2). In Bruhn u. a. (2002) wird angemerkt, welche „großen individuellen Unterschiede im Spitzensport beim Zustandekommen von Höchstleistungen bestehen können“. Einfache Gruppenvergleiche führen somit nicht weiter, sondern es müssen die individuellen Eigenschaften jedes Springers berücksichtigt werden. Dies führt zu den besonderen, sportwissenschaftlich spezifischen Aspekten bei der statistischen Analyse von Daten, die im Rahmen der Leistungssportforschung, also bei Training, Diagnostik und Wettkämpfen erhoben wurden.

1.2. Statistische Analyse von Daten aus der Leistungssportforschung

Im Rahmen der Forschung werden sowohl Quer- als auch Längsschnittdaten gewonnen. Die Analyse solcher Daten mit statistischen Methoden erfordert eine besondere Sensibilität bezüglich der Probleme bei Datensätzen aus der Leistungsdiagnostik. Diese umfassen nach Wagner (1996) generell die folgenden Punkte:

- Kleine Stichproben führen zu Problemen, da Verteilungsannahmen schlechter überprüft werden können und die Ergebnisse sensibler gegenüber Ausreißern in den Daten sind. Aus statistischer Sicht muss also insbesondere vorab sichergestellt werden, dass die Daten nicht fehlerhaft sind und die Ergebnisse nicht durch einzelne Messwerte verzerrt werden.
- Es gibt das Problem fehlender Messungen und damit verbundenen unvollständigen Datensätzen. Gerade bei Daten aus der Leistungsdiagnostik ist dieses Problem sehr ausgeprägt, da immer wieder Übungen aufgrund von Verletzungen oder Schmerzen nicht durchgeführt werden können. Mit dieser Problematik beschäftigt sich Abschnitt 2.4.
- Die Homogenität der Stichprobe ist nicht gesichert, insbesondere wenn verschiedene Datensätze zusammengeführt werden, so wie hier Daten von Skispringern und Nordischen Kombinierern.
- Es besteht das Problem der eingeschränkten Versuchsplanung, da unter den Bedingungen der Leistungsdiagnostik die Forderungen an die Versuchsplanung nur angenähert werden können. So sind balancierte Versuchsdesigns in der Regel nicht zu erreichen.

Um anhand von solchen Daten gültige und zu interpretierende Aussagen zu gewinnen, muss nach Wagner (1996) beim Einsatz von statistischen Methoden ein Maximum an Vorabinformationen aus der sportwissenschaftlichen Theorie eingesetzt werden, um sportwissenschaftlich richtige Ergebnisse zu generieren. Durch eine enge Zusammenarbeit mit dem Fachgebiet Biomechanik im Sport soll diese inhaltliche Qualitätssicherung der Ergebnisse sichergestellt werden.

Besonders bei Längsschnittuntersuchungen muss zudem die große Individualität der Athleten berücksichtigt werden. So wird in Lames (1996) angemerkt, dass bei der Beschreibung von Verläufen über die Zeit Gruppenmittelwerte möglicherweise nicht den Verlauf eines einzigen Athleten repräsentieren. Auch zur Ableitung von Trainingshinweisen ist also die Berücksichtigung von individuellen Aspekten entscheidend. Im folgenden Abschnitt werden die Fragestellungen, die diese Arbeit leiten, vorgestellt.

1.3. Fragestellungen dieser Arbeit

Im Rahmen dieser Arbeit sollen die Ergebnisse der Leistungsdiagnostik von Skispringern und Nordischen Kombinierern analysiert werden. Für einen Teil der Athleten liegen zusätzlich Ergebnisse zur Schanzendiagnostik vor, bei der Übungssprünge von der Hillsize 137-Skisprungschanze in Oberstdorf dokumentiert wurden. Beide Datenquellen haben die Form einer Längsschnittstudie, an der einzelne Athleten über einen längeren Zeitraum hinweg wiederholt teilgenommen haben. Aus Sicht der Sportwissenschaft sollen nun anhand dieser Daten die folgenden Fragestellungen untersucht werden:

1. Bereich: (Labor-)Leistungsdiagnostik (Kapitel 2 und 3)

- Wie müssen die Messergebnisse vorbereitet werden, damit die Daten verschiedener Athleten miteinander verglichen werden können? Welche Darstellungsform eignet sich zur Darstellung des (relativen) Leistungsniveaus eines Athleten?
- Wie kann mit dem Problem von fehlenden Messungen umgegangen werden?
- Welche statistischen Methoden eignen sich, um Zusammenhänge der Variablen zur Leistungsdiagnostik in einem Modell darzustellen?
- Lassen sich daraus Erkenntnisse für die Praxis ableiten, d.h. können anhand dieser Modelle Variablen identifiziert werden, die zusätzlich zu den anderen Variablen nur wenig Information liefern und daher überflüssig sind? Welcher Vorschlag kann aus statistischer Sicht für ein reduziertes Programm bei der Leistungsdiagnostik gegeben werden?
- Mit welchen statistischen Verfahren kann sicher gestellt werden, dass die Ergebnisse nicht durch einzelne Beobachtungen und Athleten stark verzerrt werden? Mit welcher Unsicherheit sind die Ergebnisse behaftet und wie können Konfidenzintervalle konstruiert werden?
- Welche Verfahren eignen sich für Gruppenvergleiche (z.B. von Skispringern und Nordischen Kombinierern), die im Gegensatz zu gängigen Verfahren wie t-Tests das longitudinale Datendesign berücksichtigen? In welchen Variablen unterscheiden sich Skispringer und Nordische Kombinierer?

2. Bereich: Schanzendiagnostik (Kapitel 4)

- In welcher Form können die Ergebnisse der Übungssprünge in der Schanzendiagnostik für statistische Zwecke genutzt werden? Wie kann man berücksichtigen, dass die Springer nicht bei jedem Übungssprung ihr volles Leistungsniveau abrufen?
- Welche Parameter messen die Leistung eines Athleten bei der Schanzendiagnostik und wie können sie dargestellt werden?

3. Bereich: Verknüpfung der Leistungsdiagnostik mit der Schanzendiagnostik (Kapitel 5)

- Wie kann ein gemeinsamer Datensatz konstruiert werden, in dem für jeden Athleten die Daten zur Leistungsdiagnostik und Schanzendiagnostik zusammengefasst werden, so dass kausale Schlüsse gezogen werden können?
- Welche verschiedenen Modellklassen eignen sich, um die Zusammenhänge der beiden Datenquellen darzustellen? Wie können die Ergebnisse interpretiert werden und welche Schlüsse können daraus für die Praxis gezogen werden?
- Welche anderen Kovariablen wie das Alter oder die Sportart sollen in diesen Modellen berücksichtigt werden?

Die weitere Arbeit gliedert sich wie folgt. Kapitel 2 beschreibt das Projektes zur „dezentralen Leistungsdiagnostik Ski Nordisch“ für Athleten aus den Bereichen Skisprung und Nordische Kombination. Dabei wird die Datenlage aufgezeigt, die erhobenen Parameter beschrieben und deskriptiv untersucht. In Kapitel 3 werden die Daten zur Leistungsdiagnostik genauer untersucht, insbesondere die Zusammenhänge der Variablen durch Faktorenmodelle beschrieben. In Kapitel 4 wird die zweite Datenquelle dieser Arbeit, die Schanzendiagnostik, beschrieben. Anschließend werden in Kapitel 5 die beiden Datenquellen zusammengeführt und gemeinsam analysiert. Kapitel 6 diskutiert die Ergebnisse dieser Arbeit und präsentiert Ableitungen auf die weitere Forschungspraxis, die sich anhand dieser Arbeit ergeben. In Kapitel 7 folgt eine Zusammenfassung der Arbeit.

2. Projekt dezentrale Leistungsdiagnostik Ski Nordisch

Dieses Kapitel beschreibt in Abschnitt 2.1 das Projektes zur leistungsdiagnostischen Betreuung von Kaderathleten aus den Bereichen Skisprung und Nordische Kombination. Es folgt eine Beschreibung der Messungen im Zeitverlauf und eine Auflistung der erhobenen Parameter, die anhand verschiedener Messungen bei den Athleten ermittelt wurden. Es liegen dabei jedoch nicht zu jedem Messzeitpunkt für alle Athleten gültige Messergebnisse vor. In Abschnitt 2.4 werden diese fehlenden Werte untersucht und Vorgehensweisen zum Umgang mit fehlenden Werte beschrieben. Abschließend folgen deskriptive Analysen der Datensätze zur Leistungsdiagnostik. Dabei werden die Korrelationen der Variablen betrachtet und Verfahren präsentiert, mit denen die einzelnen Messergebnisse vergleichbar gemacht werden sollen und Skispringer mit Nordischen Kombinierern verglichen.

2.1. Beschreibung des Projekts

Seit Anfang der 80er Jahre werden Leistungssportler der Sportarten Skisprung und Nordische Kombination durch das Institut für Sportwissenschaft der Universität Freiburg in Zusammenarbeit mit dem Olympiastützpunkt Freiburg-Schwarzwald leistungsdiagnostisch betreut. Für die Bundes- und Länderkader werden 2 bis 3 Mal pro Jahr die Kraft- und Sprungkraftparameter untersucht, wobei für einige Athleten bereits Messungen aus mehreren Jahren vorliegen (Bruhn u. a., 2002).

Seit Juli 2002 besteht ein Kooperationsvertrag zwischen dem Fachgebiet Biomechanik im Sport von der Fakultät für Sportwissenschaften der Technischen Universität München und dem Olympiastützpunkt Bayern. Diese Kooperation umfasst Dienstleistungen für den Olympiastützpunkt in folgenden Bereichen (vgl. Fachgebiet Biomechanik, 2009):

- Organisatorische Aufgaben
- Diagnostische Aufgaben (z. B. biomechanische Kraftdiagnostik, Technikanalyse)
- Beratungs- und Betreuungsaufgaben (für Trainer und Athleten)
- Wissenschaftliche Aufgaben (z. B. Weiterentwicklung von Diagnoseverfahren)

Ziel des Projekts ist es, analog zum Freiburger Modell, „eine dezentral durchführbare, einheitliche Beinkraftdiagnostik zu etablieren, die an den verschiedenen Stützpunkten unabhängig durchgeführt werden kann“ (vgl. Schwirtz u. a., 2004, S. 4). Dafür wurde an den Olympiastützpunkten in München, Freiburg und Oberhof das Trainings- und Diagnosegerät IsoMed2000 angeschafft. Alle Geräte sind baugleich und wurden 2008 miteinander verglichen.

In Schwirtz u. a. (2004) wird die Durchführung der Beinkraftdiagnostik an der IsoMed2000 detailliert beschrieben. Bei der ersten Messung wird das Gerät auf den Athleten eingestellt und diese Einstellungen werden für weitere Termine gespeichert. Für Folgetermine ist dann ein Zeitaufwand von etwa 45 Minuten (30 Minuten Beinkraftdiagnostik und 15 Minuten Sprungkraftdiagnostik) erforderlich. „Das Ziel der Neugestaltung diagnostischer Inhalte war es, in der Leistungsdiagnostik ein hohes Maß an Sportartspezifität zu erreichen, um relevante Leistungsparameter zu kontrollieren. Des Weiteren wurden eine einheitliche Parametrisierung und Auswerterroutine entwickelt“ (vgl. Schwirtz u. a., 2004, S. 6). Die gewonnenen Ergebnisse werden aufbereitet und in einer zentralen Datenbank zusammenge-

fasst. Dabei wurden die einzelnen Messungen im Rahmen dieser Arbeit auf Konsistenz überprüft und gegebenenfalls korrigiert oder Messergebnisse aus der Analyse entfernt. Der so erhaltene Datensatz wird im Folgenden beschrieben und anschließend explorativ analysiert.

2.2. Anzahl der Messungen im Zeitverlauf

Im Rahmen dieses Projekts wurden die Ergebnisse der Messungen von 99 Kaderathleten aus dem Bereich des Skisprungs und von 92 Nordischen Kombinierern aus den Jahren 2000 bis 2008 gesammelt. Das Projekt ist als Längsschnittstudie konzipiert, wobei 3 Messungen für die Athleten des A- und B-Kaders pro Jahr vorgesehen sind. Für die C bzw. D/C-Kader sind 2 bzw. 1 Messung pro Jahr angesetzt. Die Teilnahme an der Leistungsdiagnostik ist dabei verpflichtend, wobei aufgrund von Verletzungen oder Krankheit keine regelmäßige Teilnahme gewährleistet ist. Innerhalb eines Jahres ist die erste Messung für den Mai, die zweite für den Juni und die dritte für den Zeitraum September/Oktober vorgesehen. Dabei handelt es sich bei der Mai-Messung um eine Eingangsmessung im Anschluss an die Erholungsphase nach der Wettkampfsaison. Die Oktobermessung liegt hingegen vor der Wettkampfsaison, also zu einem Zeitpunkt, an dem die Athleten in bester körperlicher Verfassung sein sollten. Die Verteilung der Messungen über die Monate ist in Abbildung 2.1 zu finden. Die Athleten wurden an den Olympiastützpunkten in Freiburg, München und Oberhof, sowie bis zum Jahr 2004 in Medical Parcs untersucht. Dabei haben 140 der 191 Athleten mindestens an zwei verschiedenen Orten an der Leistungsdiagnostik teilgenommen. Die Anzahl der Messungen nach Untersuchungsort ist in Abbildung 2.1 dargestellt.

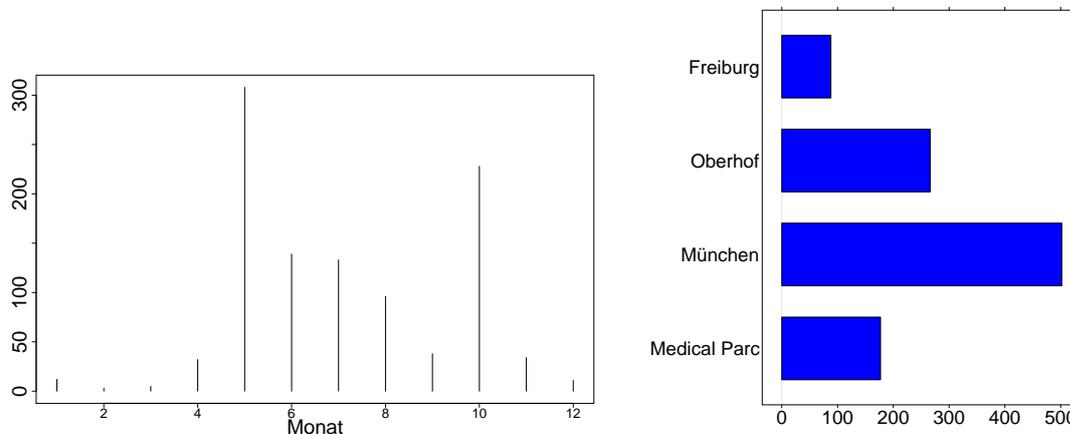


Abbildung 2.1.: Anzahl Messungen nach Monat (linke Abbildung) und nach Untersuchungsort (rechte Abbildung)

Die meisten der insgesamt 1039 Untersuchungen fanden im Mai (308) und im Oktober (228) statt. Fanden zwei Messungen im Jahr statt, so waren die Messungen dieses Jahres im Mittel 111 (Bereich: 3 bis 191) Tage von einander entfernt. Bei 3 Messungen pro Jahr lagen zwischen der ersten und der zweiten durchschnittlich 74 Tage (3 bis 181) und zwischen der zweiten und dritten 78 Tage (11 bis 165).

Für 149 Skispringer und Nordische Kombinierer liegen mehr als eine Messung vor, wobei das Mittel bei 5.45 Messungen pro Athlet liegt. Ein einzelner Nordischer Kombinierer hat sogar 22 Mal an der Leistungsdiagnostik teilgenommen. Insgesamt gab es 446 Messungen bei Skispringern und 593 bei Nordischen Kombinierern. Abbildung 2.2 zeigt die Teilnahme der Athleten an der Diagnostik getrennt

für Skispringer und Nordische Kombinierer. Bei den Nordischen Kombinierern nahmen bereits seit dem Jahr 2000 Athleten an der Leistungsdiagnostik teil, bei den Skispringern erst seit 2002. Die größte Teilnahme konnte im Jahr 2007 verzeichnet werden, wobei die mittlere Anzahl von Messungen pro Athlet im Jahr 2003 bzw. 2004 maximal war und seitdem bei beiden Sportarten wieder rückläufig ist.

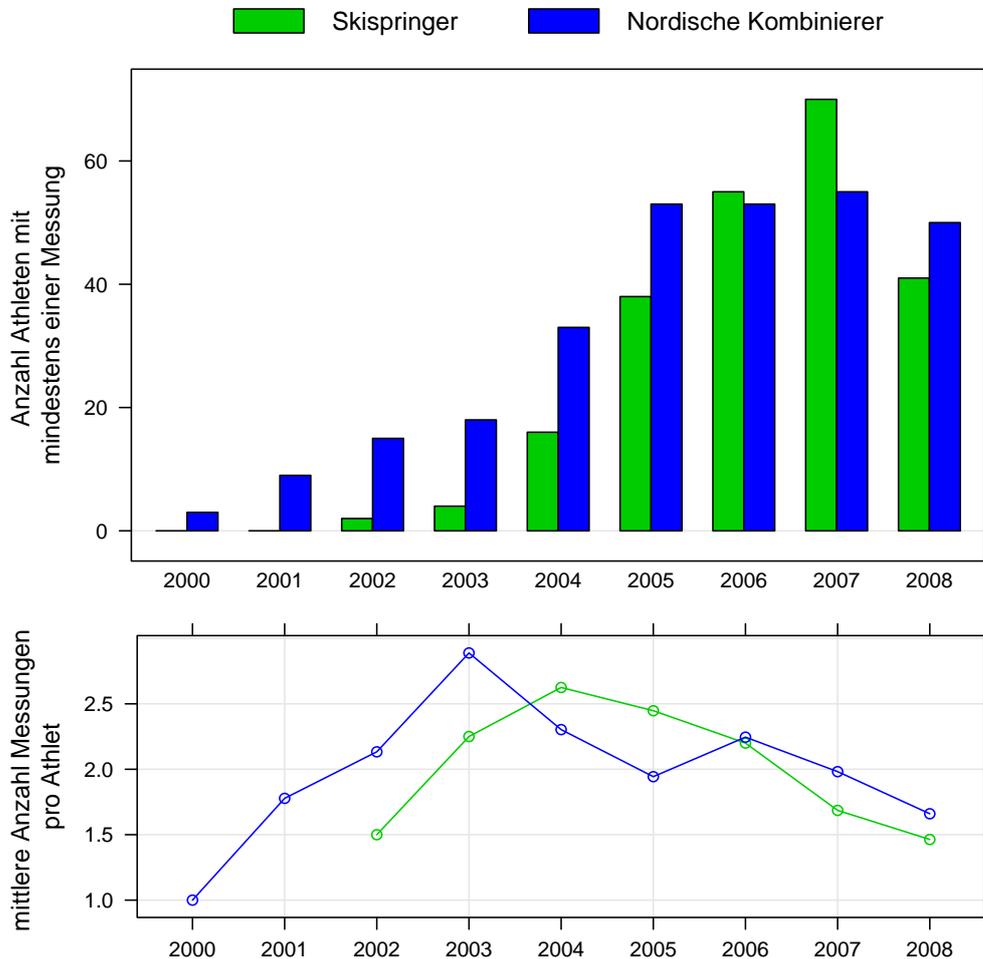


Abbildung 2.2.: Darstellung der Teilnahme an der Leistungsdiagnostik: In der oberen Abbildung ist die Anzahl der Athleten mit mindestens einer Messungen pro Jahr getrennt nach Sportart angetragen. Die untere Abbildung gibt an, wie oft die einzelnen Athleten im Mittel in diesem Jahr an der Diagnostik teilgenommen haben, wieder getrennt nach Sportart.

Zu jeder Messung wurden nun die im folgenden Abschnitt beschriebenen Leistungsparameter erhoben.

2.3. Erhobene Parameter und Beschreibung der zugehörigen Variablen

Dieser Abschnitt bietet eine Übersicht über die im Rahmen der Leistungsdiagnostik erhobenen Parameter. Beginnend mit der Sprungkraftdiagnostik absolvieren die Athleten anschließend Messungen zur isometrischen Maximalkraft und dann zur dynamischen Schnellkraft. Zwischen den Blöcken haben die Athleten die Möglichkeit zu einer kurzen Lockerungspause (Schwartz u. a., 2004). Von allen Übungen wird eine festgelegte Anzahl an Wiederholungen durchgeführt, zwischen denen vordefinierte Pausen von 1-2 Minuten liegen. Die Ergebnisse werden aufgezeichnet und unmittelbar danach von Athleten und Trainern besprochen. Die Athleten nehmen dabei zeitlich versetzt im Abstand von etwa 15 Mi-

nuten an der Leistungsdiagnostik teil, so dass keine Wartezeiten an den Messgeräten entstehen. Bei Skispringern und Nordischen Kombinierern werden im Rahmen der Sprungkraftdiagnostik teilweise unterschiedliche Messungen durchgeführt.

Die erhobenen Werte wurden in zwei getrennten Datensätzen für Skispringer und Nordische Kombinierer zusammengefasst. Diese Datensätze enthalten als *Beobachtungen* die Messergebnisse der Leistungsparameter zu einem Termin für einen Athleten und als *Variablen* die erhobenen Leistungsparameter. Die Datensätze sind dabei im *langen Datenformat* strukturiert, d.h. unterschiedliche Messungen bei dem selben Athleten entsprechen auch unterschiedlichen Beobachtungen im Datensatz. Das lange Datenformat wird in der Statistik zur Analyse von Längsschnittstudien eingesetzt (vgl. Diggle u. a., 2002). Man beachte dabei, dass es drei Athleten gibt, die im Laufe ihrer Karriere von der Nordischen Kombination zum Skisprung gewechselt sind, und daher Messungen in beiden Datensätzen haben. Da sie aber bis zum Zeitpunkt des Wechsels auch unterschiedlich trainiert haben, werden sie als unterschiedliche Athleten geführt.

Alle Messergebnisse für eine Testform bilden dann eine Variable in den Datensätzen. In diesem Abschnitt werden die gewonnenen Variablen in den Tabellen 2.1 bis 2.5 beschrieben. Die abkürzenden Variablennamen in den Tabellen dienen insbesondere dem Verständnis von Grafiken und Modellbeschreibungen. Sind in Tabellen nur die Ausprägungen für eine Sportart angegeben, so bedeutet dies, dass dieser Parameter nur bei dieser Sportart erhoben wurde.

Bemerkungen zur Vergleichbarkeit der Ergebnisse

Folgende Aspekte der Versuchsdurchführung sollen zu einer möglichst guten inter- und intraindividuellen Vergleichbarkeit der Ergebnisse führen:

- Bei allen Testformen werden mehrere Wiederholungen durchgeführt, von denen nur der beste Durchgang gewertet und dokumentiert wird. Mit diesem Vorgehen werden Ausreißer nach unten vermieden.
- Viele länger zurückliegende Sprünge wurden mit einer veralteten Methode gemessen und lieferten im Mittel bessere Ergebnisse. Um die Vergleichbarkeit der Ergebnisse zu gewährleisten, werden diese Messungen aus der Analyse entfernt und als fehlende Werte deklariert.
- Die Kniewinkelbereiche, in denen gemessen wird, sind so konzipiert, dass die Ergebnisse für Athleten unterschiedlicher Größe vergleichbar sind.

Im Rahmen dieses Projekts wurden die Messungen sowohl bei Spezial-Skispringern als auch bei Nordischen Kombinierern gemacht. Da die Nordische Kombination ein Mehrkampf ist und die Sportler neben dem Skisprung auch noch einen Skilanglauf absolvieren, erfordert diese Sportart natürlich andere körperliche Voraussetzungen als für die Skisprung-Spezialisten. So haben Nordische Kombinierer in der Regel ein höheres Körpergewicht, was sich beim Skisprung nachteilig auswirkt. Außerdem haben sie weniger Trainingszeit für das Skispringen, da parallel auch noch der Skilanglauf trainiert werden muss. Eine gemeinsame Untersuchung von Skispringern und Nordischen Kombinierern ist daher nicht immer möglich, insbesondere bei der Betrachtung der Sprungweiten bei Sprüngen von einer Skisprungschanze.

Verwendet man die Beobachtungen von Athleten beider Sportarten für statistische Analysen, so ergibt sich eine größere Datengrundlage. Allerdings können die Ergebnisse durch die unterschiedlichen Sportarten verzerrt sein. Deshalb wird im Rahmen der anschließenden Deskription unter anderem ein Vergleich beider Disziplinen vorgenommen. Danach können, soweit sinnvoll, beide Sportarten gemeinsam ausgewertet werden. Für die weiteren Analysen werden daher die beiden Datensätze der

Skispringer und Nordischen Kombinierer zusammengefasst. Hier sind dann nur noch die Variablen enthalten, die für beide Sportarten erhoben wurden.

2.3.1. Persönliche Daten

In diesem Abschnitt werden die erhobenen persönlichen Daten der Athleten beschrieben. Um die Vertraulichkeit der Ergebnisse der Leistungsdiagnostik zu gewähren, wurden die Namen der Athleten durch Identifikationsnummern ersetzt. Die Nummern 1 bis 99 kennzeichnen dabei die Skispringer und die Zahlen 100 bis 191 die Nordischen Kombinierer. Die weiteren allgemeinen Angaben zu den Athleten sind in Tabelle 2.1 zusammengefasst.

Variablenname	Variablenbeschreibung	Ausprägungen	Einheit	NA's
Athlet	Identifikationsnummer	Spr. 1,...,99		0
		Kom. 100,...,191		0
Sportart	Indikatorvariablen	1 = Skisprung		0
		2 = Nordische Kombination		0
Datum	Tag der Messung	Spr. 10.05.2002 - 14.11.2008		0
		Kom. 06.07.2000 - 23.10.2008		
Untersuchungsort		1 = Medical Parc		3 ^a
		2 = München		
		3 = Oberhof		
		4 = Freiburg		
Alter	Alter des Athleten	Spr. 12 – 30	Jahre	4
		Kom. 12 – 31		0
Größe	Körpergröße	Spr. 145 – 190	cm	3
		Kom. 144 – 194		11
Masse	Körpergewicht in kg	Spr. 29.9 – 78.8	kg	6
		Kom. 32.3 – 79.8		2
BMI	$\frac{\text{Körpergewicht}[\text{kg}]}{\text{Körpergröße}[\text{m}]^2}$	Spr. 13.5 – 23.2	$\frac{\text{kg}}{\text{m}^2}$	6
		Kom. 14.9 – 24.3		12

^aDrei Messungen an sonstigen Untersuchungsorten

Tabelle 2.1.: Erhobene persönliche Daten bei der Leistungsdiagnostik der Skispringer und der Nordischen Kombinierer: Variablenname, Beschreibung, Ausprägungen getrennt nach Sportart, wobei bei metrischen Größen Minimum und Maximum angegeben sind. Weiterhin wird die Maßeinheit genannt. In der letzten Spalte ist die jeweilige Anzahl der fehlenden Werte aufgelistet (NA).

2.3.2. Sprungkraftdiagnostik

Das zu absolvierende Sprungprogramm im Rahmen der Leistungsdiagnostik besteht im wesentlichen aus den drei Sprungformen Squat Jump, Drop Jump und Counter-Movement Jump, die in Abschnitt 1.1.2 beschrieben wurden. Die ersten beiden Varianten wurden sowohl bei Skispringern als auch bei den Nordischen Kombinierern durchgeführt, der CMJ nur bei den Skispringern. Sprünge mit Zusatzgewichten gab es nur bei den Nordischen Kombinierern. Von jeder Sprungform werden bis zu drei Wiederholungen durchgeführt und die beste ausgewählt und in den Datensatz übernommen. Die Übersicht der Variablen der Sprungkraftdiagnostik findet sich in Tabelle 2.2. Aus sportwissenschaftlicher Sicht sprechen hohe Werte bei allen Testformen für eine gute Leistung des Athleten. Lediglich bei der Bodenkontaktzeit beim Drop Jump (Variablenname: DJ32_KT) sind möglichst kleine Werte erwünscht.

Variablenname	Variablenbeschreibung	Ausprägungen	Einheit	NA's
SJ_SpH	Sprunghöhe beim SJ	Spr. 25.7 – 62.9	cm	58
		Kom. 21.3 – 49.9		207
SJ_V75E	Geschwindigkeit des Körperschwerpunktes 75 msec vor dem Abflug beim SJ	Spr. 1.86 – 2.92	$\frac{m}{s}$	113
		Kom. 1.82 – 2.88		244
SJ_S_SpH	Sprunghöhe beim SJ mit Turnschuhen	Spr. 0.2 – 50.6	cm	210
		Kom. 18.1 – 42.3		214
DJ32_SpH	Sprunghöhe beim DJ aus 32 cm Höhe	Spr. 12.7 – 61.2	cm	106
		Kom. 13.6 – 61.5		218
DJ32_KT	Bodenkontaktzeit beim DJ	Spr. 123 – 292	msec	47
		Kom. 119 – 273		32
CMJ_SpH	Spunghöhe beim CMJ	Spr. 26.3 – 62.9	cm	66
SJ_Z_SpH	Sprunghöhe beim SJ mit Zusatzgewicht (30kg)	Kom. 0.8 – 32.4	cm	435
SJ_Z_S_SpH	Sprunghöhe beim SJ mit Zusatzgewicht (30kg) und Sprungschuhen	Kom. 14.0 – 32.8	cm	442

Tabelle 2.2.: Erhobene Größen bei der Sprungkraftdiagnostik für Skispringer und Nordische Kombinerer

2.3.3. Beinkraftdiagnostik

Die Beinkraftdiagnostik wird vollständig am Gerät IsoMed2000 durchgeführt und die Kräfte und Kraftverläufe mit Hilfe von Kraftmessplatten aufgezeichnet. Die Positionierung im Testgerät ist an die sportartspezifische Haltung im Skisprung angelehnt. Ein Kniewinkel von 70° entspricht dabei der Position in der Anfahrtshocke. (Bruhn u. a., 2002). Die meisten Messungen wurden sowohl beidbeinig (**bb**) als auch getrennt nach rechtem (**re**) und linkem (**li**) Bein gemacht, wobei zuerst ein beidbeiniger Durchgang und dann abwechselnd je zwei einbeinige Messungen und zum Abschluss wieder eine beidbeinige Messung erfolgt. Die Kraftparameter werden aus der Kraft-Zeit-Kurve ermittelt, wobei das absolute Maximum die *Maximalkraft* angibt. Teilt man diese durch das Körpergewicht, so erhält man die *relative Maximalkraft* (gibt an, wie viel Kraft pro kg Körpergewicht aufgebracht werden kann). Die *Explosivkraft* ergibt sich als Steigung zwischen den Zeitpunkten zu denen 20% bzw. 70% der Maximalkraft erreicht werden. Die *relative Explosivkraft* bzw. der *relative Kraftanstieg* gibt an, welche Zeit mit der errechneten Steigung benötigt wird, um die Maximalkraft zu erreichen (Schwartz u. a., 2004). Eine Übersicht zu den isometrischen Kraftparametern gehörenden Variablen findet sich in Tabelle 2.3. Auch hier sind bei den Maximalkrafttests hohe Werte erwünscht, nur beim relativen Kraftanstieg sprechen kleine Werte für eine gute Leistung, da dann der Athlet schnell seine Maximalkraft erreicht.

An der IsoMed2000 wird die exzentrische Kraft (**Ekz**) wie folgt ermittelt: Ein Schlitten fährt den Weg von 5° Kniewinkel von 73° bis 68° ab. Die exzentrische Kraft ist nun der Mittelwert aus den Kraftwerten zwischen 69° und 71° Kniewinkel. (Schwartz u. a., 2004)

Die IsoMed2000 bietet auch die Möglichkeit zu dynamischen Schnellkrafttests. Dabei wird die Fähigkeit der Athleten überprüft, den Kraftabfall bei konzentrischer Arbeit gering zu halten und andererseits bei exzentrischen Muskelaktionsbedingungen die Kraft zu erhöhen (vgl. Schwartz u. a., 2004, S.13). Bei der praktischen Durchführung an der IsoMed2000 kann durch einen beweglichen Schlitten der Kniewinkel der Athleten variiert werden. Dazu wird das Gerät so eingestellt, dass für alle Sportler eine konstante Kniewinkelgeschwindigkeit ω (Einheit: $^\circ/\text{sec}$) erreicht wird. Wird der Kniewinkel verkleinert ($\omega < 0^\circ/\text{sec}$), so liegt eine dynamisch-exzentrische Muskelaktionsbedingung vor, bei einer Vergröße-

Variablenname	Variablenbeschreibung	Ausprägungen	Einheit	NA's
Fmax90_bb	Maximalkraft bei 90° Kniewinkel	Spr. 895 – 3505	N	188
		Kom. 993 – 3095		312
Fmax70_[bb,re,li]	Maximalkraft bei 70° Kniewinkel	Spr.	N	bb: 725 – 2535
				re: 370 – 1433
				li: 381 – 1407
		Kom.		bb: 830 – 2485
				re: 440 – 1384
				li: 412 – 1393
relKraftanstieg _[bb,re,li]	relativer Kraftanstieg	Spr.	msec	bb: 82 – 508
				re: 54 – 886
				li: 87 – 695
		Kom.		bb: 115 – 380
				re: 12 – 623
				li: 140 – 664
Ekz	Exzentrische Kraft bei 70° Kniewinkel	Spr. 688 – 2454	N	89
		Kom. 808 – 2420		234

Tabelle 2.3.: Isometrische Maximalkraftdiagnostik sitzend an der IsoMed 2000

zung des Kniewinkels eine dynamisch-konzentrische Muskelaktion. Ruht der Schlitten ($\omega = 0^\circ/\text{sec}$), so wird die Muskelkraft unter isometrischen Bedingungen überprüft (vgl. auch Tabelle 1.1 auf Seite 3). Mit Hilfe von Kraftmessplatten können *Kraft-Weg-Verläufe* (Kraft gegen Kniewinkel) aufgezeichnet werden.

Zur Ermittlung der dynamischen Schnellkraftdiagnostik werden weitere Messungen an der IsoMed2000 durchgeführt. Der Schlitten des Gerätes fährt dabei abwechselnd in kurzen Intervallen zum Athleten hin und weg, während dieser gegen die Kraftmessplatte drückt. Der Athlet muss also im schnellen Wechsel konzentrische und exzentrische Muskelarbeit verrichten. Da die Übungen koordinativ anspruchsvoll sind, werden bei den dynamischen Krafttest erst drei Wiederholungen zur Eingewöhnung an die Geschwindigkeit und dann direkt im Anschluss drei Messdurchgänge durchgeführt. Anhand dieser werden dann die folgenden Variablen bestimmt:

A60 und A180 Diese Variablen geben die verrichtete Arbeit A_ω bei einer mittleren Kniewinkelgeschwindigkeit von $\omega = 60^\circ/\text{sec}$ bzw. $\omega = 180^\circ/\text{sec}$ zwischen 80° und 110° Kniewinkel an. Dazu fährt der Schlitten der IsoMed2000 einen Kniewinkelbereich von 70° bis 130° mit der vorgegebenen Geschwindigkeit ab, während der Athlet gegen die Kraftmessplatte drückt. Die Arbeit berechnet sich als Integral des Kraft-Weg-Verlaufs F_ω :

$$A_\omega = \int_{80}^{110} F_\omega(s) ds.$$

Die Beschränkung auf den Bereich zwischen 80° und 110° erlaubt eine „möglichst hohe intra- und interindividuelle Vergleichbarkeit der Parameter“ (vgl. Schwirtz u. a., 2004, S. 15).

dynstat_[60,180] Diese Variablen geben das dynamisch-statische Kräfteverhältnis in Prozent an. Einmal wird die exzentrische Kraft bei 90° Kniewinkel und $\omega = -60^\circ/\text{sec}$ bzw. der konzentrischen Kraft bei 90° Kniewinkel und $\omega = 180^\circ/\text{sec}$ mit der isometrischen Maximalkraft bei 90° Kniewinkel verglichen. In der Theorie sollten hier Werte von größer 100% (**dynstat_60**) bzw. kleiner 100% (**dynstat_180**) auftreten.

ekV Hiermit wird das exzentrisch-konzentrische Verhältnis in Prozent angegeben. Dieses berechnet sich als der Quotient aus exzentrischer Arbeit bei $\omega = -60^\circ/\text{sec}$ und konzentrischer Arbeit bei $\omega = 60^\circ/\text{sec}$ Kniewinkelgeschwindigkeit. Da unter exzentrischen Bedingungen mehr Kraft

entwickelt werden kann, werden Werte größer als 100% erwartet.

kkV Dieser Wert gibt das konzentrisch-konzentrische Arbeitsverhältnis in Prozent an. Es wird der Quotient aus konzentrischer Arbeit bei $\omega = 180^\circ/\text{sec}$ zu konzentrischer Arbeit bei $\omega = 60^\circ/\text{sec}$ gebildet. Gemessen wird also der Unterschied zwischen schneller und langsamer Muskelaktionsbedingung. Da die Athleten unter langsamen Bedingungen größere Kräfte entwickeln können, werden in der Theorie Werte kleiner als 100% erwartet.

Eine Übersicht über die Ausprägungen und die Anzahl der fehlenden Werte bei diesen Variablen findet sich in Tabelle 2.4.

Variablenname	Variablenbeschreibung	Ausprägungen	Einheit	NA's
A60	Konzentrische Arbeit bei 60° mittlerem Kniewinkel	Spr. 102 – 447	J	188
		Kom. 108 – 474		306
A180	Konzentrische Arbeit bei 180° mittlerem Kniewinkel	Spr. 84 – 361	J	189
		Kom. 84 – 381		189
ekV	exzentrisch-konzentrisches Arbeitsverhältnis	Spr. 4.5 – 196.8	%	187
		Kom. 95 – 188		306
kkV	Konzenstisch-konzentrisches Arbeitsverhältnis	Spr. 70.3 – 135.8	%	189
		Kom. 67 – 111		308
dynstat_60	dynamisch-statisches Verhältnis bei $\omega = -60^\circ/\text{sec}$	Spr. 4 – 165	%	188
		Kom. 63 – 128		306
dynstat_180	dynamisch-statisches Verhältnis bei $\omega = 180^\circ/\text{sec}$	Spr. 47.1 – 109.2	%	189
		Kom. 47 – 87		189

Tabelle 2.4.: Dynamische Schnellkraftdiagnostik sitzend an der IsoMed 2000

2.3.4. Abgeleitete Variablen

Zusätzlich zu den bisher beschriebenen Variablen enthalten die Originaldatensätze auch noch einige Variablen, die aus den anderen Variablen abgeleitet werden. Dazu gehört das in Abschnitt 1.1.2 angesprochene Kraftdefizit. Diese und andere abgeleitete Variablen sind in Tabelle 2.5 beschrieben. Bei der Variablen **DJ32_Faktor** sprechen große Werte für eine gute Leistung (hohe Sprunghöhe und geringe Bodenkontaktzeit), bei der Kraftdifferenz **Kdif** kann nicht direkt von den Werten auf die Leistung geschlossen werden. Im Frühjahr ist die Kraftdifferenz verhältnismäßig hoch, im Sommer bei austrainierten Athleten erwartungsgemäß nahe Null.

Variablenname	Variablenbeschreibung	Ausprägungen	Einheit	NA's
DJ32_Faktor	Sprunghöhe in m / Bodenkontaktzeit in s beim DJ	Spr. 0.77 – 3.53	$\frac{\text{m}}{\text{s}}$	47
		Kom. 0.61 – 3.97		47
F_re170_[bb,re,li]	Maximalkraft bei 70° Kniewinkel / Körpergewicht	Spr.	$\frac{\text{N}}{\text{kg}}$	bb: 17.9 – 40.5
				re: 10.3 – 22.7
				li: 9.9 – 22.1
		Kom.		bb: 18.1 – 36.5
				re: 8.8 – 20.2
Kdif	Kraftdifferenz $\text{Kdif} = \frac{(\text{Ekz-Fmax70_bb})}{\text{Fmax70_bb}} \cdot 100$	Spr. -24 – 20.5	%	89
		Kom. -27.3 – 7.8		234

Tabelle 2.5.: Übersicht der abgeleiteten Variablen im Originaldatensatz zur Leistungsdiagnostik

Hinweis: Vorzeichenänderung

Bei den Variablen DJ32_KT und `relKraftanstieg_[bb,re,li]` gilt, dass ein kleiner Wert für eine gute Leistung des Athleten spricht. Um die Interpretation im weiteren Verlauf der Arbeit zu erleichtern, wurden die Vorzeichen der genannten Variablen umgedreht, so dass große Werte für eine gute Leistung des Athleten sprechen. Die Variable DJ32_Faktor berechnet sich weiterhin aus den originalen Werten von DJ32_KT, die hier im Nenner eingehen und so bereits die richtige Wirkungsrichtung hat. Diese Änderung sollte insbesondere bei den Abschnitten 2.5.4, 2.5.6 und 3.2 berücksichtigt werden.

2.4. Analyse fehlender Werte

Vor der eigentlichen Deskription der Variablen der Leistungsdiagnostik wird die Problematik der fehlenden Werte im Datensatz untersucht. Der Anteil der fehlenden Werte bei den Variablen ist in Abbildung 2.3 dargestellt. Besonders bei den dynamischen Schnellkrafttests und der beidbeinigen Maximalkraft bei 90° Kniewinkel liegen für fast die Hälfte aller Messungen keine Werte vor. Die Gründe hierfür können Verletzungen sein, die es dem Athleten nicht gestattet haben, diese Übung durchzuführen oder dieser Parameter wurde generell nicht bei diesem Athleten überprüft. Die dynamischen Schnellkrafttests werden bei den Nordischen Kombinierern nicht vom C-Kader durchgeführt und bei den Skispringern von den D/C-Kader Athleten nicht vorgenommen. Dies erklärt die Zahl der fehlenden Messungen. Da in diesen Kadern besonders die jungen Athleten sind, muss man berücksichtigen, dass die dynamischen Schnellkrafttests eher von älteren, erfahrenen Athleten durchgeführt wurden.

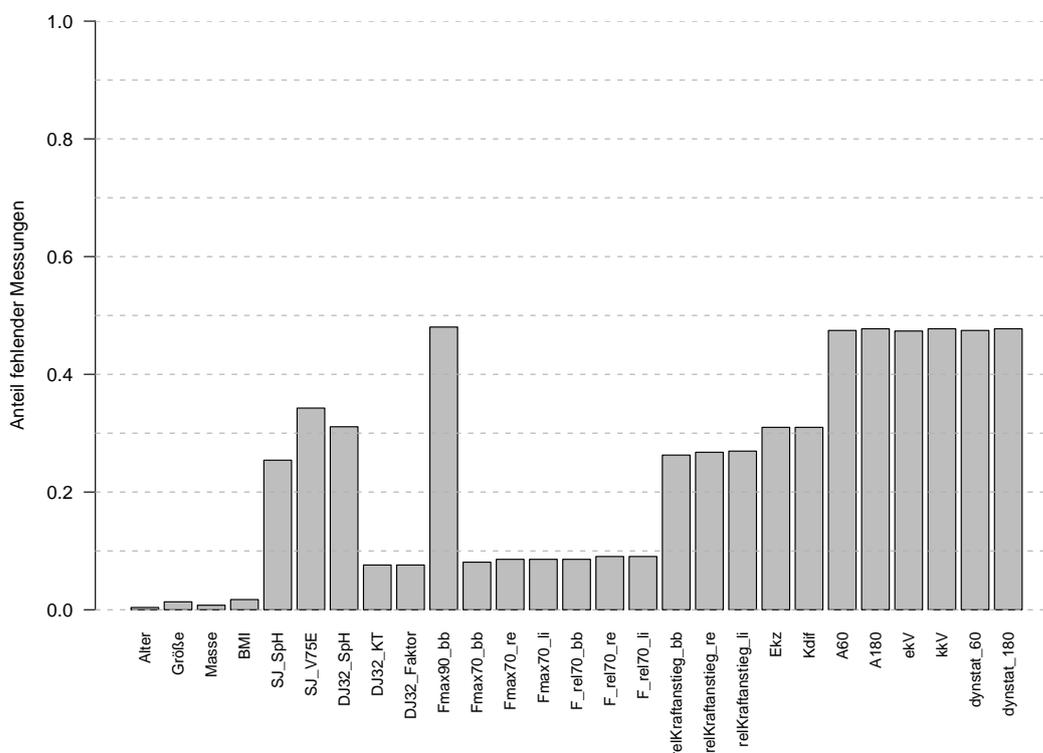


Abbildung 2.3.: Anteil der fehlenden Werte der Variablen im Datensatz zur Leistungsdiagnostik.

In Tabelle 2.6 ist der Anteil der fehlenden Messungen nach dem Untersuchungsjahr gegliedert dargestellt. Es sind nicht alle Parameter dargestellt, sondern nur erhobene Parameter und beidbeinige Messungen. Bei abgeleiteten bzw. einbeinigen Messungen sind die Anteile mit den ursprünglichen

bzw. beidbeinigen Messungen identisch. Bei einigen Sprungkraftparametern (SJ_SpH, SJ_V75E und DJ32_SpH) fehlen bis 2005 sehr viele Werte. Dies liegt daran, dass hier noch größtenteils nach einer anderen Methode gemessen wurde und diese Ergebnisse nicht vergleichbar sind (vgl. Abschnitt 2.3.2). Für die Bodenkontaktzeit DJ32_KT fehlen über die Jahre hinweg nur wenige Werte. Für die Maximalkraftdiagnostik bei 90° Kniewinkel (entsprechend Variable Fmax90_bb liegen erst Werte für die letzten drei Jahre vor. Für die Variable Fmax70_bb gibt es über den ganzen Untersuchungszeitraum hinweg nur wenige fehlende Messungen. Bei den 6 Variablen zur Schnellkraft A60, A180, ekV, kkF, dynstat_60 und dynstat_180 gibt es untereinander keine Unterschiede bei dem Anteil der fehlenden Werte. Vor dem Jahr 2003 liegen jeweils keine Werte vor. Weitere Untersuchungen können sich daher hauptsächlich nur auf die aktuelleren Jahre beziehen.

	2000	2001	2002	2003	2004	2005	2006	2007	2008
SJ_SpH	1.00	1.00	0.94	0.93	0.69	0.28	0.00	0.00	0.12
SJ_V75E	1.00	1.00	1.00	1.00	1.00	0.51	0.00	0.02	0.12
DJ32_SpH	1.00	1.00	0.94	0.95	0.70	0.40	0.06	0.07	0.15
DJ32_KT	0.00	0.12	0.03	0.10	0.19	0.05	0.00	0.07	0.15
Fmax90_bb	1.00	1.00	1.00	1.00	1.00	0.99	0.04	0.20	0.12
Fmax70_bb	0.00	0.00	0.00	0.02	0.02	0.16	0.02	0.13	0.12
relKraftanstieg_bb	1.00	1.00	1.00	1.00	0.58	0.16	0.02	0.15	0.12
Ekz	1.00	1.00	1.00	1.00	0.58	0.18	0.07	0.25	0.20
Schnellkraftparameter	1.00	1.00	1.00	1.00	0.71	0.41	0.34	0.42	0.26

Tabelle 2.6.: Anteil der fehlenden Werte ausgewählter Variablen der Leistungsdiagnostik im Jahresverlauf

Zur Untersuchung möglicher Fehlendmuster wird wie folgt vorgegangen. Für jeden Athleten werden die beiden Zustände fehlender Wert (NA) oder vorliegender Wert (\neg NA) bei einer Variable betrachtet. Nun wird für alle Athleten alle vier möglichen Übergänge gezählt und daraus für jede Variable die Übergangsmatrix dieser beiden Zustände konstruiert (Lawler, 2006). Diese werden mit einem Chi-Quadrat-Test untersucht und es ergibt sich für die meisten Variablen, dass der Übergang von einem fehlenden zu einem fehlenden Wert statistisch signifikant häufig ist. Eine Beschreibung des Chi-Quadrat-Testes findet sich z.B. in Kapitel 11 von Fahrmeir u. a. (2004). Betrachtet wird nun die Übergangsmatrix für alle Variablen gemeinsam, also die Summe aller Übergänge bei allen Variablen im Datensatz. Hier wird noch durch die entsprechenden Zeilensummen der Matrix dividiert, um Wahrscheinlichkeiten für die Übergänge zu erhalten.

		zu	
		NA	\neg NA
von	NA	0.79	0.21
	\neg NA	0.09	0.91

Tabelle 2.7.: Relative Häufigkeit der Übergänge der Zustände fehlend (NA) und nicht fehlend (\neg NA) summiert für alle Parameter (Übergangsmatrix).

In Tabelle 2.7 fällt die linke obere Zelle der Matrix auf, die angibt, wie wahrscheinlich ein fehlender Wert bei einer Messung eines Athleten ist, wenn auch bei dessen vorheriger Messung ein Wert fehlt. Dies geschieht in fast 80% der Fälle, wo hingegen auf einem nicht fehlenden Wert anschließend nur in 9% der Fälle ein fehlender Wert folgt. Dies spricht dafür, dass bestimmte Parameter bei gewissen Athleten systematisch nicht erhoben wurden und nicht von zufälligen Fehlendmechanismen auszugehen ist.

Generell bieten sich nun zwei Möglichkeiten für den Umgang mit fehlenden Werten an:

1. Einfügen von Werten zum Erhalt eines vollständigen Datensatzes (Imputationsverfahren)
2. Entfernen der Fälle mit fehlenden Werten aus der Analyse (Eliminierungsverfahren)

In der Literatur gibt es zahlreiche Vorschläge für Imputationsverfahren (Little u. Rubin, 1987). Allerdings setzen diese meist voraus, dass als Fehlmuster der Fall *Missing Completely at Random* (MCAR) vorliegt, das Fehlen eines Wertes also vollständig unabhängig von allen anderen Größen ist. Aufgrund der festgestellten Fehlmuster scheint dies hier nicht der Fall zu sein. Außerdem bringen Imputationsverfahren keine neue Information in die Analyse, da nur die bereits vorhandene Information zur Vorhersage der fehlenden Werte genutzt wird. Für Longitudinaldaten wird in Diggle u. a. (2002) vorgeschlagen, den letzten vorhandenen Wert eines Individuums bei fehlenden Werten fortzuschreiben.

Hier werden Eliminierungsverfahren eingesetzt und nur die gültigen Werte analysiert. Dabei kann zwischen einer *Complete Case Analyse* und einer *Available Case Analyse* unterschieden werden. Bei erster werden für alle Analysen nur die Beobachtungen berücksichtigt, die überhaupt keine fehlenden Werte haben. Dies hat den Vorteil, dass für alle Analysen die gleiche Grundgesamtheit vorliegt. Allerdings wird durch das Weglassen vieler Beobachtungen viel Information nicht genutzt. So verbleiben bei der Analyse der Leistungsdiagnostikdaten nur 326 vollständige Messungen bei 88 Athleten. Dies entspricht einem Datenverlust von etwa 69%.

Bei einer *Available Case Analyse* werden für die Analysen die Beobachtungen einbezogen, die bei den verwendeten Variablen keine fehlenden Werte haben. Damit variiert die Stichprobe von Analyse zu Analyse, aber es stehen insgesamt mehr Beobachtungen und damit mehr Information zur Verfügung. Bei den einzelnen Untersuchungen wird daher gegebenenfalls die Stichprobe angegeben, auf die sich die Untersuchung bezieht.

Im nächsten Abschnitt werden nun die im Rahmen der Leistungsdiagnostik erhobenen Variablen deskriptiv untersucht. Anschließend folgt eine Diskussion der bisher gewonnenen Ergebnisse.

2.5. Explorative Analysen zur Leistungsdiagnostik

Im Rahmen der Deskription werden zunächst die univariaten Verteilungen der Variablen betrachtet und anschließend bivariate Verteilungen und Korrelationen in den Daten untersucht. Da die Deskription aller Variablen des Datensatzes zu umfangreich ist, werden teilweise nur ausgewählte Variablen beschrieben. Sofern keine weiteren Anmerkungen zur Datengrundlage gemacht werden, bezieht sich die Deskription auf alle vorliegenden Werte von Skispringern und Nordischen Kombiniern (*Available-Case-Analyse*). Am Ende dieses Abschnitts folgt eine Diskussion der Ergebnisse der explorativen Analysen, die als Motivation für die Verwendung der in Kapitel 3 beschriebenen statistischen Methodik dient.

2.5.1. Univariate Verteilungen

Zunächst werden die univariaten Verteilungen einzelner Variablen untersucht und dabei die erzielten Werte von Skispringern und Nordischen Kombiniern verglichen. Das Skalenniveau aller im Rahmen der Leistungsdiagnostik gewonnenen Variablen ist metrisch, sodass für die Darstellung der Verteilung Darstellungsformen wie Boxplots oder Kerndichteschätzer gewählt werden. Für die Abbildungen werden dabei jeweils alle zur Verfügung stehenden gültigen Werte von Skispringern und Nordischen Kombiniern verwendet. In Abbildung 2.4 ist die Verteilung der Sprunghöhe beim Squat-Jump SJ_SpH dargestellt.

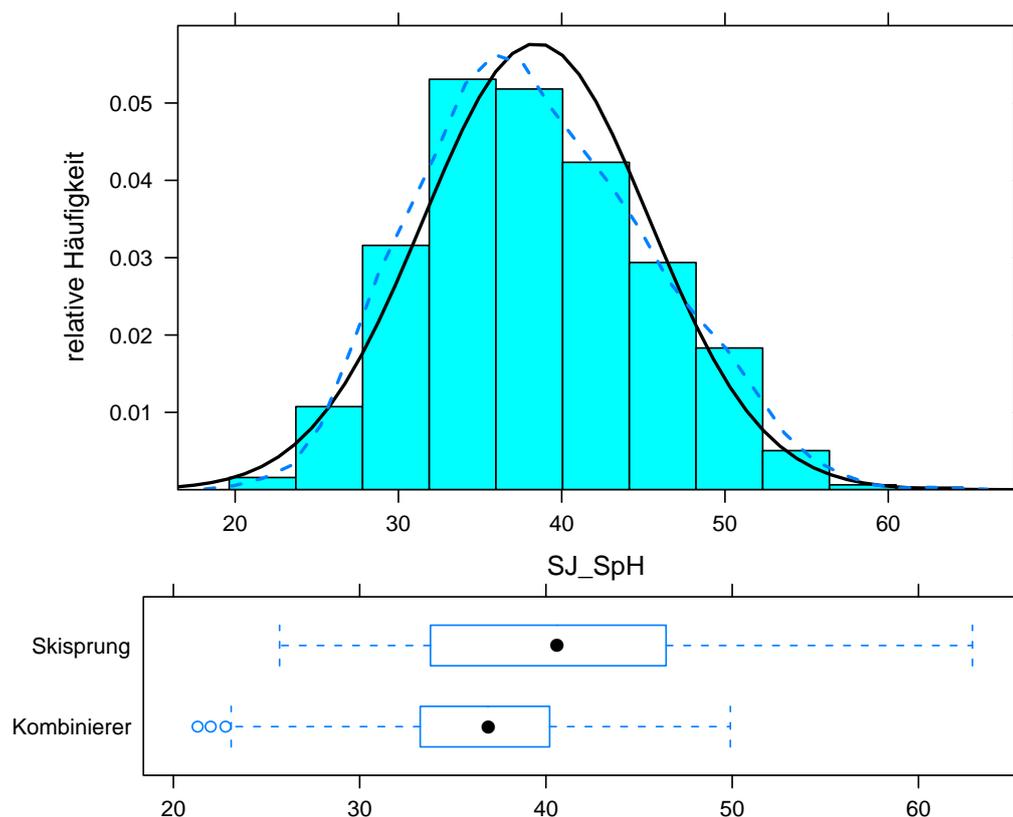


Abbildung 2.4.: Darstellung der Verteilung der Variable **SJ_SpH**: Oben ist ein Histogramm der beobachteten Werte dargestellt und als durchgezogene Linie die Dichte der Normalverteilung mit dem Mittelwert (38.45) und der empirischen Standardabweichung (6.92) der Variable, die gestrichelte Linie gibt den Kerndichteschätzer an ($n=775$).

Unten: Fünf-Punkte-Darstellung (Boxplot) der Variable **SJ_SpH** aufgeteilt nach Sportarten (Skispringer: $n=388$, Nordische Kombinierer: $n=387$).

Aus Abbildung 2.4 können zwei für die weiteren statistischen Analysen in Kapitel 3 wichtige Beobachtungen gemacht werden. Zunächst kann die empirische Häufigkeitsverteilung mit der Dichtefunktion einer Normalverteilung verglichen werden. Der Kerndichteschätzer (bzw. das Histogramm als Spezialfall eines Kerndichteschätzers) und die Dichtefunktion der Normalverteilung zeigen einen ähnlichen Verlauf. Insbesondere befinden sich die Maxima an der gleichen Stelle und die empirische Verteilung besitzt gleiche Masse an den Rändern des Wertebereichs.

Alternativ kann die Verteilungsannahme mit sogenannten *Normal-Quantil-Plots* (vgl. Fahrmeir u. a., 2004, S.95) überprüft werden. Bei einem Normal-Quantil-Plot wird die empirische Verteilung der Variablen mit der einer Normalverteilung anhand der Quantile der Verteilungen verglichen. In Abbildung 2.5 sind die Normal-Quantil-Plots einiger ausgewählter Variablen der Leistungsdiagnostik dargestellt. Liegen die Werte annähernd auf der eingezeichneten Geraden, so sind die Daten approximativ normalverteilt. Dies ist bei den dargestellten Variablen der Fall. Ein Verlauf des Normal-Quantil-Plots in Abbildung 2.5 bei der Variable **ekV** spricht für eine symmetrische Verteilung, die eine stärkere Wölbung als die Normalverteilung besitzt. Das heißt, dass es mehr extremere Werte gibt, als man unter der Normalverteilungsannahme vermuten würde. Insgesamt sind die Variablen symmetrisch verteilt, jedoch weisen einige (**relKraftanstieg_[bb,re,li]**, **Ekz**, **kkV** (nicht dargestellt) und **ekV**) stärkere Wöl-

bungen als Normalverteilungen auf. Im Folgenden wird jedoch die Normalverteilungsannahme für alle Variablen als erfüllt angesehen.

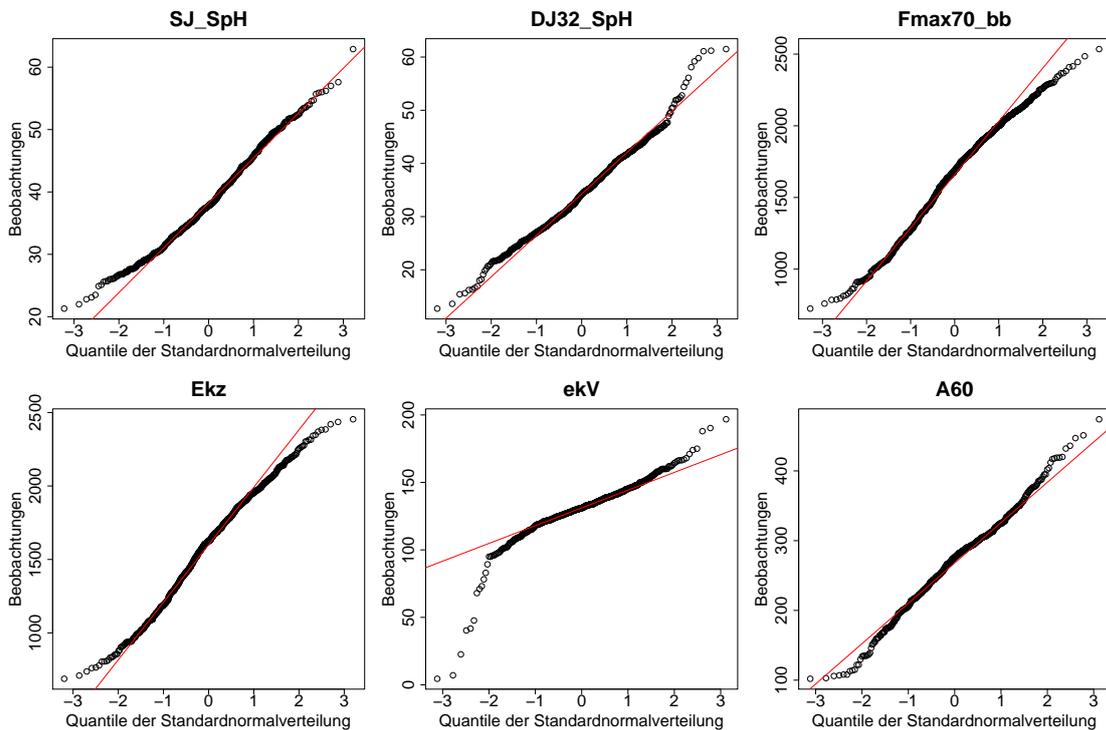


Abbildung 2.5.: Normal-Quantil-Plots der Variablen SJ_SpH, DJ32_SpH, Fmax70_bb, Ekz, ekV und A60

2.5.2. Entwicklung über die Zeit

Die Teilnahme einzelner Athleten an einer Längsschnittstudie lässt sich gut durch einen sogenannten *Traceplot* darstellen. Hier wird für einzelne Sportler der Verlauf einer Variable gegen das Datum der Messung dargestellt. In Abbildung 2.6 ist der Verlauf der Sprunghöhe beim Squat Jump und Drop Jump von 20 zufällig ausgewählten Athleten dargestellt. Man sieht, dass für viele Sportler insgesamt nur wenige Werte vorliegen. Die Skispringer nahmen im Mittel in 2.28 (Maximum: 7) Jahren an der Diagnostik teil und die Nordischen Kombinierer in 3.14 (Maximum: 9) Jahren. Dabei wurden die Athleten meist in aufeinanderfolgenden Jahren leistungsdiagnostisch untersucht. Unterbrechungen von mindestens einem Jahr (z.B. aufgrund von Verletzungen) waren eher selten (z.B. bei Athlet 36). Außerdem sieht man, dass die Messungen nicht in gleichen Abständen stattgefunden haben, sondern in unregelmäßigen Mustern.

Außerdem kann in Abbildung 2.6 die gleichzeitige Entwicklung der Ergebnisse beim Squat Jump und beim Drop Jump verfolgt werden. So kann man sehen, dass es Athleten gibt, die stets bessere Ergebnisse im Squat Jump als im Drop Jump haben. Beispiele hierfür sind die Athleten 9 und 179, während beispielsweise Athlet 111 teilweise bei Squat Jump und teilweise beim Drop Jump größere Höhen erreicht. Bei dem Athleten 64 kann man über die Zeit eine parallele Leistungssteigerung bei beiden Sprungkrafttests beobachten, wo hingegen bei Athlet 9 gegenläufige Trends zu beobachten sind. Insgesamt kann man schon beim zeitlichen Vergleich zweier Variablen feststellen, dass es große Unterschiede beim Leistungsniveau der Athleten gibt. Während bei manchen Athleten ein positiver Zusammenhang beider Variablen zu sehen scheint, ist er bei anderen gegenläufig. Vor 2004 fehlen meist die Werte, da hier nur Ergebnisse der alten Messmethode vorliegen (vgl. auch Tabelle 2.6 auf

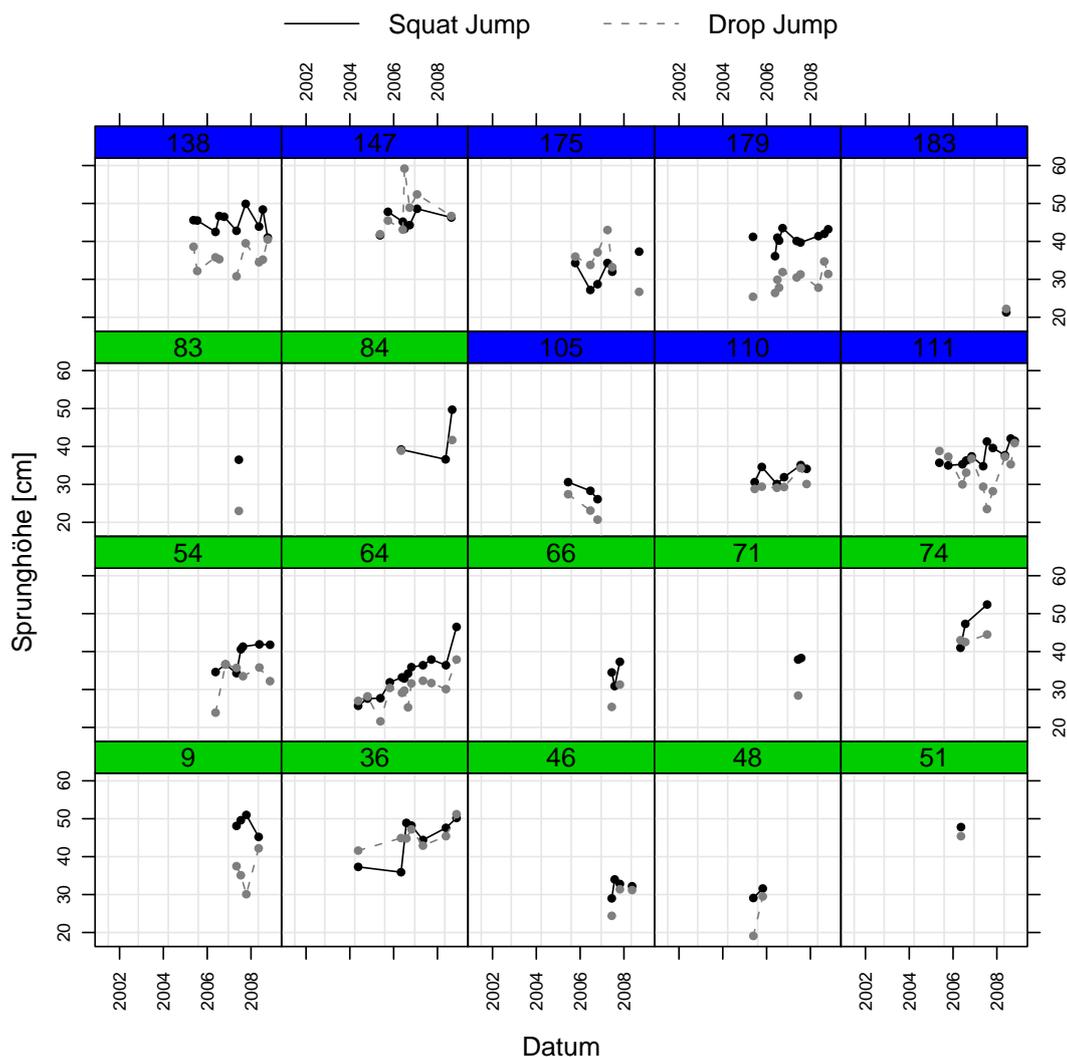


Abbildung 2.6.: Traceplot der Sprunghöhe beim Squat Jump und Drop Jump für 20 zufällig ausgewählte Athleten. Die Nummer über jeder Graphik gibt die Identifikationsnummer jedes Athleten an. Blau hinterlegte Felder kennzeichnen dabei Messungen von Nordischen Kombinierern, grüne die von Skispringern. Direkt aufeinander folgende Messungen sind mit Linien verbunden, andernfalls liegt dazwischen mindestens eine Teilnahme an der Diagnostik, bei der allerdings keine Sprunghöhe erhoben wurde.

Seite 17).

Bei dieser Darstellungsform wird noch nicht das Alter der Athleten zu den jeweiligen Messzeitpunkten betrachtet. Man sieht jedoch bei den meisten Athleten tendenziell einen Anstieg der Messergebnisse über die Zeit. Dies spricht für einen Alterseffekt bei den Variablen SJ_SpH und DJ32_SpH. Ob dieser Alterseffekt bei allen Variablen vorliegt, wird im folgenden Abschnitt betrachtet.

2.5.3. Altersspezifische Leistungsentwicklung

Die Athleten im Datensatz zur Leistungsdiagnostik sind zwischen 12 und 31 Jahre alt, was fast die gesamte Altersspannweite der sportlichen Laufbahn eines Skispringers oder Nordischen Kombinierers abdeckt. Natürlich sind große Unterschiede zwischen sehr jungen Athleten, die sich noch am Anfang ihrer sportlichen Karriere befinden, und solchen, die gerade auf ihrem sportlichen Höhepunkt sind,

und den älteren, erfahrenen Athleten zu erwarten. Hier spielen sowohl die körperliche Entwicklung als auch die Erfahrung eine Rolle. Zudem unterscheiden sich die Trainingsinhalte und Schwerpunkte. So steht das Krafttraining bei jungen Athleten nicht im Mittelpunkt des Trainings (Wolf, 1997). Bei einem Vergleich der Ergebnisse der Leistungsdiagnostik zwischen den Athleten ist es daher wichtig, das Alter der Athleten zu berücksichtigen. Im Skisprung hat sich die Klasseneinteilung in Tabelle 2.8 bewährt.

Altersklasse	1	2	3	4	5
Alter in Jahren	10-12	13-14	15-16	17-18	ab 19
Kaderzugehörigkeit	D/1-D/2	D/3-D/4	D/C-Kader	C-Kader	A/B-Kader
Trainingsform	Grundlagen	Aufbau- training	Leistungs- training	Beginn des Hochleistungs- trainings	Hochleistungs- training
Trainingseinheiten/Woche	4-5	6-7	8-9	9-11	>11
Beobachtungen im Datensatz					
Skispringer	1	78	157	115	91
Nordische Kombinierer	1	74	118	121	279

Tabelle 2.8.: Altersgemäße Strukturierung der Athleten nach Altersjahr (vgl. Wolf, 1997, S.36). Die Trainingsform und die mittlere Anzahl der Trainingseinheiten je Woche ist aus Wolf (1997) entnommen.

Mit steigendem Alter erzielen die Athleten bei den meisten Variablen bessere Ergebnisse. Der stärkste Leistungszuwachs ist ungefähr bis 18 oder 19 Jahre zu beobachten. Bei manchen Variablen (z.B. relative Maximalkraft) fallen die Werte bei den etwa 30-jährigen wieder etwas ab. Die Ergebnisse zwischen den Athleten eines Altersjahrgangs schwanken jeweils stark. In allen Altersjahren gibt es Athleten mit deutlich über- und unterdurchschnittlichen Ergebnissen.

Die Ergebnisse eines 15- und eines 25-jährigen Skispringers sind also nicht ohne Weiteres zu vergleichen. Möchte man beide bezüglich ihres Leistungsniveaus vergleichen, so kann man untersuchen, wie sie relativ zu Athleten ihres Altersjahrgangs abschneiden. Dieser Ansatz wird im folgenden Abschnitt und bei dem Modell in Kapitel 3 wieder aufgegriffen.

2.5.4. Profilbild einzelner Athleten

Das Konzept der Normierung der Ergebnisse bezüglich einer vergleichbaren Referenzgruppe ist in der Sportwissenschaft und Statistik weit verbreitet. Die Normierung entspricht dabei der Bildung sogenannter *Z-Werte*, die hier in einem kurzen Exkurs vorgestellt werden. Mit den Z-Werten sollen die einzelnen Ergebnisse einer Variablen von Athleten mit unterschiedlichen Voraussetzungen vergleichbar gemacht werden.

Exkurs: Z-Werte

Zur Bildung der Z-Werte wird für jede Beobachtung $i = 1, \dots, n$ eine Referenzgruppe von Beobachtungen $r(i)$ anhand von inhaltlichen Gesichtspunkten definiert. Setzt sich die Referenzgruppe nur aus Beobachtungen des Datensatzes zusammen, so gilt $r(i) \subseteq \{1, \dots, n\}$. Dann werden Mittelwert und Standardabweichung der Variable für alle Beobachtungen dieser Referenzgruppe $r(i)$ gebildet. Der Z-Wert z_i aus einer Beobachtung x_i bei dieser Variablen ergibt sich durch Subtraktion des Mittelwertes und Division der Standardabweichung der Werte der Referenzgruppe, d.h.

$$z_i = \frac{x_i - \mu_{r(i)}}{\sigma_{r(i)}}, \quad (2.1)$$

wobei $\mu_{r(i)}$ der Mittelwert und $\sigma_{r(i)}$ die Standardabweichung der Variable bei allen Beobachtungen der

Referenzgruppe ist. Der Z-Wert gibt also an, um wie viele Standardabweichungen nach oben bzw. unten die Beobachtung vom Mittelwert der Referenzgruppe entfernt ist. Je größer der Betrag des Z-Wertes, um so extremer ist die Beobachtung. Wenn die x_i normalverteilt sind, so liegen mehr als 99% der Z-Werte zwischen -3 und 3.

Als Referenzgruppe $r(i)$ für eine Beobachtung i werden hier alle Beobachtungen gewählt, bei denen der Athlet genauso alt oder ein Jahr jünger bzw. älter ist, jeweils bezogen auf das Geburtsjahr. Dies entspricht dem Konzept der *gleitenden Mittelwerte*. In Abbildung 2.7 ist der Effekt der Normierung für die Variable SJ_SpH dargestellt. In der linken Darstellung ist ein Ansteigen der Werte mit zunehmendem Alter zu beobachten. Bei den transformierten Z-Werten ist keine Struktur der Werte bezüglich des Alters zu erkennen.

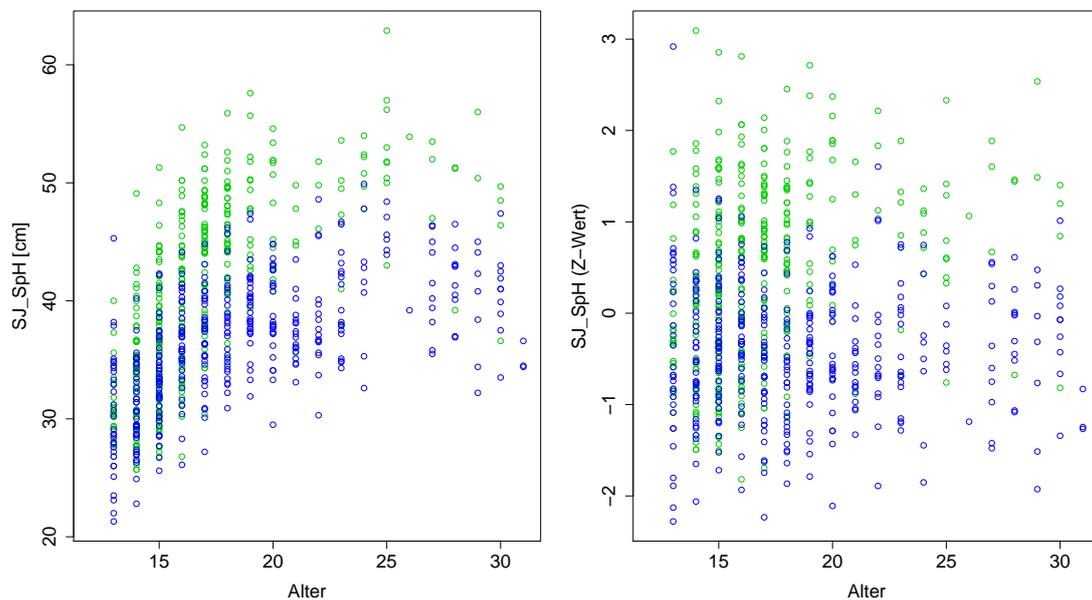


Abbildung 2.7.: Streudiagramm der Variablen SJ_SpH und der Z-transformierten Sprunghöhe gegen das Alter der Athleten zum Messzeitpunkt für alle Athleten mit verfügbaren Werten. Die grünen Kreise kennzeichnen Skispringer, die blauen Nordische Kombiniierer.

Stellt man die Z-Werte aller Variablen für eine Beobachtung eines Athleten dar, so erhält man ein sogenanntes *Profilbild* eines Athleten. In Abbildung 2.8 sind die Profile von je zwei Beobachtungen bei einem Skispringer und einem Nordischen Kombiniierer dargestellt. Man sieht, dass die Ergebnisse des Nordischen Kombiniierers jeweils relativ im Mittel zu den Ergebnissen seiner Referenzgruppe liegen. Der Skispringer hat bei einigen Variablen deutliche Stärken gegenüber seiner Referenzgruppe.

Die gezeigten Z-Werte sind auf das Alter der Athleten normiert. Daneben könnte es noch weitere Kovariablen geben, die einen Einfluss auf die diagnostischen Variablen haben. Aus sportwissenschaftlicher Sicht wird dem Gewicht und dem BMI des Athleten Bedeutung für die Ergebnisse bei der Leistungsdiagnostik beigemessen. Allerdings zeigt sich, dass dieser Einfluss bereits durch die Altersadjustierung berücksichtigt wird, wie in Abbildung 2.9 dargestellt. Demnach ist der Zusammenhang von Variablen der Leistungsdiagnostik mit *Masse* oder BMI ähnlich zu dem Zusammenhang mit dem Alter. Während man bei den nicht normierten Werten der Variablen SJ_SpH noch einen positiven Einfluss des Körpergewichts und des BMI vorliegt, so verschwindet dieser bei der Betrachtung der Z-Werte. Eine Normierung der Ergebnisse auf das Körpergewicht oder den BMI zusätzlich zu der Altersnormierung

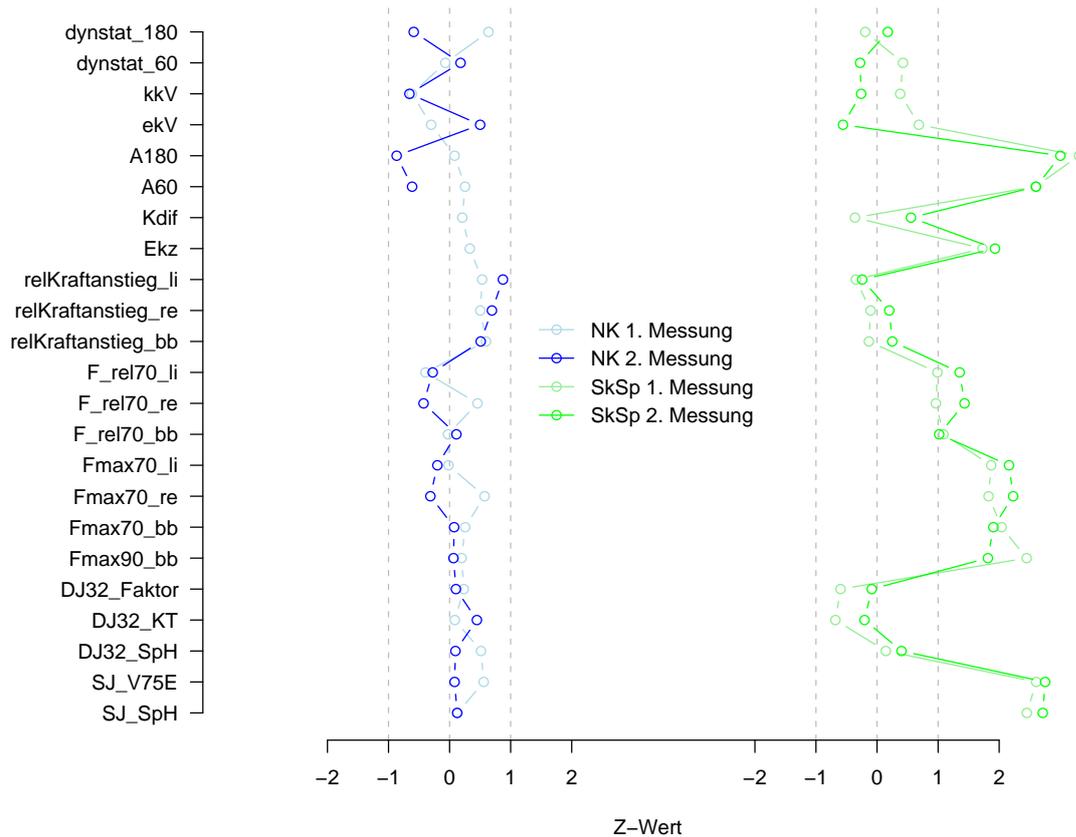


Abbildung 2.8.: Profilbild aller Variablen von je zwei Beobachtungen eines Nordischen Kombinierers (blau) und eines Skispringers (grün). Die heller eingefärbte Linie kennzeichnet dabei die frühere Messung des jeweiligen Athleten (es liegt jeweils ca. 1 Jahr Abstand zwischen den Messungen). Dabei sprechen bei allen Variablen bis auf *Kdif*, *ekV*, *kkV*, *dynstat_[60,180]* (hier kann aus inhaltlichen Gründen keine eindeutige Richtung angegeben werden) größere Z-Werte für eine bessere Leistung im Vergleich zur Referenzgruppe

ist also nicht nötig. Dies liegt daran, dass Körpergewicht und BMI mit dem Alter zunehmen.

2.5.5. Vergleich von Skispringern und Nordischen Kombinierern unter Berücksichtigung des longitudinalen Datendesigns

In diesem Abschnitt werden Skispringern und Nordischen Kombinierern mit *linearen gemischten Modellen* verglichen. Das lineare gemischte Modell ist eine Erweiterung des linearen Modells für teilweise abhängige Beobachtungen, wie z.B. Longitudinaldaten. Die Modellklasse der linearen gemischten Modelle wird hier in einem Exkurs vorgestellt, der sich an Kapitel 6 von Fahrmeir u. a. (2007) orientiert. Dort finden sich auch weitere Verteilungsannahmen und Modellvoraussetzungen des linearen gemischten Modells, die hier nicht dargestellt werden.

Exkurs: lineare gemischte Modelle

Lineare gemischte Modelle erweitern das lineare Modell um sogenannte *individuen-spezifische Effekte* zur Modellierung von Abhängigkeiten bei wiederholten Beobachtungen. Für jedes Individuum $i = 1, \dots, N$ (z.B.

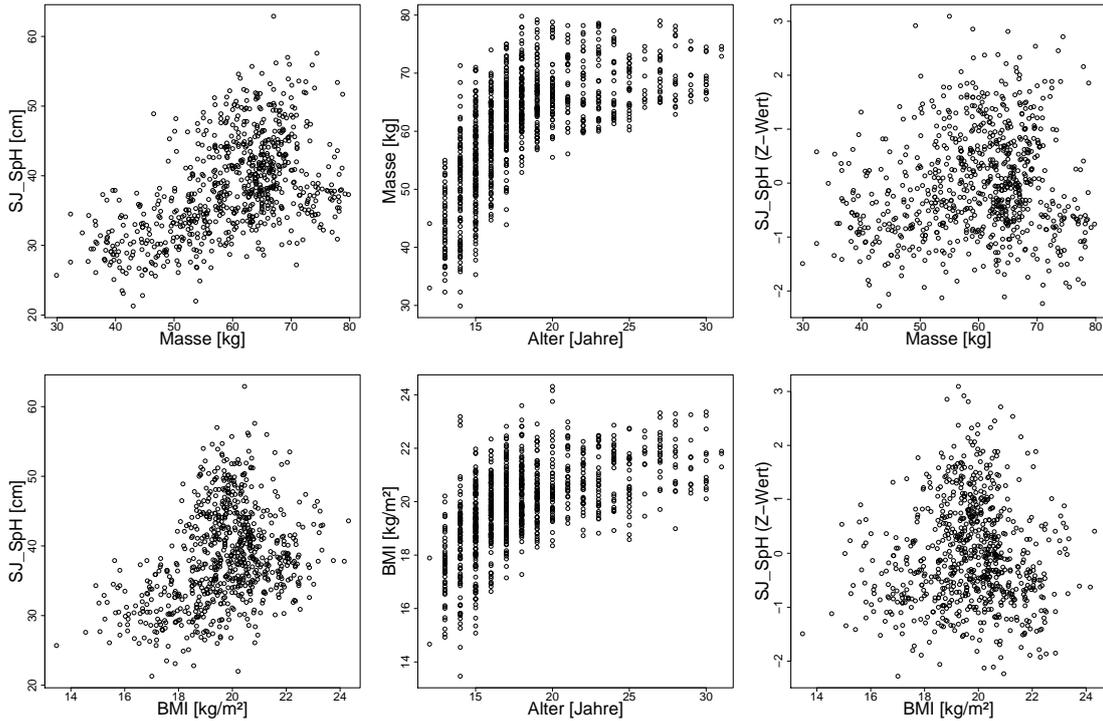


Abbildung 2.9.: Oben links: Streudiagramm von SJ_SpH gegen Masse, oben Mitte: Streudiagramm von Alter gegen Masse, oben rechts: Streudiagramm der altersnormierten Z-Werte von SJ_SpH gegen Masse. Unten links: Streudiagramm BMI SJ_SpH gegen BMI, unten Mitte: Streudiagramm von Alter gegen BMI, unten rechts: Streudiagramm der altersnormierten Z-Werte von SJ_SpH gegen BMI.

ein Athlet) liegen wiederholte Beobachtungen an unterschiedlichen Zeitpunkten $t = 1, \dots, n_i$ vor, d.h.

$$\underbrace{(y_{i1}, \dots, y_{it}, \dots, y_{in_i})}_{\mathbf{y}_i^T}, \underbrace{(x_{i1}, \dots, x_{it}, \dots, x_{in_i})}_{\mathbf{X}_i^T}, \quad (2.2)$$

wobei n_i die Anzahl der Beobachtungen von Individuum i kennzeichnet und \mathbf{y}_i den Vektor der Beobachtungen der Zielgröße und \mathbf{x}_{it} der p -dimensionale Vektor der Kovariablen für Individuum i zum Zeitpunkt t ist. Diese werden in der $n_i \times p$ -Matrix \mathbf{X}_i zusammengefasst, die die p Kovariablen zu allen n_i Zeitpunkten enthält. Die einzelnen Messwiederholungen bei einem Individuum ähneln sich in der Regel mehr als Beobachtungen bei verschiedenen Athleten. Darum wird das lineare Modell um individuen-spezifische Effekte erweitert. Das lineare gemischte Modell mit p festen Effekten und q individuen-spezifischen Effekten für Longitudinaldaten der Form (2.2) lautet

$$y_{it} = \mathbf{x}_{it}^T \boldsymbol{\beta} + \mathbf{u}_{it}^T \boldsymbol{\gamma}_i + \epsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, n_i, \quad (2.3)$$

wobei \mathbf{x}_{it} der Vektor der festen Kovariablen ist und $\boldsymbol{\beta}$ der p -dimensionale Vektor mit den festen Kovariablenn-effekten wie im linearen Modell. Hinzu kommt \mathbf{u}_{it} als der q -dimensionale Vektor der individuen-spezifischen Kovariablen und $\boldsymbol{\gamma}_i$ der q -dimensionale Vektor der individuen-spezifischen Kovariablenn-effekte für Individuum i . Mit dem Skalar ϵ_{it} wird der Fehlerterm bezeichnet, wobei $\epsilon_{it} \sim N(0, \sigma^2)$ angenommen wird. Die festen Kovariablenn-effekte gelten für alle Individuen und werden daher auch *Populationseffekte* genannt, während die individuen-spezifischen Effekte *individuen-spezifisch* sind und die Berücksichtigung der longitudinalen Datenstruktur ermöglichen. Die individuen-spezifischen Kovariablen können dabei auch ein Teil der festen Kovariablen sein. Das Modell (2.3) lässt sich auch als Beobachtungsmodell für ein Individuum schreiben

als

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{U}_i \boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i,$$

wobei $\mathbf{U}_i = (\mathbf{u}_{i1}, \dots, \mathbf{u}_{in_i})^T$ eine $n_i \times q$ -Matrix ist und $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \dots, \epsilon_{in_i})^T$ ein p -dimensionaler Vektor mit $\epsilon_i \sim N(\mathbf{0}, \sigma^2 I_{n_i})$, wobei I_{n_i} die n_i -dimensionale Einheitsmatrix bezeichnet.

Zur Schätzung solcher Modelle sind lineare Modelle nicht geeignet, da im Verhältnis zur Zahl der Beobachtungen eine zu hohe Anzahl an Parametern zu schätzen ist. Um die Identifizierbarkeit des Modells zu gewährleisten, wird angenommen, dass die individuenspezifischen Effekte einer bekannten Verteilung folgen (man spricht daher auch von *zufälligen* Effekten). Zumeist wird

$$\boldsymbol{\gamma}_i \sim N(\mathbf{0}, \mathbf{D}).$$

angenommen. Die $q \times q$ -Matrix \mathbf{D} wird dabei anhand der Daten geschätzt. Sie kann dabei mit gewissen Restriktionen belegt werden, z.B. dass \mathbf{D} die Struktur einer Diagonalmatrix haben soll. Inhaltlich entspricht das unkorrelierten individuenspezifischen Effekten. Zusätzlich nimmt man immer an, dass die individuenspezifischen Effekte verschiedener Individuen voneinander unabhängig sind.

Verbreitet sind sogenannte *Random-Intercept-Modelle*, die nur einen individuenspezifischen Intercept γ_{0i} modellieren, der die Abweichung vom globalen Intercept angibt (wegen $E(\gamma_{0i}) = 0$).

$$y_{it} = \mathbf{x}_{it}^T \boldsymbol{\beta} + \gamma_{0i} + \epsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, n_i. \quad (2.5)$$

Die individuenspezifischen Effekte werden häufig als Surrogat für unbeobachtete Kovariableneffekte interpretiert, die sich in einem unterschiedlichen Niveau der Individuen ausdrücken. Die Berücksichtigung solcher Effekte bewirkt eine höhere Schätzgenauigkeit des linearen gemischten Modells gegenüber dem linearen Modell. Die Schätzung des Modells (2.3) erfolgt mit Maximum-Likelihood bzw. restringierter Maximum-Likelihood-Schätzung, die in Fahrmeir u. a. (2007) beschrieben wird. Dabei werden die

- Populationseffekte $\boldsymbol{\beta}$
- zufälligen Effekte $\boldsymbol{\gamma}_i$
- und die Korrelationsmatrix \mathbf{D}

geschätzt. Die Software zur Schätzung der linearen gemischten Modelle ist im Anhang in Abschnitt B.1 beschrieben.

Die Verwendung von linearen gemischten Modellen ermöglicht die Betrachtung von Populationseffekten (z.B. der Unterschied zwischen Skispringern und Nordischen Kombinierern) unter der Berücksichtigung der Abhängigkeit der Messungen bei einem Athleten. Im Folgenden wird für die Daten zur Leistungsdiagnostik ein Random-Intercept-Modell der Form (2.5) angepasst, in dem für jeden Athleten ein unterschiedliches Niveau modelliert wird. Als fester Effekt wird die Variable **Sportart** aufgenommen, die wie folgt kodiert wird

$$\text{Sportart}_i = \begin{cases} 1 & \text{falls Athlet } i \text{ ein Skispringer ist} \\ 0 & \text{falls Athlet } i \text{ ein Kombinierer ist} \end{cases} \quad (2.6)$$

Dies entspricht der sogenannten *Dummy-Kodierung* (vgl. Fahrmeir u. a., 2007, S.81). Das Beobachtungsmodell für einen Athleten sieht nun wie folgt aus

$$\text{Variable}_i = \beta_0 + \beta_1 \cdot \text{Sportart}_i + \gamma_{0i} \cdot \text{Athlet}_i + \epsilon_i, \quad (2.7)$$

wobei der zufällige Effekt angibt, dass für jeden Athleten ein eigener Intercept modelliert wird. Nun sind

- Variable_i alterstandardisierte Z-Werte einer Variable der Leistungsdiagnostik für einen Athleten i (n_i -dimensionaler Vektor)
- β_0 der globale Intercept
- β_1 der Effekt der Sportart, d.h. der mittlere Unterschied von Skispringern gegenüber Nordischen Kombinierern unter allen Athleten adjustiert auf die individuenspezifischen Effekte
- γ_{0i} die Abweichung von Athlet i von β_0 über alle $j = 1, \dots, n_i$ (Niveau von Athlet i)
- ϵ_i der n_i -dimensionale Vektor der Fehlerterme für Athlet i

Die feste Kovariable `Sportart` ist dabei zeitkonstant. Zwar sind im Datensatz drei Sportler, die die Sportart gewechselt haben, die jedoch als unterschiedliche Athleten gezählt werden.

	Schätzwert	Standardfehler	n	unten	oben
SJ_SpH	0.76	0.12	775	0.53	0.99
SJ_V75E	0.68	0.12	683	0.43	0.92
DJ32_SpH	0.60	0.12	716	0.37	0.83
DJ32_KT	-0.41	0.13	960	-0.66	-0.17
DJ32_Faktor	0.23	0.13	960	-0.01	0.48
Fmax90_bb	0.34	0.14	539	0.07	0.61
Fmax70_bb	0.35	0.12	954	0.11	0.58
Fmax70_re	0.32	0.12	949	0.08	0.56
Fmax70_li	0.26	0.12	949	0.02	0.50
F_rel70_bb	0.63	0.12	949	0.40	0.87
F_rel70_re	0.54	0.12	944	0.30	0.79
F_rel70_li	0.53	0.12	944	0.29	0.77
relKraftanstieg_bb	0.19	0.12	765	-0.04	0.42
relKraftanstieg_re	0.10	0.10	760	-0.11	0.30
relKraftanstieg_li	0.03	0.11	758	-0.18	0.24
Ekz	0.27	0.12	717	0.03	0.51
Kdif	-0.05	0.10	717	-0.24	0.15
A60	0.21	0.15	545	-0.08	0.50
A180	0.21	0.15	542	-0.10	0.51
ekV	-0.21	0.15	546	-0.50	0.08
kkV	-0.01	0.12	542	-0.25	0.23
dynstat_60	-0.12	0.15	545	-0.41	0.17
dynstat_180	0.20	0.15	542	-0.10	0.50

Tabelle 2.9.: Vergleiche der Z-Werte von Skispringern und Nordischen Kombinierern mit linearen gemischten Modellen der Form (2.7). Angegeben ist der Parameterschätzer $\hat{\beta}_1$ für den Effekt der Sportart, der Standardfehler und der Stichprobenumfang, sowie Ober- und Untergrenze eines 95%-Konfidenzintervalls des Schätzers $\hat{\beta}_1$. Die fett gedruckten Konfidenzgrenzen geben dabei die Intervalle an, die nicht die Null enthalten.

Gemäß der Kodierung (2.6) gibt β_1 in (2.7) die Differenz zwischen Skispringern und Nordischen Kombinierern an. In Tabelle 2.9 sind die Schätzer für den Effekt der Sportart für jede Variable der Leistungsdiagnostik zusammengefasst. Ist der geschätzte Wert einer Variable in Tabelle 2.9 größer Null, so haben die Skispringer – adjustiert auf die zufälligen Effekte – höhere Z-Werte bei dieser Variable (man beachte wieder die umgedrehten Vorzeichen bei den Variablen `relKraftanstieg_[bb,re,li]` und `DJ32_KT`). Ein Wert $\beta_1 = 1$ würde bedeuten, dass Skispringer im Mittel bei den Z-Werten um eine Standardabweichung über den Nordischen Kombinierern liegen. Dieser Effekt ist auf dem 95%-Niveau signifikant von Null verschieden, wenn das 95%-Konfidenzintervall des Schätzers nicht die Null enthält.

Bei den Variablen SJ_SpH, SJ_V75E und DJ32_SpH sowie F_re170_[bb,re,li] haben die Skispringer signifikant höhere Werte als Nordische Kombinierer, jeweils relativ zu ihrer Altersreferenzgruppe. Nur bei der Variable DJ32_KT sind die Nordischen Kombinierer signifikant besser als die Skispringer.

2.5.6. Korrelationen der Variablen

Im Folgenden werden die Korrelationen aller Variablen betrachtet. Zunächst werden die bivariaten Verteilungen für die Variablen SJ_SpH, DJ32_SpH, Fmax70_bb, Exz, ekV und A60 aller vorliegender Werte gezeigt. Die Stärke des linearen Zusammenhangs zweier Variablen wird durch den Korrelationskoeffizienten ρ nach Bravais-Pearson gemessen (vgl. Fahrmeir u. a., 2004, Kapitel 3.4). Die Streudiagramme und Korrelationskoeffizienten sind in Abbildung 2.10 dargestellt. Hier fällt die geringe Korrelation der Variable ekV mit den anderen dargestellten Variablen auf. Die beiden Kraftparameter Exz und Fmax70_bb stehen hingegen in einem starken linearen Zusammenhang. Mittlere bis starke Korrelationen sind zwischen den Sprungkraft- und Maximalkraftparametern zu beobachten.

Die paarweisen Korrelationskoeffizienten aller Variablen lassen sich in der empirischen Korrelationsmatrix der Variablen zusammenfassen. Für alle vorliegenden Daten der Leistungsdiagnostik erhält man die Matrix in Abbildung 2.11. Alle Variablenpaare werden mit einem Korrelationstest auf Unkorreliertheit getestet (vgl. Fahrmeir u. a., 2004, S.470). Die signifikant von Null verschiedenen Korrelationen sind dann in der Abbildung markiert.

Die Korrelationsmatrix in Abbildung 2.11 bietet einen ersten Ansatz zur Identifikation von Korrelationsstrukturen in den Daten. So sieht man, dass die Variablen zum relativen Kraftanstieg relKraftanstieg_[bb,re,li], die untereinander mittlere signifikante Korrelationen aufweisen, aber mit allen anderen Variablen nur schwach bis gar nicht korrelieren. Die Sprungkraft- und Maximalkraftparameter hingegen zeichnen sich durch starke Korrelationen untereinander und mit anderen Blöcken von Variablen aus. Die beidbeinigen Messungen stehen dabei jeweils in hohem linearen Zusammenhang mit den einbeinigen Messungen. Die Variablen DJ32_KT, die die Bodenkontaktzeit beim Drop Jump angibt, korreliert nur signifikant positiv mit der Variable DJ32_Faktor, die sich als Quotient von Sprunghöhe und Bodenkontaktzeit errechnet (man beachte hier die Änderung des Vorzeichens). Die Variablen ekV, kkV, dynstat_60 und dynstat_180 weisen kaum signifikante Korrelationen mit anderen Variablen der Leistungsdiagnostik auf.

2.6. Diskussion

Vor der weiteren statistischen Analyse der Ergebnisse folgt hier eine Diskussion der bisher im Rahmen der explorativen Analysen gewonnenen Ergebnisse.

- Es konnte explorativ die Normalverteilungsannahme der Daten zur Leistungsdiagnostik nicht verworfen werden. Dies ist Voraussetzung für die Anwendung der weiteren statistischen Methoden in den folgenden Kapiteln.
- Der vorliegende Datensatz zur Leistungsdiagnostik enthält viele fehlende Daten. Imputationsverfahren scheinen besonders bei Daten aus dem Bereich der Leistungsdiagnostik nicht sinnvoll, da nichts über die Fehlendmechanismen bekannt ist. So können Messungen beispielsweise aufgrund von Verletzungen des Athleten nicht durchgeführt werden. Deshalb sind Imputationsverfahren ohne genaue Kenntnisse des konkreten Einzelfalls nicht sinnvoll. Die alternative Vorgehensweise mit Eliminationsverfahren ist vorzuziehen, auch wenn dadurch eine Vielzahl der gesammelten Information nicht genutzt wird.

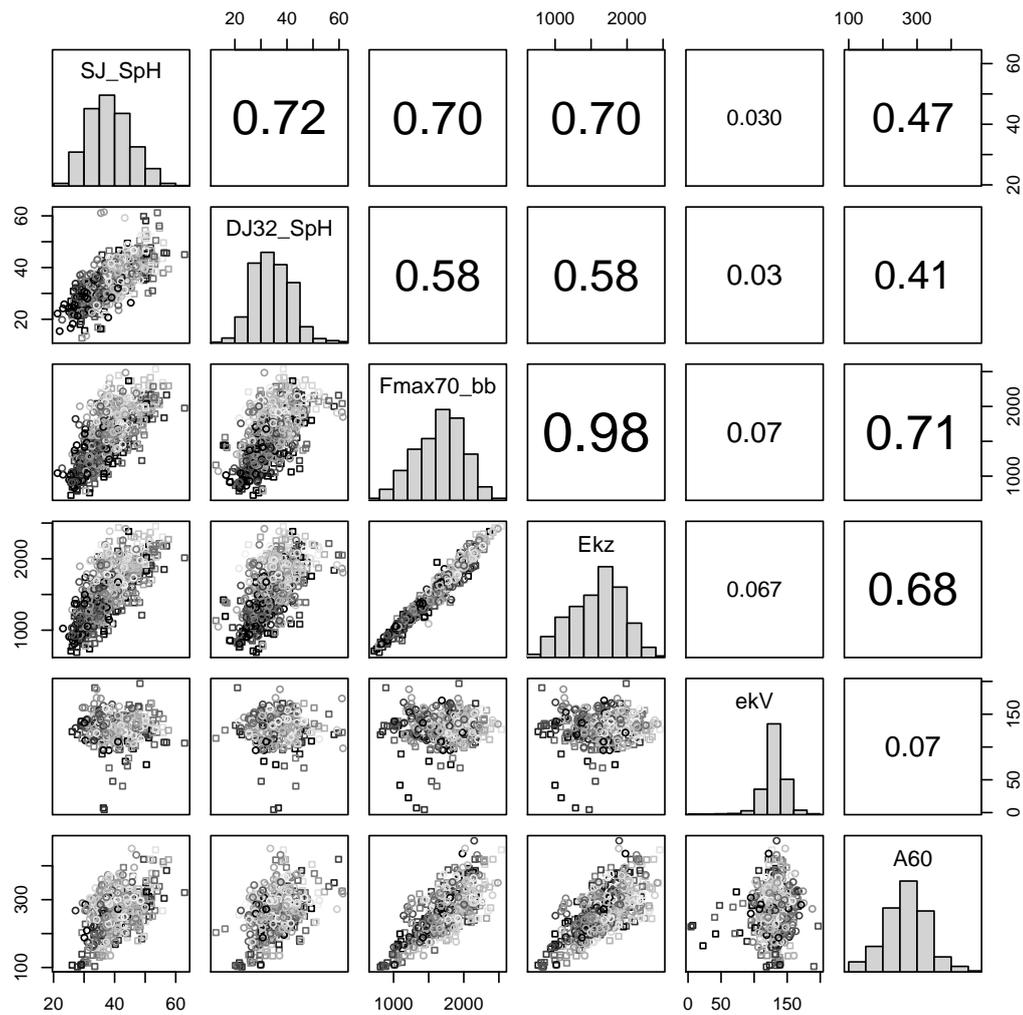


Abbildung 2.10.: Darstellung der bivariaten Verteilungen ausgewählter Variablen: In der unteren Hälfte sind die bivariaten Streudiagramme von jeweils zwei Variablen dargestellt. Die Kreise stellen dabei die Messungen von Nordischen Kombinierern und die Quadrate die von Skispringern dar. Je heller das Symbol ist, desto länger liegt die Messung des Athleten zurück. Die schwarz gefärbten Symbole bezeichnen jeweils die aktuelle Messung. Auf der Diagonale sind die univariaten Verteilungen der jeweiligen Variable durch ein Histogramm dargestellt. In der oberen Hälfte sind die Korrelationskoeffizienten der Variablenpaare angegeben.

- Um inter- und intraindividuelle Unterschiede zwischen den Athleten zu vergleichen, müssen die Messwerte mit Mittelwert und Standardabweichung einer Referenzgruppe Z-standardisiert werden. Es ist inhaltlich sinnvoll, diese Referenzgruppen anhand des Alters der Athleten zu definieren. Durch die Normierung auf das Alter wird automatisch auf die Kovariablen Gewicht und BMI adjustiert. Die standardisierten Werte lassen sich gut in Profilbildern wie in Abbildung 2.8 auf Seite 24 vergleichen. Im Kapitel 3 werden diese standardisierten Werte verwendet.
- Lineare gemischte Modelle sind bei Longitudinaldaten im Bereich der Leistungsdiagnostik eine gute Möglichkeit, um Gruppenvergleiche durchzuführen (z.B. von Skispringern und Nordischen Kombinierern) und sind t-Tests vorzuziehen, da sie die Abhängigkeit der Beobachtungen berücksichtigen und die Möglichkeit zur Berücksichtigung weiterer Kovariablen bieten. Dabei ergibt

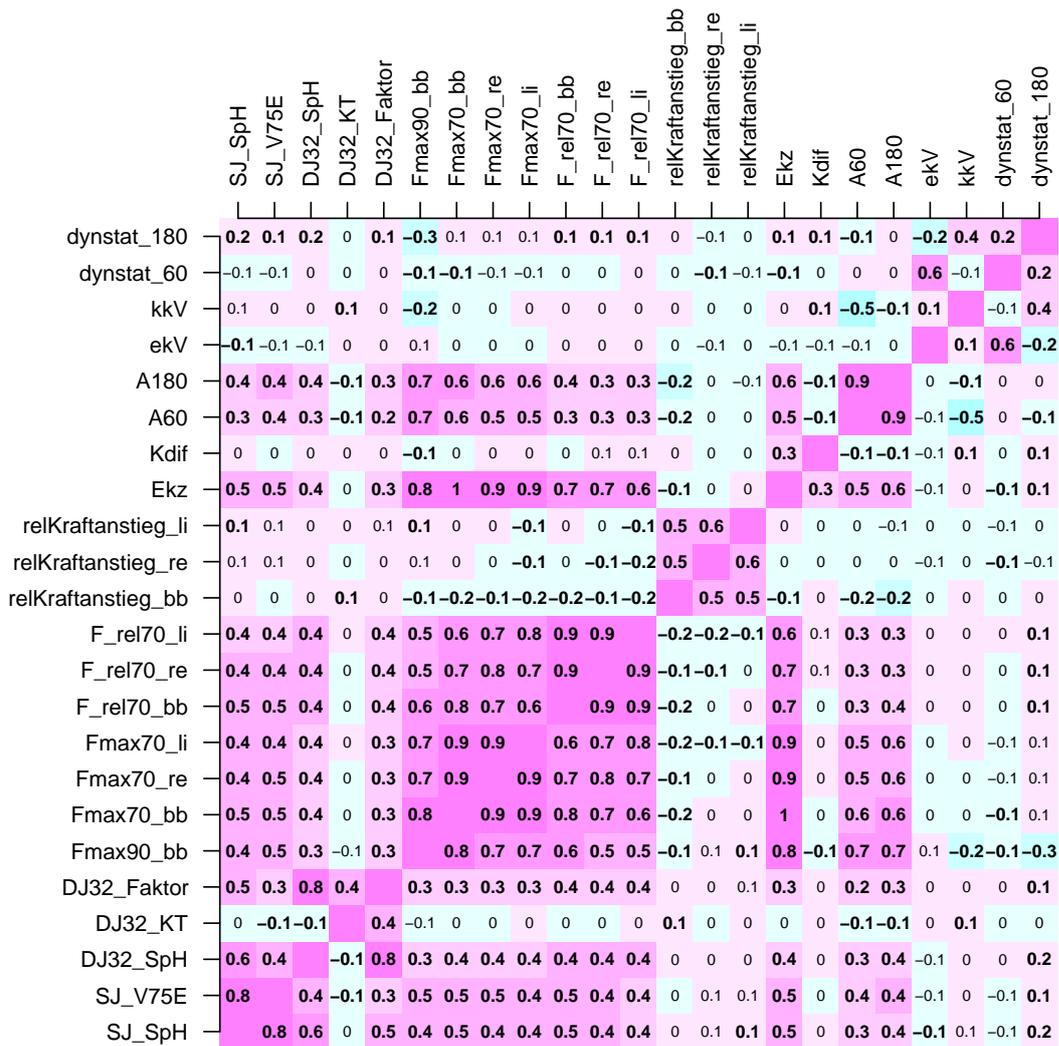


Abbildung 2.11.: Korrelationsmatrix der Parameter der Leistungsdiagnostik: Rosa eingefärbte Zellen stehen für eine positive Korrelation, türkis eingefärbt für eine negative Korrelation der Variablen. Je intensiver die Farbe, desto größer ist der Betrag des Korrelationskoeffizienten, wobei für signifikant von Null verschiedene Korrelationen der Wert des Korrelationskoeffizienten fett gedruckt ist.

sich, dass Skispringer bei den meisten Tests zur Sprungkraft und Maximalkraftdiagnostik bessere Ergebnisse erzielen als Nordischen Kombinierer.

- Anhand der empirischen Korrelationsmatrix der Variablen sieht man starke Korrelationen vieler Variablen. Das bedeutet, dass die zugehörigen Messmethoden ähnliche Leistungsmerkmale des Athleten messen. Allerdings gibt es auch einige Variablen, insbesondere bei der Schnellkraftdiagnostik, die nur schwache Korrelationen mit allen anderen Variablen aufweisen. Im folgenden Kapitel werden die Zusammenhänge der Variablen in einem Faktorenmodell dargestellt und daraus Erkenntnisse über die zukünftige Gestaltung einer Leistungsdiagnostik für Skispringer und Nordische Kombinierer gezogen.

3. Faktorenmodelle für die Daten zur Leistungsdiagnostik

In diesem Kapitel werden die im vorhergehenden Kapitel beschriebenen Daten zur Leistungsdiagnostik von 99 Skispringern und 92 Nordischen Kombinierern genauer analysiert. Dabei liegt der inhaltliche Schwerpunkt auf der Analyse der Zusammenhänge der Variablen der Leistungsdiagnostik mit Faktorenmodellen. Im Abschnitt 3.1 wird die statistische Theorie erläutert und in Abschnitt 3.2 auf die vorliegenden Daten zur Leistungsdiagnostik angewendet. In Abschnitt 3.3 werden die Ergebnisse des Kapitels diskutiert und zusammengefasst.

3.1. Statistische Methodik: Faktorenanalyse

In diesem Abschnitt wird die statistische Theorie der Faktorenanalyse beschrieben. Die Faktorenanalyse ist ein multivariates Verfahren, das sich zur Analyse der Zusammenhänge der Diagnostikdaten eignet. Einen ersten Überblick über die Zusammenhänge in einem Datensatz liefert die empirische Korrelationsmatrix der Variablen (vgl. Abschnitt 2.5.6). Allerdings lässt die Betrachtung der empirischen Korrelationsmatrix „interpretativen Spielraum“ (vgl. Ballreich u. Baumann, 1996, S.120), da z.B. Scheinkorrelationen, d.h. Korrelationen zweier Variablen über eine andere, nicht zu identifizieren sind. Die Faktorenanalyse geht von der empirischen Korrelationsmatrix aus und stellt ein Faktorenmodell auf, das die unterliegende Struktur der Diagnostikdaten darstellt. Die Methodik ist dabei für unabhängige Daten entwickelt, sodass für abhängige Beobachtungen weitere Überlegungen nötig werden.

3.1.1. Faktorenanalyse bei Longitudinaldaten

Die Leistungsdiagnostik Ski Nordisch ist als Längsschnittstudie konzipiert, sodass die Daten eine longitudinale Struktur haben wie in Abschnitt 2.5.5 beschrieben wurde. Die Beobachtungen einer Variable \mathbf{x} haben die Form

$$\mathbf{x} = (x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{Nn_N})^T$$

bei N Athleten und $\sum_{i=1}^N n_i$ Beobachtungen. Man fasst \mathbf{x} als Zufallsvektor auf, da jede Messung mit einem Zufallsfehler versehen und somit stochastisch ist. Die theoretische Kovarianzmatrix des Zufallsvektors \mathbf{x} hat aufgrund der abhängigen Beobachtungen eine blockdiagonale Struktur, d.h.

$$\text{Cov}(\mathbf{x}) = \begin{pmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & V_N \end{pmatrix}, \quad (3.1)$$

wobei V_1, \dots, V_N positiv definite Matrizen der Dimensionen $n_1 \times n_1, \dots, n_N \times n_N$ sind, die die Kovarianzstruktur zwischen den Messungen eines Athleten angeben. Die Messungen bei einem Athleten sind wegen der Messwiederholungen abhängig, während die Messungen verschiedener Athleten unabhängig sind.

Die Faktorenanalyse ist in der Theorie jedoch für unabhängige Beobachtungen entwickelt, d.h. die theoretische Kovarianzmatrix des Zufallsvektors \mathbf{x} soll eine diagonale Struktur haben. Die Faktorenanalyse betrachtet die Zusammenhänge verschiedener Variablen untereinander anhand der empirischen Kovarianzmatrix der Variablen, wo das longitudinale Design nicht berücksichtigt wird (vgl. Abschnitt 2.5.6). Im Rahmen dieser Arbeit wurde eine Simulationsstudie durchgeführt, bei der die empirische Kovarianzmatrix bei unabhängigen Daten und bei Daten mit vorgegebener Korrelationsstruktur der Beobachtungen verglichen wurde. Dabei konnten keine Unterschiede bei der Schätzung der Elemente der empirischen Kovarianzmatrix von Variablen zwischen beiden Vorgehensweisen festgestellt werden. Die einzelnen Messungen eines Athleten können also als unabhängige Beobachtungen aufgefasst werden (d.h. die longitudinale Datenstruktur wird nicht berücksichtigt), ohne dass sich Verzerrungen bei der Schätzung der Kovarianzmatrix und damit im Faktorenmodell ergeben. Die Simulationsstudie wird im Folgenden beschrieben, ist jedoch für das weitere Verständnis der Faktorenanalyse nicht nötig.

Simulationsstudie zur Kovarianzmatrix bei teilweise abhängigen Beobachtungen

In einer Simulationsstudie wurde der Einfluss der Abhängigkeit in der Stichprobe auf die Varianzen und Kovarianzen von Variablen \mathbf{x} und \mathbf{y} untersucht. Insgesamt wurden 1000 Simulationen gerechnet, wobei jedes Mal eine Stichprobe aus unabhängigen (1) und aus abhängigen Daten (2) gezogen wurde. Dabei sind die beiden Variablen \mathbf{x} und \mathbf{y} multivariat normalverteilt mit einem vorgegebenen Korrelationskoeffizienten ρ . Zur Simulation von Zufallszahlen aus einer multivariaten Normalverteilung wird in R die Funktion `rmvnorm` aus dem Paket `mvtnorm` verwendet.

- (1) In jedem Simulationsschritt werden $n=326$ (entspricht dem Stichprobenumfang für das Faktorenmodell in Abschnitt 3.2 ab Seite 46) unabhängige Zufallszahlen aus einer multivariaten Standardnormalverteilung gezogen, d.h.

$$\mathbf{x} \sim N_{326}(\mathbf{0}, \mathbf{I}_{326}).$$

Außerdem werden 326 Zufallszahlen aus der bedingten Verteilung von \mathbf{y} gegeben \mathbf{x} gemäß

$$\mathbf{y}|\mathbf{x} \sim N_{326}(\rho\mathbf{x}, \text{diag}(\sqrt{1-\rho^2}))$$

gezogen. Die so generierten Zufallsvektoren \mathbf{x} und \mathbf{y} haben einen theoretischen Korrelationskoeffizienten von ρ (Beweis mit Grimmett u. Stirzaker (2001) Example(7) auf Seite 106, und wegen (\mathbf{x}, \mathbf{y}) bivariat standardnormalverteilt $\Rightarrow \text{Cor}(\mathbf{x}, \mathbf{y})=\rho$).

- (2) In jedem Simulationsschritt werden die 326 Beobachtungen zufällig auf 100 Individuen verteilt, wobei einem Individuum auch keine Beobachtung zugeordnet werden kann, so dass insgesamt $N \leq 100$ Individuen mindestens eine Beobachtung haben. Für jedes Individuum i wird ein zufälliges τ_i aus einer Gleichverteilung auf $[u_1, u_2]$ gezogen. Dieses Intervall gibt an, in welchem Bereich die Korrelation der Messwiederholungen eines Individuums liegt. Enthält das Intervall die 0, so sind auch Individuen mit unkorrelierten Messwiederholungen möglich. Die Beobachtungen von Variable \mathbf{x} werden aus einer multivariaten Normalverteilung mit Mittelwertsvektor 0 und einer blockdiagonalen Kovarianzmatrix der Form (3.1) auf Seite 31 gezogen, wobei

$$V_i = \begin{pmatrix} 1 & & \tau_i \\ & \ddots & \\ \tau_i & & 1 \end{pmatrix}.$$

Somit sind die einzelnen Beobachtungen der Variable \mathbf{x} eines Individuums i konstant mit τ_i korreliert. Nun werden für die Variablen gemäß \mathbf{x} Zufallszahlen gezogen

$$\mathbf{x} \sim N_{326} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} V_1 & & 0 \\ & \ddots & \\ 0 & & V_N \end{pmatrix} \right).$$

Die Zufallszahlen von y werden wieder aus

$$\mathbf{y}|\mathbf{x} \sim N_{326}(\rho\mathbf{x}, \text{diag}(\sqrt{1-\rho^2}))$$

gezogen.

Es werden also jeweils Paare von korrelierten Variablen generiert, wobei die einzelnen Beobachtungen im Fall (1) unabhängig und im Fall (2) abhängig sind. In jedem Simulationsschritt werden die empirischen Werte für $\text{Var}(\mathbf{x})$ und $\text{Cov}(\mathbf{x}, \mathbf{y})$ berechnet und die Ergebnisse von (1) und (2) verglichen. Damit werden also die Einträge der empirischen Kovarianzmatrix von Variablen bei unabhängigen Daten und bei abhängigen mit vorgegebener Korrelationsstruktur der Daten verglichen. Die Verteilung der empirischen Varianzen und Kovarianzen für $\rho = 0.7$ und $u_1 = 0$ und $u_2 = 0.8$ werden in Abbildung 3.1 gegenübergestellt.

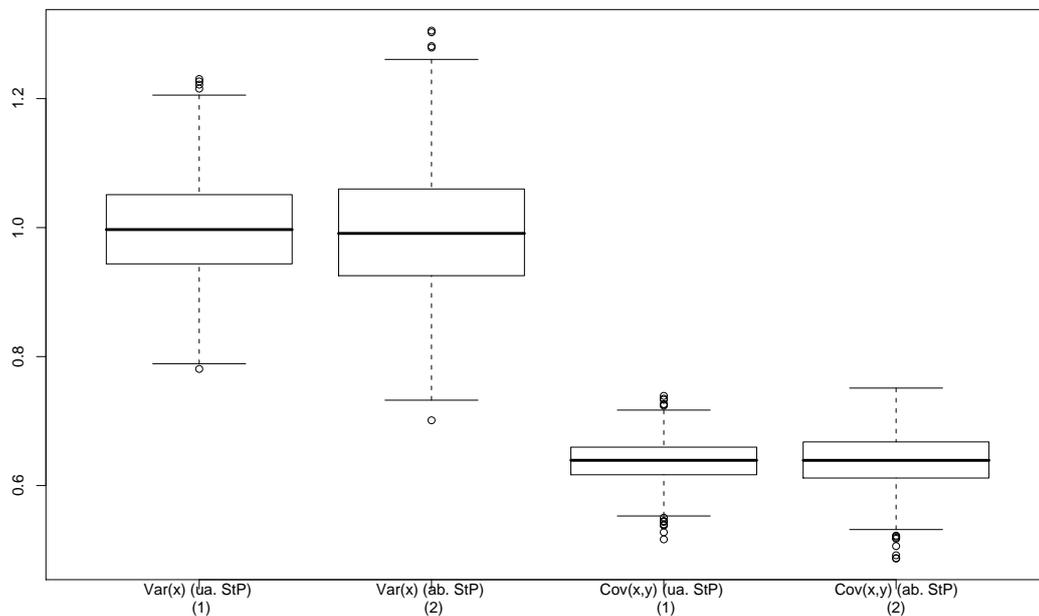


Abbildung 3.1.: Ergebnisse der Simulationsstudie (1000 Simulationendurchgänge)

Man sieht in Abbildung 3.1, dass für unabhängige und abhängige Stichproben im Mittel die gleichen Varianzen geschätzt werden. Die Streuung der Ergebnisse ist bei der abhängigen Stichprobe sowohl bei der Varianz als auch bei der Kovarianz größer. Dies bestätigt, dass die unterliegende Kovarianz der beiden Variablen gleich geschätzt wird, unabhängig davon, ob die einzelnen Messwiederholungen bei einer Variable blockweise korreliert sind. Es ist also nicht notwendig, die zur Verfügung stehende Datenmenge auf eine Beobachtung je Athlet zu reduzieren. Daher werden alle zur Verfügung stehenden Beobachtungen zur Schätzung des Faktorenmmodells verwendet, wobei die wiederholten Messungen bei

einem Athleten als unabhängig angesehen werden, wie es die Modellvoraussetzung der Faktorenanalyse besagt. Die Notation dieses Kapitels wird im nächsten Abschnitt beschrieben.

3.1.2. Notation

Ausgangspunkt ist der Datensatz zur Leistungsdiagnostik mit den Variablen, die in Kapitel 2 beschrieben wurden. Der Datensatz enthält in den Zeilen die Beobachtungen (Messergebnisse eines Athleten zu einem Untersuchungszeitpunkt) und in den Spalten die Variablen. Wiederholte Messungen bei demselben Athleten stellen unterschiedliche Zeilen und damit unterschiedliche Beobachtungen im Datensatz dar. Der Index t zur Unterscheidung der einzelnen Messwiederholungen bei einem Athleten fällt weg. Der Index i indiziert jetzt die einzelnen Beobachtungen im Gegensatz zu vorher den Athleten. In Tabelle 3.1 sind die vor der Analyse bekannten und in Tabelle 3.2 die zu schätzenden Objekte des Faktorenmodells dargestellt.

In der folgenden Beschreibung der Faktorenanalyse sind zwei Sichtweisen auf die Elemente des Faktorenmodells nötig.

1. Man geht einmal von der Sichtweise der Daten aus. Diese Sichtweise wird zur Beschreibung der Modellgleichungen benötigt. In diesem Fall wird mit \mathbf{x}_j der *empirische Datenvektor* mit n Realisationen und mit \mathbf{X} die *Datenmatrix* mit n Beobachtungen und p Variablen bezeichnet.
2. In der anderen Sichtweise sind die Elemente des Faktorenmodells Zufallsvariablen vor der Realisierung der Daten. Diese Sichtweise wird bei der Beschreibung der Modellannahmen benötigt. Hier wird mit x_j eine *univariate Zufallsvariable* und mit \mathbf{X} der *Zufallsvektor* der p univariaten Zufallsvariablen $\mathbf{x}_1, \dots, \mathbf{x}_p$ bezeichnet.

In der Beschreibung des Faktorenmodells ändert sich die Sichtweise hin von den Daten zu Zufallsvariablen, wenn auf die Modellannahmen zurückgegangen wird. Auf diesen Sichtweisenwechsel wird an entsprechender Stelle hingewiesen.

Symbol	Bedeutung	Dimension	Anmerkungen/Beispiel
n	# Beobachtungen	skalar	# Zeilen im Datensatz
p	# Variablen	skalar	# Spalten im Datensatz
\mathbf{x}_j	Alle Beobachtungen von Variable j	Vektor der Länge n	alle Messergebnisse bei Variable SJ_Sph
\mathbf{X}	Datensatz der Beobachtungen: $\mathbf{X} = (\mathbf{x}_1 \cdots \mathbf{x}_p)$	$n \times p$ - Matrix	Datenmatrix durch Zusammensetzen aller Variablen \mathbf{x}_j
x_{ij}	eine Beobachtung aus dem Datensatz \mathbf{X}	skalar	SJ_Sph bei einer Beobachtung (z.B. die Messung am 24.7.2006 bei Athlet 63)
i	Index der Beobachtungen	$i = 1, \dots, n$	
j	Index der Variablen	$j = 1, \dots, p$	
$\boldsymbol{\mu}$	empirischer Mittelwertsvektor der Daten \mathbf{X}	Vektor der Länge p	
$\boldsymbol{\Sigma}$	empirische Kovarianzmatrix der Variablen	$p \times p$ - Matrix	

Tabelle 3.1.: Notation für die Elemente der Datenmatrix (empirische Sichtweise). Diese Elemente sind alle vor der Analyse bekannt.

Symbol	Bedeutung	Dimension	Anmerkungen/Beispiel
q	# Faktoren	skalar	Es gilt: $q < p$ (Für eindeutige Schätzer muss $p \geq 2q + 1$ gelten (Fahrmeir u. a., 1996))
f_k	Faktorenwerte aller Beobachtungen für Faktor k	Vektor der Länge n	
F	Matrix der Faktorenwerte: $F = (f_1 \cdots f_q)$	$n \times q$ - Matrix	Faktorenwerte aller Beobachtungen
f_{ik}	Eintrag von Matrix F	skalar	Faktorenwert von Beobachtung i bezüglich Faktor k
k	Index der Faktoren	$k = 1, \dots, q$	
L	Ladungsmatrix	$p \times q$ - Matrix	
l_{jk}	Ladungskoeffizient von Variable j für Faktor k	skalar	
e_j	Fehlerterm	Vektor der Länge n	Restfehler für Variable j
e	Matrix der Fehlerterme $e = (e_1 \cdots e_p)$	$n \times p$ - Matrix	
e_{ij}	Eintrag von e	skalar	Fehlerterm von Beobachtung i bei Variable j

Tabelle 3.2.: Notation für die Komponenten der Faktorenanalyse (empirische Sichtweise). Alle Elemente dieser Tabelle sind unbekannt und sollen mit der Faktorenanalyse geschätzt werden.

3.1.3. Ausgangslage und Zielsetzung

Die *Faktorenanalyse* ist nicht nur die Bezeichnung für ein statistisches Verfahren, sondern ein Sammelbegriff für verschiedene Ansätze und Schätzverfahren. Allen Verfahren ist jedoch gemeinsam, dass sie versuchen, die vorhandene Information auf einige wenige sogenannte *latente Faktoren* zurückzuführen. Latente Faktoren sind Konstrukte, die beobachtbare Vorgänge beeinflussen, jedoch selbst nicht beobachtbar sind. Häufig trifft man in der Psychologie auf solche Situationen, wenn beispielsweise der Faktor „Intelligenz“ von Interesse ist, der nicht direkt zu beobachten ist, aber anhand verschiedener Tests abgefragt werden kann. Bei den Diagnostikdaten ist beispielsweise denkbar, dass die erhobenen Variablen latente Faktoren wie die allgemeinen Sprungkraftfähigkeit oder Schnellkraftfähigkeit messen. Es können zwei Varianten der Faktorenanalyse unterschieden werden, die sich bezüglich der Herangehensweise unterscheiden (Fahrmeir u. a., 1996).

- Bei der *explorativen Faktorenanalyse* sollen aus den Daten Faktoren extrahiert werden, wobei über die genaue Art und Anzahl der Faktoren nichts bekannt ist. Die Faktoren ergeben sich dabei aus der Korrelationsstruktur der beobachteten Variablen. Die konstruierten Faktoren werden im Rahmen der explorativen Faktorenanalyse inhaltlich gedeutet. Hieraus können Hypothesen generiert werden, z.B. der Form: „Die Variablen A und B bilden gemeinsam einen latenten Faktor“. Diese Hypothesen sind jedoch nicht statistisch zu prüfen.
- Bei der *konfirmatorischen Faktorenanalyse* geht es um die statistische Absicherung von Hypothesen in einem bereits aufgestellten Faktorenmodell. Die Anzahl und Art der Faktoren wird dabei als bekannt angesehen. Anhand neuer Daten sollen diese Hypothesen überprüft werden.

Bei dieser Analyse liegt der Ausgangspunkt in der explorativen Faktorenanalyse, da zunächst noch kein Wissen über die Struktur eines Faktorenmodells vorliegt. Das bedeutet, dass weder Anzahl noch Art der Faktoren bekannt ist.

Ausgangspunkt der Faktorenanalyse ist ein Datensatz \mathbf{X} mit n Beobachtungen und p Variablen. Die Variablen können z.B. alle im Rahmen der Leistungsdiagnostik erfassten Variablen sein oder eine Teilmenge davon. In der Regel sind diese Variablen nicht unabhängig, sondern korreliert (vgl. Abschnitt 2.5.6). Hohe Korrelationen zwischen Variablen deuten darauf hin, dass hinter den Variablen möglicherweise eine kleine Zahl von Konstrukten bzw. latenten Faktoren steht, die nicht direkt erhoben werden können, aber von den Variablen teilweise gemessen werden. Jede Variable wiederum misst dann einen oder mehrere dieser Faktoren. In diesem Fall können die beobachteten Variablen als Linearkombinationen der Faktoren angenähert werden. Der Grad der Annäherung wird durch den Anteil der Varianz an der Gesamtvarianz der Daten, den die Faktoren erklären, ausgedrückt. Bei hoch korrelierten Variablen können die Variablen anhand weniger Faktoren gut dargestellt werden. Sind die Variablen jedoch unkorreliert, so kann keine sinnvolle Darstellung durch Faktoren gefunden werden.

Mit der Faktorenanalyse sollen diese latenten Faktoren identifiziert werden. Die Faktorenanalyse kann als ein strukturentdeckendes Verfahren bezeichnet werden, mit dem Ziel, die vorhandene Information vieler Variablen auf einige wenige Faktoren zu reduzieren (Hartung u. Elpelt, 1999). Damit die Ergebnisse leicht zu interpretieren sind, sind wenige Faktoren wünschenswert. Die Angabe „wenig“ bezieht sich darauf, dass die Anzahl der Faktoren gegenüber der Anzahl der Variablen klein ist, so dass die Modellkomplexität sinkt. Da wenige Faktoren jedoch auch weniger Varianz erklären, muss hier ein Kompromiss zwischen Modellanpassung und Modellkomplexität gefunden werden. Für diese Entscheidung ist es in der Anwendung sinnvoll, auch inhaltliches Vorwissen einfließen zu lassen.

Stellt man die Beobachtungen im p -dimensionalen Raum dar, so streuen die Daten möglicherweise hauptsächlich nur in $q < p$ Dimensionen, da einige Variablen ähnliche Konstrukte messen. In diesem Fall ist es ausreichend, die Information durch q Faktoren darzustellen. Die Faktoren können so konstruiert werden, dass sie untereinander unabhängig sind, also im Raum orthogonal aufeinander stehen. So können sie in einer folgenden Regressionsanalyse problemlos als Kovariablen verwendet werden, und dann sind nicht wie bei Verwendung der ursprünglichen, korrelierten Variablen Kollinearitätsprobleme zu erwarten.

Voraussetzungen für die Durchführbarkeit einer Faktorenanalyse sind stetige Daten und mehr Beobachtungen als Variablen, d.h. $n > p$. Da die Ergebnisse einer Faktorenanalyse skalierungsabhängig sind, wird die Datenmatrix häufig vorab standardisiert zu \mathbf{X}^{st} . Die einzelnen Einträge x_{ij}^{st} ergeben sich als

$$x_{ij}^{st} = \frac{x_{ij} - \bar{x}_j}{\sqrt{\text{Var}(\mathbf{x}_j)}}. \quad (3.2)$$

Nach der Standardisierung haben die Variablen \mathbf{x}_j^{st} Mittelwert 0 und Standardabweichung 1. Im Folgenden wird von standardisierten Daten ausgegangen, aber zugunsten der Übersichtlichkeit auf das Kürzel „ st “ verzichtet. Die empirische Korrelation zwischen zwei Variablen j und j' ergibt sich dann als

$$r_{jj'} = \frac{1}{n-1} \sum_{i=1}^n x_{ij} x_{ij'} \quad j, j' = 1, \dots, p; \quad j \neq j'$$

Im folgenden Abschnitt wird das Modell der Faktorenanalyse aufgestellt und anschließend beschrieben, wie die Faktoren ausgehend von der empirischen Kovarianzmatrix geschätzt und interpretiert werden können.

3.1.4. Modell der Faktorenanalyse

Das Modell der Faktorenanalyse wird aus der empirischen Sichtweise beschrieben. Den Ausgangspunkt bilden die Variablen $\mathbf{x}_1, \dots, \mathbf{x}_p$ mit Mittelwertsvektor $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)^T = (\bar{x}_1, \dots, \bar{x}_p)^T$ und der

empirischen Kovarianzmatrix der Variablen

$$\Sigma = \frac{1}{n-1} \begin{pmatrix} \sum_{i=1}^n (x_{i1} - \mu_1)^2 & \sum_{i=1}^n (x_{i1} - \mu_1)(x_{i2} - \mu_2) & \cdots & \sum_{i=1}^n (x_{i1} - \mu_1)(x_{ip} - \mu_p) \\ \sum_{i=1}^n (x_{i2} - \mu_2)(x_{i1} - \mu_1) & \sum_{i=1}^n (x_{i2} - \mu_2)^2 & \cdots & \sum_{i=1}^n (x_{i2} - \mu_2)(x_{ip} - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n (x_{ip} - \mu_p)(x_{i1} - \mu_1) & \sum_{i=1}^n (x_{ip} - \mu_p)(x_{i2} - \mu_2) & \cdots & \sum_{i=1}^n (x_{ip} - \mu_p)^2 \end{pmatrix}.$$

Geht man von der standardisierten Datenmatrix aus, so entspricht Σ auch gleichzeitig der Korrelationsmatrix \mathbf{R} .

Das Modell der Faktorenanalyse nimmt an, dass die p Variablen durch q latente Faktoren bis auf einen Restfehler erklärt werden können. Alle Beobachtungen x_{ij} bzw. alle Variablen \mathbf{x}_j können bis auf diesen Fehlerterm als Linearkombination von Faktorenwerten $f_{i1}, f_{i2}, \dots, f_{iq}$ bzw. Faktoren $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_q$ dargestellt werden, d.h. es gilt

$$\text{Beobachtungsebene: } x_{ij} = l_{j1}f_{i1} + l_{j2}f_{i2} + \dots + l_{jq}f_{iq} + e_{ij} \quad \text{mit } i = 1, \dots, n, \quad j = 1, \dots, p$$

$$\text{Variablenebene: } \mathbf{x}_j = l_{j1}\mathbf{f}_1 + l_{j2}\mathbf{f}_2 + \dots + l_{jq}\mathbf{f}_q + \mathbf{e}_j \quad \text{mit } j = 1, \dots, p. \quad (3.3)$$

Man beachte hierbei, dass (3.3) die Form eines Regressionsmodells hat, aber keines ist, da die Faktorenwerte \mathbf{f}_k , $k = 1, \dots, q$ nicht beobachtbar sind, sondern im Anschluss an die Faktorenanalyse geschätzt werden. Der Koeffizient l_{jk} gibt die *Ladung* der j -ten Variable auf den k -ten Faktor an und wird als *Faktorladung* bezeichnet. Die Faktorladungen sind unbekannt und sollen geschätzt werden. Die Faktoren \mathbf{f}_k werden häufig als *gemeinsame* Faktoren bezeichnet, während die Fehlerterme \mathbf{e}_j *spezifische* Faktoren genannt werden. Die gemeinsamen Faktoren wirken auf mindestens zwei Variablen, während die spezifischen Faktoren immer nur auf einer Variable wirksam sind.

In Matrixschreibweise ergibt sich das Faktorenmodell aus Gleichung (3.3) für alle Variablen $j = 1, \dots, p$ als

$$\mathbf{X} = \mathbf{F}\mathbf{L}^T + \mathbf{e}, \quad (3.4)$$

wobei die $p \times q$ - Matrix \mathbf{L} als *Ladungsmatrix* bezeichnet wird und den Zusammenhang zwischen Variablen und Faktoren angibt. Dabei hat \mathbf{L} die folgende Form:

$$\mathbf{L} = \begin{pmatrix} l_{11} & \cdots & l_{1q} \\ \vdots & \ddots & \vdots \\ l_{p1} & \cdots & l_{pq} \end{pmatrix}.$$

Bei unkorrelierten Faktoren hat diese Matrix nur Einträge zwischen -1 und 1 , die dann die Korrelation zwischen Variablen und Faktoren angeben. Bei korrelierten Faktoren sind in seltenen Fällen auch betragsmäßig größere Ladungen möglich.

Die $n \times q$ - Matrix \mathbf{F} stellt die Verbindung zwischen Beobachtungen und Faktoren her (vgl. auch Abbildung 3.2). Der Eintrag f_{ik} gibt die Ausprägung des k -ten Faktors bei der i -ten Beobachtung an und wird auch als *Faktorenwert* bezeichnet. Die Matrix der Faktorenwerte \mathbf{F} ist so standardisiert, dass die Spaltenmittelwerte Null und die Spaltenvarianzen 1 sind. Hat somit eine Beobachtung i einen großen Faktorwert f_{ik} , so bedeutet dies, dass ihr Wert von Faktor \mathbf{f}_k überdurchschnittlich ist.

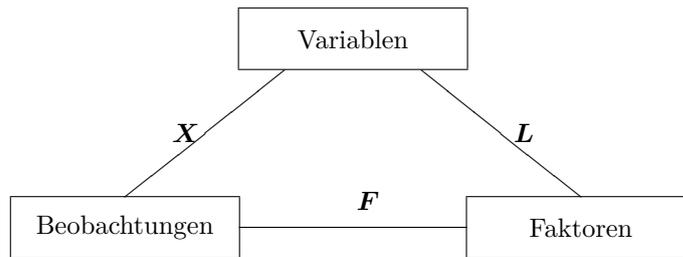


Abbildung 3.2.: Schematische Darstellung der Elemente der Faktorenanalyse: Beobachtungen, Variablen und Faktoren und deren Zusammenhänge ausgedrückt durch die Matrizen aus Gleichung (3.4).

3.1.5. Modellannahmen

Bisher wurde das Modell der Faktorenanalyse anhand der Datenmatrix beschrieben. Für die Modellannahmen geht man von der empirischen Sichtweise zur theoretischen Sichtweise über. Die Modellannahmen beziehen sich auf Zufallsvariablen vor der Realisierung der Beobachtungen. Die zusammengehörigen Gegenstücke der beiden Sichtweisen werden mit den gleichen Symbolen gekennzeichnet. Die abweichenden Bedeutungen in der Sichtweise der Zufallsvariablen sind in Tabelle 3.3 zusammengefasst.

Symbol	Bedeutung	Dimension
f_k	Faktor k	univariate Zufallsvariable
e_j	Fehlerterm j	univariate Zufallsvariable
x_j	Variable j	univariate Zufallsvariable
Σ	theoretische Kovarianzmatrix	$p \times p$ - Matrix
\mathbf{X}	Zufallsvektor der Variablen $j = 1, \dots, p$	p - dimensionaler Zufallsvektor
\mathbf{F}	Zufallsvektor der Faktoren $k = 1, \dots, q$	q - dimensionaler Zufallsvektor
l_{jk}	Ladungskoeffizient von Variable j für Faktor k	skalar
\mathbf{L}	Ladungsmatrix	$p \times q$ - Matrix
$\boldsymbol{\mu}$	Erwartungswertvektor der Variablen $j = 1, \dots, p$	p - dimensionaler Zufallsvektor

Tabelle 3.3.: Notation für die Komponenten der Faktorenanalyse (Sichtweise der Zufallsvariablen).

Für die gemeinsamen und spezifischen Faktoren wird eine multivariate Normalverteilung angenommen, d.h.

$$(\mathbf{f}_1, \dots, \mathbf{f}_q)^T \sim N_q(0, \boldsymbol{\Phi}) \quad (3.5)$$

$$(\mathbf{e}_1, \dots, \mathbf{e}_p)^T \sim N_p(0, \boldsymbol{\Psi}). \quad (3.6)$$

Für die Fehlerterme $\mathbf{e}_1, \dots, \mathbf{e}_p$ wird Erwartungswert $E(\mathbf{e}_j) = 0$ für alle j und $\text{Cov}(\mathbf{e}_j, \mathbf{e}_{j'}) = 0$ für alle Paare von spezifischen Faktoren $j, j' = 1, \dots, p$ mit $j \neq j'$ angenommen. Also ist die Kovarianzmatrix in Gleichung (3.6) eine Diagonalmatrix mit

$$\text{Cov}(\mathbf{e}_1, \dots, \mathbf{e}_p) = \text{diag}(\sigma_1^2, \dots, \sigma_p^2) =: \boldsymbol{\Psi}. \quad (3.7)$$

Weiterhin werden die gemeinsamen und spezifischen Faktoren im Faktorenmodell als unkorreliert

angenommen, d.h. für die theoretische Kovarianz der Zufallsvariablen \mathbf{f}_k und \mathbf{e}_j gilt

$$\text{Cov}(\mathbf{f}_k, \mathbf{e}_j) = 0, \quad (3.8)$$

für alle k und j . Im orthogonalen Fall gilt für die Kovarianz zwischen Variablen und Faktoren

$$\text{Cov}(\mathbf{x}_j, \mathbf{f}_k) = l_{jk},$$

so dass die Ladungsmatrix der Korrelationsmatrix zwischen Variablen und Faktoren entspricht. Wird auch für die gemeinsamen Faktoren untereinander Unabhängigkeit angenommen, also $\text{Cov}(\mathbf{f}_1, \dots, \mathbf{f}_q) = \Phi = \mathbf{I}$ (man spricht hier auch von einem *orthogonalen* Faktorenmodell), so gilt

$$\text{Var}(\mathbf{x}_j) = \underbrace{\sum_{k=1}^q l_{jk}^2}_{=: k_j^2} + \sigma_j^2 =: k_j^2 + \sigma_j^2 \quad (3.9)$$

Die Größen k_j^2 heißen *Kommunalitäten* und geben im Fall von standardisierten Merkmalen (d.h. $\text{Var}(\mathbf{x}_j) = 1$) an, welcher Anteil der Varianz von Variable j durch die gemeinsamen Faktoren erklärt wird, d.h. $k_j^2 = 1 - \sigma_j^2 = l_{j1}^2 + \dots + l_{jq}^2$ für $j = 1, \dots, p$. Die Varianz einer Variable lässt sich also in den Anteil, der durch die Faktoren erfasst wird (Kommunalität k_j^2), und einen variablenspezifischen Restfehler σ_j^2 (engl. als „uniquenesses“ bezeichnet) zerlegen.

Nun lässt sich die Kovarianzmatrix zerlegen zu

$$\begin{aligned} \Sigma &= \underset{\mathbf{X} = \mathbf{LF} + \mathbf{e}}{\text{E}} \left((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T \right) \underset{\boldsymbol{\mu} = \mathbf{0}}{=} \text{E}(\mathbf{X}\mathbf{X}^T) \\ &= \text{E}((\mathbf{LF} + \mathbf{e})(\mathbf{LF} + \mathbf{e})^T) \\ &= \text{E}(\mathbf{LF}^T \mathbf{F}\mathbf{L}^T) + \text{E}(\mathbf{L}\mathbf{F}\mathbf{e}^T) + \text{E}(\mathbf{e}\mathbf{F}^T \mathbf{L}^T) + \text{E}(\mathbf{e}\mathbf{e}^T) \\ &\stackrel{(3.6), \Phi = \mathbf{I}, (3.8), (3.7)}{=} \mathbf{L}\mathbf{L}^T + 0 + 0 + \Psi \\ \Rightarrow \Sigma &= \mathbf{L}\mathbf{L}^T + \Psi. \end{aligned} \quad (3.10)$$

Die Gleichung (3.10) wird auch als *Fundamentaltheorem* der Faktorenanalyse bezeichnet. Sie besagt, dass alle Kovarianzen in den Daten nur durch die Faktoren erklärt werden, da Ψ eine Diagonalmatrix ist. Je besser die Anpassung des Faktorenmodells ist, desto kleiner sind die Einträge in Ψ . Die Ladungsmatrizen können durch Zerlegung von Σ in die Form $\Sigma = \mathbf{L}\mathbf{L}^T + \Psi$ ermittelt werden. Allerdings ist diese Zerlegung nicht eindeutig, da beliebige orthogonale Rotationen der Faktoren zu äquivalenten Lösungen führen, sodass die Faktorladungen also nur bis auf orthogonale Transformationen eindeutig bestimmt sind. Um eine eindeutige Schätzung der Ladungsmatrix zu ermöglichen, müssen weitere Nebenbedingungen gefordert werden. Insgesamt enthält das Modell der Faktorenanalyse aufgrund der hohen Anzahl an freien Parametern Unbestimmtheiten und Identifizierbarkeitsprobleme, die empirisch nicht überprüfbar sind. In Fahrmeir u. a. (1996) wird das genannte Rotationsproblem sowie das Kommunalitätenproblem beschrieben.

3.1.6. Schätzung der Ladungsmatrix

Im Folgenden wird beschrieben, wie die Ladungsmatrix und die Matrix der Faktorenwerte aus den Daten geschätzt werden können. Es existieren eine Vielzahl von Schätzverfahren, wobei hier die *Maximum-Likelihood-Methode* (Lawley u. Maxwell, 1971) und die *Hauptkomponentenmethode* (Ho-

telling, 1936) vorgestellt werden.

Maximum-Likelihood-Methode Diese Methode fordert, dass $\mathbf{L}^T \mathbf{\Psi} \mathbf{L}$ eine Diagonalmatrix ist und nimmt an, dass

$$(\mathbf{x}_1, \dots, \mathbf{x}_p)^T \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

gilt. Unter dieser Nebenbedingung führt die Optimierung zu einer eindeutigen Lösung. Ausgehend von Gleichung (3.10) muss die Ladungsmatrix \mathbf{L} mit $p \cdot q$ unbekanntem Parametern geschätzt werden sowie Schätzer für die p Diagonalelemente von $\mathbf{\Psi}$ gefunden werden. Die gesuchten Schätzer $\hat{\mathbf{L}}$ für \mathbf{L} und $\hat{\sigma}_1^2, \dots, \hat{\sigma}_p^2$ für $\sigma_1^2, \dots, \sigma_p^2$ sollen die Likelihoodfunktion der Daten, die sich mit der Normalverteilungsannahme ergibt als

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{np}{2}} \det(\boldsymbol{\Sigma})^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})\right)$$

maximieren. Als Maximum-Likelihood-Schätzer für die erste Komponente ergibt sich das Stichprobenmittel, also $\hat{\boldsymbol{\mu}}_{ML} = \bar{\mathbf{x}}$. Nun muss noch $L(\boldsymbol{\Sigma})$ maximiert werden, das sich nach (3.10) auch durch $L(\mathbf{L}, \mathbf{\Psi})$ darstellen lässt. Die Schätzung umfasst die folgenden Schritte:

1. Festlegung der Anzahl der gemeinsamen Faktoren q .
2. Festlegung von Startwerten für das iterative Maximieren.
3. Schaukelalgorithmus zur Maximierung der Likelihood $L(\mathbf{L}, \mathbf{\Psi})$. Dabei werden abwechselnd die folgenden beiden Schritte ausgeführt und die Ergebnisse des einen als neuen Startwert für den anderen gewählt, bis ein Konvergenzkriterium erfüllt ist.
 - Für einen festen Wert $\mathbf{\Psi}_0$ wird $L(\mathbf{L}, \mathbf{\Psi} = \mathbf{\Psi}_0)$ maximiert. Eine Lösung hierfür kann analytisch durch ein Eigenwertproblem gefunden werden. Dies ergibt den neuen Startwert \mathbf{L}_0 für $\hat{\mathbf{L}}$.
 - Für festes \mathbf{L}_0 wird mit numerischen Verfahren $L(\mathbf{L} = \mathbf{L}_0, \mathbf{\Psi})$ maximiert. Dies ergibt den neuen Startwert $\mathbf{\Psi}_0$ für $\hat{\mathbf{\Psi}}$.

Eine genauere Beschreibung des Maximum-Likelihood-Verfahrens findet sich in Kapitel 11 von Fahrmeir u. a. (1996) oder Kapitel 8 von Hartung u. Elpelt (1999). Man beachte, dass für das Verfahren keine Konvergenz gewährleistet ist. In diesem Fall müssen die Startwerte variiert werden. Ist die Normalverteilungsannahme verletzt, so spricht man bei den Schätzern von *Quasi-Maximum-Likelihood-Schätzern*, die weiterhin konsistent und asymptotisch normal sind. Auch die Aussagen zu Vertrauensintervallen gelten asymptotisch (Fahrmeir u. a., 1996).

Hauptkomponentenmethode Diese Methode basiert auf der *Hauptkomponentenanalyse* (Hotelling, 1936), die zunächst in einem Exkurs vorgestellt wird. Bei der Hauptkomponentenmethode werden nicht wie bei der Maximum-Likelihood-Methode Verteilungsannahmen gestellt, sondern es wird versucht, die Daten direkt durch *Komponenten* möglichst gut anzunähern. Komponenten sind Linearkombinationen der ursprünglichen Variablen. Die Hauptkomponentenanalyse stellt eine alternative Möglichkeit zur Faktorenschätzung dar. Die Eindeutigkeit der Lösung wird hier durch die Nebenbedingung erzielt, dass die Faktoren sukzessive weniger Varianz erklären.

Exkurs: Hauptkomponentenanalyse

Die Hauptkomponentenanalyse ist wie die Faktorenanalyse eine variablenorientierte Methode, um die korrelierten, ursprünglichen Variablen durch wenige Komponenten zu erklären. Der wesentliche Unterschied der Hauptkomponentenanalyse zur Faktorenanalyse besteht darin, dass zunächst

p Komponenten konstruiert werden, von denen dann q ausgewählt werden. Da p Komponenten die ursprünglichen Variablen komplett erfassen, werden keine Fehlerterme benötigt. Der Fehler kommt erst durch Weglassen von Komponenten hinzu. Bei der Faktorenanalyse werden von vornherein nur q gemeinsame Faktoren und p spezifische Faktoren angenommen (Fahrmeir u. a., 1996).

Es werden Komponenten konstruiert, die sukzessive weniger der Gesamtvarianz der Daten erklären. Die Hauptkomponentenanalyse hat nicht zum Ziel, interpretierbare Faktoren zu konstruieren (deshalb werden die Ergebnisse im Anschluss an eine Hauptkomponentenanalyse in der Regel auch nicht rotiert), sondern die p Variablen durch p orthogonale Komponenten darzustellen. Die erste Komponente z_1 ergibt sich als

$$\mathbf{z}_1 = a_{11}\mathbf{x}_1 + a_{21}\mathbf{x}_2 + \dots + a_{p1}\mathbf{x}_p = \mathbf{X}\mathbf{a}_1,$$

wobei $\mathbf{a}_1 = (a_{11}, a_{21}, \dots, a_{p1})^T$ mit der Nebenbedingung $\mathbf{a}_1^T \mathbf{a}_1 = 1$. Die erste Komponente wird so bestimmt, dass die Varianz von z_1 maximal wird. Dabei gilt

$$\text{Var}(\mathbf{z}_1) = \text{Var}(\mathbf{X}\mathbf{a}_1) = \mathbf{a}_1^T \boldsymbol{\Sigma} \mathbf{a}_1 = \lambda,$$

wieder unter der Nebenbedingung $\mathbf{a}_1^T \mathbf{a}_1 = 1$. Dazu muss das Gleichungssystem

$$(\boldsymbol{\Sigma} - \lambda \mathbf{I}_p) \mathbf{a}_1 = 0 \tag{3.11}$$

gelöst werden. Lösungen des Eigenwertproblems in (3.11) können nur Eigenwerte $\lambda_1 > \lambda_2 > \dots > \lambda_p \geq 0$ der Matrix $\boldsymbol{\Sigma}$ sein, die alle größer oder gleich Null sind, da die Kovarianzmatrix positiv semidefinit ist. Da die Varianz maximiert werden soll, wird als Lösung von (3.11) der größte Eigenwert λ_1 gewählt. Aus dem Eigenwertproblem ergibt sich, dass \mathbf{a}_1 der zugehörige Eigenvektor ist. Die erste Komponente erklärt den größten Anteil der Varianz der Daten. Die zweite Komponente $\mathbf{z}_2 = \mathbf{X}\mathbf{a}_2$ wird so konstruiert, dass $\text{Var}(\mathbf{z}_2)$ maximal ist unter den Nebenbedingungen $\mathbf{a}_2^T \mathbf{a}_2 = 1$ und $\mathbf{a}_2^T \mathbf{a}_1 = 0$. Die zweite Nebenbedingung besagt, dass die zweite Komponente orthogonal auf der ersten steht, d.h. $\text{Cov}(\mathbf{z}_1, \mathbf{z}_2) = 0$. Somit werden sukzessive die weiteren Komponenten konstruiert, die jeweils orthogonal auf den vorherigen stehen und unter diesen Nebenbedingungen maximale Varianz erklären. Die Lösung dieses Problems liefert eine Spektralzerlegung der Kovarianzmatrix in

$$\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Lambda}\mathbf{A}^T, \tag{3.12}$$

wobei die orthogonale $p \times p$ -Matrix \mathbf{A} der Eigenvektoren genau die gesuchten Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_p$ enthält und $\boldsymbol{\Lambda}$ eine $p \times p$ -Diagonalmatrix mit den gesuchten Eigenwerten $\lambda_1, \dots, \lambda_p$ in absteigender Größe auf der Diagonalen ist. Diese können als Varianz der Hauptkomponenten $\mathbf{z}_1, \dots, \mathbf{z}_p$ interpretiert werden. Nun gilt mit Gleichung (3.12) folgendes:

$$\sum_{j=1}^p \text{Var}(\mathbf{z}_j) = \sum_{j=1}^p \lambda_j = \text{Spur}(\boldsymbol{\Lambda}) = \text{Spur}(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T) = \text{Spur}(\boldsymbol{\Sigma}\mathbf{A}^T\mathbf{A}) = \text{Spur}(\boldsymbol{\Sigma}) = \sum_{j=1}^p \text{Var}(\mathbf{x}_j). \tag{3.13}$$

Gleichung (3.13) besagt, dass die Summe der Varianzen der ursprünglichen Variablen und der Hauptkomponenten identisch ist. Damit lässt sich sagen, dass die k -te Komponente den Anteil $\frac{\lambda_k}{\sum_{j=1}^p \lambda_j}$ der Gesamtvarianz erklärt. Eine Hauptkomponente mit einem Eigenwert größer als 1 erklärt mehr Varianz als eine durchschnittliche ursprüngliche Variable, eine Hauptkomponente mit Eigenwert kleiner als 1 erklärt weniger Varianz als eine durchschnittliche ursprüngliche Variable. Erklärt eine Anzahl von q Komponenten die Daten bereits gut, so wird angenommen, dass es höchstens q latente Faktoren gibt. In der Regel ist es das Ziel, die Dimension der Daten zu reduzieren und sich auf wenige Kom-

ponenten zu beschränken. Die Frage, wie viele Komponenten zu verwenden sind, ist dabei subjektiv. Es existieren aber eine Reihe von möglichen Entscheidungshilfen:

- Beim *Scree-Test* (Cattell, 1966) werden die Eigenwerte λ_j gegen den Index j angetragen. Hat dieser Graph einen deutlichen Knick, so kann man die Eigenwerte rechts von diesem Knick als nur zufällig von Null verschieden auffassen. Diese Komponenten werden dann weggelassen.
- Kaiser-Kriterium: Man verwendet so viele Hauptkomponenten, wie es Eigenwerte größer als 1 gibt, d.h. man verwendet nur die Komponenten, die mehr Varianz als eine ursprüngliche Variable erklären.
- Man verwendet so viele Hauptkomponenten, bis deren kumulierte Varianz mindestens einen vorgegebenen Anteil κ der Gesamtvarianz erklärt, z.B. $\kappa = 0.80$.
- Es wird inhaltliches Vorwissen eingesetzt.

Liefern die genannten Verfahren unterschiedliche Ergebnisse für die Anzahl der Faktoren, so wird häufig die kleinste Faktorenzahl gewählt. Für jedes Individuum des Datensatzes \mathbf{X} können dann seine Ausprägungen bezüglich der ausgewählten Komponenten berechnet werden.

Zur Schätzung der Ladungsmatrix werden bei der Hauptkomponentenmethode die q ausgewählten Komponenten

$$\hat{\mathbf{L}}_{(q)} = (\mathbf{a}_1^* \cdots \mathbf{a}_q^*),$$

mit $\mathbf{a}_j^* = \mathbf{a}_j \sqrt{\lambda_j}$ zu der Ladungsmatrix zusammengesetzt. Analog zu Gleichung (3.10) ergibt sich dann

$$\mathbf{\Sigma} = \hat{\mathbf{L}}_{(q)} \hat{\mathbf{L}}_{(q)}^T + \hat{\mathbf{\Psi}}_{(q)},$$

wobei $\hat{\mathbf{\Psi}}_{(q)}$ die geschätzte Fehlermatrix ist, die aber im Gegensatz zu (3.10) nun keine diagonale Struktur mehr aufweist.

3.1.7. Rotation der Faktoren

In Anschluss an eine Faktorenanalyse wird die geschätzte Ladungsmatrix häufig rotiert, damit sie leichter zu interpretieren ist. Dabei wird zwischen *orthogonalen* und *schiefwinkligen* Rotationen unterschieden (Hartung u. Elpelt, 1999). Man beachte, dass die Schätzverfahren zunächst orthogonale Faktoren konstruieren. Bei einer orthogonalen Rotation wird dann die Ladungsmatrix so transformiert, dass die Unkorreliertheit der Faktoren erhalten bleibt, während diese Forderung bei schiefwinkligen Rotationen aufgegeben wird. Bei einer orthogonalen Rotation wird die Ladungsmatrix mit einer orthonormalen $q \times q$ Transformationsmatrix \mathbf{A} multipliziert, für die gilt $\mathbf{A}\mathbf{A}^T = \mathbf{I}$. Dabei werden alle Faktoren um denselben Winkel rotiert. Die rotierte Ladungsmatrix erhält man nun als

$$\hat{\mathbf{L}}_{\text{rot}} = \hat{\mathbf{L}}\mathbf{A}.$$

Geometrisch gesprochen wird durch die Drehung erreicht, dass die Faktoren im Raum näher am Koordinatensystem der Variablen liegen. Inhaltlich gesehen soll die Ladungsmatrix die so genannte *Einfachstruktur* erfüllen (Thurstone, 1947). Diese besagt, dass die beobachteten Variablen mit einigen Faktoren hoch laden und mit den anderen annähernd gar nicht (Nullladungen). Dadurch werden die Ergebnisse einer Faktorenanalyse leichter interpretierbar, da jeder Faktor sich durch einige wenige Variablen ausdrücken und darstellen lässt. Ein Verfahren zur Näherung an die Einfachstruktur ist die orthogonale Rotation nach der *Varimax-Methode* (Kaiser, 1958). Die iterative Maximierung für

dieses Verfahren wird in Hartung u. Elpelt (1999) beschrieben. Der Anteil der Varianz, der durch die Faktoren erklärt wird, ändert sich durch die Rotation nicht, da orthonormale Transformationsmatrizen verwendet werden.

Die *Promax-Rotation* geht von einer Varimax-Rotation aus und versucht, das Ergebnis weiter zu verbessern, indem die Orthogonalitätsforderung für die Faktoren aufgegeben wird und für jeden Faktor ein eigener, optimaler Drehwinkel bestimmt wird. Den Faktoren kann dann allerdings kein Anteil an erklärter Varianz mehr zugeordnet werden. Schiefwinklige Rotationen sind dann sinnvoll, wenn nicht von einer Unabhängigkeit der Faktoren ausgegangen werden kann oder weitere, hinter den Faktoren stehende Muster ermittelt werden sollen. Aus den korrelierten Faktoren können dann in einem zweiten Schritt *Faktoren zweiter Ordnung* ermittelt werden, indem die Faktoren der ersten Stufe als Variablen für eine zweite Faktorenanalyse verwendet werden. Bei orthogonalen Faktoren kann keine erneute Faktorenanalyse durchgeführt werden. Ob korrelierte Faktoren sinnvoll sind, muss immer auch nach inhaltlichen Gesichtspunkten beurteilt werden.

3.1.8. Interpretation der Ergebnisse

Die Ergebnisse einer orthogonalen Rotation sind in der Regel einfacher zu interpretieren, da bei schiefwinkligen Rotationen berücksichtigt werden muss, dass die Faktoren nicht unabhängig sind. Die Interpretation der Zusammenhänge zwischen den ursprünglichen Variablen und den Faktoren geschieht anhand der Ladungsmatrix. Ist die Ladung zwischen einem Faktor und einer Variablen betragsmäßig hoch, so wird diese Variable dem Faktor zugeordnet. So können variablenübergreifende Konstrukte gebildet werden (z.B. die allgemeine Sprungkraftfähigkeit). Die Variablen, die mit dem entsprechenden Faktor hoch laden, messen dieses Konstrukt. Hat eine Variable hohe Nebenladungen auf weiteren Faktoren, so bedeutet dies, dass sie noch einen zusätzlichen Faktor erfasst und verschiedene Konstrukte misst. Hat eine Variable auf allen Faktoren nur geringe Ladungen, so misst diese Variable etwas anderes als die Faktoren. Variablen, die einem Faktor zugeordnet werden, stehen untereinander in höherem Zusammenhang als solche, die verschiedenen Faktoren zugeordnet werden.

3.1.9. Schätzung der Faktorenwerte

Bisher wurde die Anzahl der Faktoren q festgelegt und Punktschätzer für die Parameter \mathbf{L} und $\mathbf{\Psi}$ des Faktorenmodells gefunden. Nun soll die Matrix der Faktorenwerte \mathbf{F} gefunden werden, d.h. die Ausprägungen jeder Beobachtung $i = 1, \dots, n$ für die q identifizierten Faktoren geschätzt werden. Für die Schätzung der Faktorenwerte gibt es verschiedene Ansätze, wobei zwei davon die *Maximum-Likelihood-Methode* und die *Regressionsmethode* sind. Bei ersterer soll die Likelihood $L(\mathbf{X}|\mathbf{F})$ bezüglich \mathbf{F} maximiert werden. Als Schätzer ergibt sich bei standardisierten Daten

$$\hat{\mathbf{F}}_{ML}^T = (\hat{\mathbf{L}}^T \hat{\mathbf{\Psi}}^{-1} \hat{\mathbf{L}})^{-1} \hat{\mathbf{L}}^T \hat{\mathbf{\Psi}}^{-1} \mathbf{X}^T. \quad (3.14)$$

Es gilt dabei, dass $\hat{\mathbf{F}}_{ML}$ erwartungstreu für \mathbf{F} ist. Bei der Regressionsmethode wird der Schätzer $\hat{\mathbf{F}}_{Reg}$ durch eine lineare Regression von \mathbf{F} auf \mathbf{X} ermittelt. Mit $\hat{\mathbf{\Phi}}$ als geschätzte Kovarianzmatrix von \mathbf{F} ergibt sich der Schätzer als

$$\hat{\mathbf{F}}_{Reg}^T = \hat{\mathbf{\Phi}}(\mathbf{I} + \hat{\mathbf{L}}^T \hat{\mathbf{\Psi}}^{-1} \hat{\mathbf{L}} \hat{\mathbf{\Phi}})^{-1} \hat{\mathbf{L}}^T \hat{\mathbf{\Psi}}^{-1} \mathbf{X}^T, \quad (3.15)$$

wiederum für standardisierte Daten \mathbf{X} . Dabei gilt, dass der Regressionsschätzer nicht unverzerrt ist. Dafür bietet er unter allen in \mathbf{X} linearen Schätzern die geringste Fehlervarianz (Fahrmeir u. a.,

1996). Für jedes Individuum können dann die Ausprägungen bezüglich der unterschiedlichen Faktoren geschätzt werden. Dies entspricht den Einträgen in der Faktorenwertmatrix F . Große Werte bedeuten hier, dass die Variablenausprägungen einer Beobachtung bezüglich eines Faktors sehr ausgeprägt sind, also beispielsweise eine hohe allgemeine Sprungkraftfähigkeit vorliegt.

3.1.10. Quantifizierung der Unsicherheit der Ergebnisse (Inferenzstatistik)

In den vorherigen Abschnitten wurden stets Punktschätzer der Parameter des Faktorenmodells betrachtet. In der *Inferenzstatistik* soll die Unsicherheit der Punktschätzer durch Angabe von Standardfehlern und Konfidenzintervallen quantifiziert werden. Die Frage ist, wie anhand der vorliegenden Stichprobe (alle Athleten, die ins Faktorenmodell aufgenommen werden) auf die Grundgesamtheit (alle Skispringer und Nordischen Kombiniierer) geschlossen werden kann. Da die Bildung der Stichprobe einem Zufallsprozess unterliegt, so sind auch die aus der Stichprobe gewonnenen Punktschätzer Zufallsvariablen, die mit einer Verteilung versehen sind.

Für diese Verteilung können entweder parametrische Annahmen getroffen werden oder die Verteilung wird direkt aus der Stichprobe geschätzt. Die letztere Herangehensweise ist die Idee der sogenannten *nichtparametrischen Bootstrap-Verfahren*. Diese verzichten auf konkrete Verteilungsannahmen, sondern konstruieren die Verteilung aus der vorliegenden Stichprobe. Im folgenden wird die Idee der Bootstrap-Schätzung zur Konstruktion von Standardfehlern und Konfidenzintervallen erläutert. Die Darstellung orientiert sich dabei an Efron u. Tibshirani (1993).

Bootstrap-Methoden werden eingesetzt, um die Verteilung von statistischen Kennwerten zu approximieren, wenn keine parametrischen Annahmen gemacht werden können oder sollen. Ausgangspunkt ist die Schätzung eines Parameters θ . Dies ist in der Faktorenanalyse eine Faktorladung l_{jk} oder ein Eigenwert λ_j . Dieser Parameter wird anhand der vorliegenden Stichprobe geschätzt und der Schätzwert wird mit $\hat{\theta}$ bezeichnet. Kennwerte der Verteilung von $\hat{\theta}$ werden mit Bootstrap-Verfahren geschätzt. Dies umfasst die folgenden Schritte:

1. Es wird B mal eine Stichprobe des Umfangs n aus der vorliegenden Stichprobe *mit Zurücklegen* gezogen. Jede der $b = 1, \dots, B$ gezogenen Stichproben wird als eine Bootstrap-Stichprobe bezeichnet. Damit kann eine einzelne Beobachtung 0 bis n -mal in einer Bootstrap-Stichprobe vorhanden sein.
2. In jeder Bootstrap-Stichprobe wird eine Schätzung für den interessierenden Parameter θ_b^* durchgeführt, d.h. das Faktorenmodell aufgestellt und die Eigenwerte und Ladungskoeffizienten bestimmt.
3. Es werden Kennwerte der Verteilung des Punktschätzers aus den Bootstrap-Schätzern $\theta^* = (\theta_1^*, \dots, \theta_B^*)$ ermittelt.

Aus der Verteilung von θ^* kann auf die Verteilung von $\hat{\theta}$ geschlossen werden (Bootstrap-Schätzung). Hier geht der Einfluss der Stichprobenzusammensetzung ein. Folgende Kennwerte von Verteilungen der Punktschätzer werden betrachtet:

Bootstrap-Schätzung eines Standardfehlers: Der theoretische Standardfehler $se(\theta)$ wird geschätzt durch

$$\hat{se}_B = \left(\frac{1}{B-1} \sum_{b=1}^B (\theta_b^* - \theta_{(\cdot)}^*)^2 \right)^{\frac{1}{2}}, \quad \text{wobei } \theta_{(\cdot)}^* = \frac{1}{B} \sum_{b=1}^B \theta_b^*.$$

Bootstrap-Schätzung eines Konfidenzintervalls: Zur Angabe von Konfidenzintervallen für $\hat{\theta}$ mit Bootstrap gibt es verschiedene Ansätze, von denen drei im Folgenden beschrieben werden.

- Der einfachste ist die klassische Konstruktion eines zweiseitigen $(1 - \alpha)$ -Konfidenzintervalles mit der Bootstrap-Schätzung des Standardfehlers $\hat{s}e_B$

$$\hat{\theta} \pm t^{n-1, 1-\frac{\alpha}{2}} \cdot \hat{s}e_B,$$

wobei $t^{n-1, 1-\frac{\alpha}{2}}$ das $(1 - \frac{\alpha}{2})$ -Quantil einer t-Verteilung mit $n - 1$ Freiheitsgraden ist. Dieses Verfahren hat jedoch verschiedene Nachteile. So können die Grenzen dieses Konfidenzintervalls außerhalb des Definitionsbereichs von θ liegen und es können nur symmetrische Konfidenzintervalle konstruiert werden.

- Bei der *Bootstrap-Perzentil-Methode* wird das Konfidenzintervall direkt aus der empirischen Verteilung von θ^* konstruiert. Dazu werden die θ_b^* der Größe nach sortiert und die Grenzen des Konfidenzintervalls geschätzt durch

$$[\theta^{*(\alpha/2)}, \theta^{*(1-\alpha/2)}],$$

wobei $\theta^{*(\alpha/2)}$ die $B \cdot \alpha/2$ -te Stelle in der geordneten Liste von $(\theta_1^*, \dots, \theta_B^*)$ und $\theta^{*(1-\alpha/2)}$ die $B \cdot (1 - \alpha/2)$ -te Stelle in der geordneten Liste angibt. Die Perzentil-Methode führt automatisch zu Konfidenzintervallen, deren Grenzen im Definitionsbereich von θ liegen (z.B. $\lambda_j \geq 0$), jedoch unterdecken die Intervalle oft das wahre Konfidenzintervall, d.h. sie haben eine kleinere tatsächliche Überdeckungswahrscheinlichkeit.

- In Efron u. Tibshirani (1993) wird als Alternative zu den obigen Konfidenzintervallen das sogenannte *bias-corrected and accelerated*-Konfidenzintervall (kurz BC_α) eingeführt. Mit einer Korrektur für die Verzerrung („bias“) und die sogenannte Beschleunigung („acceleration“) sollen die Probleme der Perzentil-Methode vermieden werden. Das Verfahren korrigiert die Konfidenzgrenzen für verzerrte Bootstrap-Schätzer im Gegensatz zur Perzentil-Methode automatisch. Das $(1 - \alpha)$ - BC_α Konfidenzintervall ergibt sich als

$$[\theta^{*(\alpha_1)}, \theta^{*(\alpha_2)}],$$

wobei $\theta^{*(\alpha_{[1,2]})}$ wieder die $B \cdot \alpha_{[1,2]}$ -te Stelle in der geordneten Liste von θ_b^* bezeichnet und gilt

$$\begin{aligned} \alpha_1 &= \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha/2)}}{1 - \hat{\alpha}(\hat{z}_0 + z^{(\alpha/2)})} \right) \\ \alpha_2 &= \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha/2)}}{1 - \hat{\alpha}(\hat{z}_0 + z^{(1-\alpha/2)})} \right), \end{aligned} \quad (3.16)$$

wobei mit Φ die Verteilungsfunktion der Standardnormalverteilung bezeichnet wird und $z^{(\alpha)}$ das $100 \cdot \alpha$ -Quantil der Standardnormalverteilung ist. Man beachte, dass bei einer Wahl von $\hat{z}_0 = 0$ und $\hat{\alpha} = 0$ in (3.16) das BC_α -Konfidenzintervall wegen $\Phi(z^{(\alpha)}) = \alpha$ dem Perzentil-Intervall entspricht. Diese beiden Parameter geben damit die Korrektur an. Sie berechnen sich als

$$\hat{z}_0 = \Phi^{-1} \left(\frac{\#(\theta_b^* < \hat{\theta})}{B} \right),$$

wobei Φ^{-1} die Quantilsfunktion der Standardnormalverteilung bezeichnet und $\#(\theta_b^* < \hat{\theta})$ den Anteil der Bootstrap-Replikationen, die kleiner als der originale Schätzer sind. Die

Beschleunigung wird geschätzt zu

$$\hat{\alpha} = \frac{\sum_{i=1}^n (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^2}{6(\sum_{i=1}^n (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^{\frac{3}{2}})},$$

wobei $\hat{\theta}_{(i)}$ die Schätzung für θ ohne die i -te Beobachtung der Stichprobe ist und $\hat{\theta}_{(\cdot)} = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{(i)}$. Damit können die Intervallgrenzen mit den Gleichungen aus (3.16) bestimmt werden. Das so konstruierte Konfidenzintervall weist einige theoretische Vorteile auf und ist den anderen Methoden vorzuziehen (vgl. Efron u. Tibshirani, 1993, S.187).

Für die Wahl von B gibt es Faustregeln, die anhand theoretischer Überlegungen besagen, dass $B = 200$ zur Schätzung von Standardfehlern ausreichend ist, wohingegen $B = 2000$ zur Schätzung von Konfidenzintervallen benötigt wird. Insgesamt gibt es $\binom{2n-1}{n}$ mögliche Bootstrap-Stichproben. Wählt man B als Anzahl der mögliche Stichproben, so spricht man von einer *idealen Bootstrap-Schätzung*. In der Praxis ist dies jedoch nicht zu realisieren, so dass man sich auf die obigen Faustregel beschränkt. Wichtig ist auch eine gute Qualität der Stichprobe, d.h. sie soll die Grundgesamtheit gut abbilden. Dabei wird angenommen, dass die einzelnen Elemente der Stichprobe unabhängig sind. Als Alternative bei Zeitreihen kommt der *Moving Block Bootstrap* in Frage, der jedoch nicht unumstritten ist (Hall u. Horowitz, 1996, vgl.) und auf den hier verzichtet wird.

Bootstrap bietet eine Möglichkeit zur Konstruktion von Konfidenzintervallen für Parameter des Faktorenmodells wie die Eigenwerte λ_j , $j = 1, \dots, p$, für die keine theoretische Verteilung bekannt ist. In der Praxis werden solche Unsicherheitsabschätzungen nur selten gemacht. Als Beispiel wird in Shikano (2006) die Verteilung der Eigenwerte bei einem Faktorenmodell aus der Politikwissenschaft untersucht. Hier wird das Faktorenmodell von Lijphart (1999) kritisiert, der nach dem Kaiser-Kriterium zwei Faktoren zu Dimensionen der Demokratie auswählt. Dort ist der Punktschätzer für den dritten Eigenwert nur knapp kleiner 1, aber Shikano (2006) zeigt, dass der dritte Eigenwert im Mittel der Bootstrap-Stichproben größer als 1 ist.

Zusammenfassung der Faktorenanalyse

Die Faktorenanalyse erlaubt also die Ermittlung von latenten (und unabhängigen) Faktoren, die die Information der ursprünglichen Variablen auf wenige Faktoren reduzieren und einen großen Teil der Originalvarianz erklären können. Im Folgenden wird versucht, anhand der Daten zur Leistungsdiagnostik einige wenige Faktoren zu konstruieren, die einen großen Teil der Unterschiede der Athleten erklären und aus sportwissenschaftlicher Sicht gut zu interpretieren sind. Die Unsicherheit der Punktschätzer wird mit Bootstrap-Verfahren untersucht.

3.2. Auswertung und Ergebnisse

In diesem Abschnitt wird ein Faktorenmodell für die Daten zur Leistungsdiagnostik angepasst. Durch Betrachten der Ladungsmatrix können die Zusammenhänge zwischen Variablen und Faktoren analysiert werden. Zunächst folgt eine Motivation, warum dieses Verfahren in der konkreten Situation angewendet wird. Anschließend werden die weiteren inhaltlichen und technischen Schritte der Datenvorbereitung erläutert. Vor der Schätzung der Ladungsmatrix muss die Anzahl der Faktoren festgelegt werden, womit sich Abschnitt 3.2.3 beschäftigt. Anschließend wird die Ladungsmatrix geschätzt und interpretiert und in Abschnitt 3.2.5 die Faktorenwerte der einzelnen Beobachtungen geschätzt und dargestellt. Es folgen Unsicherheitsabschätzungen der gewonnenen Punktschätzer von Eigenwerten

und Ladungskoeffizienten mit Bootstrap-Verfahren in Abschnitt 3.2.6.

3.2.1. Motivation zur Verwendung einer Faktorenanalyse in der Leistungsdiagnostik

Im Rahmen der Leistungsdiagnostik von Kaderathleten aus den Bereichen Skisprung und Nordische Kombination wurde eine Reihe von Variablen anhand verschiedener Sprungkraft- und Beinkrafttests erhoben. Möglicherweise lässt sich die Leistung der Athleten jedoch durch einige wenige Faktoren erklären. So ist es denkbar, dass es einen Faktor der allgemeinen Sprungkraftfähigkeit gibt, der sich sowohl bei den Ergebnissen beim Squat Jump als auch beim Drop Jump äußert. Die zugehörigen Variablen SJ_SpH, DJ32_SpH und möglicherweise SJ_V75E und DJ32_KT würden also dasselbe Konstrukt „Sprungkraftfähigkeit“ messen. Es wäre dabei möglich, dass verschiedene Sprungkrafttests nötig sind, um diesen Faktor zu operationalisieren oder im anderen Fall, dass nur eine Sprungform ausreichend ist, um die Sprungkraftfähigkeit eines Athleten zu messen (bei sehr hohen Korrelationen der ursprünglichen Variablen).

Ein anderer praktischer Gewinn durch die Faktorenanalyse ist die Reduktion der Modellkomplexität mit dem Effekt, dass die vorhandenen Daten der Athleten schneller anhand von den Faktorenwerten verglichen werden können. Anstelle detaillierter Profile aller p Variablen wie der in Abschnitt 2.5.4, verwendet man dann die reduzierten Profile der $q < p$ Faktoren. Somit können Stärken und Defizite von Athleten in allgemeiner Form festgestellt und so Hinweise zur Trainingssteuerung geliefert werden. Denkbar wäre, dass eine Feststellung, dass die Sprungkraftfähigkeit allgemein trainiert werden soll, den Trainern mehr Information liefert, als dass beispielsweise die hinter einer speziellen Variablen stehende Fähigkeit, z.B. die Absprunggeschwindigkeit beim Squat Jump, trainiert werden soll. Im Folgenden wird anhand der vorliegenden Daten versucht, solche Faktoren zu konstruieren. Da über die Anzahl und Art der Faktoren zunächst nichts bekannt ist, hat die hier durchgeführte Faktorenanalyse explorativen Charakter. Die gewonnenen Ergebnisse haben zunächst einmal nur Gültigkeit für die Gruppe an Sportlern, die analysiert wurden. Eine Verallgemeinerung auf zukünftige Athleten ist nur dann zulässig, wenn keine substantziellen Unterschiede zu vermuten sind, z.B. durch neue Trainingsmethoden.

3.2.2. Vorbereitung der Analyse

Vor der Durchführung der Faktorenanalyse wird der Datensatz zur Leistungsdiagnostik weiter vorbereitet. Dies umfasst die folgenden Schritte:

1. Anstelle der absoluten Werte werden die bezüglich des Alters normierten Z-Werte aus Abschnitt 2.5.4 verwendet. Damit werden jeweils die relativen Stärken und Schwächen eines Athleten betrachtet. Für die Bildung der Referenzwerte werden dabei alle zur Verfügung stehenden Werte verwendet.
2. Im Rahmen der Faktorenanalyse werden alle Variablen simultan miteinander verglichen, so dass nur die Beobachtungen, bei denen alle Messergebnisse vorliegen, berücksichtigt werden können (Complete Case Analyse). Insgesamt können also nur die 326 vollständigen Beobachtungen von 48 Skispringern und 40 Nordischen Kombinierern ausgewertet werden (vgl. Abschnitt 2.4). Die Menge dieser Beobachtungen wird im Folgenden als reduzierter Datensatz bezeichnet. Die Verteilung der Anzahl der Beobachtungen je Athlet im reduzierten Datensatz ist in Tabelle 3.4 dargestellt.

Man sieht, dass durch Reduktion auf die vollständigen Beobachtungen für die meisten Athleten nur wenige Beobachtungen zur Analyse verbleiben. Im Mittel liegen noch 3.7 Beobachtungen je

Beobachtungen je Athlet	1	2	3	4	5	6	7	8	9	10
Häufigkeit	17	21	13	7	9	8	3	5	4	1

Tabelle 3.4.: Verteilung der Athleten nach der Anzahl ihrer Messungen im reduzierten Datensatz.

Athlet vor.

- In einem nächsten Schritt werden aus den Variablen des Datensatzes die interessierenden ausgewählt. Dies sind alle erhobenen und abgeleiteten Größen zu Sprungkraft- und Krafttest der Leistungsdiagnostik, entsprechend allen Variablen aus den Tabellen 2.2 bis 2.5. Nicht berücksichtigt werden Untersuchungsort und Zeitpunkt (keine stetigen Größen und nicht relevant), das Alter (bereits durch die Skalierung berücksichtigt), Größe und Gewicht (indirekt in den Variablen zur relativen Maximalkraft $F_{re170_}[bb, re, li]$).

Der so formatierte Datensatz bildet die Grundlage für die weitere Faktorenanalyse. Er enthält $n=326$ Beobachtungen von $p=23$ Variablen. Die weitere Notation gestaltet sich analog zum vorherigen Abschnitt bzw. Tabellen 3.1 und 3.2. Im nächsten Schritt muss die Anzahl der zu schätzenden Faktoren im Faktorenmodell (3.3) festgelegt werden.

3.2.3. Festlegung der Anzahl der Faktoren

Für die Wahl der Anzahl der Faktoren gibt es keine objektiv beste Lösung. Stattdessen wird hier versucht auch mit inhaltlichen Argumenten ein Anzahl von Faktoren zu wählen, die einen Kompromiss aus guter Datenanpassung und geringer Komplexität bzw. guter Interpretierbarkeit liefert. Sollen anhand der Faktorenanalyse für die Trainer Rückschlüsse gewonnen werden, so sind gut zu interpretierende Faktoren mit einer eindeutigen Wirkungsrichtung (z.B. „große Faktorenwerte entsprechen guter Leistung“) wichtig. Einen ersten Eindruck, wie gut die Anpassung eines Faktorenmodells an die Daten ist, liefert eine Eigenwertzerlegung der Kovarianzmatrix Σ in p Eigenwerte und p zugehörige Eigenvektoren. Die ersten fünf Eigenwerte der insgesamt 23 sind dabei $\lambda_1=8.5$, $\lambda_2=2.23$, $\lambda_3=2.17$, $\lambda_4=1.6$ und $\lambda_5=1.5$. In Abbildung 3.3 sind alle 23 Eigenwerte und der jeweilige Anteil der erklärten Varianz dargestellt.

Der Scree-Test weist einen Knick nach dem ersten und dritten Eigenwert auf. Dementsprechend würde man einen bzw. drei Faktoren wählen, womit bereits 37.1% bzw. 56.2% der Varianz der ursprünglichen Variablen erklärt werden. Mit vier Faktoren würde man 63.3%, mit fünf 69.9% und mit sechs 75.4% der Originalvarianz erklären. Durch Hinzunehmen von weiteren Faktoren lassen sich jeweils nur noch wenige Prozentpunkte mehr an erklärter Gesamtvarianz erreichen. Für die folgenden Analysen wird festgelegt, dass drei Faktoren geschätzt werden sollen. Es wurden Faktorenmodelle für verschiedene Anzahlen von Faktoren berechnet und dabei konnten Faktorenmodelle mit drei Faktoren inhaltlich gut interpretiert werden. Die Entscheidung für drei Faktoren entspricht dem Ergebnis des Scree-Tests.

3.2.4. Schätzung der Ladungsmatrix

Im nächsten Schritt wird das Faktorenmodell mit $q=3$ Faktoren aufgestellt und die Ladungsmatrix L mit Hilfe des Maximum-Likelihood-Verfahrens geschätzt. Anschließend werden die Faktoren einmal mit der Varimax-Methode und einmal mit der Promax-Methode rotiert und jeweils die geschätzte Ladungsmatrix gezeigt und gegenübergestellt. Bei der schiefwinkligen Rotation können in einem zweiten Schritt Faktoren der zweiten Ordnung geschätzt werden, da die Faktoren nicht nach Definition unkorreliert sind. Die Ergebnisse dieser Faktoren zweiter Ordnung werden in Abschnitt 3.2.7 gezeigt.

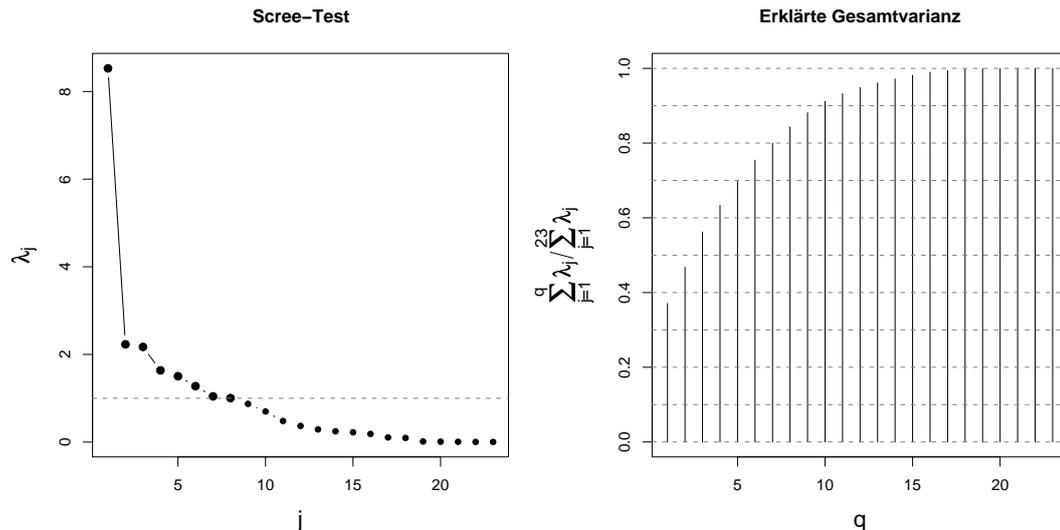


Abbildung 3.3.: Bestimmung der Anzahl der Faktoren: Die linke Abbildung zeigt die Eigenwerte $\lambda_1, \dots, \lambda_{23}$ (Scree-Test). Dabei sind 8 Eigenwerte größer als 1. Die rechte Abbildung stellt den kumulierten Anteil der erklärten Varianz $\sum_{j=1}^k \lambda_j / \sum_{j=1}^{23} \lambda_j$ durch $k=1, \dots, 23$ orthogonale Faktoren dar.

Orthogonale Rotation der Faktoren

Die geschätzte Ladungsmatrix nach einer Rotation mit der Varimax-Methode ist in Tabelle 3.5 dargestellt. Man beachte, dass die Varimax-Rotation wieder orthogonale Faktoren generiert und der Ladungskoeffizient l_{jk} die Korrelation der j -ten Variable mit dem k -ten Faktor angibt. Die Faktoren können getrennt interpretiert werden, da sie untereinander unabhängig sind.

Die Faktoren werden, wie in Abschnitt 3.1.8 beschrieben, interpretiert. Man sieht, dass nicht alle der $p=23$ Variablen mit einem der drei Faktoren eine Korrelation größer als 0.3 haben. Der Wert von 0.3 wird als bedeutsame Korrelation angesehen. Die Variablen `DJ32_KT`, `relKraftanstieg_bb`, `Kdif`, `ekV`, `kkV`, `dynstat_60` und `dynstat_180` haben keine bedeutsame Ladung mit einem der Faktoren. Der Grund liegt darin, dass manche dieser Variablen keine eindeutige Wirkungsrichtung im Gegensatz zu den anderen Variablen haben. Manche Variablen drücken sich hingegen in mehreren Faktoren aus, was bedeutet, dass diese Variablen mehrere Faktoren messen. Dies trifft auf die Variablen zur Maximalkraft zu, d.h. `Fmax90_bb`, `Fmax70_[bb, re, li]`, `F_re170_[bb, re, li]` und `Ekz` zu, die jeweils mit dem ersten und zweiten und teilweise dem dritten Faktor laden.

Bei den Kommunalitäten aus Tabelle 3.5 sieht man, dass für die Variablen, die von keinem der Faktoren erfasst werden, nahezu keine Varianz durch das Faktorenmodell erklärt werden kann. Die Richtungen großer Streuung zwischen den Beobachtungen bei diesen Variablen stimmen also nicht mit der Richtung bei den anderen Variablen überein. Hingegen werden besonders die Variablen zur Maximalkraft fast vollständig durch das Faktorenmodell erfasst. Ein verhältnismäßig großer Anteil an unerklärter Streuung verbleibt bei den Variablen zur Sprungkraftdiagnostik. Insbesondere die Bodenkontaktzeit beim Drop Jump wird nicht durch das Faktorenmodell erfasst.

Schiefwinklige Rotation der Faktoren

Mit einer schiefwinkligen Rotation wird versucht, eine bessere Annäherung der Ladungsmatrix an die Einfachstruktur zu erreichen. Dazu wird die Orthogonalitätsforderung für die Faktoren aufgegeben

Variable j	Faktor 1 l_{j1}	Faktor 2 l_{j2}	Faktor 3 l_{j3}	Kommunalität $k_j^2 = l_{j1}^2 + l_{j2}^2 + l_{j3}^2$
SJ_SpH	0.59			0.41
SJ_V75E	0.47			0.35
DJ32_SpH	0.46			0.23
DJ32_KT				0.03
DJ32_Faktor	0.40			0.17
Fmax90_bb	0.42	0.64	0.30	0.67
Fmax70_bb	0.54	0.76	0.34	0.99
Fmax70_re	0.51	0.77		0.86
Fmax70_li	0.54	0.82		0.99
F_rel70_bb	0.92	0.31		0.99
F_rel70_re	0.84	0.35		0.83
F_rel70_li	0.87	0.40		0.99
relKraftanstieg_bb				0.04
relKraftanstieg_re			0.30	0.09
relKraftanstieg_li			0.36	0.14
Ekz	0.52	0.74	0.30	0.90
Kdif				0.01
A60		0.55		0.35
A180		0.62		0.44
ekV				0.02
kkV				0.01
dynstat_60				0.04
dynstat_180				0.02

Tabelle 3.5.: Geschätzte Ladungsmatrix bei 3 Faktoren und Varimax-Rotation. Dargestellt sind nur die Ladungskoeffizienten, die betragsmäßig größer gleich 0.3 sind. Zudem wird für jede Variable die Kommunalität angegeben, also der Anteil der Varianz, der durch die drei Faktoren erklärt wird.

und mit Hilfe der Promax-Rotation schiefwinklige Faktoren geschätzt. Man beachte, dass die Faktoren nicht mehr unabhängig sind und damit nicht getrennt zu interpretieren sind. Die geschätzte Ladungsmatrix nach der Promax-Rotation ist in Tabelle 3.6 dargestellt.

Insgesamt ähnelt die Verteilung der Variablen auf die Faktoren der Ladungsmatrix nach einer Varimax-Rotation. Allerdings werden die Variablen zur Maximalkraft **Fmax90_bb** und **Fmax70_[bb, re, li]** nun dem zweiten Faktor zugeordnet (mit Nebenladungen auf Faktor 3) und die Variablen zur relativen Maximalkraft **F_rel70_[bb, re, li]** eindeutig dem ersten Faktor zugeordnet. Ein Vergleich der beiden Rotationen ist in Abbildung 3.4 dargestellt. Dort sind die Ladungen der Variablen auf den ersten beiden Faktoren gegenübergestellt.

Man sieht in Abbildung 3.4, dass sich bei der Varimax-Rotation viele Variablen im ersten Quadranten finden, d.h. sie haben eine hohe Ladung mit beiden Faktoren, wie im vorherigen Abschnitt dargestellt wurde. Bei der Promax-Rotation wird die Verteilung der Variablen auf die Faktoren deutlicher. Daher wird für die inhaltliche Interpretation der Faktoren die Promax-Rotation zugrunde gelegt. Dabei lassen sich die folgenden übergeordneten Begriffe vergeben:

Faktor 1: Allgemeine Sprungkraftfähigkeit und relative Maximalkraft Mit diesem Faktor laden vier der fünf Variablen zur Sprungkraftdiagnostik und nur die Bodenkontaktzeit **DJ32_KT** hat keine bedeutende Ladung mit diesem Faktor. Weiterhin laden die drei Variablen zur relativen Maximalkraft **F_rel70_[bb, re, li]** hoch mit diesem Faktor. Insgesamt fasst dieser Faktor die Variablen der Sprungkraft und relativen Maximalkraft zusammen. Athleten, die bezüglich Faktor 1 ausgeprägte Werte haben, kennzeichnet also, dass sie – jeweils relativ zur ihrer Altersreferenz-

Variable j	Faktor 1 l_{j1}	Faktor 2 l_{j2}	Faktor 3 l_{j3}
SJ_SpH	0.60		
SJ_V75E	0.39		
DJ32_SpH	0.47		
DJ32_KT			
DJ32_Faktor	0.44		
Fmax90_bb		0.55	0.35
Fmax70_bb		0.65	0.39
Fmax70_re		0.73	
Fmax70_li		0.85	
F_re170_bb	0.90		
F_re170_re	0.83		
F_re170_li	0.90		-0.32
relKraftanstieg_bb			
relKraftanstieg_re			0.33
relKraftanstieg_li			0.39
Ekz		0.64	0.34
Kdif			
A60		0.57	
A180		0.65	
ekV			
kkV			
dynstat_60			
dynstat_180			

Tabelle 3.6.: Geschätzte Ladungsmatrix bei 3 Faktoren und Promax-Rotation. Dargestellt sind nur die Ladungskoeffizienten, die betragsmäßig größer gleich 0.3 sind. Eine Angabe von Kommunalitäten ist bei der Promax-Rotation nicht sinnvoll.

gruppe – hohe Werte bei den Sprungkrafttests und zugleich eine hohe relative Maximalkraft haben. Das bedeutet auch, dass mit der relativen Maximalkraft auch die Sprungkraftfähigkeit eines Athleten gemessen wird.

Faktor 2: Allgemeine Maximalkraft und Schnellkraft Mit diesem Faktor laden die vier Variablen zur Maximalkraft $F_{\max 90_bb}$ und $F_{\max 70_ [bb, re, li]}$, sowie die exzentrische Maximalkraft Ekz und die Arbeit bei 60° bzw 180° Kniewinkelgeschwindigkeit. Dieser Faktor fasst also die Maximalkraftparameter und die Arbeit unter dynamischen Bedingungen zusammen. Athleten, die bezüglich diesem Faktor große Werte haben, kennzeichnet besonders die im Vergleich zur Referenzgruppe hohe Maximal- und Schnellkraft.

Faktor 3: Kraftanstiegsverhalten Dieser Faktor fasst die beidbeinigen Messungen zur Maximalkraft $F_{\max 90_bb}$ und $F_{\max 70_ [bb]}$ sowie die beiden einbeinigen Messungen zum Kraftanstieg $relKraftanstieg_ [re, li]$ zusammen. Athleten, die bezüglich diesem Faktor ausgeprägte Werte haben, kennzeichnen also neben der hohen Maximalkraft auch ein schneller Kraftanstieg.

Inhaltlich gesprochen sind bei allen Faktoren möglichst große Werte wünschenswert. Für jede Beobachtung können nun anhand der Variablenausprägungen die Faktorenausprägungen geschätzt werden. Die so konstruierten Faktoren sind nicht unabhängig. Aus sportwissenschaftlicher Sicht ist dies sinnvoll, da Korrelationen zwischen den Faktoren zu vermuten sind.

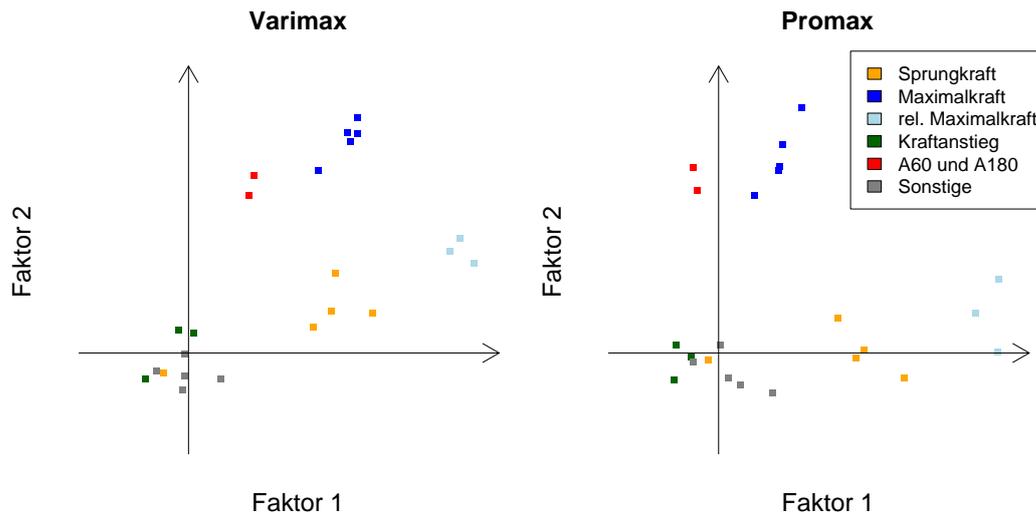


Abbildung 3.4.: Darstellung der Variablen im neuen Koordinatensystem der Faktoren 1 und 2. Die Position der farbigen Punkte in x- bzw. y-Richtung entspricht dabei den Ladungskoeffizienten der Variable mit Faktor 1 bzw. 2. Links ist die Varimax-Rotation und rechts die Promax-Rotation dargestellt.

3.2.5. Darstellung der Faktorenwerte

In einem nächsten Schritt werden für jede der 326 Beobachtungen die Ausprägungen bezüglich der drei Faktoren mit der Maximum-Likelihood-Methode geschätzt und daraus die Faktorenwertmatrix \mathbf{F} zusammengesetzt. Dabei wird für jede von mehreren Beobachtungen bei einem Athleten ein eigener Faktorenwert geschätzt. Das bedeutet inhaltlich, dass sich die zugrundeliegenden Faktorenwerte eines Athleten von Messung zu Messung ändern können. Die Spalten von \mathbf{F} sind dabei normiert auf Mittelwert 0 und Standardabweichung 1. Zur grafischen Darstellung von jeweils zwei Faktoren eignet sich der so genannte *Biplot* (Gabriel, 1971). Hier werden die Ausprägungen zweier Faktoren von allen Beobachtungen gegeneinander dargestellt. Jeder Punkt des Streudiagramms stellt dabei eine Beobachtung im Datensatz dar. In Abbildung 3.5 sind die beiden ersten Faktoren gegeneinander dargestellt.

Man sieht in Abbildung 3.5, dass sich die Beobachtungen bei Skispringern und die Beobachtungen bei Nordischen Kombinierern in der Randdichte zum ersten Faktor deutlich unterscheiden. Nordische Kombinierer haben im Mittel kleiner Werte beim ersten Faktor und damit schlechtere Resultate in den Sprungkraftübungen bzw. bei der relativen Maximalkraft. Inhaltlich interessant sind dabei die Athleten mit Faktorenwerten in der Gruppe der anderen Sportart. Bei dem zweiten Faktor zeigen sich keine größeren Unterschiede in der Verteilung der Faktorenwerte zwischen den Sportarten. Gleiches gilt auch für den dritten Faktor (nicht dargestellt). Weiterhin ist eine positive Korrelation zwischen Faktor 1 und 2 zu erkennen, da die Werte um die Winkelhalbierende herum streuen.

Wie in Abschnitt 3.2.1 angesprochen, liefern die Faktorenwerte eines Athleten einen schnellen Überblick über sein Leistungsniveau. Anstelle eines Profilbilds der ursprünglichen Variablen (vgl. Abbildung 2.8 auf Seite 24) tritt nun eine Darstellung der Faktorenwerte, die wie die Z-Werte normiert sind. In Abbildung 3.6 sind die Profile der Faktorenwerte von denselben Beobachtungen aus Abbildung 2.8 dargestellt. Da bei der neueren Messung des Nordischen Kombinierers die Messungen zu den Variablen Ekz und $Kdif$ fehlen, können für diese Beobachtung keine Faktorenwerte ermittelt werden.

Anhand der Profile der Faktorenwerte in Abbildung 3.6 kann direkt das Leistungsprofil der Athleten in aggregierter Form betrachtet werden. Der Nordische Kombinierer hat dabei keine ausgeprägten Stärken oder Schwächen bei einem der Faktoren. Die beiden Beobachtungen bei dem Skispringer

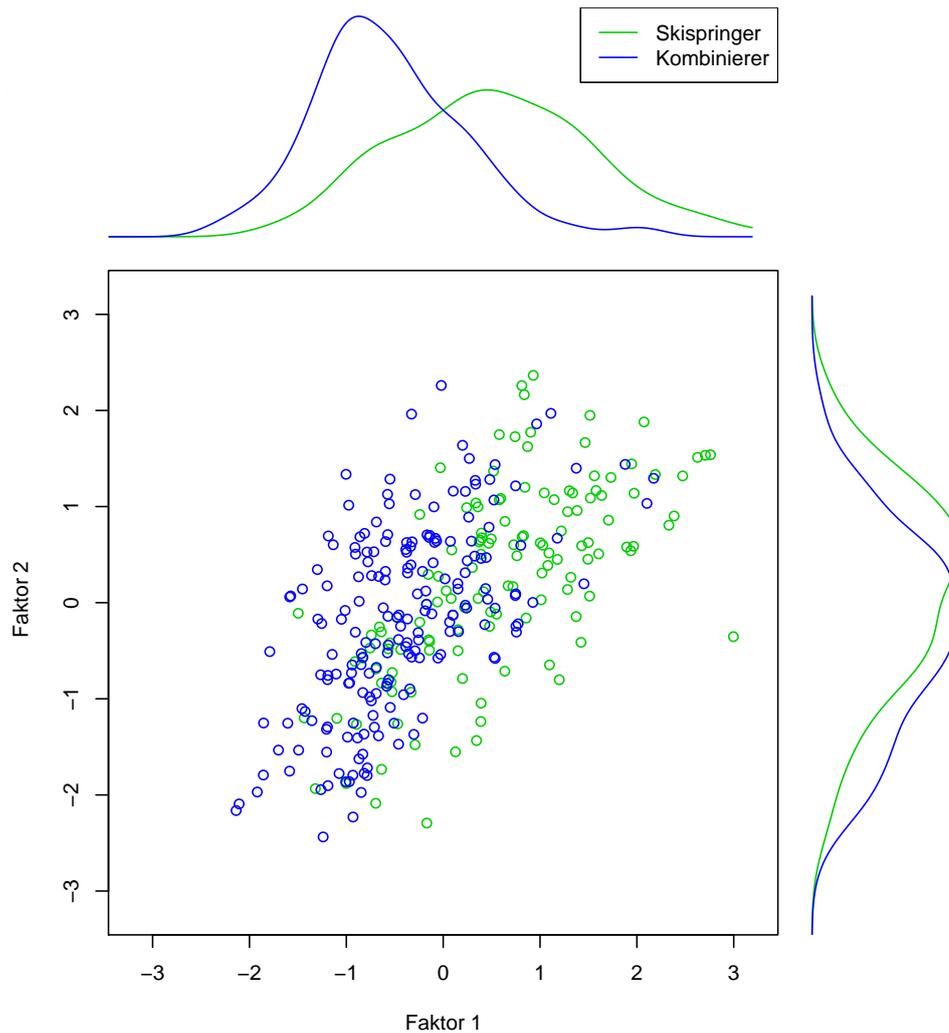


Abbildung 3.5.: Biplot der Faktoren 1 (Allgemeine Sprungkraftfähigkeit und relative Maximalkraft) und 2 (Allgemeine Maximalkraft und Schnellkraft) mit den Randdichten der beiden Faktoren, jeweils getrennt nach Sportart. Die Faktorenwerte wurden nach Gleichung (3.14) und der Ladungsmatrix aus Tabelle 3.6 (Promax-Rotation) ermittelt. Die Randdichte wird dabei mit Hilfe eines Kerndichteschätzers geschätzt.

zeigen jeweils überdurchschnittliche Ergebnisse bei allen drei Faktoren. Allerdings hat sein relatives Leistungsvermögen bezüglich seiner Referenzgruppe bei dem dritten Faktor von der ersten Messung zur zweiten abgenommen. Anhand dieser Profile der Faktorenwerte können individuell für jede Beobachtung das aktuelle Leistungsvermögen in aggregierter Form dargestellt werden und Hinweise an Trainer und Athleten geliefert werden. Eine Darstellung des zeitlichen Verlaufs der Faktorenwerte für eine Gruppe von Athleten ist in Abbildung 3.7 dargestellt. Man sieht, dass sich die Faktorenwerte zwischen den Messungen eines Athleten meist nur gering ändern.

3.2.6. Unsicherheitsabschätzungen

Die meisten Untersuchungen, in denen Faktorenmodelle zum Einsatz kommen, beschränken sich auf die Angabe von Punktschätzern bei Ladungskoeffizienten und Eigenwerten. Im Rahmen dieser Analyse soll auch die Unsicherheit, mit der die Schätzer behaftet sind, betrachtet werden. Insbesondere soll der Frage nachgegangen werden, ob die Ergebnisse durch einzelne Beobachtungen oder Athleten verzerrt

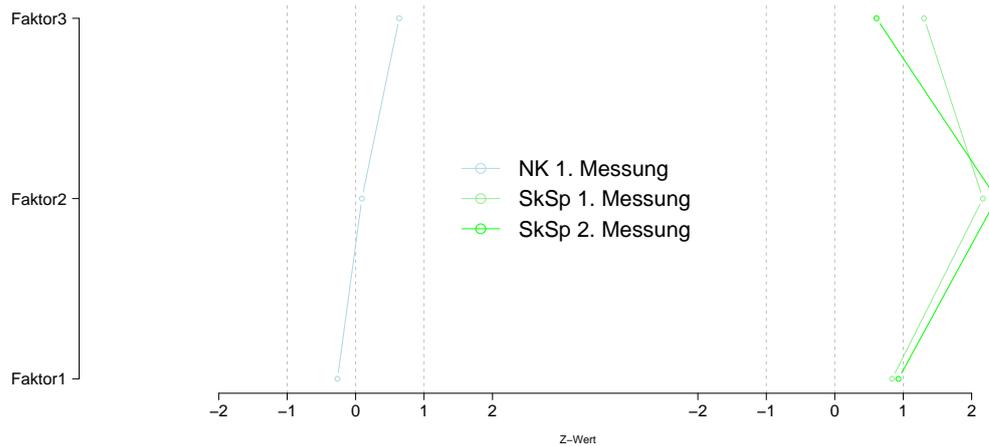


Abbildung 3.6.: Profil der Faktorenwerte von Messungen eines Nordischen Kombinierers (blau) und eines Skispringers (grün). Die heller eingefärbte Linie kennzeichnet dabei die frühere Messung des jeweiligen Athleten.

werden. Die in den vorherigen Abschnitten dargestellten Punktschätzer sind dabei jeweils mit einer Unsicherheit behaftet, die sich aus der Stichprobenzusammensetzung ergibt. Wie in Abschnitt 3.1.10 beschrieben, wird die Verteilung der Punktschätzer mit nichtparametrischem Bootstrap approximiert. Für die Analysen werden dabei $B=2000$ Bootstrap-Stichproben generiert.

Als erstes wird mit Bootstrap die Verteilung der Eigenwerte $\lambda_1, \dots, \lambda_{23}$ geschätzt. Dies ist bedeutend für die Wahl der Anzahl der Faktoren, da sich die Entscheidung an den Eigenwerten orientiert. In Abbildung 3.8 ist die Verteilung der ersten 10 Eigenwerte in Boxplots dargestellt. Dabei sieht man, dass der erste Eigenwert in allen Bootstrap-Stichproben deutlich oberhalb der anderen liegt. Auch bei der Verteilung der Bootstrap-Schätzer lässt sich ein Sprung zwischen drittem und vierten Eigenwert ausmachen. Das bestätigt die Wahl von $q=3$ Faktoren. Die Varianz der Bootstrap-Schätzer ist für alle Eigenwerte gering, was auf einen geringen Einfluss der Stichprobenzusammensetzung schließen lässt.

Anhand der Bootstrap-Schätzer aus Abbildung 3.8 kann auch der jeweilige erklärte Anteil der Varianz durch drei Faktoren angegeben werden. Der Mittelwert und Median liegen dabei bei 56.9%, das Minimum bei 53.3% und das Maximum bei 60.2%. Durch Hinzunahme eines vierten Faktors könnten noch im Mittel 7.4 (6.2 bis 8.7) Prozentpunkte an zusätzlicher Varianz erklärt werden.

In Abbildung 3.9 ist die Verteilung der Ladungskoeffizienten l_{jk} , $j = 1, \dots, 23$ und $k = 1, 2, 3$ dargestellt. Dazu wurden anhand der $B=2000$ Bootstrap-Replikationen 95%-Konfidenzintervalle nach der BC_α -Methode erstellt. Man beachte, dass die resultierenden Konfidenzintervalle nicht symmetrisch sein müssen und auch Werte außerhalb von $[-1, 1]$ möglich sind. Die Größe der Konfidenzintervalle gibt Aufschluss darüber, wie sehr die Ladungskoeffizienten zwischen den Bootstrap-Replikationen schwanken, d.h. wie sensibel die Ladungen bezüglich der Stichprobenzusammensetzung sind.

Man sieht, dass die Zuteilung der Variablen zur Sprungkraftdiagnostik SJ_SpH, DJ32_SpH und DJ32_Faktor zum ersten Faktor bestätigt wird. Die 95%-Konfidenzintervalle der zugehörigen Ladungskoeffizienten liegen überhalb der 0.3 und den entsprechenden Konfidenzintervallen des zweiten Faktors. Allerdings gibt es Überschneidungen mit dem dritten Faktor. In einigen Bootstrap-Replikationen haben diese Variablen auch Nebenladungen auf dem dritten Faktor.

Die Bootstrap-Konfidenzintervalle bestätigen außerdem die Zuordnung der Variablen zur Maximalkraft Fmax90_bb, Fmax70_[bb,re,li] und Ekz zum zweiten Faktor. Die Aufteilung der Variablen zur relativen Maximalkraft ist weniger klar, da sich die jeweiligen Konfidenzintervalle der Faktoren

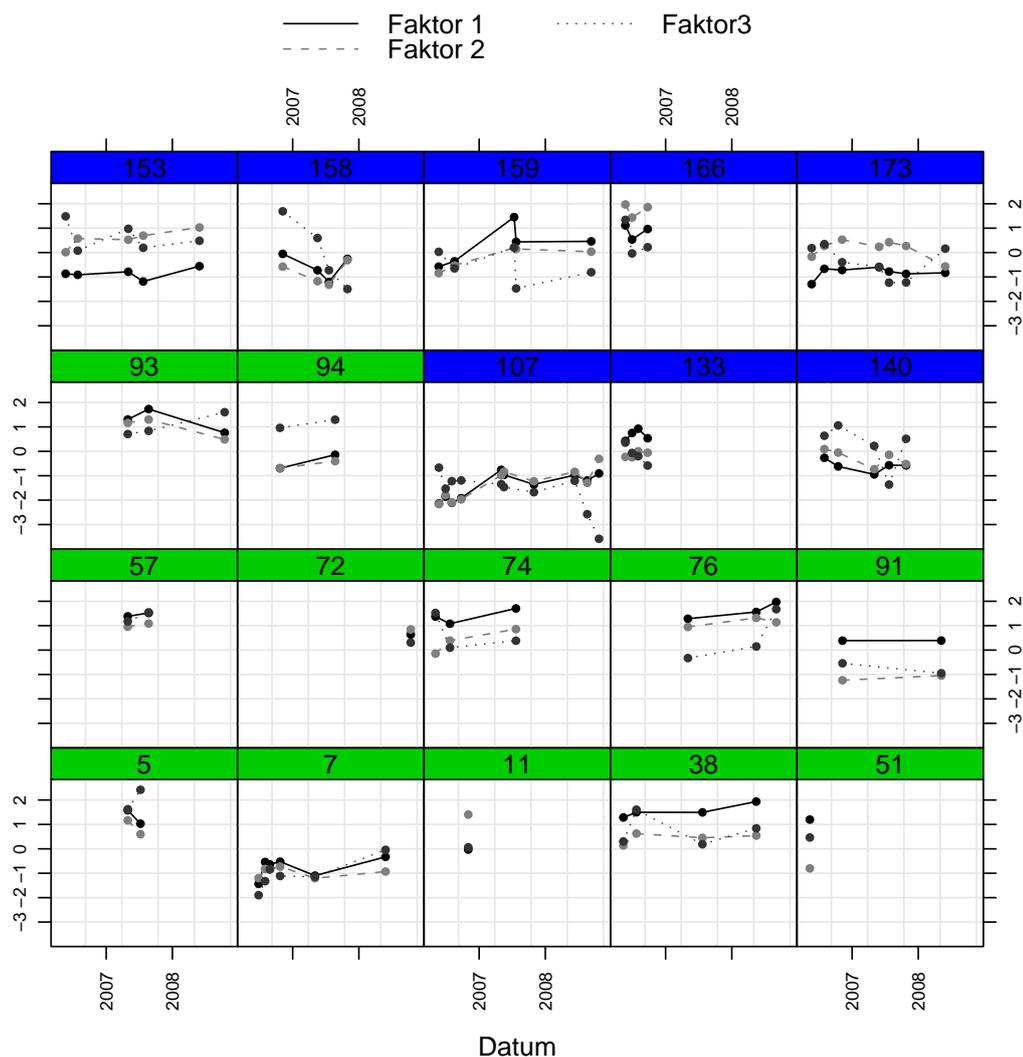


Abbildung 3.7.: Traceplot der geschätzten Faktorenwerte (Promax-Rotation) für 20 zufällig ausgewählte Athleten.

überschneiden und auch Werte kleiner 0.3 enthalten. Gleiches gilt für die Variablen zum Kraftanstieg `relKraftanstieg_[bb,re,li]`. Die Variablen `DJ32_KT`, `Kdif`, `ekV`, `kkV` und `dynstat_[60,180]`, die in der Ladungsmatrix aus Tabelle 3.6 keinem der Faktoren zugeordnet werden konnten, liegen auch in den meisten Bootstrap-Replikation zwischen -0.3 und 0.3 und werden keinem der Faktoren zugeordnet. Unabhängig von der Stichprobenszusammensetzung gilt, dass diese Variablen keinen der drei Faktoren messen.

Die Ergebnisse der Bootstrap-Schätzungen lassen sich wie folgt zusammenfassen. Die Bootstrap-Schätzungen der Eigenwerte in Abbildung 3.8 bestätigen die Notwendigkeit von drei Faktoren. Die Betrachtung der Ladungskoeffizienten bestätigt die Aufteilung der Variablen auf den ersten und zweiten Faktor. Dies bestätigt insbesondere die Aufteilung der Variablen zur Sprungkraft und Maximalkraft in unterschiedliche Faktoren. Der dritte Faktor wird insgesamt ungenauer geschätzt, trägt aber zur Verbesserung der Modellanpassung um einige Prozentpunkte bei.

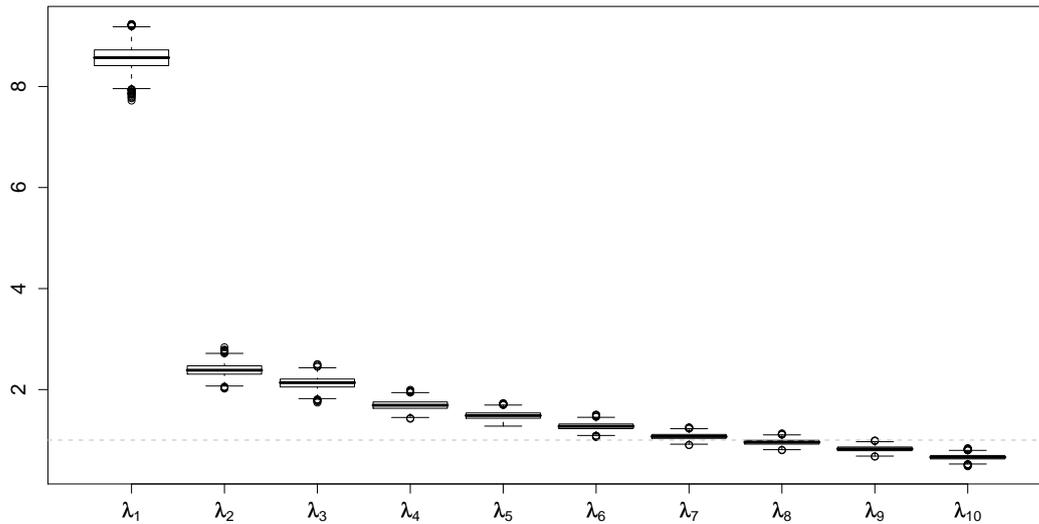


Abbildung 3.8.: Verteilung der Bootstrap-Schätzer für die Eigenwerte $\lambda_1, \dots, \lambda_{10}$ ($B=2000$).

3.2.7. Schätzung von Faktoren zweiter Ordnung

Bislang wurden nur die Beziehungen zwischen Variablen und Faktoren betrachtet. Im Anschluss an eine schiefwinklige Rotation können zusätzlich die Zusammenhänge der Faktoren untereinander betrachtet werden. Dazu werden in einem weiteren Faktorenmodell Faktoren zweiter Ordnung ermittelt. Im Modell (3.3) auf Seite 37 stehen die Faktoren zweiter Ordnung dann auf der rechten Seite und die bereits ermittelten Faktoren erster Ordnung auf der linken. Gesucht ist also das Faktorenmodell für die Faktoren erster Ordnung

$$\mathbf{f}_k = \tilde{l}_{k1}\tilde{\mathbf{f}}_1 + \dots + \tilde{l}_{kr}\tilde{\mathbf{f}}_r, \quad k = 1, 2, 3 \quad (3.17)$$

mit $r < 3$ Faktoren $\tilde{\mathbf{f}}_1, \dots, \tilde{\mathbf{f}}_r$ zweiter Ordnung. Um festzustellen, ob (3.17) sinnvoll ist, wird die empirische Korrelationsmatrix der drei Faktoren erster Ordnung betrachtet. Diese ist in Tabelle 3.7 dargestellt. Man sieht, dass Faktor 1 und 2 positiv korreliert sind und Faktor 3 eine mittlere Korrelation

		Faktor		
		1	2	3
Faktor	1	1	0.62	0.38
	2	0.62	1	0.32
	3	0.38	0.32	1

Tabelle 3.7.: Korrelationsmatrix der drei Faktoren aus Tabelle 3.6 (Promax-Rotation). Faktor 1: „Allgemeine Sprungkraftfähigkeit und relative Maximalkraft“, Faktor 2: „Allgemeine Maximalkraft und Schnellkraft“ und Faktor 3: „Kraftanstiegsverhalten“

mit den beiden anderen Faktoren aufweist. Rechnet man das Faktorenmodell (3.17) mit einem Faktor zweiter Ordnung, so ergibt sich die folgende Ladungsmatrix

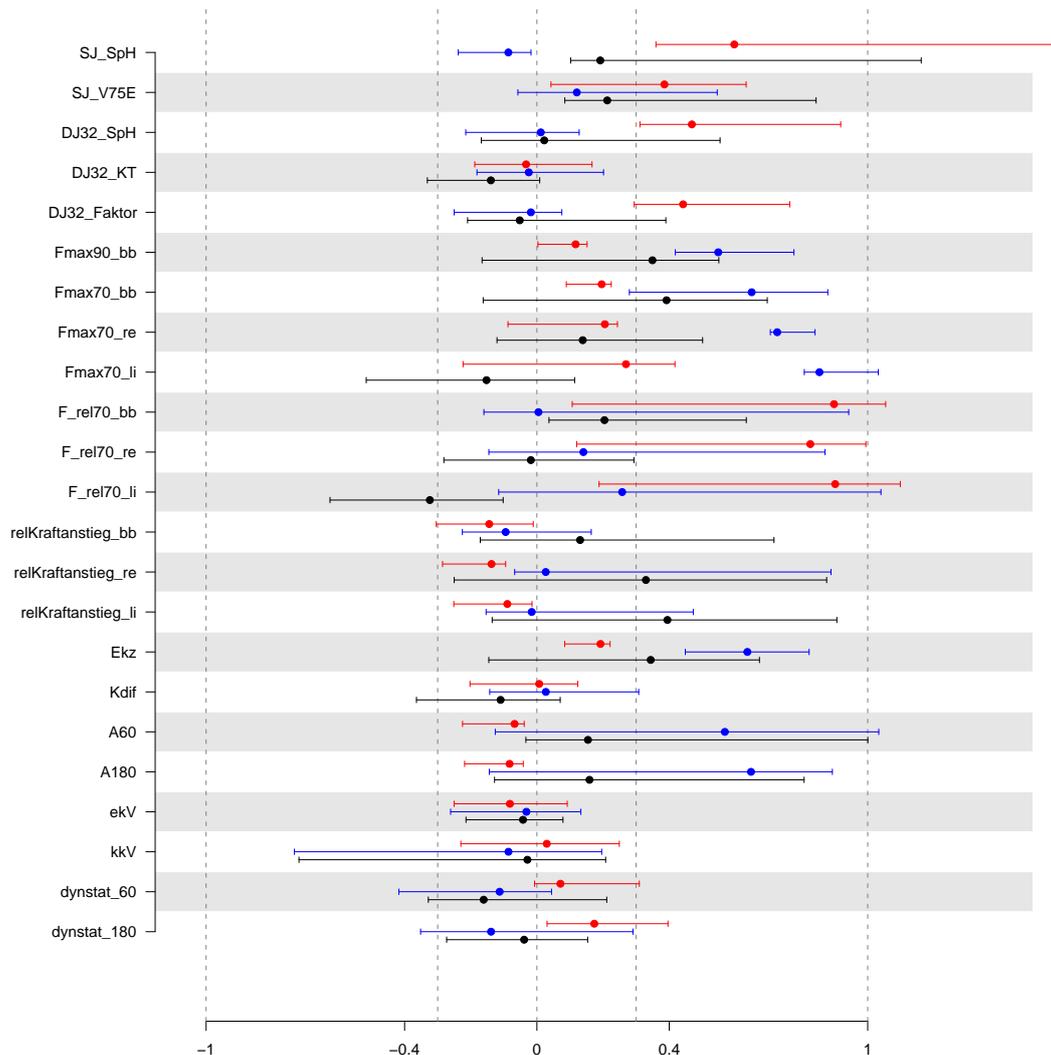


Abbildung 3.9.: Konfidenzintervalle für die Elemente der Ladungsmatrix aus Tabelle 3.6 (Promax-Rotation, siehe Seite 51). Als Punkte sind die Punktschätzer für die Ladungskoeffizienten dargestellt. Diese entsprechen genau den in Tabelle 3.6 dargestellten Ladungskoeffizienten. Das Intervall gibt das 95%-Konfidenzintervall nach der BC_α -Methode ($B=2000$) an. In jeder Variablenzeile gibt das oberste Konfidenzintervall das für Faktor 1 (rot) an, das mittlere Faktor 2 (blau) und das untere Faktor 3 (schwarz).

	Faktor 2. Ordnung	Kommunalität
	\tilde{l}_{k1}	\tilde{l}_{k1}^2
Faktor 1	0.86	0.74
Faktor 2	0.72	0.52
Faktor 3	0.44	0.2

Ein Faktor der zweiten Ordnung korreliert also positiv mit allen drei Faktoren erster Ordnung. Inhaltlich gesprochen, könnte er als das gesamte Leistungsniveau eines Athleten bezeichnet werden, wobei hier insbesondere die ersten beiden Faktoren eingehen. Mehr als ein Faktor zweiter Ordnung lässt sich aus Identifizierbarkeitsgründen nicht bestimmen.

3.3. Diskussion

- Die Faktorenanalyse ist eine geeignete Methode, um die Zusammenhänge der Variablen in der Leistungsdiagnostik darzustellen. Da viele Variablen hoch miteinander korrelieren, können Modelle mit latenten Faktoren konstruiert werden, die die Modellkomplexität reduzieren. Die Ergebnisse bei Faktorenmodellen beruhen auf Festlegungen, die der Anwender zu treffen hat. Aus statistischer Sicht kann nicht beurteilt werden, ob die gewählten Modelle diejenigen sind, die die Daten am besten wiedergeben. So lässt sich keine Modellwahl, wie beispielsweise bei Regressionsmodellen, durchführen. Die präsentierten Modelle haben daher zunächst nur explorativen Charakter. Im Anschluss an diese Arbeit können auf dieser Basis Hypothesen generiert werden, die anhand neuer Daten im Rahmen einer konfirmatorischen Faktorenanalyse überprüft werden. Hier bieten sich beispielsweise sogenannte Strukturgleichungsmodelle an, mit denen die Zusammenhänge zwischen latenten Faktoren auf Signifikanz überprüft werden können (vgl. Jöreskog (1973) oder Kline (2005)).
- Voraussetzung für die Faktorenanalyse sind unabhängige Beobachtungen. Das Faktorenmodell wird dabei anhand der empirischen Korrelationsmatrix der Variablen aufgestellt. Im Rahmen einer Simulationsstudie wurde festgestellt, dass sich durch die teilweise abhängigen Beobachtungen, wie sie bei longitudinalen Daten vorliegen, keine Verzerrungen bei der Schätzung der empirischen Korrelationsmatrix ergeben. Daher scheint es sinnvoll, alle zur Verfügung stehenden Beobachtungen in das Faktorenmodell aufzunehmen, um eine breitere Datengrundlage zu erhalten, anstelle von einer Beobachtung je Athlet. Hier wäre zudem unklar, wie die Auswahl einer Beobachtung erfolgen sollte.
- Es wurden drei Faktoren konstruiert, wobei der erste die allgemeine Sprungkraftfähigkeit und relative Maximalkraft, der zweite die allgemeine Maximalkraft und Schnellkraft und der dritte das Kraftanstiegsverhalten zusammenfasst. Diese Faktoren erlauben eine allgemeine Darstellung der relativen Stärken und Schwächen der Athleten und erklären bereits einen großen Teil der Unterschiede zwischen den einzelnen Beobachtungen und Athleten. Die Modellkomplexität wird dabei von 23 Variablen auf drei Faktoren reduziert. Auf individueller Ebene können damit kompakte Profile erstellt werden, die in allgemeiner Form Aufschluss über die relative Leistung des Athleten geben und für Sportwissenschaftler und Trainer bzw. Athleten Hinweise zur Trainingsteuerung liefern. Die Faktoren sind jeweils so konstruiert, dass größere Werte für eine bessere Leistung des Athleten sprechen. Bei dem ersten Faktor zur allgemeinen Sprungkraftfähigkeit und relativen Maximalkraft haben die Skispringer im Mittel größere und damit bessere Werte als die Nordischen Kombinierer. Bei den Faktoren 2 und 3 gibt es keine Unterschiede zwischen den Sportarten.
- Anhand der Ladungsmatrix wurde festgestellt, dass die Variablen zu jedem der Blöcke Sprungkraft, Maximalkraft und Schnellkraft jeweils demselben Faktor zugeordnet werden und daher in einem engen Zusammenhang stehen. Daher müssten aus statistischer Sicht nicht jeweils alle Übungen zu jedem Block durchgeführt werden, da die verschiedenen Variablen jeweils dasselbe Konstrukt messen. So misst beispielsweise auch die relative Maximalkraft den latenten Faktor der Sprungkraft. Die Variablen DJ32_KT, Kdif, ekV, kkV und dynstat_[60,180] können keinem der drei Faktoren zugewiesen werden. Sie werden also von keinem der drei Faktoren erfasst. Ein großer Teil der unerklärten Varianz im Faktorenmodell lässt sich auf diese Variablen zurückführen. Daher ist es sinnvoll, diese Variablen weiter zu erheben, da sie nicht durch ein anderes Konstrukt erfasst werden. In Kapitel 6 wird ein Vorschlag für eine reduzierte Leistungsdiagnostik vorgestellt.

- Die Rotation der Ladungsmatrix mit der Promax-Rotation bringt eine bessere Anpassung an die Einfachstruktur der Ladungsmatrix und bewirkt damit eine leichtere Interpretation der Faktoren. Die so konstruierten drei Faktoren korrelieren jeweils mäßig bis stark positiv miteinander. Daher kann ein Faktor zweiter Ordnung ermittelt werden, der das allgemeine Leistungsniveau eines Athleten misst. Er setzt sich insbesondere aus den Faktoren 1 und 2 erster Ordnung zusammen.
- Mit Bootstrap-Verfahren kann die Unsicherheit der Ergebnisse quantifiziert werden. Bootstrap-Verfahren werden bei Faktorenmodellen in der Praxis eher selten eingesetzt, ermöglichen es aber, den Einfluss der Stichprobenszusammensetzung zu berücksichtigen und Konfidenzintervalle für die geschätzten Parameter zu konstruieren. Sind die Ergebnisse relativ unabhängig von der Stichprobenszusammensetzung, d.h. sind die Konfidenzintervalle eher klein, so gelten diese Ergebnisse für die Mehrzahl der Athleten. Insbesondere kann damit der Frage nachgegangen werden, wie die Punktschätzer von der Stichprobenszusammensetzung beeinflusst werden. Es wurde festgestellt, dass sich die Konfidenzintervalle für die Ladungskoeffizienten der ersten beiden Faktoren nur bei wenigen Variablen überschneiden. Dies bestätigt die Zusammensetzung dieser beiden Faktoren und bedeutet, dass die Schätzungen von Parametern der Faktorenanalyse nicht durch einzelne Athleten stark beeinflusst werden.
- Für jede Beobachtung wurden im Anschluss an die Faktorenanalyse Schätzer mit der Bartlett-Methode für die Faktorenwerte der Beobachtungen gefunden. Damit ändern sich die Faktorenwerte bei Athleten mit mehreren Beobachtungen über die Zeit. Inhaltlich wäre es auch denkbar gewesen, konstante Faktoren zu konstruieren, wenn man annimmt, dass das relative Leistungsvermögen des Athleten über die Zeit konstant bleibt. Für einige der Athleten aus Abbildung 3.7 auf Seite 55 würde diese Annahme bestätigt.

4. Auswertung der Schanzenparameter

In der Leistungssportforschung ist neben der Analyse von Trainings- und Diagnostikdaten, wie in den Kapiteln 2 und 3 vorgestellt, auch die Analyse von Wettkampfdaten von großem Interesse. Gerade im Skisprung lässt sich die Wettkampfleistung eines Athleten sehr genau an einigen wenigen Parametern, wie der Sprungweite, messen. Zur Messung der Leistung werden Übungssprünge von der Schattenschanze in Oberstdorf (Hillsize 137) ausgewertet. Hierfür lagen zu Beginn der Arbeit noch keine ausgewerteten Ergebnisse in Form eines Datensatzes vor. Im Rahmen dieser Arbeit wurden die maschinell erstellten Sprungprotokolle der Übungssprünge ausgelesen, aufbereitet und ausgewertet. Die Schritte zur Aufbereitung der Daten werden in Abschnitt 4.1 beschrieben. Der gewonnene Datensatz wird deskriptiv in Abschnitt 4.2 untersucht und abschließend die Ergebnisse diskutiert. Aufbauend auf diesem Kapitel werden in Kapitel 5 die Ergebnisse der Leistungsdiagnostik aus dem vorherigen Kapitel und die Ergebnisse der Schanzendiagnostik aus diesem Kapitel in Verbindung gebracht.

4.1. Datenlage und Datenaufbereitung

4.1.1. Erster Schritt: Einlesen der Rohdaten

Seit dem Jahr 2004 werden Übungssprünge von Skispringern und Nordischen Kombinierern mit dem *SIMI Reality Motion System* (SIMI Reality Motion Systems GmbH, Unterschleißheim) aufgezeichnet und archiviert. Für jeden Sprung werden dabei in einer Report-Datei Informationen zu Datum und Uhrzeit, Athlet und Sprunginformationen, wie Sprungweite und Geschwindigkeit gespeichert. Daneben wird der Sprung durch eine Videoaufnahme dokumentiert. Zusätzlich sind auf den 10 letzten Metern der Sprungschanze im Abstand von jeweils 50cm Kraftmessplatten angebracht, die die Kraft messen, die der Athlet während des Absprungvorgangs auf die Schanze überträgt.

Insgesamt liegen aktuell etwa 1500 verschiedene Sprungprotokolle vor. Jedes Sprungprotokoll repräsentiert dabei einen Sprung eines Athleten. Dabei wurden für jeden Sprung die in Tabelle 4.1 dokumentierten Variablen ausgelesen und in einem Datensatz zusammengefasst. Dabei sind 1226 der Sprünge von Athleten, die auch an der Leistungsdiagnostik teilgenommen haben. Im Folgenden beschränken sich die Analysen auf diese Teilmenge, da im Kapitel 5 die Daten der Leistungsdiagnostik mit denen der Schanzendiagnostik in Verbindung gebracht werden sollen. Die Athletennamen werden daher nach demselben Schlüssel wie in der Leistungsdiagnostik anonymisiert, sodass später eine Zusammenführung mit den Diagnostikdaten möglich ist. Insgesamt gibt es Daten von 59 Athleten, die sowohl an der Leistungsdiagnostik als auch an der Schanzendiagnostik teilgenommen haben. Der so gewonnene Datensatz enthält als Beobachtungen die einzelnen Sprünge und die in Tabelle 4.1 beschriebenen Variablen. Der gewonnene Datensatz wurde vorab um fehlende Beobachtungen und unplausible Variablenausprägungen bereinigt.

4.1.2. Zweiter Schritt: Aggregation der Rohdaten

Im zweiten Schritt werden die Rohdaten aggregiert. Ziel ist es, die Ergebnisse von allen Durchgängen eines Tages bzw. eines Termins (teilweise fanden mehrtägige Lehrgänge statt) zu einer Beobachtung zu

Variablenname	Beschreibung	Ausprägungen		Einheit
Athlet	Identifikationsnummer	Spr.	Teilmenge von 1,...,99	
		Komb.	Teilmenge von 100,...,191	
Sportart	Indikatorvariable		1 = Skisprung 2 = Nordische Kombination	
Datum	Datum des Sprungs		27.7.2005 – 22.10.2008	
Durchgang	mehrere Sprünge an einem Tag sind mit einer Durchgangsnummer gekennzeichnet		1,...,9	
Weite	Sprungweite	Spr.	87.2 – 142	m
		Kom.	100 – 138	
Geschwindigkeit	absolute Geschwindigkeit des Athleten bei Absprung (offizielle Geschwindigkeit)	Spr.	90 – 118	km/h
		Kom.	92.9 – 98	
Genauigkeit	Absprunggenauigkeit, > 0 zu spät abgesprungen, < 0 zu früh abgesprungen	Spr.	-80.6 – 102.5	cm
		Kom.	-66.1 – 75.8	
vertikaleGeschw	vertikale Absprunggeschwindigkeit (senkrecht zum Schanzentisch)	Spr.	1.54 – 6.59	m/s
		Kom.	2.01 – 3.01	

Tabelle 4.1.: Ausgelesene Variablen aus den Reportdateien zu den Übungssprüngen (Rohdaten)

aggregieren. Diese Aggregation ist einerseits notwendig, da die Athleten nicht alle Sprünge an einem Tag voll ausspringen und die einzelnen Sprünge eines Termins somit nicht direkt zu vergleichen sind und andererseits, da bei der Zusammenführung von Diagnostik- und Schanzendaten, für die Modelle eine Beobachtung je Termin benötigt wird. Insgesamt liegen an zehn Terminen Schanzmessungen in den Rohdaten vor.

	Termin	Beschreibung	Sportart	# Sprünge	# Athleten
1	27.7.-31.7.2005	TL Skisprung	Skispringer	178	17
2	10.9.-11.9.2005	Training NK	Kombinierer	70	6
3	10.10.-12.10.2005	Training Skisprung	Skispringer	142	7
4	23.6.-26.6.2006	Training Skisprung	Skispringer	80	4
5	26.7.-30.7.2006	Trainingswettkampf	Skispringer	124	18
6	24.8.2006	TL Nordische Kombination	Kombinierer	8	4
7	4.9.-8.9.2006	Trainingslehrgang	beide	244	11
8	11.10.-16.10.2006	Trainingslehrgang	Skispringer	83	18
9	17.7.-20.7.2007	Basislehrgang	Skispringer	237	30
10	21.10.-22.10.2008	Trainingslehrgang	Skispringer	60	8
	Σ			1226	123

Tabelle 4.2.: Termine der Schanzendiagnostik von Skispringern und Nordischen Kombinierern. Die Anzahl der Sprünge und Athleten bezieht sich nur auf Athleten, die auch an der Leistungsdiagnostik teilgenommen haben und auf Sprünge bei denen alle Variablen aus Tabelle 4.1 gemessen wurden. Einzelne Springer haben dabei an mehreren Terminen teilgenommen.

Für jeden der Termine soll für jeden der teilnehmenden Athleten *eine* interessierende Kenngröße für jede der vier Variablen **Weite**, **Geschwindigkeit**, **Genauigkeit** und **vertikaleGeschw** bestimmt werden. Die Variable von größtem inhaltlichen Interesse ist dabei die **Weite**. Zur Aggregation der Variablen eines Athleten an dem oder den Tagen der Termine bieten sich dabei die folgenden Vorgehensweisen an

1. Es wird für die vier Variablen der Mittelwert aller Durchgänge eines Termins gewählt. Für dieses Vorgehen spricht, dass Ausreißer in den Ergebnissen weniger ins Gewicht fallen. Mit diesem Vorgehen wird das mittlere Leistungsniveau des Athleten erfasst.
2. Es wird für jede der vier Variablen separat das Maximum aus allen Durchgängen gewählt, also die maximale Weite, die maximale Geschwindigkeit, die maximale Genauigkeit und die maximale vertikale Absprunggeschwindigkeit. Damit wird das maximale Leistungsniveau eines Athleten zum entsprechenden Termin erfasst. Allerdings können die einzelnen Beobachtungen dann aus unterschiedlichen Durchgängen stammen.
3. Es wird der Sprung mit der maximalen Sprungweite aller Durchgänge ausgewählt zusammen mit der zugehörigen Geschwindigkeit, Genauigkeit und vertikalen Absprunggeschwindigkeit diese Sprungs. Mit diesem Vorgehen wird die maximale Sprungleistung des Athleten betrachtet und zu diagnostischen Zwecken die anderen Variablen genau dieses Sprungs.

Im Folgenden wird die dritte Vorgehensweise gewählt. Dafür spricht auch, dass die Athleten häufig die ersten Sprünge nur zum Einspringen nutzen und lediglich wenige voll ausspringen. Außerdem entspricht die Vorgehensweise auch der Kraftdiagnostik, bei der das beste Ergebnis aus mehreren Durchgängen gewählt wird. Bei der dritten Vorgehensweise können auch Zusammenhänge der Variablen `Weite`, `Geschwindigkeit`, `Genauigkeit` und `vertikaleGeschw` betrachtet werden, da alle Werte jeder Beobachtung aus demselben Sprung stammen. Der aggregierte Datensatz enthält dann die Variablen aus Abschnitt 4.1, wobei je Athlet und je Termin nur ein Sprung mit vier Variablen verbleibt. Insgesamt enthält der aggregierte Datensatz 123 Beobachtungen (Sprünge) von 47 Skispringern und 12 Nordischen Kombiniern an zehn verschiedenen Terminen.

4.2. Explorative Analysen

Im Folgenden werden die aggregierten Daten explorativ analysiert. Da alle Sprünge von der selben Schanze ausgeführt wurden, können sie miteinander verglichen werden. Allerdings können die Athleten zwischen verschiedenen Luken wählen, von denen sie an der Schanze starten, wodurch sich unterschiedliche Anlaufstrecken bis zum Schanzentisch ergeben. In den Report-Dateien ist diese Information allerdings nicht notiert. Sie drückt sich jedoch indirekt in der `Geschwindigkeit` aus, die umso höher ist, je länger die Anfahrtsstrecke ist. Daher kann eine normierte Weite `Weite_norm` definiert werden, die sich wie folgt ergibt

$$\text{Weite_norm} = \frac{\text{Weite}[m]}{\text{Geschwindigkeit}[km/h]} \quad [mh/km].$$

In Abbildung 4.1 sind die univariaten und bivariaten Verteilungen der Variablen `Weite`, `Geschwindigkeit`, `Genauigkeit`, `vertikaleGeschw` und `Weite_norm` dargestellt. Die `Weite`, `Genauigkeit` und `Weite_norm` sind leicht rechtssteil verteilt, die `Geschwindigkeit` symmetrisch in einem Bereich von 93 bis 98 km/h. Auch die `vertikaleGeschw` ist symmetrisch verteilt. Im Folgenden wird für alle vier Variablen eine approximative Normalverteilung angenommen. Die mittlere maximale Sprungweite eines Athleten an einem Termin beträgt 125.2 m. Dabei sprangen die Athleten bei 17 % der Sprünge zu früh und bei 81 % zu spät ab.

Betrachtet man die bivariaten Verteilungen, so sind bei den meisten Kombinationen nur schwache Korrelationen der fünf Variablen zu beobachten. Auch die Korrelationskoeffizienten geben an, dass die Variablen nahezu unkorreliert sind. Ein sehr starker Zusammenhang besteht zwischen `Weite` und `Weite_norm`. Dies wirft die Frage auf, ob die Normierung der Sprungweite überhaupt notwendig ist.

Das Streudiagramm zeigt einen leicht positiven Zusammenhang zwischen *Weite* bzw. *Weite_norm* und *vertikaleGeschw*, wobei der Korrelationskoeffizient aufgrund zweier Ausreißer nur gering ist ($\rho=0.16$ bzw. 0.21).

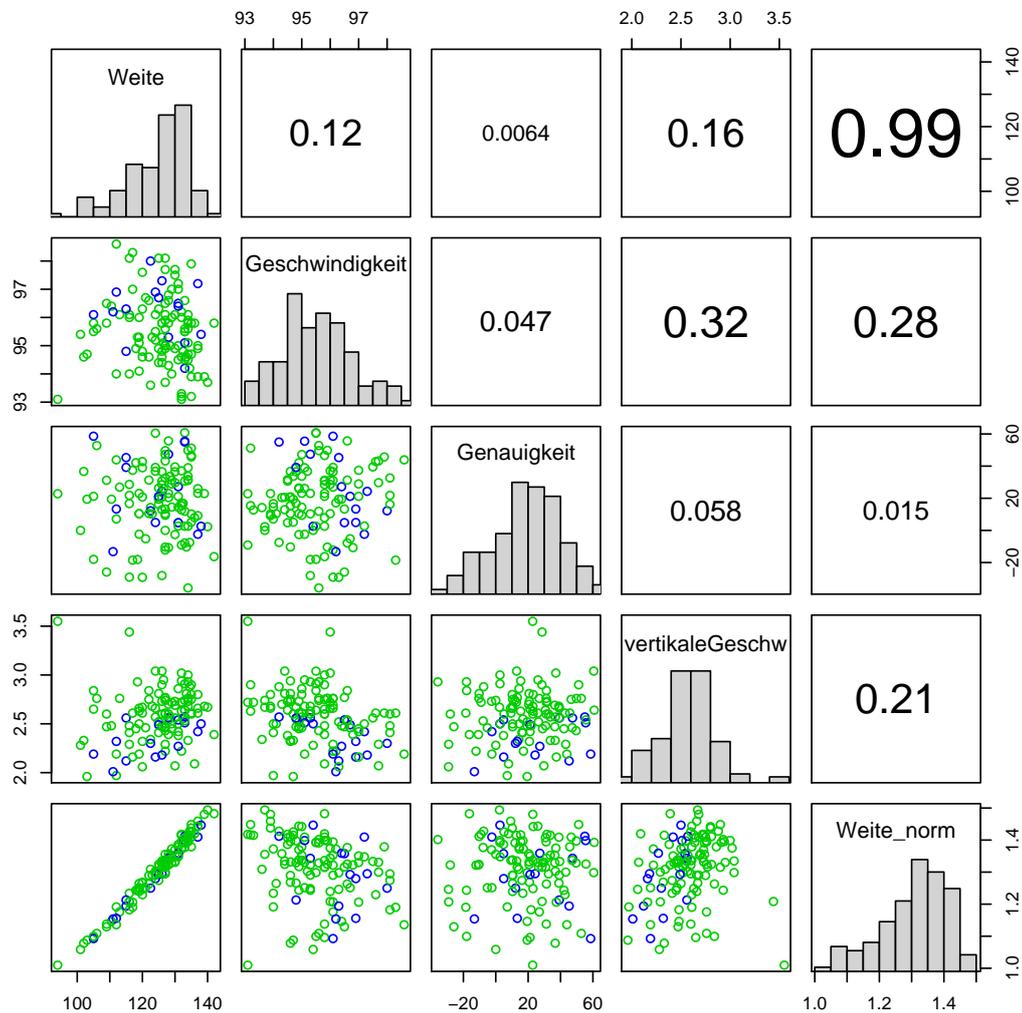


Abbildung 4.1.: Darstellung der univariaten Verteilungen der Variablen *Weite*, *Geschwindigkeit*, *Genauigkeit*, *vertikaleGeschw* und *Weite_norm* der aggregierten Schanzendaten als Histogramme auf der Diagonalen, die paarweisen Korrelationskoeffizienten in der oberen Hälfte ($n=123$) und die bivariaten Streudiagramme in der unteren Hälfte. Dabei kennzeichnen die blauen Kreise Beobachtungen bei Nordischen Kombinierern, die grünen Beobachtungen bei Skispringern.

In Abbildung 4.2 sind die maximalen Sprungweiten an einem Termin von 32 zufällig ausgewählten Athleten dargestellt. Zunächst sieht man, dass insgesamt nur wenige Werte vorliegen. Für viele Athleten gibt es nur eine (25 Athleten) oder zwei Beobachtungen (20 Athleten). Das Maximum liegt bei 7 Beobachtungen je Athlet. Betrachtet man die individuellen Verläufe, so gibt es sowohl Athleten, die sich kontinuierlich verbessern (z.B. Athlet 39), stagnieren (z.B. Athleten 63 und 138) als auch verschlechtern (z.B. Athleten 38 und 93).

Analog zu den Untersuchungen zur Leistungsdiagnostik in Kapitel 2 werden auch die Ergebnisse der Schanzendiagnostik zwischen Skispringern und Nordischen Kombinierern verglichen. In Abbildung 4.3 sind die Werte der Athleten getrennt nach Sportart als Boxplots dargestellt. Bei der *Weite* liegen die

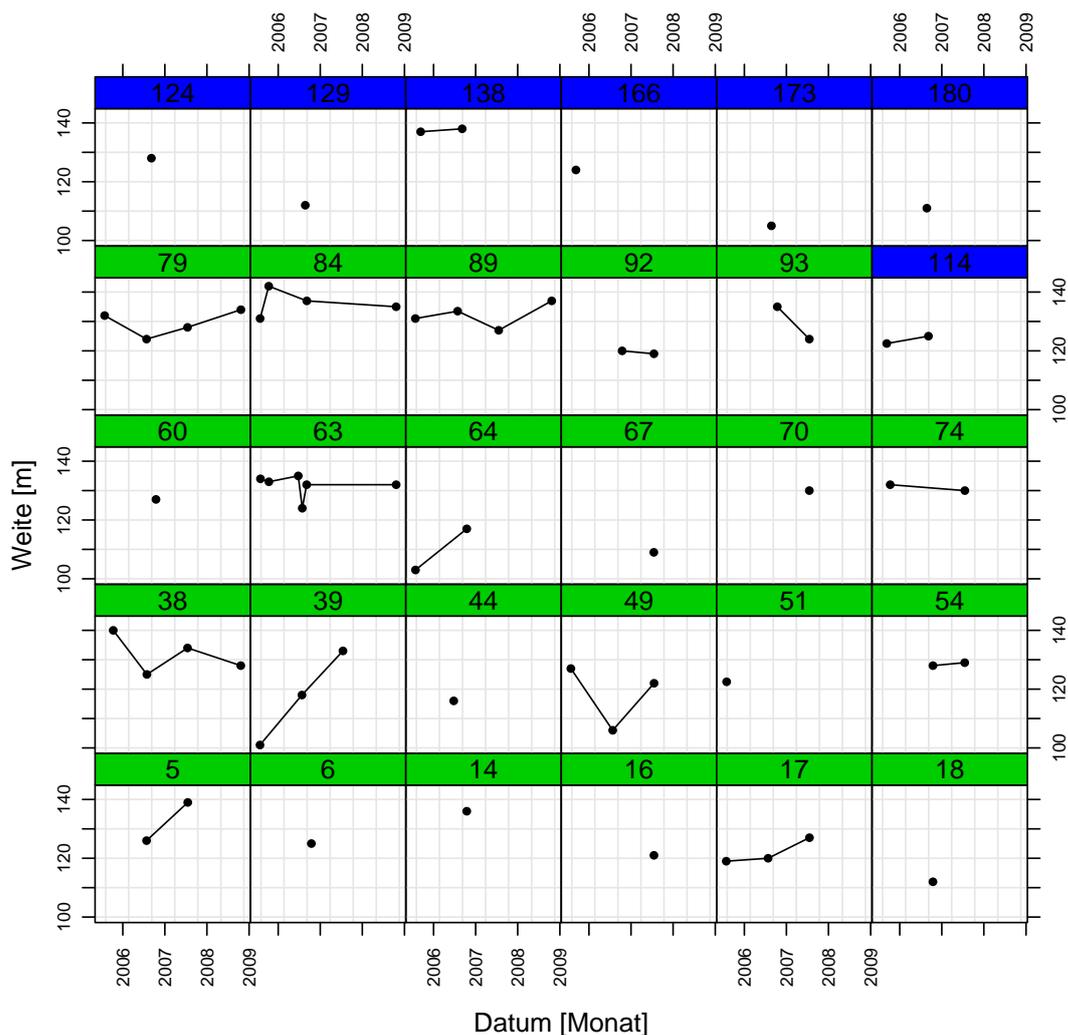


Abbildung 4.2.: Traceplot von 32 zufällig ausgewählten Skispringern und Nordischen Kombinierern für die Variable *Weite* aus den aggregierten Schanzen Daten.

Skispringer knapp vor den Nordischen Kombinierern. Die mediane *Geschwindigkeit* der Kombinierer liegt allerdings ca. 1 km/h über dem Median der Skispringer. Dies spricht dafür, dass die Nordischen Kombinierer eine längere Anlaufstrecke gewählt haben. Damit ist der Vorsprung der Skispringer bei der normierten Sprungweite etwas deutlicher als bei der absoluten Sprungweite. Trotz einer kürzeren Anlaufstrecke springen die Skispringer also weiter als die Nordischen Kombinierer. Die Skispringer haben auch deutlich höhere vertikale Absprunggeschwindigkeiten. Bei der *Genauigkeit* haben Athleten beider Sportarten im Median etwa gleiche Werte, allerdings ist die Streuung bei den Nordischen Kombinierern etwas größer.

4.3. Diskussion

- In diesem Kapitel wird ein Datensatz zur Schanzendagnostik aus den Report-Dateien von Übungssprüngen von der Hillsize 137-Schanze in Oberstdorf erstellt. Dazu wurde aus mehreren Sprüngen an einem Tag nur derjenige mit der maximalen Sprungweite ausgewählt. Des Weiteren wurde die vertikale Absprunggeschwindigkeit, die Anlaufgeschwindigkeit und die Ge-

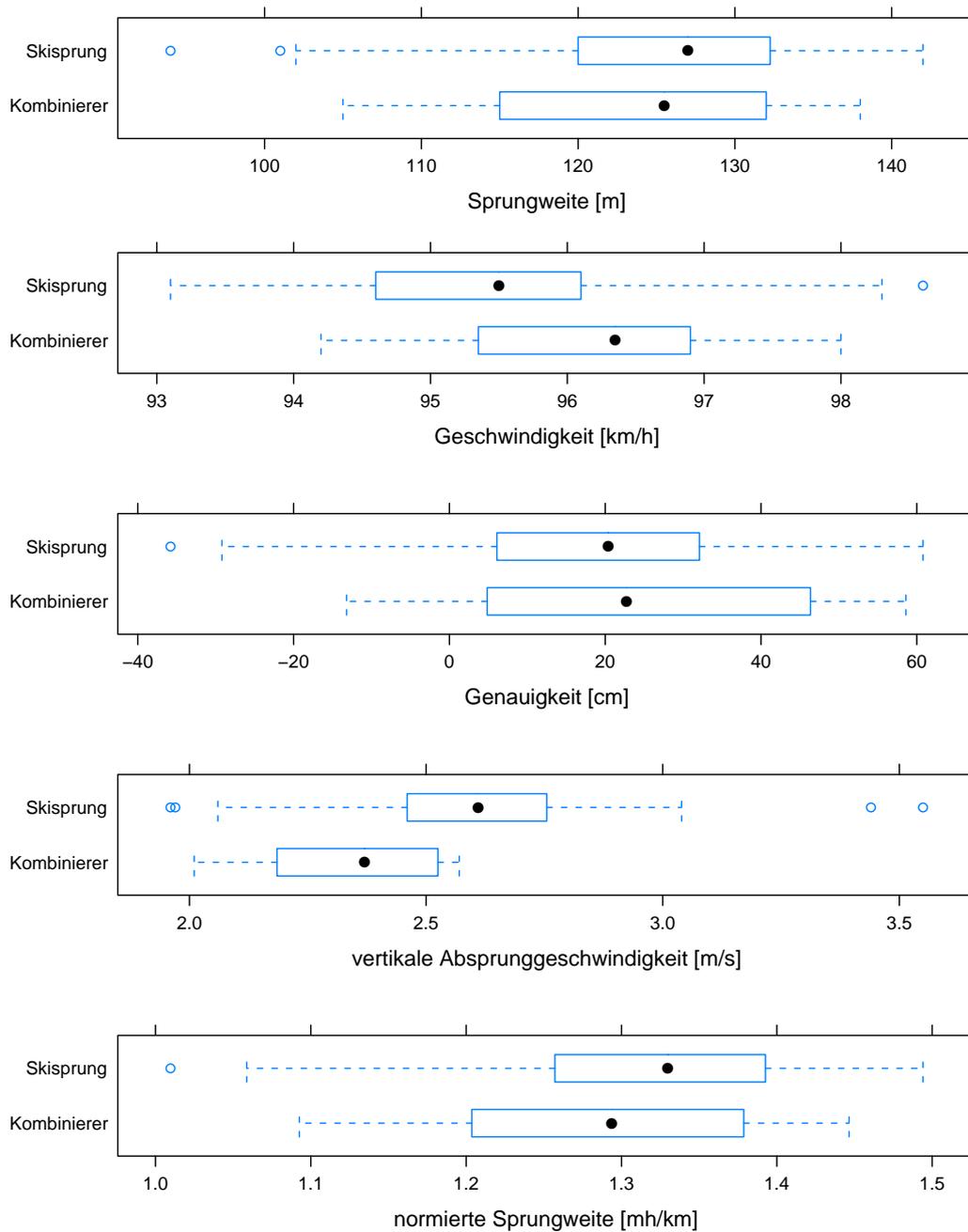


Abbildung 4.3.: Vergleich der aggregierten Schanzendaten bezüglich der Sportart. Von oben nach unten sind dabei für die Variablen *Weite*, *Geschwindigkeit*, *Genauigkeit*, *vertikaleGeschw* und *Weite_norm* die Verteilung für Skispringer ($n=107$) und Nordische Kombiniierer ($n=16$) als Boxplot angetragen.

nauigkeit erfasst. Diese Variablen stehen alle in einem positiven, aber vergleichsweise geringen Zusammenhang. Im Rahmen der Schanzendiagnostik wurden für Spezial-Skispringer bessere Ergebnisse bei allen Variablen als für Nordische Kombiniierer festgestellt.

- Im Rahmen der Schanzendiagnostik konnten bei der Modellierung keine Kovariablen wie Anlaufstrecke oder Wind berücksichtigt werden. Für zukünftige genauere Analysen bietet sich eine Erfassung dieser Merkmale an.

- Bei der Aggregation der Schanzendaten hätte anstelle der maximalen Sprungweite aller Durchgänge eines Tages auch die mittlere Weite gewählt werden können. Dies erscheint jedoch nicht sinnvoll, wenn die Athleten nicht alle Sprünge voll ausspringen. Bei der Wahl der maximalen Sprungweite müssten jedoch auch die Athleten wissen, dass sie einen Sprung an jedem Termin voll ausspringen sollen.
- Da die Sprünge teilweise mit unterschiedlichen Anlaufstrecken ausgeführt werden, bietet es sich an, auch die normierte Weite zu betrachten, die die Sprungweite auf die unterschiedlichen Anlaufstrecken normiert. Allerdings zeigt sich in der hohen Korrelation der Variablen `Weite` und `Weite_norm`, dass der Einfluss der Geschwindigkeit nur gering ist. Daher wird die normierte Sprungweite im Folgenden nicht weiter betrachtet.
- Aus sportwissenschaftlicher Sicht sind insbesondere die Variablen `Weite` und `vertikaleGeschw` von Interesse, wenn die Leistung des Athleten im Rahmen der Schanzendiagnostik beurteilt werden soll. Im nächsten Kapitel wird der Zusammenhang dieser Variablen mit denen der Laborleistungsdiagnostik analysiert.

5. Gemeinsame Auswertung der Leistungsdiagnostik und Schanzendiagnostik

In diesem Kapitel wird die Verbindung zwischen den Daten zur Leistungsdiagnostik und denen zur Schanzendiagnostik hergestellt. Dabei sollen die Variablen der Leistungsdiagnostik identifiziert werden, die in einem hohen Zusammenhang mit wichtigen Variablen zur Schanzendiagnostik stehen. Mit der Leistungsdiagnostik wird die Sprung-, Maximal- und Schnellkraft erfasst und die Schanzendiagnostik misst die Leistung des Athleten unter Wettkampfbedingungen. Aus sportwissenschaftlicher Sicht ist dieser Transfer interessant, da die Leistungsdiagnostik Variablen misst, die sich durch Training beeinflussen lassen und die Schanzendiagnostik Aufschluss über das Wettkampfpotenzial eines Athleten gibt. Eine gemeinsame Betrachtung beider Diagnostiken gewährt im besten Falle Hinweise auf eine Trainingssteuerung zur positiven Beeinflussung der Wettkampfergebnisse. Damit dieser Schluss möglich ist, müssen Bedingungen erfüllt sein, die im folgenden Abschnitt beschrieben werden. In Abschnitt 5.3 werden Modellansätze zur Zusammenführung der beiden Datenquellen beschrieben und die Ergebnisse gezeigt. In Abschnitt 5.4 werden die Ergebnisse des Kapitels diskutiert.

5.1. Datenvorbereitung

Für gemeinsame Modelle müssen für einen Athleten je eine Messung zur Schanzendiagnostik und eine zur Leistungsdiagnostik zugeordnet werden. In der Regel finden Leistungsdiagnostik und Schanzendiagnostik an unterschiedlichen Tagen statt. Damit die Ergebnisse verglichen werden können, sollte der Zeitabstand klein sein, wobei hier festgelegt wird, dass nur Diagnostiken, die weniger als zwei Monate auseinander liegen, miteinander verglichen werden. Die zeitliche Verteilung der vorhandenen, vollständigen Beobachtungen zu Leistungs- und Schanzendiagnostik pro Athlet ist in Abbildung 5.1 dargestellt.

Der neue, gemeinsame Datensatz wird wie folgt generiert:

1. Für jede vorliegende Beobachtung aus der Schanzendiagnostik wird gesucht, ob derselbe Athlet im Zeitraum von 60 Tagen davor oder danach eine Messung zur Leistungsdiagnostik hatte.
2. Ist dies der Fall, so werden die Variablen zur Schanzendiagnostik aus Tabelle 4.1 auf Seite 61 und die Variablen zur Leistungsdiagnostik aus den Tabelle 2.2 bis 2.5 auf den Seiten 13 bis 15 zu einer neuen Beobachtung zusammengefasst. Diese Beobachtung enthält zwei Datumsangaben für die beiden Diagnostiktermine. Zudem enthält diese Beobachtung Alter, Größe oder Gewicht des Athleten aus den Ergebnissen der Leistungsdiagnostik, was bislang in der Schanzendiagnostik nicht bekannt war.
3. Sollten innerhalb eines Zeitraums von \pm zwei Monaten um die Schanzendiagnostik für einen Athleten mehr als eine Leistungsdiagnostik liegen, so werden mehrere neue Beobachtungen gebildet.

Der neue gemeinsame Datensatz enthält 107 Beobachtungen von 46 Athleten. Für jede Beobachtung liegen Ergebnisse der Schanzendiagnostik und Leistungsdiagnostik vor, die zeitlich so nah liegen, dass inhaltliche Zusammenhänge plausibel sind. Bei einer Complete-Case-Analyse verbleiben 75 Beobach-

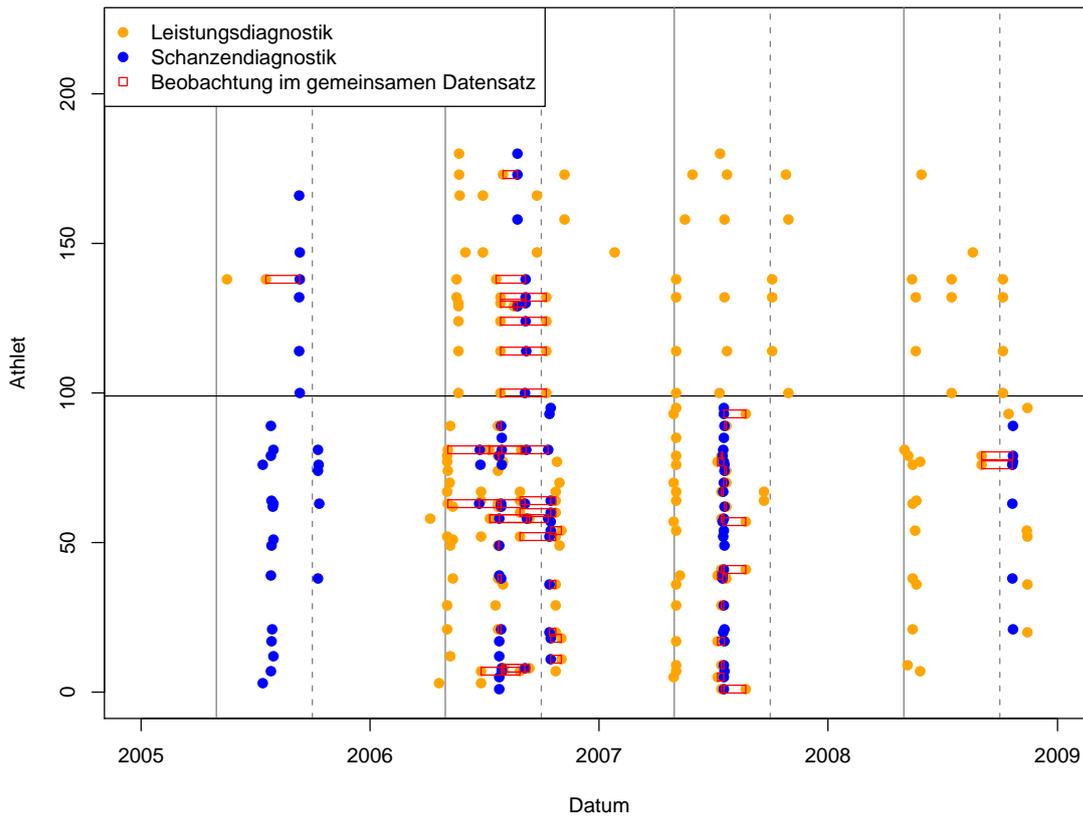


Abbildung 5.1.: In jeder Zeile ist der Termin der vollständigen Beobachtungen (Complete-Cases) von Leistungsdiagnostik und Schanzendiagnostik für jeden Athleten, repräsentiert durch seine Identifikationsnummer, dargestellt. Die horizontale Linie trennt Skispringer unten und Nordische Kombinierer oben. Die durchgezogene vertikale Linie kennzeichnet den 1. Mai jeden Jahres, die gestrichelte den 1. Oktober. Die Messungen, die weniger als 2 Monate voneinander entfernt sind, sind durch ein Rechteck verbunden und bilden im Folgenden eine Beobachtung im gemeinsamen Datensatz (2005: $n=1$, 2006: $n=52$, 2007: $n=20$, 2008: $n=2$).

tungen von 33 Skispringern und 8 Nordischen Kombinierern. Für die meisten Athleten sind ein oder zwei Beobachtungen in dem gemeinsamen Datensatz enthalten. Für je einen Athleten bleiben 6 bzw. 7 Beobachtungen in dem gemeinsamen Datensatz.

5.2. Deskription

In Abbildung 5.2 sind die univariaten und bivariaten Verteilungen der interessierenden Variablen aus der Schanzendiagnostik und der geschätzten Faktorenwerte stellvertretend für die Variablen der Leistungsdiagnostik dargestellt. Dabei korreliert die Variable `Weite` im gemeinsamen Datensatz vergleichsweise schwach mit den Faktorenwerten. Bei der Variablen `vertikaleGeschw` kann jeweils eine mittlere, positive Korrelation mit den drei Faktoren verzeichnet werden. Jede der beiden Variablen der Schanzendiagnostik korreliert mit Faktor 3 (Kraftanstiegsverhalten) stärker als mit Faktor 1 (Allgemeine Sprungkraftfähigkeit und relative Maximalkraft) bzw. Faktor 2 (Allgemeine Maximalkraft und Schnellkraft). Die beiden Variablen der Schanzendiagnostik haben einen Korrelationskoeffizienten von $\rho=0.237$.

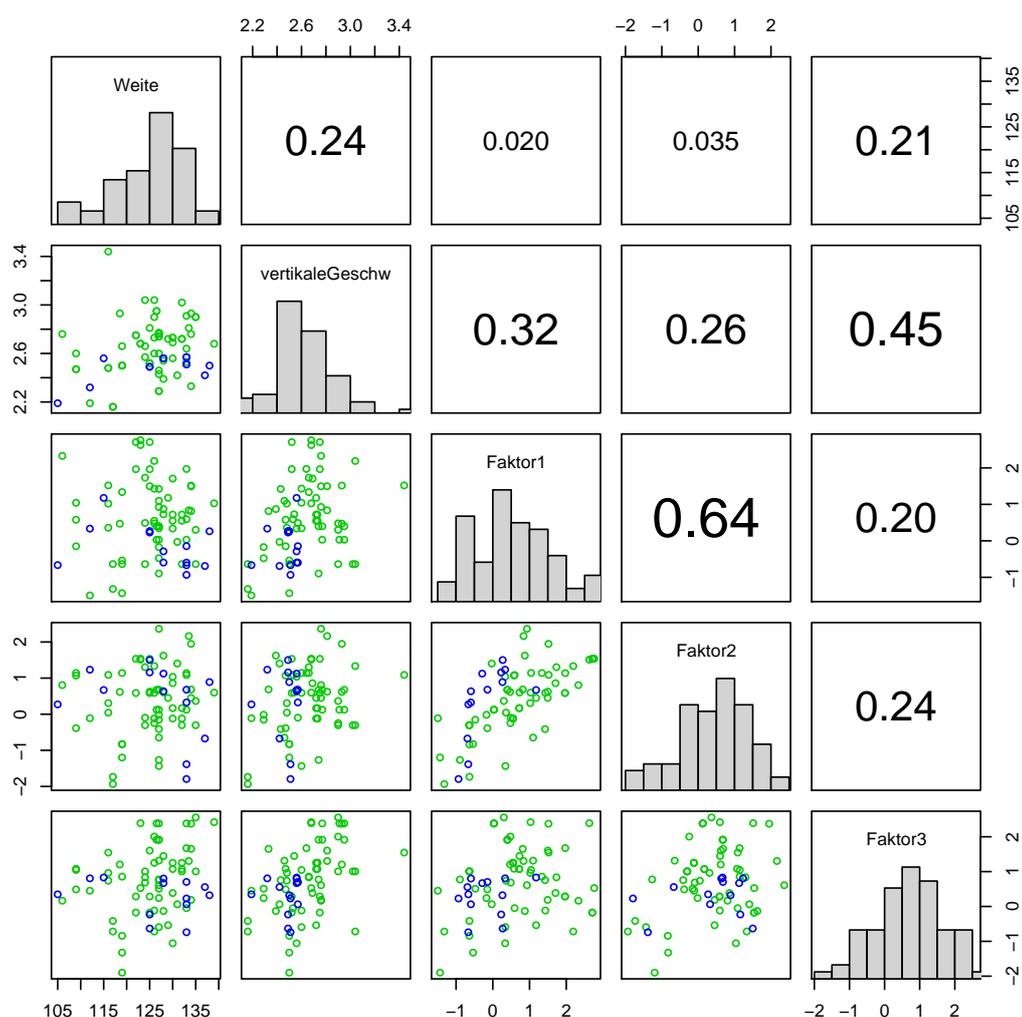


Abbildung 5.2.: Darstellung der univariaten Verteilungen der Variablen `Weite` und `vertikaleGeschw` aus der Schanzendiagnostik und der geschätzten Faktorenwerte aus der Leistungsdiagnostik als Histogramme auf der Diagonalen, die paarweisen Korrelationskoeffizienten sind in der oberen Hälfte und die bivariaten Streudiagramme in der unteren Hälfte dargestellt. Es werden für die Darstellung nur die vollständigen Beobachtungen verwendet (Complete-Case-Analyse mit $n=75$). Dabei kennzeichnen die blauen Kreise Beobachtungen bei Nordischen Kombinierringern und die grünen Beobachtungen bei Ski-springern.

5.3. Analysen und Ergebnisse

Anhand des gemeinsamen Datensatzes soll der Zusammenhang zwischen Leistungsdiagnostik und Schanzendiagnostik hergestellt werden. Basierend auf den Ergebnissen der Kapitel 3 und 4 werden hier drei Modelle präsentiert, die den Zusammenhang zwischen Leistungsdiagnostik und Schanzendiagnostik modellieren. Die Zielgrößen sind die Variablen `Weite` und `vertikaleGeschw` aus der Schanzendiagnostik. Das erste Verfahren ist eine Kombination von Faktorenanalyse und Regressionsanalyse (FARA) zur Berechnung von linearen Strukturmodellen (Wagner, 1996). Dieses Verfahren wird bereits bei der Analyse von Daten aus der Sportwissenschaft eingesetzt. Hier werden zusätzlich zwei Modellklassen präsentiert, die bislang noch nicht in der Leistungssportforschung eingesetzt wurden.

Das erste ist ein *Latente-Variablen-Modell* (Everitt, 1984), das die Idee der latenten Faktoren aus Kapitel 3 aufgreift und um direkte und indirekte Kovariablennachte erweitert. Das zweite Modell gehört zur allgemeinen Klasse der *strukturiert additiven Regressionsmodelle* (Brezger u. Lang, 2006) und setzt Bayesianische Regularisierung zur Modellwahl ein.

5.3.1. Faktoren- und Regressionsanalyse

Die FARA-Methode kombiniert Faktorenanalyse und Regressionsanalyse. Der zentrale Unterschied zwischen einer Faktorenanalyse und einer Regressionsanalyse ist, dass in der Regression nach gerichteten Zusammenhängen gesucht wird. Das bedeutet, dass eine *Zielgröße* definiert ist und der Einfluss von *Kovariablen* auf diese Zielgröße untersucht wird. Nach Wagner (1996) umfasst die FARA-Methode die folgenden Schritte:

1. Aufstellen eines Faktorenmodells
2. Bestimmen der Ladungsmatrix
3. Schätzen der Faktorenwerte nach der Bartlett-Methode
4. Rechnen einer multiplen Regression mit den geschätzten Faktorenwerte als Kovariablen und einer Zielgröße

Die Zielgröße ist hier einmal die Variable `Weite` und einmal die `vertikaleGeschw`. Diese werden nicht standardisiert, sondern das Modell wird um die Kovariablen `Alter` und `Sportart` erweitert. Als Faktorenmodell wird das aus Abschnitt 3.2 mit drei latenten Faktoren gewählt. Das Modell der FARA für eine Beobachtung sieht wie folgt aus:

$$y_i = \beta_0 + \beta_1 f_{i1} + \beta_2 f_{i2} + \beta_3 f_{i3} + \beta_4 \text{Alter}_i + \beta_5 \text{Sportart}_i + \epsilon_i, \quad (5.1)$$

wobei y_i die Zielgröße ist und f_{iq} , $q = 1, 2, 3$ die Einträge in der Faktorenwertmatrix von Beobachtung i sind und ϵ_i der Fehlerterm mit $\epsilon_i \sim N(0, \sigma^2)$ ist. Die Parameterschätzer der zwei gerechneten Modelle sind in Tabelle 5.1 zusammengefasst. Die Variable `Alter` hat einen signifikanten positiven

Zielgröße	Weite			vertikaleGeschw		
	$\hat{\beta}$	SD	p-Wert	$\hat{\beta}$	SD	p-Wert
Faktor 1	-1.17	1.2	0.335	0.01	0.03	0.859
Faktor 2	0.9	1.27	0.48	0.04	0.03	0.171
Faktor 3	0.88	0.96	0.364	0.06	0.02	0.018
<code>Alter</code>	0.66	0.21	0.002	0.02	0.01	0
<code>Sportart</code>	1.96	2.77	0.481	0.23	0.07	0.002
R^2	0.18			0.42		

Tabelle 5.1.: Ergebnisse der FARA-Modelle: Angegeben sind jeweils Parameterschätzer, Standardabweichung und p-Wert, sowie das Bestimmtheitsmaß R^2 der beiden Modelle (jeweils $n = 75$).

Effekt auf die `Weite` und die `vertikaleGeschw`. Skispringer erzielen bei allen drei Variablen bessere Ergebnisse als Nordische Kombinierer, allerdings ist der Effekt nur bei `vertikaleGeschw` signifikant. Hier zeigt sich auch ein signifikanter positiver Effekt von Faktor 3. In den anderen Modellen haben die Faktoren keine signifikanten Effekte. Man sieht aus Tabelle 5.1 anhand der Bestimmtheitsmaße, dass für jede der Zielgrößen `Weite` und `vertikaleGeschw` die Modellanpassung nicht gut ist, wobei die Anpassung im Modell mit `vertikaleGeschw` noch besser ist. Es gilt also, dass die Faktoren gut die Zusammenhänge der Variablen zur Leistungsdiagnostik erklären, allerdings nur einen geringen

Erklärungswert für die Variablen zur Schanzendiagnostik haben. Dies liegt möglicherweise daran, dass man an Schätzgenauigkeit verliert, da die Faktorenwerte bereits geschätzt sind und ihr Einfluss im FARA-Modell erneut geschätzt wird. Daher verwenden die weiteren Modelle wieder die ursprünglichen Variablen der Leistungsdiagnostik als erklärende Kovariablen, um Zusammenhänge zwischen Variablen der Schanzendiagnostik und einzelnen Variablen der Leistungsdiagnostik festzustellen.

5.3.2. Latente-Variablen-Modell

Das Latente-Variablen-Modell (LVM) greift den Gedanken von einem oder mehreren latenten Faktoren, die hinter den beobachteten Variablen stehen (vgl. Faktorenanalyse aus Kapitel 3), auf und erweitert den Ansatz um weitere Kovariableneffekte, die sowohl auf die beobachteten Variablen als auch auf die latenten Faktoren wirken. Der Vorteil ist, dass man die Daten nicht wie im Faktorenmodell vorab auf Kovariableneffekte durch Standardisierung adjustieren muss, sondern unmittelbar die Effekte der Kovariablen untersuchen kann. Daher werden im Folgenden die nicht standardisierten Variablen aus der Leistungsdiagnostik verwendet. Konstruiert man das LVM so, dass die beobachteten Variablen den Zielgrößen entsprechen (hier also die Variablen zur Schanzendiagnostik) und lässt die möglichen Kovariablen (hier also die Variablen zur Leistungsdiagnostik) auf den latenten Faktor wirken, so kann mit dieser Modellklasse der Zusammenhang der beiden Datenquellen parametrisiert und geschätzt werden.

In diesem Abschnitt werden Latente-Variablen-Modelle (LVM) vorgestellt und auf den gemeinsamen Datensatz angewandt. Im Gegensatz zu den FARA-Modellen des vorherigen Abschnitts gehen hier wieder die ursprünglichen Variablen der Leistungsdiagnostik als erklärende Kovariablen anstelle der Faktorenwerte ein, da sich im vorherigen Modell der geringe Zusammenhang zwischen den Faktoren und den Variablen zur Schanzendiagnostik als Zielgrößen gezeigt hat. Zunächst werden die beiden im LVM enthaltenen Modelle vorgestellt, ein Modell für den gemeinsamen Datensatz formuliert und die Ergebnisse präsentiert. Hinweise zur verwendeten Software finden sich im Anhang B.1 ab Seite 95.

Modellgleichungen

Das hier präsentierte Latente-Variablen-Modell orientiert sich an der Darstellung von Raach (2005). Allerdings wird hier ein LVM mit nur einem latenten Faktor formuliert, da die Schanzendiagnostik zwei interessierende Zielgrößen enthält und nur Modelle mit weniger Faktoren als Zielgrößen sinnvoll sind. Die Erweiterung auf mehrere Faktoren wie in Raach (2005) bei mehr Zielgrößen ist einfach möglich. Wie in der Regressionsanalyse wird auch im Latente-Variablen-Modell die Unterscheidung zwischen Zielgrößen und Kovariablen getroffen. Die Kovariablen können im LVM direkt und indirekt über Faktoren auf die Zielgrößen wirken. Das Latente-Variablen-Modell erweitert das Faktorenmodell um Kovariableneffekte, die die beobachteten Zielgrößen zusätzlich zu dem latenten Faktor beeinflussen. Diese Effekte werden als *direkte* Kovariableneffekte bezeichnet. Sie sind in der inhaltlichen Interpretation nur von sekundärem Interesse, tragen aber zu einer besseren Anpassung des Modells an die Daten bei. Außerdem werden *indirekte* Kovariableneffekte modelliert. Diese wirken auf die latenten Faktoren und beeinflussen so indirekt die beobachteten Zielgrößen. Ein Überblick über die Elemente des Latente-Variablen-Modells ist in Tabelle 5.2 dargestellt.

Das Latente-Variablen-Modell besteht aus zwei Teilen. Zum einen das *Messmodell* (M)

$$y_{ij} = \lambda_0 + \mathbf{w}_{ij}^T \boldsymbol{\beta}_j + \lambda_j f_i + \epsilon_{ij}, \quad (5.2)$$

welches das Faktorenmodell mit einem latenten Faktor für die Zielgröße y_{ij} um die festen Kovaria-

Variablenname	Modell	Beschreibung	Laufindex
<i>Zielgrößen</i>			
n	M, S	Anzahl der Beobachtungen	$i = 1, \dots, n$
p	M	Anzahl der beobachteten Zielgrößen	$j = 1, \dots, p$
y_{ij}	M	beobachtete Ausprägung der Beobachtung i bei Zielgröße j	
<i>direkte Kovariableneffekte</i>			
λ_0	M	Interceptterm	
d	M	Anzahl der direkten Kovariablen	$c = 1, \dots, d$
\mathbf{w}_{ij}	M	direkte Kovariablen von Beobachtung i für Zielgröße j	
β_j	M	Vektor der direkten Regressionskoeffizienten für die Zielgröße j	
<i>Latente Variablen</i>			
f_i	M, S	Ausprägung des latenten Faktors bei Beobachtung i	
λ_j	M	Ladungskoeffizient des latenten Faktors f_i auf die Zielgröße j	
<i>indirekte Kovariableneffekte</i>			
g	S	Anzahl der indirekten Kovariableneffekte, die nichtparametrisch modelliert werden	$h = 1, \dots, g$
x_{ig}	S	indirekte Kovariable für Beobachtung i , die nichtparametrisch modelliert werden	
f_h	S	nichtlineare Funktionen, die den Einfluss der indirekten Kovariable \mathbf{x}_h modelliert	
w	S	Anzahl der indirekten Kovariableneffekte, die linear modelliert werden	$v = 1, \dots, w$
\mathbf{u}_i	S	linear modellierte indirekte Kovariablen für Beobachtung i	
γ	S	Regressionskoeffizient der indirekten Kovariablen	

Tabelle 5.2.: Übersicht über die Variablen im Latente-Variablen-Modell. Jeder Variable wird zugeordnet, ob sie im Messmodell (M) oder Strukturmodell (S) verwendet wird.

bleneffekte \mathbf{w}_{ij} und den zugehörigen d -dimensionalen Vektor der Regressionskoeffizienten β_j erweitert. Mit λ_0 wird der Interceptterm des Messmodells bezeichnet. Der Fehlerterm des Messmodells wird mit ϵ_{ij} bezeichnet und es wird

$$\epsilon_{ij} \sim N(0, \sigma_j^2)$$

angenommen. Das Modell (5.2) stellt den linearen Zusammenhang zwischen Zielgrößen und dem latenten Faktor einerseits und den direkten Kovariablen andererseits her. Damit hat das LVM im Wesentlichen die Struktur eines linearen Modells. Weiterhin besteht das Latente-Variablen-Modell aus einem *Strukturmodell* (S), das den latenten Faktor modelliert. Das Strukturmodell hat folgende additive Struktur:

$$f_i = \mathbf{u}_i^T \boldsymbol{\gamma} + f_1(x_{i1}) + \dots + f_g(x_{ig}) + \xi_i, \quad (5.3)$$

wobei ξ_i den Restfehler im Strukturmodell bezeichnet und es wird

$$\xi_i \sim N(0, 1)$$

angenommen. Außerdem sind \mathbf{u}_i die Kovariablen, die linear modelliert werden und $\boldsymbol{\gamma}$ der w -dimensionale Vektor der zugehörigen Regressionskoeffizienten. Die f_h , $h = 1, \dots, g$ sind glatte Funktionen zur Model-

lierung von nichtparametrischen Zusammenhängen. Eine Darstellung der Funktionen ist beispielsweise durch *P-Splines* möglich. Eine Beschreibung von P-Splines folgt im nächsten Abschnitt 5.3.3. Das nur die indirekten Kovariablen nichtparametrisch modelliert werden können, begründet Raach (2005) damit, dass die direkten Kovariableneffekte nur von sekundärem Interesse sind und das Augenmerk eher auf den latenten Faktoren liegt. Wie im Faktorenmodell kann die Varianz der beobachteten Zielgröße mit Gleichungen (5.2) und (5.3) zerlegt werden in einen Teil der durch den latenten Faktor erklärt wird und einen Restfehler, d.h. (Fahrmeir u. Raach, 2007)

$$\text{Var}(y_{ij}) = \lambda_j^2 + \sigma_j^2.$$

Latente-Variablen-Modelle können durch sogenannte *Pfaddiagramme* dargestellt werden. In Abbildung 5.3 ist exemplarisch ein Modell mit drei Zielgrößen y_1 , y_2 und y_3 , einer direkten Kovariablen w_1 , einem latenten Faktor, der hier mit z_1 bezeichnet wird, und zwei indirekten Kovariablen x_1 und x_2 gezeigt. Die beobachtbaren Variablen sind durch rechteckige Rahmen gekennzeichnet und der latente Faktor durch einen ovalen Rahmen. Die Pfeile zeigen jeweils von der unabhängigen Variable zur abhängigen Variable. Die Pfeile ohne Ursprung stehen für Fehlerterme. Man sieht, dass für die direkten und indirekten Kovariablen kein Messfehler angenommen wird. Hingegen wird für die beobachteten Zielgrößen angenommen, dass ihre Messung fehlerbehaftet ist.

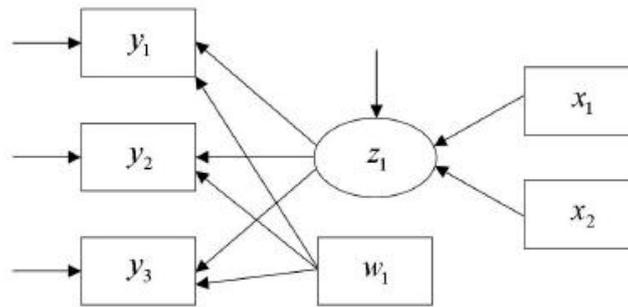


Abbildung 5.3.: Latentes Variablenmodell mit Kovariableneffekten (übernommen aus Raach (2005)).

Inferenz

Die unbekannt Parameter des Latente-Variablen-Modells von Raach (2005) werden mit *Bayesianischer Inferenz* geschätzt. In der Bayesianischen Inferenz nimmt man an, dass alle unbekannt Parameter $\theta \in \Theta$ Zufallsvariablen mit einer unbekannt Verteilung sind. Über diese Verteilung werden vorab Verteilungsannahmen getroffen, die sogenannte *Priori-Verteilung* $p(\theta)$. Anhand der beobachteten Daten \mathbf{x} wird dann die Likelihood der Daten $L(\mathbf{x}|\theta)$ aufgestellt. In ihr drückt sich die Information über die unbekannt Parameter anhand der Daten aus. Nachdem Daten beobachtet wurden, wird die Priori-Verteilung mit der Likelihood zur *Posteriori-Verteilung* aktualisiert. Mit dem Satz von Bayes (vgl. Grimmett u. Stirzaker (2001)) ergibt sich die Posteriori-Verteilung als

$$p(\theta|\mathbf{x}) = \frac{p(\theta, \mathbf{x})}{p(\mathbf{x})} = \frac{L(\mathbf{x}|\theta)p(\theta)}{p(\mathbf{x})} = \frac{L(\mathbf{x}|\theta)p(\theta)}{\int_{\Theta} L(\mathbf{x}|\theta)p(\theta)d\theta} \propto L(\mathbf{x}|\theta)p(\theta).$$

Dies besagt, dass die Posteriori-Verteilung proportional zum Produkt von Likelihood und Priori-Verteilung ist und sich so berechnen lässt. Auf dieser Posteriori-Verteilung beruht die Bayesianische Inferenz. Aus ihr können Punktschätzer und Kreditabilitätsintervalle für unbekannt Parameter ermit-

telt werden.

Vor der Analyse müssen daher Priori-Verteilungen für alle Parameter des Modells festgelegt werden. Im Messmodell müssen die folgenden Parameter geschätzt werden:

$$\boldsymbol{\theta} = (\lambda_0, \boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_p, \lambda_1, \dots, \lambda_p, \sigma_1^2, \dots, \sigma_p^2)^T.$$

Für alle Parameter werden als Priori-Verteilungen die Standardeinstellungen des Programms gewählt. Dabei handelt es sich um flache nichtinformative Prioris, die auch als *diffuse* Priori-Verteilungen bezeichnet werden. Inhaltlich gesprochen bedeutet dies, dass a priori kein Vorwissen in Form einer Verteilung für die unbekannt Parameter in das Modell eingeht. Für die Varianzen der Fehlerterme werden inverse Gamma-Verteilung (IG) als Priori-Verteilung angenommen, d.h.

$$\sigma_j^2 \sim \text{IG} \left(\frac{\nu}{2}, \frac{\nu s^2}{2} \right), j = 1, \dots, p,$$

mit zwei Hyperparametern $\nu > 0$ und $s^2 > 0$. Kleine Hyperparameter führen zu einer schwach informativen Priori-Verteilung. Auch für die Elemente des Strukturmodells werden nichtinformativ Priori-Verteilungen verwendet. Dabei wird angenommen, dass die beiden Teile des Latente-Variablen-Modells stochastisch unabhängig sind. Anhand dieser Priori-Annahmen kann dann die gemeinsame Posteriori-Verteilung aller unbekannt Parameter konstruiert werden.

Da die gemeinsame Posteriori-Verteilung nicht analytisch berechnet werden kann, werden in der Bayesianischen Inferenz mit Markov-Chain-Monte-Carlo (MCMC)-Verfahren eine Markov-Kette von Stichproben aus der gemeinsamen Posteriori-Verteilung der unbekannt Parameter generiert. Zur Ziehung der Stichproben wird der sogenannte *Generalisierte Gibbs Sampler* (GGs) verwendet, der in Raach (2005) beschrieben wird. Die Ziehung erfolgt anhand der vollständig bedingten Dichten, die in Fahrmeir u. Raach (2007) beschrieben werden. Damit die Stichproben als Zufallsziehungen aus der Posteriori-Verteilung angesehen werden können, muss die Markov-Kette nach einer Konvergenzzeit (die sogenannte „Burn-In-Phase“) einen stationären Zustand erreichen. Die gesamte Markov-Kette hat hier eine Länge von 7000 Stichproben, und dabei wird eine Burn-In-Phase von 2000 Stichproben gewählt. Häufig sind aufeinanderfolgende Stichproben korreliert. Darum wählt man zusätzlich ein Ausdünnungsintervall für die Markov-Kette. Hier wird nur jede 10. Stichprobe ausgewertet. Insgesamt beruhen die Posteriori-Schätzer damit auf 500 MCMC-Stichproben. Im Anschluss an die Schätzung eines Modells muss diese Konvergenz anhand der *Samplingpfade* überprüft werden. In einem Samplingpfad werden die Schätzer einzelner Parameter aus jeder Stichprobe nach Burn-In-Phase und Ausdünnung dargestellt. Eine längere Markov-Kette ließ sich aufgrund eines Fehlers in dem vorliegenden Programmcode nicht realisieren.

In Latente-Variablen-Modellen ist man häufig mit Identifizierbarkeitsproblemen konfrontiert (vgl. Fahrmeir u. Khatab, 2008, S.9). Jede Transformation des Ladungskoeffizienten von λ_j zu $\lambda_j^* = \lambda_j \cdot c$ und des latenten Faktors f_i zu $f_i^* = f_i \cdot \frac{1}{c}$ führt zu demselben Effekt $\lambda_j f_i = \lambda_j^* f_i^*$ im Messmodell (5.2) auf die Zielgröße y_{ij} . Man erkennt Identifizierbarkeitsprobleme in den Samplingpfaden genau dann, wenn die geschätzten Werte von λ_j und f_i nicht konvergieren, sondern sich in entgegengesetzte Richtungen entwickeln. Daher ist es zweckmäßig, den Samplingpfad des Produktes $\lambda_j f_i$ zu betrachten. Inhaltlich gesprochen ist dies der „Nettoeffekt“ des latenten Faktors auf die Zielgrößen $j = 1, \dots, p$ im Messmodell.

In diesem Zusammenhang ist auch der sogenannte *Heywood-Case* bekannt, der in Raach (2005) beschrieben wird. In diesem Fall wird eine Zielgröße j' vollständig von einem latenten Faktor erfasst und die Varianz des Restfehlers $\sigma_{j'}^2$ geht gegen Null. Gleichzeitig konvergiert der Ladungskoeffizienten $\lambda_{j'}$

gegen unendlich, was anhand der Samplingpfade zu erkennen ist. In diesem Fall trifft die Annahme eines latenten Faktors nicht zu. Untersuchungen haben ergeben, dass die Wahrscheinlichkeit eines Heywood-Cases mit steigendem Stichprobenumfang n sinkt.

Latente-Variablen-Modell für die Diagnostikdaten

Im Folgenden wird ein Latente-Variablen-Modell für die gemeinsamen Daten zur Leistungsdiagnostik und Schanzendiagnostik aufgestellt. Um eine breitere Datengrundlage zu erhalten, werden alle vollständigen Beobachtungen des gemeinsamen Datensatzes zur Schätzung verwendet. Als Zielgrößen werden die interessierenden Größen der Schanzendiagnostik, also `Weite` und `vertikaleGeschw` gewählt. Als direkte Kovariablen wirken auf die Zielgrößen `Alter` und `Sportart`. Das Latente-Variablen-Modell enthält einen latenten Faktor, auf den je eine Variablen der Leistungsdiagnostik linear wirkt. Eine größere Anzahl an indirekten Kovariablen führte zu Schätzproblemen, da keine konvergenten Markov-Ketten erzeugt werden können. Dieses Problem tritt häufig auf, wenn die Anzahl der unbekannt Parameter im Modell im Verhältnis zur Stichprobenanzahl groß ist. Bei einem größeren Stichprobenumfang ließen sich mehr Kovariablen im Strukturmodell realisieren. Eine schematische Darstellung des Modells für die Diagnostikdaten findet sich in Abbildung 5.4.

Inhaltlich ist dieses Latente-Variablen-Modell wie folgt zu interpretieren. Die Leistung eines Athleten bei den Übungssprüngen drückt sich in der Sprungweite und der vertikalen Absprunggeschwindigkeit aus. Hinter diesen beiden messbaren Größen steht ein latenter Faktor, der so etwas wie das Leistungsniveau des Athleten bei Übungssprüngen ausdrückt und sich in den Variablen `Weite` und `vertikaleGeschw` manifestiert. Neben diesem latenten Faktor haben auch `Alter` und `Sportart` einen direkten Effekt auf die Weite. So vermutet man bessere Ergebnisse bei erfahrenen Athleten als bei jungen und bessere Ergebnisse bei Skispringern als bei Nordischen Kombiniern. Auf den latenten Faktor wirkt nun je eine der Variablen der Leistungsdiagnostik. Damit untersucht man, wie das Leistungsniveau bei den Übungssprüngen von dieser Variable indirekt über den latenten Faktor beeinflusst wird. Somit vereint dieses Modell zwei Stufen. Zum einen wird der Zusammenhang der beobachteten Variablen `Weite` und `vertikaleGeschw` durch einen latenten Faktor dargestellt im Messmodell dargestellt. Zum anderen wird der Zusammenhang von Kovariablen mit dem latenten Faktor im Strukturmodell hergestellt.

Ergebnisse

Im Folgenden wird das obige Latente-Variablen-Modell an die vorliegenden Daten angepasst und die unbekannt Parameter geschätzt. Es werden exemplarisch die Ergebnisse von drei Latente-Variablen-Modellen mit jeweils den gleichen Zielgrößen und direkten Kovariablen und einer der drei Variablen `SH_V75E`, `F_re170_bb` und `A60` als indirekte Kovariable präsentiert. Die Posteriori-Schätzer der $d=2$ direkten Kovariablen auf die beiden Zielgrößen in den drei Modellen sind in Tabelle 5.3 dargestellt.

direkte Kovariable	SJ_V75E		F_re170_bb		A60	
	Sportart	Alter	Sportart	Alter	Sportart	Alter
<code>Weite</code> (β_1^T)	-0.21 (0.27)	0.21 (0.12)	0.05 (0.26)	0.42 (0.1)	0.13 (0.25)	0.35 (0.1)
<code>vertikaleGeschw</code> (β_2^T)	0.79 (0.22)	0.14 (0.1)	1.14 (0.23)	0.42 (0.08)	1.4 (0.19)	0.29 (0.08)

Tabelle 5.3.: Posteriori-Mittelwerte (Standardabweichungen) für die Effekte $\beta_1 = (\beta_{11}, \beta_{12})^T$ und $\beta_2 = (\beta_{21}, \beta_{22})^T$ der direkten Kovariablen in den drei Latente-Variablen-Modellen in einem LVM mit je einer der indirekten Kovariablen `SJ_V75E`, `F_re170_bb` und `A60`.

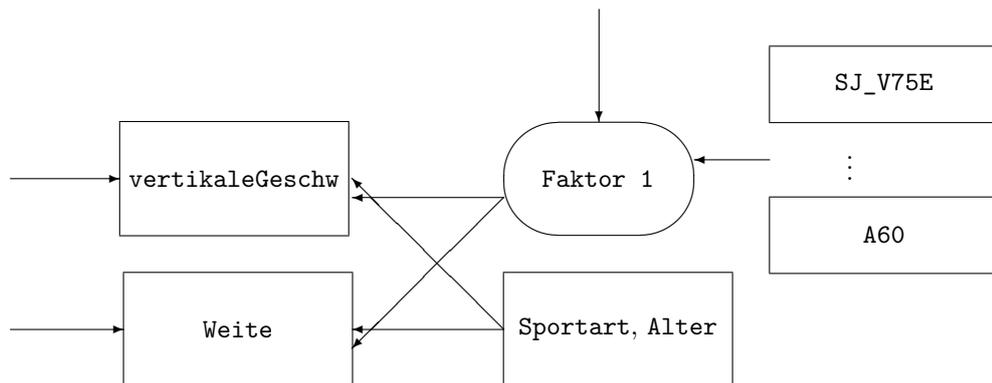


Abbildung 5.4.: Pfaddiagramm des Latente-Variablen-Modells für die Daten zur Leistungsdiagnostik und Schanzendiagnostik.

Die Variable **Sportart** ist wie in Gleichung (2.6) auf Seite 26 kodiert. Man sieht nur bei **vertikaleGeschw** einen bedeutsamen Unterschied der Sportarten. Hier wird für die Skispringer in allen drei Modellen ein positiver Effekt geschätzt. Zwischen **Sportart** und **Weite** kann in keinem der Modelle ein signifikanter Effekt festgestellt werden. Die direkte Kovariable **Alter** hat in den drei Modellen einen positiven Effekt auf die **vertikaleGeschw** und die **Weite**. Dies bestätigt die Annahmen, die man vorab über die Wirkungsrichtung der direkten Kovariablen hatte.

Von besonderem Interesse ist die Schätzung des Effekts der indirekten Kovariable, da diese den Einfluss einer Variable der Leistungsdiagnostik auf die Variablen der Schanzendiagnostik angibt. Das zugehörige Strukturmodell hat die Form $f_i = u_{i1}\gamma_1 + \xi_i$, wobei u_{i1} die Ausprägung der indirekten Kovariable bei Beobachtung i ist. Zusätzlich sind die Posteriori-Schätzer für die Ladungskoeffizienten $\hat{\lambda}_1$ und $\hat{\lambda}_2$ wichtig, da diese multipliziert mit dem Regressionskoeffizient $\hat{\gamma}_1$ im Messmodell (5.2) den Nettoeffekt der indirekten Kovariable auf die **Weite** bzw. **vertikaleGeschw** angeben. In Tabelle 5.4 sind die Posteriori-Schätzer der Ladungskoeffizienten und des Regressionskoeffizienten für alle drei Modelle dargestellt.

ind. Kovariable	SJ_V75E		F_re170_bb		A60	
	MW (SD)	95%-KI	MW (SD)	95%-KI	MW (SD)	95%-KI
$\hat{\gamma}_1$	4.51(3.28)	[0.82,10.78]	0.38(0.14)	[0.12,0.71]	0.97(0.69)	[0.37,2.96]
$\hat{\lambda}_1$	0.11(0.08)	[0.01,0.3]	0.18(0.1)	[0,0.38]	0.14(0.09)	[-0.01,0.32]
$\hat{\lambda}_2$	0.23(0.18)	[0.04,0.65]	0.69(0.12)	[0.38,0.85]	0.52(0.2)	[0.11,0.78]
$\hat{\lambda}_1 \cdot \hat{\gamma}_1$	0.27 (0.12)	[0.04,0.52]	0.07 (0.05)	[0,0.2]	0.11 (0.08)	[-0.01,0.3]
$\hat{\lambda}_2 \cdot \hat{\gamma}_1$	0.53 (0.09)	[0.35,0.71]	0.26 (0.09)	[0.07,0.42]	0.38 (0.09)	[0.21,0.55]

Tabelle 5.4.: Posteriori-Mittelwert, Standardabweichung und 95%-Kreditabilitätsintervall für die drei LVM mit je einer der indirekten Kovariablen **SJ_V75E**, **F_re170_bb** und **A60**. Dargestellt sind die Schätzer für den Regressionskoeffizienten γ_1 im Strukturmodell, die Ladungskoeffizienten λ_1 und λ_2 und das Produkt beider Parameter (Nettoeffekt).

Man sieht in Tabelle 5.4, dass in allen drei Modellen ein positiver Effekt der jeweiligen indirekten

Kovariable auf den latenten Faktor geschätzt wird. Dabei umfasst das 95%-Kredibilitätsintervall jeweils nicht die Null. Mit mehr als 95%-Posteriori-Wahrscheinlichkeit wird also ein positiver Effekt geschätzt. Bei den Ladungskoeffizienten gilt jeweils in den drei Modellen $\hat{\lambda}_1 < \hat{\lambda}_2$. Der latente Faktor wirkt also jeweils stärker auf die `vertikaleGeschw` als auf die `Weite`. Der stärkste Effekt einer indirekten Kovariable ist der von `SJ_V75E` im ersten Modell. Allerdings stellt man bei der Betrachtung der Samplingpfade fest, dass der Effekt in den MCMC-Simulationen groß ist, in denen die Ladungskoeffizienten klein sind. In diesem Fall liegt ein Identifizierbarkeitsproblem zwischen Effekt und Ladungskoeffizient vor. Die Samplingpfade der Modelle sind in Abbildung D.1 im Anhang auf Seite 102 dargestellt. Betrachtet man den Nettoeffekt, so verschwindet der starke Effekt der Variable `SJ_V75E` auf die Zielgrößen.

Mit den Latente-Variablen-Modellen konnte festgestellt werden, dass die indirekte Kovariable jeweils einen positiven Effekt auf den latenten Faktor und damit die Zielgrößen hat. Jede der drei Variablen `SJ_V75E`, `F_re170_bb` und `A60` aus der Leistungsdiagnostik hat einen positiven Effekt auf das Leistungsniveau bei der Schanzendiagnostik. Bessere Ergebnisse in der Leistungsdiagnostik haben also einen positiven Effekt auf die Ergebnisse bei der Schanzendiagnostik. Also lassen sich auch anhand der Leistungsdiagnostik Aussagen über das Leistungsniveau eines Athleten bei Übungssprüngen machen. Der latente Faktor erfasst dabei hauptsächlich die Zielgröße `vertikaleGeschw`. Ein Latente-Variablen-Modell erlaubt es, den Einfluss einer Kovariable simultan auf zwei Zielgrößen zu untersuchen. Möchte man den Einfluss mehrerer Kovariablen auf eine Zielgröße untersuchen, und dabei die wichtigsten Kovariablen identifizieren, so bieten sich Regressionsmodelle an, wie sie im folgenden Abschnitt beschrieben werden.

5.3.3. Strukturierte additive Regressionsmodelle

Die *strukturierten additiven Regressionsmodelle* (STAR) bezeichnen eine sehr flexible Modellklasse zur Analyse der Zusammenhänge zwischen einer Zielgröße und einer Anzahl an Kovariablen. In dieser Modellklasse kann der Effekt einer Gruppe von Kovariablen linear und von einer Gruppe nichtparametrisch modelliert werden. Weiterhin können auch räumliche Effekte und Interaktionen modelliert werden. Die Schätzung erfolgt wie bei den Latente-Variablen-Modellen mit Bayesianischer Inferenz und MCMC-Simulationen. Eine Einführung in STAR-Modelle findet sich in Brezger u. Lang (2006).

Beim Auswerten von Daten, z.B. aus der Sportwissenschaft, steht der Anwender vor der Entscheidung, welche Gestalt das Modell haben soll, mit dem er die Zusammenhänge modellieren will. Diese Entscheidungsfindung wird als *Modellwahl* bezeichnet. In der Modellwahl wird festgelegt, welche der vorliegenden Kovariablen (*Variablenselektion*) in welcher Form (linear vs. nichtparametrisch) im Modell berücksichtigt werden. Bei dem Aufstellen von Modellen konkurrieren dabei die beiden gegensätzlichen Anforderungen an das Modell, *Modellanpassung* und *Modellkomplexität*, zwischen denen ein Kompromiss gefunden werden muss. Die Lösung soll ein Modell mit möglichst wenigen Variablen, die aber die wichtigsten Effekte identifizieren, sein.

Eine Regularisierung der Effekte im linearen Modell bietet sich an, wenn viele Kovariablen im Modell sind, die untereinander stark korreliert sind. Regularisieren bedeutet dabei, dass die Regressionskoeffizienten in ihrem Wertebereich beschränkt werden. Dabei bietet die Regularisierung der Effekte häufig eine Möglichkeit Schätzer zu erhalten, wenn die Kleinste-Quadrate-Schätzer nicht existieren. Zudem haben die regularisierten Effekte eine kleinere Varianz als die Kleinste-Quadrate-Schätzer und damit bessere Prognoseeigenschaften.

In der Bayesianischen Inferenz entspricht das Regularisieren dem Einführen von informativen Prioriverteilungen (sogenannten *Shrinkage-Prioris* bzw. *Schrumpfung-Prioris*) für die Regressionskoeffizi-

enten. Dieser Vorgang wird auch als *Bayesianische Regularisierung* bezeichnet, da die Effekte einiger Kovariablen *regularisiert* werden. Regularisieren bedeutet, dass der Effekt einer Kovariable fast auf Null geschrumpft wird und dass das Modell somit effektiv einen Parameter weniger enthält. Damit kann eine Variablenselektion durchgeführt werden. Dabei ist es eine gewünschte Eigenschaft von Schrumpfungsprioris, dass kleine und unbedeutende Effekte auf Null geschrumpft werden, aber große Effekte nur moderat geschrumpft werden. In Kneib u. a. (2009) werden die STAR-Modelle um die Möglichkeit der Bayesianischen Regularisierung erweitert. STAR-Modell mit Bayesianischer Regularisierung können mit der Software *BayesX* (Brezger u. a., 2009) geschätzt werden.

Modellformulierung

Das STAR-Modell stellt den Zusammenhang zwischen einer Zielgröße und einer Reihe von Kovariablen her. Der Effekt einer Kovariable wird auf verschiedene Art und Weise im Modell berücksichtigt. Das folgende Modell enthält lineare, nichtparametrische und regularisierte lineare Effekte, die jeweils *additiv* miteinander verknüpft werden. Da die Zielgrößen normalverteilt sind, werden die Kovariablen linear mit der Zielgröße verknüpft. Das Modell lautet dann in allgemeiner Form:

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{u}_i^T \boldsymbol{\gamma} + f_1(z_{i1}) + \dots + f_q(z_{iq}) + \epsilon_i \quad , i = 1, \dots, n, \quad (5.4)$$

wobei y_i die Ausprägung der Zielgröße bei Beobachtung i bezeichnet. Mit \mathbf{x}_i wird der Vektor der p Kovariablen bezeichnet, die regularisiert werden und mit $\boldsymbol{\beta}$ die entsprechenden (regularisierten) Regressionskoeffizienten. Zusätzlich wird eine Anzahl von r Kovariablen \mathbf{u}_i nicht regularisiert in das Modell aufgenommen. Mit $\boldsymbol{\gamma}$ wird der zugehörige r -dimensionale Vektor von Regressionskoeffizienten bezeichnet. Weiterhin enthält das Modell nichtparametrische Effekte der Kovariablen z_{i1}, \dots, z_{iq} , deren funktionale Form des Effekt f_1, \dots, f_q durch P-Splines geschätzt wird. Für den Fehlerterm des Modells wird

$$\epsilon_i \sim N(0, \sigma^2)$$

angenommen. Der Vektor der zu regularisierenden Regressionskoeffizienten $\boldsymbol{\beta} = (\beta_1, \dots, \beta_j, \dots, \beta_p)^T$ wird mit einer Schrumpfungspriori versehen. Eine Variante ist die sogenannte NMIG-Priori („normal mixture of inverse gamma distributions“), die von Ishwaran u. Rao (2005) vorgeschlagen wurde. Für einen Regressionskoeffizienten wird folgende Priori-Verteilung angenommen

$$\beta_j \sim N(0, \tau_j^2 := I_j \Psi_j^2) \quad , j = 1, \dots, p, \quad (5.5)$$

wobei für den Varianzparameter der Normalverteilung τ_j^2 eine Mischverteilung von zwei inversen Gamma-Verteilungen (IG) spezifiziert wird. Die Mischverteilung wird anhand des Produkts des Indikators I_j und des Varianzparameter Ψ_j^2 konstruiert (Konrath u. a., 2008). Die Schrumpfungspriori lässt sich wie folgt darstellen durch:

$$I_j \sim (1 - \omega) \delta_{\nu_0}(I_j) + \omega \delta_{\nu_1}(I_j) \quad (5.6)$$

$$\Psi_j^2 \sim \text{IG}(a, b). \quad (5.7)$$

Die erste Komponente (5.6) ist eine binäre Indikatorvariable mit Punktmassen bei den Werten $\nu_0 > 0$ und $\nu_1 > 0$, dargestellt durch das Kroneckersymbol δ . Dabei gilt $\delta_{\nu_1}(I_j) = 1$ falls $I_j = \nu_1$ und 0 sonst. Dabei wird für ν_0 ein positiver Wert nahe Null und für ν_1 der Wert 1 gewählt. Der *Komplexitätsparameter* ω steuert wie wahrscheinlich die Indikatorvariable die Ausprägung ν_1 bzw. ν_0 annimmt. A priori wird kein Vorwissen über die Verteilung des Komplexitätsparameters angenommen, d.h. es wird

eine Gleichverteilung auf $[0,1]$ als Priori-Verteilung für ω angenommen

$$\omega \sim U(0, 1).$$

Durch die Annahmen in (5.6) und (5.7) wird für die Varianzkomponente $\tau_j^2 = I_j \Psi_j^2$ eine bimodale Mischverteilung von zwei inversen Gamma-Verteilungen spezifiziert. Die Hyperparameter a, b, ν_0 und ν_1 steuern die Gestalt der Verteilung des Varianzparameters τ_j^2 . Die Mischverteilung der inversen Gamma-Verteilungen hat zwei Modi an den Stellen:

$$\text{mod}_{\nu_0} = \frac{\nu_0 b}{a + 1}, \quad \text{mod}_{\nu_1} = \frac{\nu_1 b}{a + 1}.$$

Wird der Effekt β_j einer Variable regularisiert, so gilt $I_j = \nu_0$ in Gleichung (5.6) und damit sind für τ_j^2 nur kleine Werte möglich. Für die Regressionskoeffizienten in (5.5) führen kleine Varianzparameter τ_j^2 zu einer Verteilung, die stark um die Null konzentriert ist. Mit wachsendem Varianzparameter wird die Schätzung betragsmäßig größerer Effekte ermöglicht und so wird bei $I_j = \nu_1$ der Effekt nur schwach regularisiert. Variablenselektion wird mit der NMIG-Priori wie folgt durchgeführt. Anhand der relativen Häufigkeit von $I_j = \nu_1$ in den MCMC-Simulationen wird der Posteriori-Schätzer für $P(I_j = \nu_1 = 1)$ für jede Variable $j = 1, \dots, p$ gefunden. Die Variablen mit einem großen Wert von $P(I_j = 1)$ werden nur schwach regularisiert und sind daher bedeutsam für das Modell. Die Beurteilung, ob ein Wert von $P(I_j = 1)$ „groß“ ist, kann relativ zu den anderen Werte getroffen werden. Eine intuitive Möglichkeit ist die Auswahl der Variablen mit $P(I_j = 1) > 0.5$.

Neben den NMIG-Prioris gibt es in der Bayesianischen Inferenz weitere Priori-Verteilungen zur Schrumpfung von Effekten. Dazu gehören die Bayesianischen Varianten von *Lasso* und *Ridge*, wobei letztere hier beschrieben wird. Die Ridge-Priori ist gegeben durch folgende Hierachiestufen (Konrath u. a., 2008)

$$\begin{aligned} \beta_j &\sim N(0, \tau_j^2) \quad , j = 1, \dots, p, \\ \tau_j^2 &\sim \delta_{\frac{1}{2\lambda}}(\tau_j^2) \\ \lambda &\sim \text{Ga}(a, \frac{b}{2}). \end{aligned}$$

Dabei bezeichnet δ wieder das Kronecker-Symbol und $\text{Ga}()$ die Gammaverteilung mit den Hyperparametern $a > 0$ und $b > 0$. Der Erwartungswert dieser Gamma-Verteilung ist $\frac{2a}{b}$ und die Varianz ist $\frac{4a}{b^2}$. Der Parameter λ wird auch als *Schrumpfungparameter* bezeichnet. Mit ihm wird die Fensterbreite der Priori-Verteilung und somit der Wertebereich der Regressionskoeffizienten β_j festgelegt. Je größer λ , desto kleiner wird τ_j^2 und damit der Wertebereich für β_j . Die Darstellung von λ mit einer Gamma-Verteilung erlaubt eine flexible Modellierung des Schrumpfungparameters mit den Hyperparametern a und b . Bei der Bayesianischen Ridge-Priori wird Variablenselektion wie folgt durchgeführt. Anhand der MCMC-Stichproben wird der Regressionskoeffizient durch den Posteriori-Mittelwert geschätzt. Zur Auswahl bedeutsamer Kovariablen bieten sich die folgenden Vorgehensweisen an:

- Eine Kovariable wird aus dem Modell entfernt, wenn das 95%-Kredibilitätsintervall ihres Regressionskoeffizienten die Null umfasst.
- Eine Kovariable wird aus dem Modell entfernt, wenn in dem Intervall $[\hat{\beta}_j - \hat{se}(\hat{\beta}_j), \hat{\beta}_j + \hat{se}(\hat{\beta}_j)]$ die Null enthalten ist.
- Es werden die Kovariablen aus dem Modell entfernt, für die der geschätzte Regressionskoeffizient $|\hat{\beta}_j| < c$ ist („Hard Shrinkage“).

Für die nicht regularisierten Effekte $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_j, \dots, \gamma_r)^T$ werden sogenannte *diffuse Prioris*

$$\gamma_j \propto 1 \quad , j = 1, \dots, r,$$

angenommen. Mit diffusen Prioris wird kein Vorwissen über den Parameter zum Ausdruck gebracht. Da die Varianz einer diffusen Priori nicht begrenzt ist, sind auch die Regressionskoeffizienten a priori nicht in ihrem Wertebereich beschränkt, so wie es bei den Schrumpfungsprioris der Fall ist.

Vergleich von NMIG- und Ridge-Priori

Für die Regressionskoeffizienten kann jeweils die marginale Verteilung angegeben werden, die nur noch von den Hyperparametern abhängt. Dabei gilt (Konrath u. a., 2008):

$$\text{NMIG} : \beta_j \sim \frac{1}{2} t \left(\beta_j | df = 2a, \text{scale} = \sqrt{\frac{\nu_0 b}{a}} \right) + \frac{1}{2} t \left(\beta_j | df = 2a, \text{scale} = \sqrt{\frac{\nu_1 b}{a}} \right) \quad (5.8)$$

$$\text{Ridge} : \beta_j \sim t \left(\beta_j | df = 2a, \text{scale} = \sqrt{\frac{b}{2a}} \right), \quad (5.9)$$

wobei $t()$ die Dichtefunktion einer skalierten t-Verteilung bezeichnet. Als Hyperparameter für die NMIG-Prioris werden gewählt:

$$\nu_0 = 2.5e^{-5}, \nu_1 = 1, a = 5, b = 25.$$

Diese Wahl der Hyperparameter bewirkt eine schwächere Regularisierung kleiner Effekte. Bei der Ridge-Priori werden die Hyperparameter

$$a = 0.01, b = 0.01$$

gewählt. Ein Vergleich der beiden Priori-Verteilungen der Regressionskoeffizienten für beide Varianten der Regularisierung mit den gewählten Hyperparametern ist in Abbildung 5.5 dargestellt. Man sieht, dass beide Priori-Verteilungen eine große Wahrscheinlichkeitsmasse auf die Null legen. Sehr starke Effekte ab ca. einem Betrag von 8 sind unter der Ridge-Priori wahrscheinlicher als unter der NMIG-Priori.

Inferenz

Die Schätzung des Modells erfolgt Bayesianisch (vgl. dazu auch den Abschnitt Inferenz zum Latente-Variablen-Modell). Die Inferenz beruht auf der gemeinsamen Posteriori-Verteilung der unbekannt Parameter aus Gleichung (5.4):

$$\boldsymbol{\theta} := (\boldsymbol{\beta}, \boldsymbol{\gamma}, f_1, \dots, f_q, \sigma^2)^T.$$

Dazu wird die gemeinsame Posteriori-Verteilung aller unbekannter Parameter maximiert. Da die gemeinsame Posteriori-Verteilung nicht analytisch berechnet werden kann, werden mit MCMC-Verfahren Stichproben aus ihr gezogen. Dazu werden die vollständig bedingten Dichten aller unbekannter Parameter benötigt, die in Kneib u. a. (2009) beschrieben werden. Aus diesen Verteilungen werden mit dem Gibbs-Sampler Stichproben gezogen. Hier werden für die Schätzung der Posteriori-Verteilung 30000 MCMC-Stichproben gezogen, wobei eine Burn-In-Phase von 10000 gewählt wird und zur Vermeidung von Autokorrelation ein Ausdünnungsintervall von 20. Damit beruht die Inferenz auf 1000 MCMC-Samples der Posterioriverteilung.

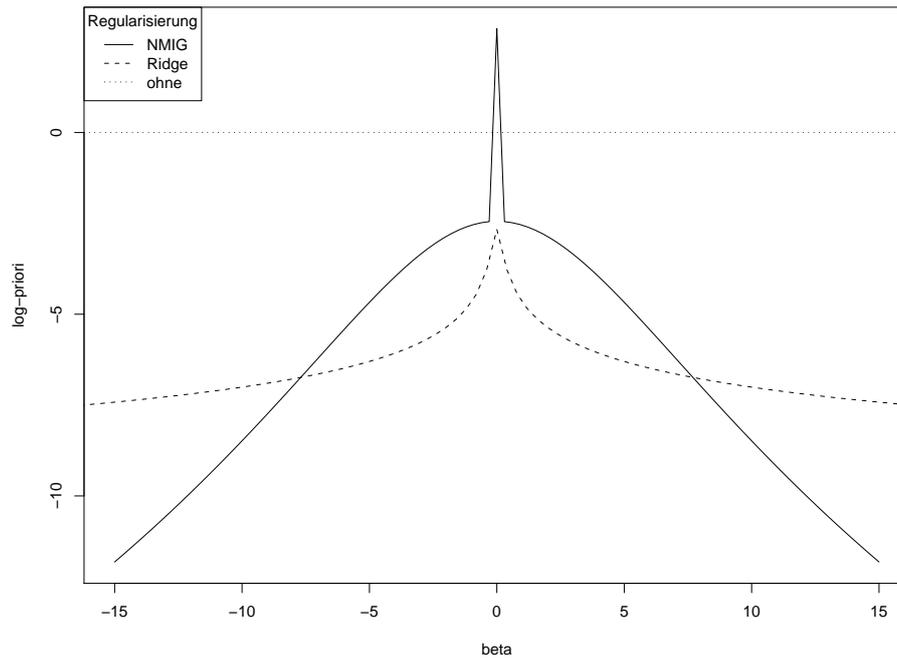


Abbildung 5.5.: Logarithmierte Dichten der Schrumpfungsprioris bei festgelegten Hyperparametern. Man beachte, dass die Priori-Verteilung im Fall von keiner Regularisierung keine reguläre Dichte ist.

Die unbekannt, glatten Funktionen f_j , $j = 1, \dots, q$ werden durch um Null zentrierte Polynom-Splines zu dem Grad l_j und einer vorgegebenen Knotenmenge $\kappa_{j1} < \dots < \kappa_{jm}$ dargestellt. Ein Polynom-Spline wird dann durch eine Linearkombination von $M_j = m_j + l_j - 1$ B-Spline-Basisfunktionen vom Grad l_j berechnet, d.h.

$$f_j(x) = \sum_{m=1}^{M_j} \beta_{jm} B_{jm}(x).$$

Je höher Grad und Anzahl der Knoten gesetzt werden, desto komplexere Funktionen können erzeugt werden. Bei der Schätzung wird der Koeffizientenvektor $\boldsymbol{\beta}_j = (\beta_{j1}, \dots, \beta_{jM_j})^T$ geschätzt und die Basisfunktionen, ausgewertet an der Stelle x , $B_{j1}(x), \dots, B_{jM_j}(x)$ damit skaliert. Zur Vermeidung eines zu variablen Funktionsverlaufes werden die Koeffizienten $\beta_{j3}, \dots, \beta_{jM_j}$ ($\beta_{j1} \propto \beta_{j2} \propto 1$) mit einer Random-Walk-Priori zweiter Ordnung belegt (Brezger u. Lang, 2006). Diese ist definiert als

$$\beta_{jm} = 2\beta_{j,m-1} - \beta_{j,m-2} + u_{jm}, \quad m = 3, \dots, M_j$$

und $u_{jm} \sim N(0, \tau_j^2)$ und $\tau_j^2 \sim IG(a_j, b_j)$. Damit werden große Abweichungen zwischen benachbarten Koeffizienten regularisiert. Weitere Informationen zur nichtlinearen Funktionsanpassung finden sich in Hastie u. Tibshirani (1990) oder Fahrmeir u. a. (2007).

Ergebnisse

Mit dieser Modellklasse soll anhand der Ergebnisse der beiden vorherigen Modelle ein Modell konstruiert werden, das die wichtigsten Variablen der Leistungsdiagnostik identifiziert, die in einem Zusammenhang mit den Variablen zur Schanzendiagnostik stehen. Dabei werden folgende Ideen verfolgt:

- In das Modell gehen alle 75 vollständigen Beobachtungen ein und werden als unabhängig angesehen. Eine Modellierung individuenspezifischer Effekte wie im lineare gemischten Modell wäre im Rahmen von STAR-Modellen möglich, ist aber aufgrund der geringen Anzahl an Messwiederholungen nicht sinnvoll.
- Statt der Faktorenwerte gehen die ursprünglichen Variablen der Leistungsdiagnostik als Kovariablen in das Modell ein. Anhand vorheriger Untersuchungen hat sich ergeben, dass die Effekte von `relKraftanstieg_bb` und `DJ32_Faktor` nichtparametrisch geschätzt werden sollen. Die übrigen Variablen gehen linear in das Modell ein. Um eine Variablenselektion durchzuführen, werden die Effekte der Variablen der Leistungsdiagnostik regularisiert. Ziel ist die Identifikation von den Variablen, die im Rahmen einer Modellwahl als bedeutsam angesehen werden, weil sie in engem Zusammenhang mit der Zielgröße stehen. Um die Modellkomplexität bereits vorab zu reduzieren, werden von den Messungen, die beidbeinig und einbeinig durchgeführt wurden, nur die beidbeinigen ausgewählt.
- Die Variablen zur Leistungsdiagnostik gehen ohne eine Standardisierung auf das Alter in das Modell ein. Daher werden als zusätzliche Kovariablen `Alter` (nichtparametrisch) und `Sportart` in die Modelle aufgenommen. Diese werden nicht regularisiert.

Um die Anwendbarkeit von NMIG-Prioris zu gewährleisten, werden alle Variablen der Leistungsdiagnostik vor der Analyse auf Mittelwert Null und Standardabweichung 1 standardisiert. Damit können auch die Absolutbeträge der Effekte direkt verglichen werden, da alle Kovariablen auf derselben Skala vorliegen. Zur Schätzung des Modells wird die Methode `regress` aus `BayesX` verwendet. Die Effekte der Variablen `Alter`, `relKraftanstieg_bb` und `DJ32_Faktor` werden nichtparametrisch mit P-Splines der Ordnung 3 und 10 Knoten geschätzt. Die Variable `Sportart` wird linear in das Modell genommen und die Variablen zur Leistungsdiagnostik werden mit NMIG-Prioris zur Regularisierung der Effekte versehen. Als Zielgröße wird zunächst die `Weite` gewählt und anschließend vergleichend die Ergebnisse eines Modells mit `vertikaleGeschw` als Zielgröße vorgestellt.

Die geschätzten Kurven für die nichtparametrischen Effekte sind in Abbildung 5.6 dargestellt. Man sieht, dass der Effekt des Alters denselben Verlauf zeigt wie in der Leistungsdiagnostik. Bei jungen Athleten ist der Effekt negativ, steigt dann stark an und bleibt bei älteren Athleten weiter konstant. Der geschätzte Effekt des `relKraftanstieg_bb` auf die `Weite` ist bei mittleren Ausprägungen positiv und bei extremen Ausprägungen negativ. Der Verlauf bei der Variable `DJ32_Faktor` ist genau invers zum relativen Kraftanstieg. Hier ist der Effekt bei extremen Ausprägungen positiv.

Die geschätzten Effekte der regularisierten Effekte sind in Tabelle 5.5 dargestellt. Die 95%-Kreditätsintervalle geben in der Bayesianischen Inferenz den Bereich an, in dem der Parameter mit 95%-iger (Posteriori-)Wahrscheinlichkeit liegt. Für alle Parameter enthält dieses Kreditätsintervall die Null. Man hat also auf dem 5%-Signifikanzniveau keine signifikanten Effekte auf die `Weite`. Betrachtet man die Werte von $P(I_j=1)$, so stellt man fest, dass die meisten Variablen nur mit maximal 3% Posteriori-Wahrscheinlichkeit nicht regularisiert werden. Diese Variablen haben keinen bedeutsamen Effekt auf die Zielgröße `Weite` und müssen daher nicht im Modell berücksichtigt werden. Die Ausnahme ist die Variable `SJ_V75E`, die mit knapp 20% Wahrscheinlichkeit nicht geschrumpft wird und damit deutlich vor den anderen Variablen liegt.

Keine der 15 Variablen wird in der Mehrzahl der Stichproben ausgewählt. Man würde daher sagen, dass anhand dieses Modells keiner der Variablen aus der Leistungsdiagnostik ein bedeutsamer Effekt auf die `Weite` zugeschrieben wird. Am ehesten sticht noch die Variable `SJ_V75E` hervor, da ihr Wert von $P(I_j=1)$ deutlich größer als die anderen Werte ist.

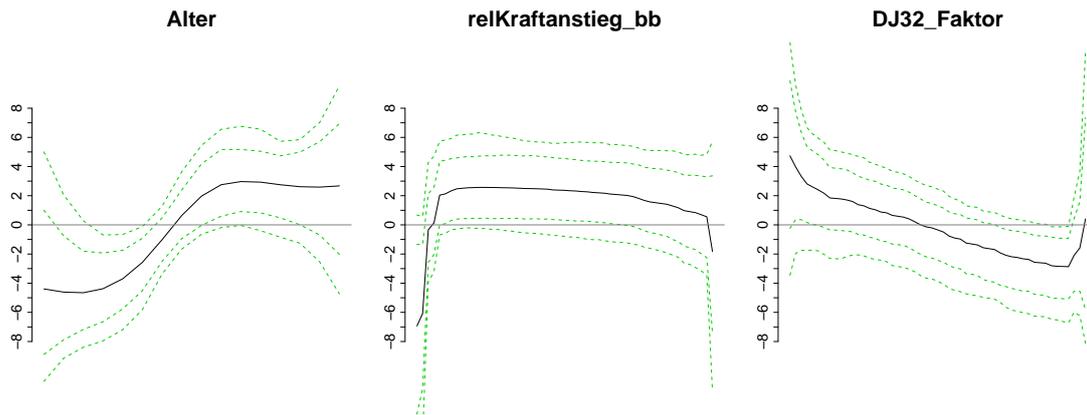


Abbildung 5.6.: Geschätzte nichtparametrische Effekte für die Variablen **Alter**, **relKraftanstieg_bb** und **DJ32_Faktor** bei der Zielgröße **Weite** mit P-Splines. Die mittlere Linie gibt den mittleren Funktionsverlauf wieder, die äußeren Linien kennzeichnen ein punktweises 80% bzw. 95%-Kredibilitätsintervall.

Der Effekt der nicht regularisierten Variable **Sportart** wird als 2.56 (SD=3.13) geschätzt. Dies bestätigt wieder den Vorteil der Skispringer gegenüber den Nordischen Kombinierern. Der globale Intercept wird als 120.2 geschätzt. Der Komplexitätsparameter ω wird in diesem Modell geschätzt als 0.082 (SD=0.076).

Variable	Mittelwert	SD	95%-Kredibilitätsintervall	$P(I_j=1)$
SJ_SpH	-0.03	0.67	[-0.22,0.21]	0.030
SJ_V75E	0.73	1.61	[-0.16,5.1]	0.197
DJ32_SpH	0.01	0.21	[-0.15,0.19]	0.011
DJ32_KT	-0.01	0.12	[-0.19,0.18]	0.008
Fmax90_bb	0.05	0.31	[-0.15,0.35]	0.023
Fmax70_bb	0.01	0.16	[-0.18,0.18]	0.009
F_rel170_bb	0	0.14	[-0.16,0.18]	0.008
Ekz	0.01	0.19	[-0.16,0.19]	0.008
Kdif	0	0.1	[-0.16,0.17]	0.004
A60	0	0.19	[-0.19,0.18]	0.011
A180	0	0.13	[-0.16,0.17]	0.008
ekV	0.06	0.33	[-0.15,1.09]	0.035
kkV	0.01	0.16	[-0.17,0.17]	0.009
dynstat_60	0.03	0.19	[-0.14,0.22]	0.011
dynstat_180	-0.03	0.22	[-0.22,0.15]	0.021

Tabelle 5.5.: Geschätzte regularisierte Effekte im STAR-Modell mit der Zielgröße **Weite**. Dargestellt sind Posteriori-Mittelwert, Posteriori-Standardabweichung und ein 95%-Kredibilitätsintervall. Außerdem ist die Posteriori-Wahrscheinlichkeit dafür angegeben, dass der Effekt nicht geschrumpft wird, d.h. $P(I_j=1)$.

Zusätzlich zur **Weite** wird ein weiteres Modell mit **vertikaleGeschw** als Zielgröße und den gleichen Kovariablen und Einstellungen wie oben gerechnet. Die Effekte der Variablen zur Leistungsdiagnostik sind in Tabelle 5.6 dargestellt. Insgesamt werden im Modell mit der Zielgröße **vertikaleGeschw** wesentlich größere Posteriori-Wahrscheinlichkeiten von $P(I_j=1)$ der Effekte geschätzt. Dabei haben die Variablen **SJ_SpH**, **SJ_V75E**, **DJ32_SpH** und **A180** Werte von größer als 10%. Als besonders bedeutsam sticht hier die Variable **SJ_SpH** mit dem Wert von 0.682 hervor (vgl. Abbildung 5.7).

Variable	Mittelwert	SD	95%-Kreditabilitätsintervall	$P(I_j=1)$
SJ_SpH	0.36	0.28	[-0.01,0.8]	0.682
SJ_V75E	0.11	0.21	[-0.02,0.62]	0.269
DJ32_SpH	0.03	0.12	[-0.01,0.46]	0.103
DJ32_KT	-0.01	0.04	[-0.12,0.02]	0.05
Fmax90_bb	0.01	0.04	[-0.01,0.17]	0.043
Fmax70_bb	0	0.06	[-0.01,0.04]	0.04
F_rel70_bb	0	0.02	[-0.02,0.02]	0.018
Ekz	0.01	0.06	[-0.01,0.1]	0.049
Kdif	0.01	0.04	[-0.02,0.08]	0.031
A60	0	0.09	[-0.04,0.16]	0.057
A180	0.07	0.17	[-0.01,0.55]	0.206
ekV	0.01	0.05	[-0.01,0.18]	0.074
kkV	0	0.03	[-0.01,0.03]	0.031
dynstat_60	0	0.02	[-0.02,0.02]	0.018
dynstat_180	0	0.02	[-0.03,0.01]	0.025

Tabelle 5.6.: Geschätzte regularisierte Effekte im STAR-Modell mit der Zielgröße `vertikaleGeschw`. Dargestellt sind Posteriori-Mittelwert, Posteriori-Standardabweichung und ein 95%-Kreditabilitätsintervall. Außerdem ist die Posteriori-Wahrscheinlichkeit dafür angegeben, dass der Effekt nicht geschrumpft wird, d.h. $P(I_j=1)$.

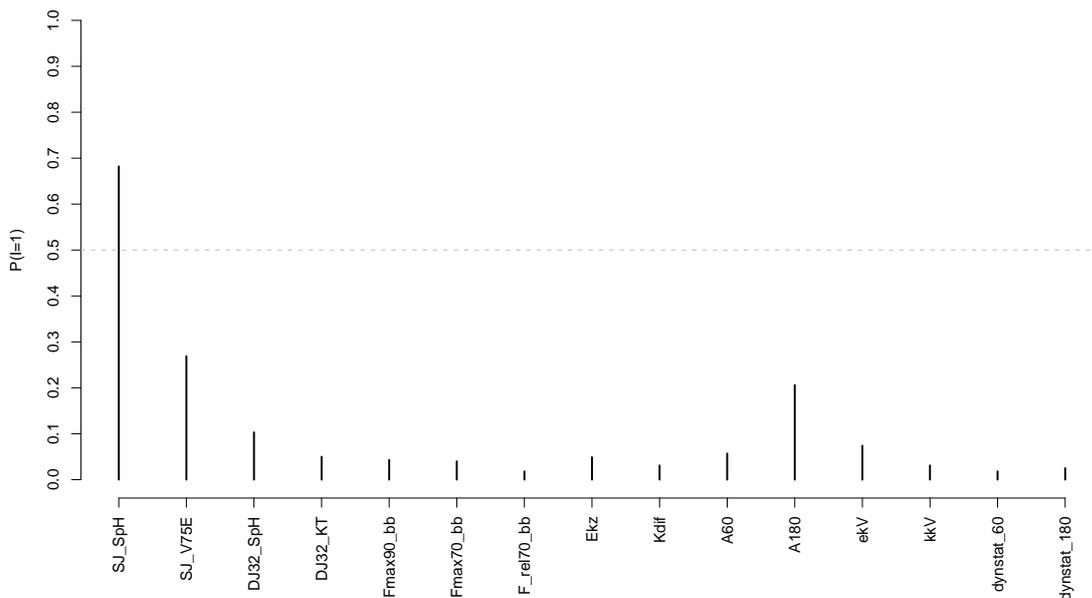


Abbildung 5.7.: Posteriori-Wahrscheinlichkeit $P(I_j = 1)$ der regularisierten Effekte im Modell mit der Zielgröße `vertikaleGeschw`.

Bedeutame Kovariablen für die `vertikaleGeschw` sind also an erster Stelle die Variablen zur Sprungkraftdiagnostik. Insgesamt stehen die Variablen zur Leistungsdiagnostik mehr mit der Variablen `vertikaleGeschw` als `Weite` in Zusammenhang. Bei den nichtparametrischen Effekten wird ein linearer positiver Einfluss von `Alter` und `DJ32_Faktor` geschätzt. Der Verlauf des Effekts von `relKraftanstieg_bb` ist wie im obigen Modell. Die geschätzten Kurven sind in Abbildung 5.8 dargestellt. Der Komplexitätsparameter ω wird in diesem Modell als 0.158 (SD=0.101) geschätzt. Hier wird also in etwa doppelt so vielen Fällen wie bei der `Weite` ein Regressionskoeffizient aus der Komponente der Mischverteilung mit ν_1 gezogen.

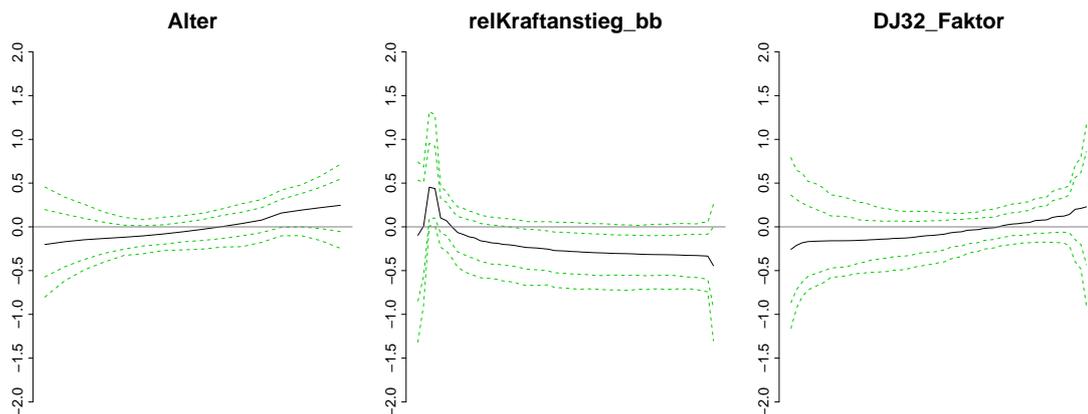


Abbildung 5.8.: Geschätzte nichtparametrische Effekte für die Variablen `Alter`, `relKraftanstieg_bb` und `DJ32_Faktor` bei der Zielgröße `vertikaleGeschw` mit P-Splines. Die mittlere Linie gibt den mittleren Funktionsverlauf wieder, die äußeren Linien kennzeichnen ein punktweises 80% bzw. 95%-Kredibilitätsintervall.

Zusätzlich zu den STAR-Modellen mit der NMIG-Priori wurden die beiden Modelle auch mit Ridge-Prioris für die Kovariableneffekte gerechnet. Die Ergebnisse sind im Anhang C ab Seite 100 dargestellt. Die Beurteilung der Ergebnisse der regularisierten Effekte kann hier direkt anhand der Effektschätzer oder der Lage der Null im 95%-Kredibilitätsintervall bzw. im Intervall von \pm einer Standardabweichung erfolgen. Im ersten Modell mit der Zielgröße `Weite` umfassen sowohl die 95%-Kredibilitätsintervalle als auch die Intervalle mit \pm einer Standardabweichung um den Mittelwert die Null. Allerdings sticht wieder die Variable `SJ_v75E` aufgrund des betragsmäßig größten Effekts hervor. Auch im Modell mit der Zielgröße `vertikaleGeschw` umfassen die 95%-Kredibilitätsintervalle aller Kovariablen die Null. Auch hier sticht die Variable `SJ_v75E` hervor, da das zugehörige Intervall mit \pm einer Standardabweichung nicht die Null enthält. Die geschätzten nichtparametrischen Effekte entsprechen in etwa denen der Modelle mit NMIG-Priori. Mit der Verwendung einer Ridge-Priori werden im Wesentlichen die Ergebnisse der STAR-Modelle mit NMIG-Priori bestätigt.

Modelldiagnose

Die Ergebnisse in der Bayes-Inferenz sind nur valide, wenn die simulierten Markov-Ketten der Parameter eine stationäre Verteilung erreicht haben. Dazu wurden die Samplingpfade aller Parameter des Modells betrachtet, wobei keine Auffälligkeiten festgestellt werden konnten. Als wichtiges Beispiel ist der Samplingpfad der Devianz (nach Burn-In und Ausdünnung) in Abbildung 5.9 dargestellt. Die Devianz ist ein Maß für die Anpassung des Modells an die Daten (Fahrmeir u. a., 2007). Man sieht, dass keine Struktur in dem Samplingpfad zu erkennen ist und die Annahme der stationären Verteilung damit nicht verworfen werden kann. Die Samplingpfade weiterer Parameter für das Modell mit der Zielgröße `Weite` sind in Abbildung D.2 im Anhang D auf Seite 103 dargestellt. Samplingpfade für das Modell mit der Zielgröße `vertikaleGeschw` finden sich in Abbildung D.3 auf Seite 104.

Des Weiteren werden die Residuen des Modells betrachtet. Das Residuum für Beobachtung i ergibt sich als

$$r_i = y_i - \hat{y}_i,$$

wobei y_i die tatsächliche und \hat{y}_i die anhand des Modells vorhergesagte Sprungweite ist. In Abbildung 5.10 sind die Residuen der 75 Beobachtungen dargestellt. Die meisten Residuen liegen zwischen -5

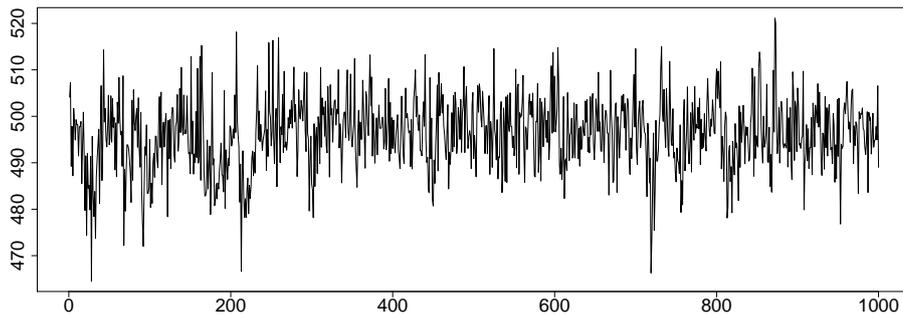


Abbildung 5.9.: Samplingpfad der Devianz im STAR-Modell nach Burn-In-Phase und Ausdünnung.

und 5 Metern. Der maximale Abstand zwischen tatsächlichem und vorhergesagtem Wert liegt zwischen -16 und 14.2 Metern. Insgesamt kann keine Struktur in den Residuen festgestellt werden, was die Gültigkeit des Modells bestätigt.

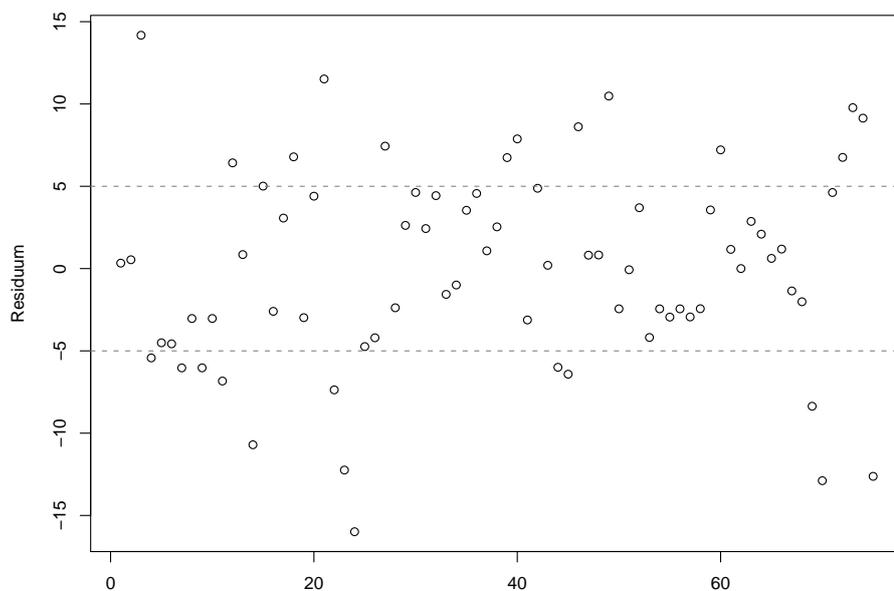


Abbildung 5.10.: Geschätzte Residuen des Modells.

5.4. Diskussion

In diesem Abschnitt werden verschiedene Ergebnisse dieses Kapitels aus methodischer und sportwissenschaftlicher Sicht weiter diskutiert:

- Anhand der beiden vorliegenden Datenquellen wurde ein neuer gemeinsamer Datensatz generiert, bei dem jeweils für einen Athleten die Messungen zur Schanzen- und Leistungsdiagnostik zugeordnet werden, die weniger als zwei Monate voneinander entfernt liegen. Die Wahl dieses Zeitraums hat pragmatische Gründe, da bei kleineren Intervallen nicht mehr genügend Beobachtungen zur Schätzung von Modellen verbleiben würden. Insgesamt verbleiben hier 75 Beob-

achtungen für eine Complete-Case-Analyse.

- Die Faktorenwerte aus Kapitel 3 bieten einen schnellen Überblick über das Leistungsniveau eines Athleten. Allerdings stehen sie nicht in einem engen Zusammenhang mit den Variablen der Schanzendiagnostik. Es tritt ein Effizienzverlust auf, da die Effekte zweimal geschätzt werden (die Faktorenwerte werden im Anschluss an die Faktorenanalyse mit der Bartlett-Methode geschätzt und der Effekt der Faktorenwerte wird im FARA-Modell geschätzt). Nach jeder Schätzung wird der Effekt mit einem neuen Messfehler versehen. Daher ist es sinnvoller, bei der gemeinsamen Betrachtung auf die ursprünglichen Variablen der Leistungsdiagnostik zurückzugehen, wie es bei den Latente-Variablen-Modellen und den strukturiert additiven Regressionsmodellen gemacht wird.
- Latente-Variablen-Modelle und strukturierte additive Regressionsmodelle sind beides geeignete Methoden zur Modellierung der Zusammenhänge zwischen den Variablen der Leistungsdiagnostik und Schanzendiagnostik. Welche Modellklasse gewählt wird, sollte anhand der Problemstellung entschieden werden. Latente-Variablen-Modelle eignen sich besonders, wenn der Effekt von wenigen Kovariablen auf mehrere, untereinander korrelierte Zielgrößen untersucht werden soll. STAR-Modelle untersuchen den Einfluss mehrerer Kovariablen auf eine Zielgröße und bieten mit Bayesianischer Regularisierung die Möglichkeit zur Modellwahl.
- Bei der Analyse der Zusammenhänge von Schanzendiagnostik und Leistungsdiagnostik zeigte sich die Variable `SJ_V75E` als besonders bedeutsam für die `Weite`. Der Zusammenhang der Sprungweite von der Schanze und `SJ_V75E` lässt sich auch inhaltlich erklären. So ist es für den Athleten zweckmäßig, während des Anlaufs möglichst lange in der Kauerstellung zu verweilen, da diese den geringsten Luftwiderstand bietet, und sich dann kurz vor dem Absprung möglichst schnell aufzurichten. Hierbei kommt dem Athleten eine schnelle Absprunggeschwindigkeit zu Gute. Aus sportwissenschaftlicher Sicht sind die Kraftfähigkeiten eines Athleten eine notwendige Bedingung für eine gute Sprungweite, während eine schnelle Absprunggeschwindigkeit eine hinreichende Bedingung ist.
- Die `vertikaleGeschw` steht besonders in einem Zusammenhang mit den Variablen `SJ_SpH` und `SJ_V75E`. Beide Variablen zeigen einen positiven Effekt auf die `vertikaleGeschw`. Dieser Zusammenhang ist wesentlich deutlicher als mit den Variablen zur Maximalkraft- und Schnellkraftdiagnostik.
- Bei allen Modellen hat sich gezeigt, dass der Zusammenhang der Variablen der Leistungsdiagnostik mit der Variable `vertikaleGeschw` stärker als der Zusammenhang mit der `Weite` ist. Gute Ergebnisse in der Leistungsdiagnostik spiegeln sich daher deutlicher in der vertikalen Absprunggeschwindigkeit von der Schanze als in der Sprungweite wieder. Auf die Sprungweite wirken also noch wesentlich mehr Faktoren als im Rahmen der Leistungsdiagnostik erhoben werden können. Die bestehenden Unterschiede lassen sich wohl zu einem gewissen Teil auf die Tagesform des Athleten oder die äußeren Bedingungen wie den Einfluss des Windes zurückführen.

6. Diskussion und Ableitungen für die weitere Forschungspraxis

Im Rahmen dieser Arbeit wurde eine Reihe von statistischen Methoden auf die Daten von Skispringern und Nordischen Kombinerern angewandt und Ergebnisse gewonnen. Aus sportwissenschaftlicher Sicht ist interessant, welche Schlüsse für die weitere Forschungspraxis gezogen werden können. Dieses Kapitel unterscheidet zwischen dem Bereich zur Leistungsdiagnostik mit Hinweisen zur zukünftigen Gestaltung einer Leistungsdiagnostik für Skispringer und Nordische Kombiniere und dem Bereich der Verknüpfung der Leistungsdiagnostik mit der Schanzendiagnostik und Hinweisen zur Wahl geeigneter Modellklassen und der Interpretation der Ergebnisse.

6.1. Leistungsdiagnostik

In deskriptiven Untersuchungen hat sich gezeigt, dass die Daten zur Leistungsdiagnostik in einem Zusammenhang mit dem Alter des Athleten stehen. Dies liegt sowohl an der körperlichen Entwicklung, als auch an der Wahl der Trainingsschwerpunkte. Sollen Athleten mit statistischen Verfahren verglichen werden, so muss eine Adjustierung auf das Alter erfolgen, damit sich keine Verzerrung der Ergebnisse z.B. durch einen unterschiedlichen Altersdurchschnitt bei Gruppenvergleichen ergibt. Hier wurde jede Beobachtung mit Mittelwert und Standardabweichung einer Referenzgruppe Z-standardisiert. Die Referenzgruppe bildeten dabei jeweils die Athleten, die zum Zeitpunkt der Diagnostik maximal ein Jahr jünger bzw. älter sind. Neben der hier vorgenommenen Standardisierung sind auch größere Referenzgruppen denkbar, z.B. alle unter 15-Jährigen, die 16 bis 17-Jährigen, die 18 bis 19-Jährigen und die über 20-Jährigen. Eine solche größere Einteilung ist dann sinnvoll, wenn insgesamt weniger Daten vorliegen oder einzelne Altersjahre nur schwach oder gar nicht vertreten sind.

Mit dem Faktorenmodell aus Kapitel 3 können Ableitungen für eine zukünftige Zusammensetzung der Leistungsdiagnostik gemacht werden. Da insbesondere die einbeinigen Messungen stark mit den beidbeinigen korrelieren und die Variablen aus den Sprungkraft- und Maximalkrafttests in engem Zusammenhang stehen, wäre aus statistischer Sicht beispielsweise die folgende Auswahl an Variablen aus der Leistungsdiagnostik ausreichend, um die wesentlichen Unterschiede zwischen den Athleten zu erfassen:

DJ32_SpH	Ekz
DJ32_KT	ekV
F_rel70_bb	kkV
Fmax70_bb	dynstat_60
relKraftanstieg_bb	dynstat_180

In dieser Leistungsdiagnostik enthält jeder der drei latenten Faktoren aus Kapitel 3 mindestens zwei Variablen, die den Faktor messen. Die anderen würden wegfallen, da sie stark mit diesen korreliert sind und nur wenig zusätzliche Information bieten. Außerdem sind alle Variablen enthalten, die nicht durch einen der Faktoren erfasst werden. Diese Diagnostik würde in der Praxis eine Zeitersparnis bedeuten, da der Squat Jump entfallen würde, sowie die Maximalkraftmessung bei 90° Kniewinkel und die beiden Tests zur Schnellkraft A60 und A180. Außerdem würden noch die beiden einbeinigen

Messungen beim Kraftanstieg und bei der Maximalkraft bei 70° Kniewinkel entfallen. Das Resultat wäre eine Leistungsdiagnostik mit nur 10 statt 23 Variablen, für die der Athlet weniger Zeit benötigt, die aber weiterhin einen großen Teil der Varianz zwischen den Beobachtungen erfasst.

Die Faktorenmodelle eignen sich auch, um aus statistischer Sicht zu beurteilen, ob neue Tests im Rahmen der Leistungsdiagnostik sinnvoll sind. Um zu überprüfen, ob ein neuer Test einen neuen Aspekt in dem Leistungsniveau eines Athleten misst, oder aber nur ähnliche Ergebnisse wie bereits vorhandene Verfahren liefert, muss eine größere Anzahl an Beobachtungen des neuen Tests parallel mit den bestehenden Tests durchgeführt werden. Für die resultierenden Daten wird dann eine Faktorenanalyse gerechnet und die Ladungsmatrix bestimmt. Wird die Variable, die anhand des neuen Tests bestimmt wurde, einem der bestehenden Faktoren zugeordnet, so kann man sagen, dass der neue Test dasselbe Konstrukt misst und kaum neue Information liefert. Steht er hingegen nicht in einem Zusammenhang mit den bestehenden latenten Faktoren, so misst er eine neue Dimension im Leistungsprofil eines Skispingers oder Nordischen Kombinierers. Denkbar ist so ein Vorgehen für die inzwischen neu eingeführten Sprungkrafttests mit Zusatzgewicht.

6.2. Verknüpfung der Leistungsdiagnostik mit der Schanzendiagnostik

Im Rahmen der Analysen in dieser Masterarbeit wurden Messungen der Leistungsdiagnostik mit denen der Schanzendiagnostik zusammengeführt, die weniger als zwei Monate voneinander entfernt sind. Eine Beschränkung der Analyse auf einen noch kleineren Abstand ist bei der bisherigen Datenlage nicht zu verwirklichen, da ansonsten der Stichprobenumfang zu gering wird. Möglicherweise erscheint dieser Abstand bereits zu groß, um inhaltliche Schlüsse zu ziehen. So kann sich der Athlet beispielsweise in der Zwischenzeit eine Verletzung zugezogen haben. Für die statistische Modellierung der Zusammenhänge erscheint eine gemeinsame Koordination von Leistungsdiagnostik und Schanzendiagnostik sinnvoll. Würde jeder Athlet zeitnah zu einer Schanzendiagnostik auch eine Leistungsdiagnostik absolvieren, so könnte schnell ein gemeinsamer Datensatz ohne das Problem der zeitlichen Differenz generiert werden.

In Kapitel 5 hat sich gezeigt, dass sich sowohl Latente-Variablen-Modelle als auch strukturierte additive Regressionsmodelle zur gemeinsamen Analyse der Zusammenhänge beider Datenquellen eignen. Die hier vorgestellten Latente-Variablen-Modelle ermöglichen es, den Effekt einer Variable aus der Leistungsdiagnostik simultan auf die beiden Variablen `Weite` und `vertikaleGeschw` der Schanzendiagnostik zu betrachten. Dabei sind auch Erweiterungen des Modells denkbar. So könnten simultan weitere Variablen der Schanzendiagnostik wie die `Genauigkeit` als Zielgrößen aufgenommen werden. Bei einem größeren Stichprobenumfang ließen sich auch mehrere indirekte Kovariablen in das Modell aufnehmen und nichtparametrische Effekte betrachten.

Mit strukturierten additiven Regressionsmodellen kann der simultane Einfluss mehrerer Kovariablen auf eine Zielgröße untersucht werden. Die Bayesianische Regularisierung bietet die Möglichkeit relevante Effekte von nicht relevanten zu trennen und so Variablenselektion durchzuführen. Die Indikatorwahrscheinlichkeiten bei der Regularisierung mit der NMIG-Priori bieten dabei einen intuitiven Ansatz zur Variablenselektion. Außerdem ist die Möglichkeit der Modellierung von nichtlinearen Effekten gegeben. Bislang ist es mit der vorhandenen Software noch nicht möglich, gleichzeitig lineare und nichtlineare Effekte zu regularisieren. Wenn dies möglich ist, kann simultan eine Modellwahl von nichtlinearen und linearen Effekten durchgeführt werden.

Die insgesamt geringe Zahl an identifizierten Effekten von Variablen der Leistungsdiagnostik lässt sich auch darauf zurückführen, dass ein Großteil der Messwerte nur in einem kleinen Wertebereich liegt

und deshalb die Streuung zwischen den Athleten gering ist. Alle Athleten haben also ein gewisses Kraft- und Sprungkraftniveau, das notwendig für gute Sprungweiten bei Schanzensprüngen ist. Eine bessere Abschätzung des Effekts der Variablen aus der Leistungsdiagnostik ließe sich gewinnen, wenn man eine Kontrollgruppe von Personen hätte, die kein sportartspezifisches Training absolvieren. In diesem Fall hat man stärkere Schwankungen bei den Ergebnissen zwischen den Kader-Athleten und der Kontrollgruppe und könnte damit den Effekt des sportartspezifischen Trainings abschätzen.

Zur gemeinsamen Auswertung der Daten zur Schanzen- und Leistungsdiagnostik wären auch Verfahren der *funktionalen Datenanalyse* denkbar. Eine Einführung in das Konzept der funktionalen Datenanalyse findet sich in Ramsay u. Silverman (2005). Bei der funktionalen Datenanalyse werden die zeitlichen Verläufe der Variablenausprägungen bei einem Athleten als funktional angesehen. Anhand der vorliegenden Beobachtungen wird beispielsweise mit P-Splines die unterliegende Funktion geschätzt. In Yao u. a. (2005) wird ein funktionaler Regressionsansatz für Longitudinaldaten vorgeschlagen, der sich insbesondere zur Analyse von Fragestellungen eignet, bei denen Zielgrößen und Kovariablen zu unterschiedlichen Zeitpunkten erhoben werden. Mit Hilfe der Hauptkomponentenanalyse werden die Regressionskoeffizienten und das Bestimmtheitsmaß in Abhängigkeit von der Zeit geschätzt. Damit könnten für jeden Athleten Kurven der Zielgröße vorhergesagt werden und damit beispielsweise Vorhersagen für die Wettkampfsaison getroffen werden. Zudem könnten auch saisonale Effekte berücksichtigt werden. Allerdings liegen in diesem Datensatz für diese Methodik zu wenig Beobachtungen je Athlet und Jahr zur Schätzung valider Kurven vor. Dafür wäre es nötig, für jeden Athleten mindestens fünf bis zehn Beobachtungen je Jahr zur Leistungsdiagnostik und Schanzendiagnostik über mehrere Jahre hinweg zu sammeln. Außerdem könnten in den Datensatz die Ergebnisse des Athleten bei Wettkämpfen aufgenommen werden. Damit kann man versuchen, anhand der Ergebnisse der Leistungsdiagnostik und Schanzendiagnostik im Sommer Vorhersagen für die Wettkampfsaison zu treffen.

7. Zusammenfassung

In dieser Masterarbeit wurden umfangreiche Daten zu Training und Leistung von Skispringern und Nordischen Kombinierern analysiert. Es liegen für beide Sportarten Längsschnittdaten zu Leistungs- und Schanzendiagnostik über mehrere Jahre von knapp 200 Athleten vor. Die beiden Datenquellen wurden sowohl getrennt, als auch gemeinsam analysiert.

In der Leistungsdiagnostik wurden eine Reihe von Messungen zur Sprung-, Maximal- und Schnellkraft mit den Athleten durchgeführt. Bei den resultierenden Variablen kann keine Unterscheidung zwischen abhängigen und unabhängigen Variablen getroffen werden. Mit einem Faktorenmodell wurden latente Faktoren konstruiert, die sich in den beobachteten Variablen manifestieren und die Korrelationsstruktur der Daten erfassen. Es wurde festgelegt, dass drei Faktoren zu schätzen sind. Anhand der Ladungsmatrix wurde der erste Faktor als „Allgemeine Sprungkraftfähigkeit und relative Maximalkraft“, der zweite als „Allgemeine Maximalkraft und Schnellkraft“ und der dritte als „Kraftanstiegsverhalten“ bezeichnet. Nur beim ersten Faktor haben Skispringer bessere Werte als Nordische Kombinierer. Mit Bootstrap-Verfahren konnte die Zusammensetzung der Faktoren bestätigt werden. Die geschätzten Faktorenwerte liefern eine schnelle Übersicht über das Leistungsniveau eines Athleten, wobei jeweils gilt, dass hohe Faktorenwerte wünschenswert sind. Anhand der Faktoren wurde ein Vorschlag für eine reduzierte Leistungsdiagnostik mit weniger Übungen erarbeitet.

Als zweite Datenquelle wurden im Rahmen dieser Arbeit diverse automatisch erstellte Sprungprotokolle von Übungssprüngen von der Schattenbergschanze (Hillsize 137) in Oberstdorf ausgewertet. Dazu wurden die verschiedenen Durchgänge während eines Lehrgangs zu einer Beobachtung mit der maximalen Sprungweite zusammengefasst. Es wurden zwei für die weitere Analyse wichtige Größen der Schanzendiagnostik festgelegt: Die Sprungweite und die vertikale Absprunggeschwindigkeit. Bei beiden Variablen konnten die Skispringer im Mittel bessere Ergebnisse als die Nordischen Kombinierer erzielen.

Im letzten Schritt wurden die beiden Datenquellen zu einem gemeinsamen Datensatz zusammengeführt. Dabei wurden einander nur Schanzen- und Leistungsdiagnostikdaten zugeordnet, die der Athlet innerhalb eines Zeitraums von maximal zwei Monaten durchgeführt hat. Zur Analyse der Zusammenhänge zwischen den Daten wurde einmal eine Regression der Faktorenwerte auf die Variablen zur Schanzendiagnostik durchgeführt. Dabei konnten keine signifikanten Effekte festgestellt werden. Außerdem wurden Latente-Variablen-Modelle mit einem latenten Faktor und den zwei Zielgrößen Weite und vertikale Absprunggeschwindigkeit aufgestellt. Dabei wurde jeweils eine Variable der Leistungsdiagnostik als indirekte Kovariable in das Modell aufgenommen. In jedem der drei präsentierten Modelle konnte ein positiver Effekt der indirekten Kovariable festgestellt werden. Mit strukturiert additiven Regressionsmodellen und Bayesianischer Regularisierung wurde die Absprunggeschwindigkeit beim Squat Jump als bedeutsamste Kovariable sowohl für die Sprungweite von der Schanze als auch für die vertikale Absprunggeschwindigkeit vom Schanzentisch identifiziert. Insgesamt stehen die Variablen der Leistungsdiagnostik mit der vertikalen Absprunggeschwindigkeit in einem engeren Zusammenhang als mit der Sprungweite.

Anhang

A. Abkürzungsverzeichnis

- # Anzahl
- bb** beidbeinig
- CMJ** Counter-Movement Jump
- det** Determinante
- DJ** Drop Jump
- DVZ** Dehnungs-Verkürzungs-Zyklus
- I_p Indikatormatrix der Dimension p
- J** Joule (Maßeinheit der Arbeit $1\text{J} = 1 \frac{\text{kg} \cdot \text{m}^2}{\text{s}^2}$)
- K/K-Punkt** Konstruktionspunkt
- L** Likelihoodfunktion
- li** links bzw. linkes Bein
- msec** Millisekunde
- ML** Maximum-Likelihood
- N** Newton (Maßeinheit der Kraft, $1\text{N} = 1 \frac{\text{kg} \cdot \text{m}}{\text{s}^2}$)
- NA** *not available*, Bezeichnung für einen fehlenden Wert
- NK** Nordische Kombination bzw. Nordischer Kombinierer
- re** rechts bzw. rechtes Bein
- SJ** Squat Jump
- Sp, Spr** Skisprung bzw. Skispringer
- TL** Trainingslehrgang

B. Hinweise zur statistischen Software

In diesem Kapitel des Anhangs folgen Hinweise zu Funktionen, die im Rahmen dieser Arbeit eingesetzt wurden. Dabei wird zwischen bereits implementierten und eigenen Funktionen unterschieden.

B.1. Vorliegende Funktionen aus dem Basis-R-Paket oder Erweiterungen

Diese Funktionen sind bereits im statistischen Programmpaket R (R Development Core Team, 2009) implementiert und dokumentiert. Hier folgt lediglich eine kurze Erläuterung für die praktische Anwendung der Funktionen und weitere Details sind in der entsprechenden Online-Hilfe (Befehl: `help()`) zu den Funktionen zu finden. In geschweiften Klammern ist dabei das Paket, in dem die Funktion enthalten ist, angegeben.

factanal{stats} Mit Hilfe dieser Funktion kann eine explorative Faktorenanalyse durchgeführt werden. Neben einem Datensatz oder einer Kovarianzmatrix muss dabei die Anzahl der zu schätzenden Faktoren übergeben werden. Die Ladungsmatrix wird nach der Maximum-Likelihood-Methode geschätzt und das Ergebnis kann mit `rotation="varimax"` nach der Varimax-Methode und mit `rotation="promax"` schiefwinklig rotiert werden. Standardmäßig werden keine Faktorenwerte für die Beobachtungen geschätzt. Mit der Einstellung `scores="Bartlett"` werden die Faktorwerte nach der Maximum-Likelihood-Methode geschätzt, mit `scores="regression"` nach der Regressionsmethode. Teilweise tritt die Fehlermeldung auf, dass Optimierung bei den vorgegebenen Startwerten fehlschlägt. In diesem Fall müssen mit der Einstellung `control = list(nstart = x)` neue Startwerte probiert werden.

lmer{lme4} Mit dieser Funktion lassen sich lineare gemischte Modelle in R rechnen. Als Argument `formula` wird die Modellformel übergeben. Diese hat die Form

$$y \sim x + (u | \text{Gruppierungsvariable}),$$

wobei y die Zielgröße und x die festen und u die individuenspezifischen Kovariablen sind. Soll nur ein individuenspezifischer Intercept modelliert werden, so hat der Ausdruck für die individuenspezifischen Effekte lediglich die Form `(1|Gruppierungsvariable)`. Mit dem Argument `data` wird der Name des Datensatzes angegeben. Die Schätzung der Effekte kann mit der Maximum-Likelihood oder der restringierten Maximum-Likelihood-Methode erfolgen. Die Art der Schätzung kann mit dem Argument `REML` gesteuert werden.

MCMCm{MCMCpack} Diese Funktion wurde von Raach (2005) entwickelt und von Steinert (2006) erweitert und ist in einer speziell angepassten Version von `MCMCpack`, Version 0.6-6 enthalten. Eine genaue Dokumentation findet sich in Seiler (2008). Mit dieser Funktion lassen sich Latente-Variablen-Modelle schätzen. Mit dem Argument `ind.form` werden die Zielgrößen übergeben. Nach jedem Variablennamen wird in Klammern die Art der Variablen angegeben. Dabei steht `o` für eine ordinale, `p` für eine Poissonverteilte und `c` für eine stetige Variable. Mit dem Argument `covar.form` werden die

indirekten Kovariablen spezifiziert. Nach dem Variablennamen kann mit dem Ausdruck `(p,i,d)` angegeben werden, dass der Effekt mit P-Splines von Grad `d` und `i` Intervallen nichtparametrisch geschätzt wird. Mit dem Argument `fixed.form` werden die direkten Kovariablen an die Funktion übergeben. Mit `factors` wird die Anzahl der latenten Faktoren festgelegt. Bei den MCMC-Optionen wird mit dem Argument `burnin` die Länge der Burn-In-Phase, mit `mcmc` die Anzahl der MCMC-Stichproben nach der Burn-In-Phase und mit `thin` das Ausdünnungsintervall angegeben.

B.2. Eigene Funktionen

In diesem Abschnitt folgt eine Beschreibung der R-Funktionen, die für die Berechnungen in dieser Arbeit geschrieben wurden. Die Dokumentationen wurden in dem R Dokumentationsformat `Rd` geschrieben und in das `LATEX`-Format konvertiert. Da die Daten zur Leistungsdiagnostik und Schanze dem Datenschutz unterliegen, werden die Funktionen im Abschnitt „Examples“ auf Datensätze angewendet, die mit `R` mitgeliefert werden.

`profilPlot`

Funktion zum Erstellen eines Profilplots

Description

Diese Funktion zeichnet einen Profilplot für je zwei Beobachtungen eines Skispringers und eines Nordischen Kombinierers in eine Abbildung.

Usage

```
profilPlot(data, colsNK, colsSkSp, lpar = 8, ...)
```

Arguments

<code>data</code>	(standardisierter) Datensatz mit allen Variablen, für die ein Profilbild erstellt werden soll
<code>colsNK</code>	Vektor der Länge 2, der die beiden Zeilen des Datensatzes mit den Beobachtungen des Nordischen Kombinierers angibt
<code>colsSkSp</code>	Vektor der Länge 2, der die beiden Zeilen des Datensatzes mit den Beobachtungen des Skispringers angibt
<code>lpar</code>	Breite des linken Grafikrandes, bei langen Variablennamen größer setzen
<code>...</code>	weitere Grafikparameter

Details

`colsNK` und `colsSkSp` dürfen auch `NA`-Werte enthalten. In diesem Fall werden die Beobachtungen nicht gezeichnet.

Value

Eine beschriftete Grafik mit dem Profil eines Nordischen Kombinierers links und eines Skispringers rechts.

Author(s)

Valentin Wimmer

bootFA

Bootstrap für Faktorenanalyse

Description

Bootstrap-Schätzung für Faktorenmodelle mit der Funktion `factanal`. Es werden Bootstrap-Replikationen für Eigenwerte der Kovarianzmatrix und Ladungskoeffizienten erstellt und die Standardfehler der Ladungskoeffizienten geschätzt.

Usage

```
bootFA(data, B = 100, factors = 3, rotation = "varimax", ...)
```

Arguments

<code>data</code>	n x p Datensatz, für den die Faktorenanalyse gerechnet werden soll
<code>B</code>	Anzahl der Bootstrap-Replikationen
<code>factors</code>	Anzahl der zu schätzenden Faktoren
<code>rotation</code>	Rotation der Ladungsmatrix
<code>...</code>	Weitere Argumente, die <code>factanal</code> übergeben werden

Details

Es wird nichtparametrischer Bootstrap gerechnet. Alle `B/10` Schritte erfolgt eine Ausgabe des Iterationsfortschrittes.

Value

Eine Liste mit folgenden Elementen

<code>ret</code>	Matrix, in der die geschätzten Ladungsmatrizen der <code>B</code> Iterationsschritte aneinander gesetzt sind. Die Spalte <code>var</code> gibt die Variable und die Spalte <code>B</code> Nummer der Bootstrapreplikation an.
<code>sdLoad</code>	Matrix die die Bootstrap-Schätzung der Standardfehler der Ladungskoeffizienten enthält.
<code>eV</code>	Matrix mit <code>B</code> Zeilen und <code>p</code> Spalten, die die Eigenwerte der jeweiligen Bootstrap-Replikation enthalten.

B Verwendete Zahl der Bootstrap-Replikationen.

Author(s)

Valentin Wimmer

See Also

plotKILoadings

Examples

```
data(swiss)
(FA.swiss <- factanal(swiss,3))
boot.swiss <- bootFA(swiss,2000,factors=3)
# anzeigen der geschätzten Standardabweichungen für die Ladungskoeffizienten des ersten Faktors
with(boot.swiss,tapply(ret[,1],list(var=ret[,4]),sd))
# Darstellung der Eigenwerte
boxplot(as.data.frame(boot.swiss$eV),names=NULL)
```

plotKILoadings

Funktion zum Zeichnen von Bootstrap-Konfidenzintervallen

Description

Mit dieser Funktion können Bootstrap-Konfidenzintervalle für die Ladungskoeffizienten einer Faktorenanalyse berechnet und dargestellt werden.

Usage

```
plotKILoadings(FA, FA.boot, type = c("Percentil","Bootstrap-t","BCa"),
alpha = 0.05, data,...)
```

Arguments

FA	Objekt der Klasse <code>factanal</code>
FA.boot	Rückgabeobjekt eines Aufrufs von <code>boot.FA</code>
type	String, der angibt, welche Art von Konfidenzintervall gezeichnet werden soll, möglich sind <code>Percentil</code> , <code>Bootstrap-t</code> und <code>BCa</code>
alpha	Signifikanzniveau
data	Datensatz, für den auch <code>FA</code> und <code>FA.boot</code> gerechnet wurden
...	weitere Grafikparameter

Details

Die Berechnung von Konfidenzintervallen der Art BCa erfordert eine zusätzliche Jackknife-Schätzung. Die Berechnung ist deswegen sehr zeitintensiv. Es erfolgt eine Angabe des Iterationsfortschrittes.

Value

Eine Abbildung mit Konfidenzintervallen für die Ladungskoeffizienten jeder Variable und jeden Faktor.

Author(s)

Valentin Wimmer

See Also

bootFA

Examples

```
data(swiss)
FA.swiss <- factanal(swiss,3)
boot.swiss <- bootFA(swiss,2000,factors=3,rotation="varimax")

plotKILoadings(FA.swiss,boot.swiss,"Percentil",0.05,swiss)
```

C. Ergebnisse der Regressionsmodelle mit Ridge-Regularisierung

	Variable	Mittelwert	Standardabweichung	95%-Kredibilitätsintervall
1	SJ_SpH	0.17	0.82	[-1.52,1.82]
2	SJ_V75E	0.88	0.92	[-0.57,3.03]
3	DJ32_SpH	0.05	0.84	[-1.64,1.70]
4	DJ32_KT	-0.17	0.70	[-1.60,1.25]
5	Fmax90_bb	0.36	0.78	[-1.09,2.14]
6	Fmax70_bb	0.05	0.81	[-1.54,1.69]
7	F_re170_bb	-0.09	0.75	[-1.76,1.31]
8	Ekz	0.14	0.82	[-1.57,1.74]
9	Kdif	-0.04	0.72	[-1.57,1.52]
10	A60	-0.05	0.84	[-1.91,1.54]
11	A180	0.18	0.78	[-1.43,1.74]
12	ekV	0.34	0.68	[-1.06,1.66]
13	kkV	0.47	0.67	[-0.78,2.00]
14	dynstat_60	0.43	0.75	[-0.91,2.14]
15	dynstat_180	-0.45	0.68	[-1.96,0.72]

Tabelle C.1.: Geschätzte regularisierte Effekte mit einer Ridge-Priori im Modell mit der Zielgröße *Weite* (vgl. Abschnitt 5.3.3 ab Seite 77).

	Variable	Mittelwert	Standardabweichung	95%-Kredibilitätsintervall
1	SJ_SpH	0.11	0.11	[-0.10,0.33]
2	SJ_V75E	0.13	0.10	[-0.06,0.36]
3	DJ32_SpH	0.05	0.12	[-0.17,0.29]
4	DJ32_KT	-0.08	0.08	[-0.25,0.08]
5	Fmax90_bb	0.02	0.10	[-0.18,0.22]
6	Fmax70_bb	0.04	0.11	[-0.18,0.27]
7	F_re170_bb	-0.03	0.09	[-0.21,0.14]
8	Ekz	0.05	0.11	[-0.17,0.28]
9	Kdif	0.04	0.08	[-0.12,0.21]
10	A60	0.06	0.11	[-0.14,0.29]
11	A180	0.10	0.11	[-0.11,0.33]
12	ekV	0.07	0.09	[-0.09,0.26]
13	kkV	0.06	0.08	[-0.08,0.23]
14	dynstat_60	-0.01	0.09	[-0.19,0.16]
15	dynstat_180	-0.04	0.08	[-0.21,0.10]

Tabelle C.2.: Geschätzte regularisierte Effekte mit einer Ridge-Priori im Modell mit der Zielgröße *vertikaleGeschw* (vgl. Abschnitt 5.3.3 ab Seite 77).

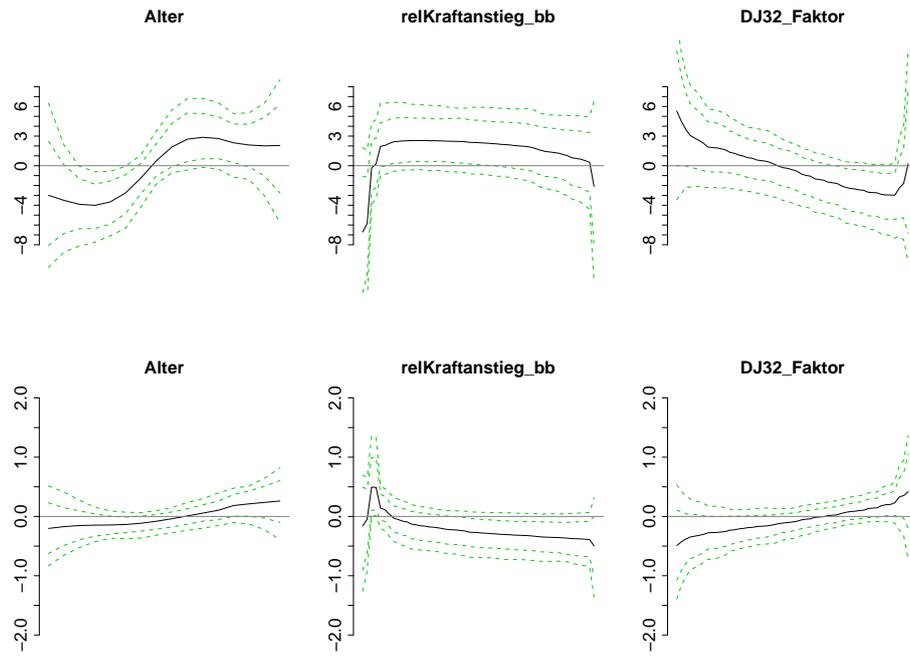


Abbildung C.1.: Geschätzte Kurven der nichtlinearen Effekte im Modell der strukturierten additiven Regression (vgl. Abschnitt 5.3.3 ab Seite 77). Die Abbildungen in der oberen Zeile sind die Effekte auf die Zielgröße *Weite*, die Abbildungen in der unteren Zeile die Effekte auf die Zielgröße *vertikaleGeschw*.

D. Samplingfade

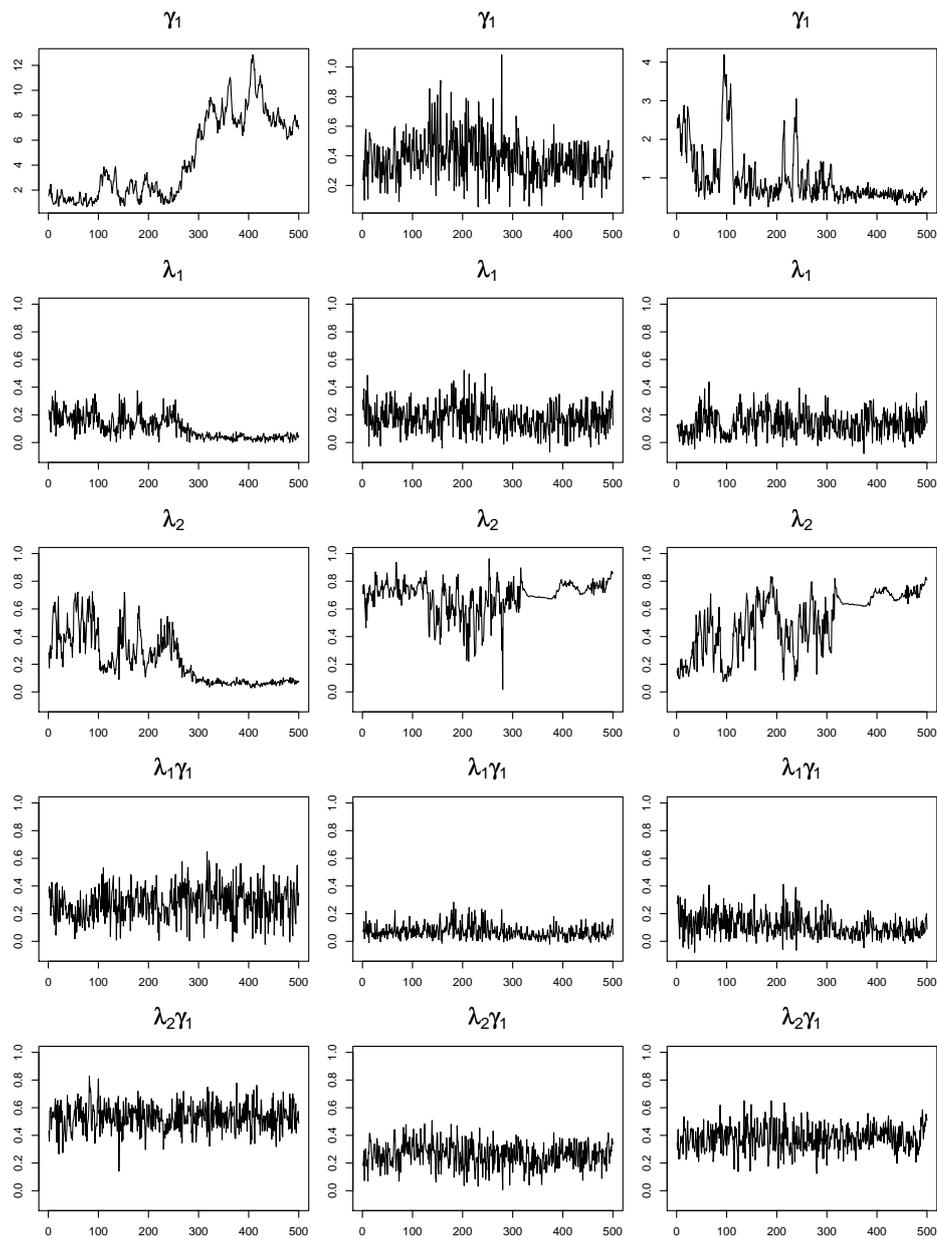


Abbildung D.1.: Samplingfade des Regressionskoeffizienten γ_1 und der Ladungskoeffizienten λ_1 und λ_2 nach Burn-In-Phase und Ausdünnung in den Latente-Variablen-Modellen. In der ersten Spalte sind die Samplingfade des LVM mit der indirekten Kovariable SJ_V75E, in der zweiten die von F_re170_bb und in der dritten die von A60 dargestellt. In den untersten beiden Zeilen sind die Samplingfade der Produkte von $\gamma_1\lambda_1$ bzw. $\gamma_1\lambda_2$ dargestellt. Man sieht in allen Modellen Identifizierbarkeitsprobleme zwischen Ladungs- und Regressionskoeffizient. Betrachtet man hingegen die Produkte (Nettoeffekte), so haben die MCMC-Samples einen konvergenten Zustand erreicht.

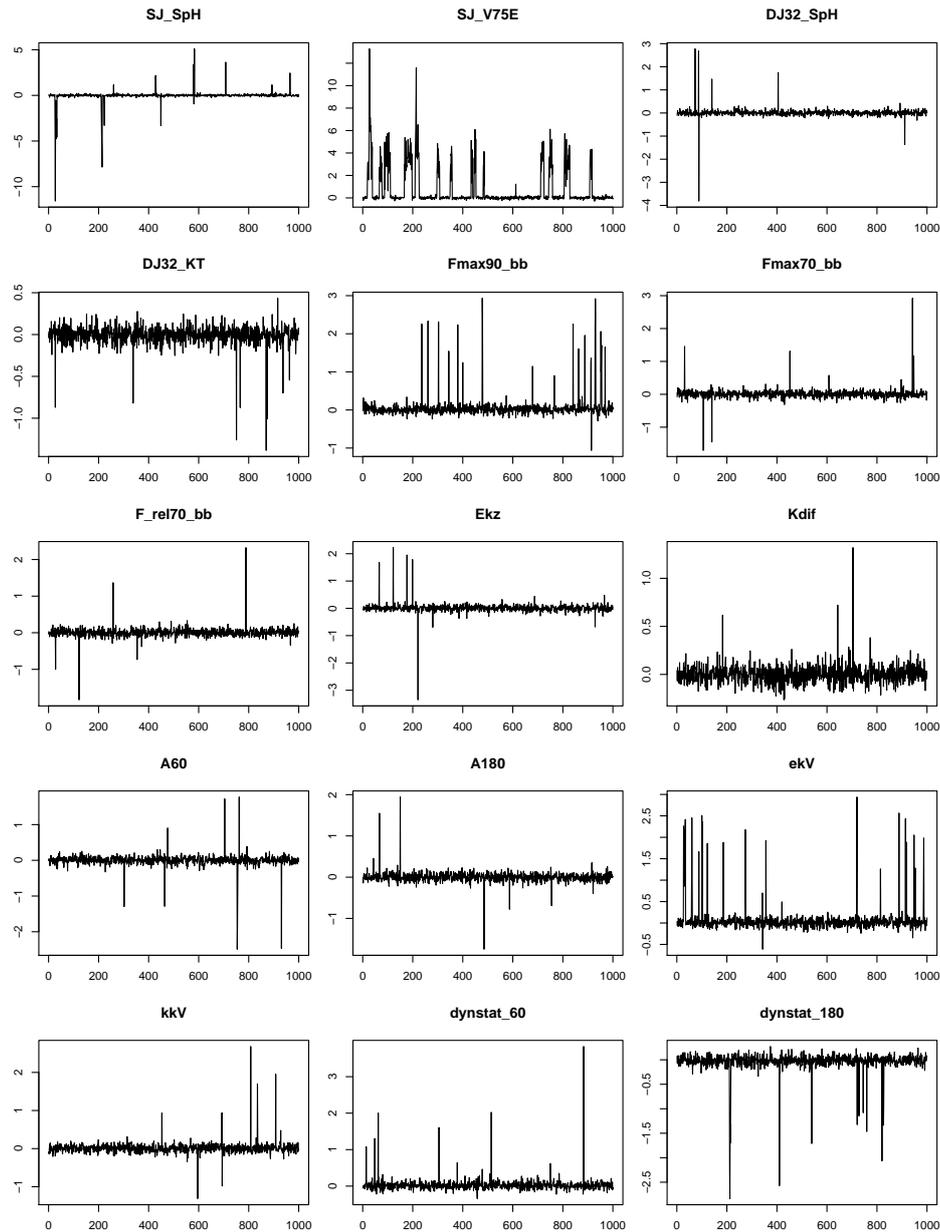


Abbildung D.2.: Samplingpfade der regularisierten Effekte im STAR-Modell im Abschnitt 5.3.3 mit Zielgröße *Weite*. Die Samplingpfade haben einen stationären Zustand erreicht. Die Ausreißer in den Samplingpfaden kennzeichnen die Samples, in denen der Effekt nicht regularisiert wurde und daher betragsmäßig wesentlich größer ist. Bei der Variablen *SJ_V75E* sieht man, dass der Effekt wesentlich öfter nicht regularisiert wurde als bei den anderen Variablen und dass nur positive Effekte geschätzt wurden.

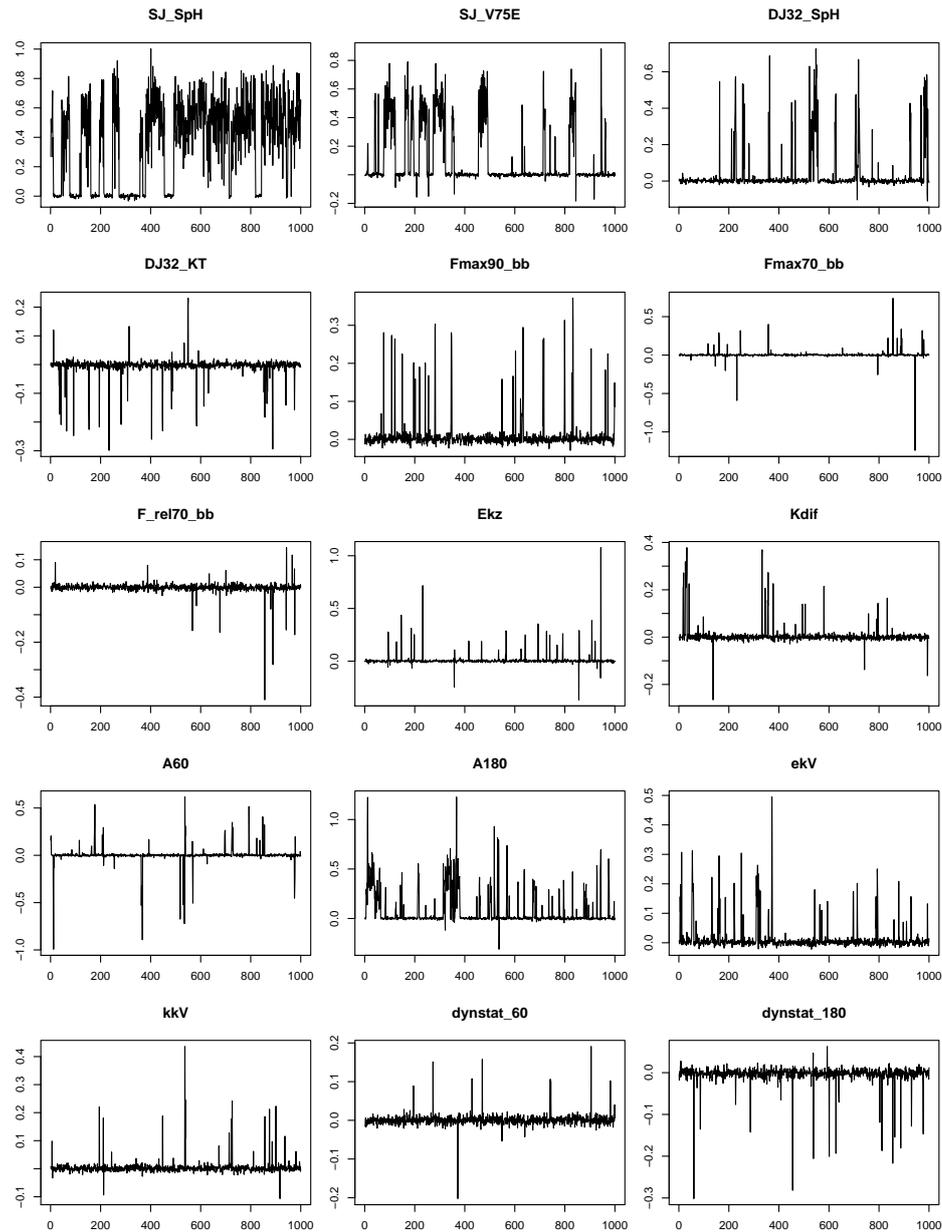


Abbildung D.3.: Samplingpfade der regulierten Effekte im STAR-Modell im Abschnitt 5.3.3 mit Zielgröße *vertikaleGeschw*. Die Samplingpfade haben einen stationären Zustand erreicht. Die Ausreißer in den Samplingpfaden kennzeichnen die Samples, in denen der Effekt nicht reguliert wurde und daher betragsmäßig wesentlich größer ist. Bei der Variablen *SJ_SpH* sieht man, dass der Effekt meist nicht reguliert wurde.

E. Hinweise zur Literatur

In diesem Teil des Anhangs werden einige Literaturhinweise zu den verschiedenen Themengebieten der Arbeit gegeben. Einerseits soll diese Literatur methodisch interessierten Lesern aus dem Bereich der Sportwissenschaft und interessierten Trainern und Athleten weitere Details zur eingesetzten Methodik liefern. Dies umfasst Beweise, Modellvoraussetzungen und weitere Ansätze und Schätzverfahren, auf die in dieser Arbeit aus Gründen der Übersichtlichkeit verzichtet wurde. Solche Literaturverweise sind unter dem Punkt „Methodik“ zusammengefasst. Methodisch versierten Lesern, die jedoch mit der Datenlage und den Sportarten Skisprung und Nordische Kombination weniger vertraut sind, sei die Lektüre von Kapitel 1 dieser Arbeit und die Literaturhinweise unter dem Punkt „Inhaltliches“ empfohlen.

Methodik

- Die Anhänge von Toutenburg (2002) und Fahrmeir u. a. (2007) bieten eine Beschreibung von Verfahren der Matrixrechnung, die für die Methodik dieser Arbeit benötigt wird.
- Fahrmeir u. a. (2004) führt in die Grundlagen der Statistik ein. Gezeigt werden hier wichtige methodische Grundlagen, sowie gängige Darstellungsformen.
- Diggle u. a. (2002) führt in die Darstellung und Analyse von longitudinalen Daten ein.
- In Fahrmeir u. a. (1996) findet sich eine gute Darstellung von multivariaten statistischen Verfahren und grafischen Darstellungsformen.

Inhaltliches

- Für Detailinformation zum Regelwerk bei Skisprung und Nordischer Kombination sei auf FIS (2009c) verwiesen.
- In Schwirtz u. a. (2004) finden sich detaillierte Informationen, wie die Daten zur Leistungsdiagnostik erhoben werden.

Literaturverzeichnis

- [Ballreich u. Baumann 1996] BALLREICH, Rainer ; BAUMANN, Wolfgang: *Grundlagen der Biomechanik des Sports*. Stuttgart : Ferdinand Enke Verlag, 1996
- [Brezger u. a. 2009] BREZGER, Andreas ; KNEIB, Thomas ; LANG, Stefan: *BayesX – Software for Bayesian Inference in Structured Additive Regression Models*. Version 2.00, Mai 2009. <http://www.stat.uni-muenchen.de/~bayesx/bayesx.html>. – Methodology Manual
- [Brezger u. Lang 2006] BREZGER, Andreas ; LANG, Stefan: Generalized additive regression based on Bayesian P-splines. In: *Computational Statistics and Data Analysis* (2006), Nr. 50, S. 967–991
- [Bruhn u. a. 2002] BRUHN, Sven ; SCHWIRTZ, Ansgar ; GOLLHOFER, Albert: Diagnose von Kraft- und Sprungkraftparametern zur Trainingssteuerung im Skisprung. In: *Leistungssport* 32 (2002), Nr. 5, S. 34–37
- [Cattell 1966] CATTELL, Raymond B.: The scree test for the number of factors. In: *Multivariate Behavioral Research* 1 (1966), S. 245–276
- [Diggle u. a. 2002] DIGGLE, Peter J. ; HEAGERTY, Patrick J. ; LIANG, Kung-Yee ; ZEGER, Scott L.: *Analysis of Longitudinal Data*. Oxford New York : Oxford University Press, 2002. – ISBN 0–19–852484–6
- [DSV 2009] DSV, Deutscher Ski V.: *Athleten und Betreuer*. http://www.ski-online.de/1909-Athleten_und_betreuer.htm. Version: 2009. – [Online; Stand 4. Juli 2009]
- [Efron u. Tibshirani 1993] EFRON, Bradley ; TIBSHIRANI, Robert: *An introduction to the bootstrap*. New York : Chapman & Hall, 1993
- [Everitt 1984] EVERITT, Brian S.: *An Introduction to Latent Variable Models*. London, New York : Chapman and Hall, 1984
- [Fachgebiet Biomechanik 2009] FACHGEBIET BIOMECHANIK: *Kooperation mit dem Olympiastützpunkt (OSP) Bayern*. http://www.sp.tum.de/cms/index.php?option=com_content&task=view&id=28&Itemid=305. Version: 2009. – [Online; Stand 29. April 2009]
- [Fahrmeir u. a. 1996] FAHRMEIR, Ludwig ; HAMERLE, Alfred ; TUTZ, Gerhard: *Multivariate statistische Verfahren*. Berlin, New York : de Gruyter, 1996
- [Fahrmeir u. Khatab 2008] FAHRMEIR, Ludwig ; KHATAB, Khaled: Geoadditive Latent Variable Modelling of Child Morbidity and Malnutrition in Nigeria / Department of Statistics, University of Munich. 2008. – Forschungsbericht. – Technical Report No. 20
- [Fahrmeir u. a. 2007] FAHRMEIR, Ludwig ; KNEIB, Thomas ; LANG, Stefan: *Regression: Modelle, Methoden und Anwendungen*. Berlin Heidelberg : Springer-Verlag, 2007 (Statistik und ihre Anwendungen). – ISBN 3–540–33932–9
- [Fahrmeir u. a. 2004] FAHRMEIR, Ludwig ; KÜNSTLER, Rita ; PIGEOT, Iris ; TUTZ, Gerhard: *Statistik - Der Weg zur Datenanalyse*. Berlin Heidelberg : Springer Verlag, 2004
- [Fahrmeir u. Raach 2007] FAHRMEIR, Ludwig ; RAACH, Alexander: A Bayesian semiparametric latent variable model for mixed responses. In: *Psychometrika* (2007), Nr. 72, S. 327–346
- [FIS 2009a] FIS, Internationaler Ski V.: *FIS-Ski-Resultate*. <http://www.fis-ski.com/de/disciplines/skispringen/resultate.html>. Version: 2009. – [Online; Stand 9. Juni 2009]

- [FIS 2009b] FIS, Internationaler Ski V.: *Internationale Skiwettkampfordnung - Nordische Kombination*. http://www.fis-ski.com/data/document/iwo_nc_2008.pdf. Version: 2009. – [Online; Stand 9. Juni 2009]
- [FIS 2009c] FIS, Internationaler Ski V.: *Internationaler Ski Verband - Skispringen*. <http://www.fis-ski.com/de/disciplines/skispringen.html>. Version: 2009. – [Online; Stand 9. Juni 2009]
- [FIS 2009d] FIS, Internationaler Ski V.: *Skisprungschancen - Baunormen 2008*. http://www.fis-ski.com/data/document/skisprungschancen_bau-normen2008.pdf. Version: 2009. – [Online; Stand 9. Juni 2009]
- [Frick u. a. 1991] FRICK, Ulrich ; SCHMIDTBLEICHER, Dietmar ; WÖRN, Claus: Vergleich biomechanischer Messverfahren zur Bestimmung der Sprunghöhe bei Vertikalsprüngen. In: *Leistungssport* (1991), Nr. 2, S. 48–53
- [Gabriel 1971] GABRIEL, K. R.: The biplot graphic display of matrices with application to principal component analysis. In: *Biometrika* 58 (1971), Nr. 3, S. 453–467
- [Grimmett u. Stirzaker 2001] GRIMMETT, Geoffrey ; STIRZAKER, David: *Probability and Random Processes*. Oxford, New York : Oxford University Press, 2001
- [Hall u. Horowitz 1996] HALL, Peter ; HOROWITZ, Joel L.: Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators. In: *Econometrica* 64 (1996), S. 891–916
- [Hartung u. Elpelt 1999] HARTUNG, Joachim ; ELPALT, Bärbel: *Multivariate Statistik*. München Wien : Oldenbourg Wissenschaftsverlag, 1999. – ISBN 3–486–25287
- [Hastie u. Tibshirani 1990] HASTIE, Trevor ; TIBSHIRANI, Robert: *Generalized Additive Models*. London : Chapman and Hall, 1990
- [Hotelling 1936] HOTELLING, Harold: Relations between two sets of variates. In: *Biometrika* 28 (1936), S. 321–377
- [Ishwaran u. Rao 2005] ISHWARAN, H. ; RAO, S. J.: Spike and Slab Variable Selection: Frequentist and Bayesian Strategies. In: *The Annals of Statistics* (2005), Nr. 33, S. 730–773
- [Jöreskog 1973] JÖRESKOG, Karl G.: A general method for estimating a linear structural equation system. In: *Structural equation models in the social sciences* (1973), S. 85–112
- [Kaiser 1958] KAISER, Henri F.: The varimax criterion for analytic rotation in factor analysis. In: *Psychometrika* 23 (1958), S. 187–200
- [Kline 2005] KLINE, Rex B.: *Principles and Practice of Structural Equation Modeling*. The Guilford Press, 2005
- [Kneib u. a. 2009] KNEIB, Thomas ; KONRATH, Susanne ; FAHRMEIR, Ludwig: High-dimensional Structured Additive Regression Models: Bayesian Regularisation, Smoothing and Predictive Performance / Department of Statistics, University of Munich. 2009. – Forschungsbericht. – Technical Report No. 46
- [Konrath u. a. 2008] KONRATH, Susanne ; KNEIB, Thomas ; FAHRMEIR, Ludwig: Bayesian Regularisation in Structured Additive Regression Models for Survival Data / Department of Statistics, University of Munich. 2008. – Forschungsbericht. – Technical Report No. 35
- [Lames 1996] LAMES, Martin: Zeitreihenanalyse: Anwendung in der Trainingswissenschaft. In: *Berichte und Materialien des Bundesinstituts für Sportwissenschaft* 4 (1996), S. 45–57
- [Lawler 2006] LAWLER, Gregory F.: *Introduction to Stochastic Processes*. Chapman & Hall, 2006
- [Lawley u. Maxwell 1971] LAWLEY, Derrick N. ; MAXWELL, Albert E.: *Factor Analysis as a Statistical Method*. London : Butterworths, 1971

- [Leisch 2002] LEISCH, Friedrich: Sweave, Part I: Mixing R and L^AT_EX. In: *R News* 2 (2002), December, Nr. 3, 28–31. <http://CRAN.R-project.org/doc/Rnews/>
- [Letzelter u. a. 1990] LETZELTER, Helga ; LETZELTER, Manfred ; STEINMANN, Werner: Meßfehler in der Kraftdiagnostik. In: *Leistungssport* 20 (1990), Nr. 2, S. 46–52
- [Lijphart 1999] LIJPHART, Arend: *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven : Yale University Press, 1999
- [Little u. Rubin 1987] LITTLE, Roderick J. ; RUBIN, Donald B.: *Statistical Analysis with Missing Data*. New York : Wiley, 1987
- [Martin 1980] MARTIN, Dietrich: *Grundlagen der Trainingslehre, Teil II*. Berlin Heidelberg : Hofmann Karl GmbH + Co, 1980. – ISBN 9783778047729
- [Müller u. Wachter 1989] MÜLLER, E. ; WACHTER, E.: Trainingsmethoden zur Verbesserung der speziellen Sprungkraft von Skispringern. In: *Spectrum Sportwissenschaft* 1 (1989), S. 47–71
- [Neumaier u. a. 1997] NEUMAIER, August ; BABIEL, Stephan ; GEIST, Jens ; SCHARFENBERG, Max Rieder A. ; MESTER, Joachim: Aspekte des Techniktrainings im Skisprung. In: *Techniktraining - Beiträge zu einem interdisziplinären Ansatz, Schriftenreihe des Bundesinstituts für Sportwissenschaft* 94 (1997), S. 369–434
- [Nigg u. Herzog 1994] NIGG, Benno M. ; HERZOG, Walter: *Biomechanics of the Musculo-skeletal System*. John Wiley & Sons, 1994. – ISBN 0 471 94444 0
- [Pyrka u. a. 2008] PYRKA, Patricia ; VEITH, Michael ; SCHÜLLER, Katharina ; ERLBECK, Frank ; BÖSL, Peter ; SCHWIRTZ, Ansgar: *Performance Testing Through Isokinetic Strength and Jump Height Measurements in Nordic Combined*. 2008
- [R Development Core Team 2009] R DEVELOPMENT CORE TEAM: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2009. <http://www.r-project.org>
- [Raach 2005] RAACH, Alexander W.: *A Bayesian semiparametric latent variable model for binary, ordinal and continuous response*. Ludwig-Maximilians- Universität München : Dissertation, 2005
- [Ramsay u. Silverman 2005] RAMSAY, James O. ; SILVERMAN, Bernard W.: *Functional data analysis*. New York : Springer, 2005
- [Schwartz u. a. 2004] SCHWIRTZ, Ansgar ; HAHN, Daniel ; HUBER, A. ; BÖSL, Peter: *Projekt – Dezentrale Leistungsdiagnostik Ski Nordisch 2002/2003*. München: Fachgebiet Biomechanik im Sport, Fakultät für Sportwissenschaften, Technische Universität München, 2004. – Version 01/04/2004
- [Seiler 2008] SEILER, Christian: *Bayesianische Latente-Variablen-Modelle: Statistische Methoden und empirische Analysen*. Ludwig-Maximilians- Universität München : Diplomarbeit, 2008
- [Shikano 2006] SHIKANO, Susumu: Bootstrap and Jackknife. In: *Methoden der Politikwissenschaft* (2006), S. 69–79
- [Starischka 1981] STARISCHKA, Stephan: Überlegungen zur Leistungsdiagnostik aus sportwissenschaftlicher Sicht. In: *Leistungssport* 11 (1981), Nr. 5, S. 340–349
- [Steinert 2006] STEINERT, Sven: *Semiparametrische Latente-Variablen-Modelle*. Ludwig-Maximilians- Universität München : Diplomarbeit, 2006
- [Thurstone 1947] THURSTONE, Louis L.: *Multiple factor analysis*. Chicago : University of Chicago Press, 1947
- [Toutenburg 2002] TOUTENBURG, Helge: *Lineare Modelle*. Physica Verlag, 2002
- [Vdovičenko 1980] VDOVIČENKO, V. P.: Individuelle Besonderheiten der Entwicklung der Schnellkraft bei Skispringern. In: *Leistungssport* 10 (1980), Nr. 5, S. 382–384

-
- [Wagner 1996] WAGNER, Klaus: Multiple mathematisch-statistische Verfahren in der Trainingswissenschaft. In: *Berichte und Materialien des Bundesinstituts für Sportwissenschaft* 4 (1996), S. 85–99
- [Wolf 1997] WOLF, Jürgen: *Skisprung und Nordische Kombination - Rahmentrainingsplan*. DSV-Trainierschule, 1997. – Schriftenreihe des Deutschen Skiverbandes, Heft 31
- [Yao u. a. 2005] YAO, Fang ; MÜLLER, Hans-Georg ; WANG, Jane-Ling: Functional linear regression analysis for longitudinal data. In: *The Annals of Statistics* 33 (2005), Nr. 6, S. 2873–2903

Hiermit versichere ich, dass ich die vorliegende Masterarbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

München, den 3. November 2009

Valentin Wimmer