

Statistik und Datenanalyse: Aufbau

9. Übung – Dimensionsreduktion

Benjamin Fretwurst

▶ PDF-Version der Folien



Inhalt

- Take Home – Ausblick – Vokabeln
- LEF 8
- Literatur

Orga



Technische Vorbereitungen auf die Klausur «On-site»

- Vorbereitung des «Safe Exam Browser»/ Funktionstest verfügbar
- Probeprüfung am 20. Dezember 7:00–22:00 (Rekursrecht bei technischen Problemen)
- Prüfung 3. Januar 14:30 bis 15:30 in Raum Y24-G-45 (der von Methoden Einführung)

Lernziele

Faktorenanalyse in R-Studio umsetzen

- Sie wissen, wie man eine Faktorenanalyse in R macht
- Sie können identifizieren, wie viele Faktoren eine gute Lösung ergeben
- Sie können Faktorladungen (einer rotierten Faktorenlösung) in R berechnen lassen und interpretieren
- Sie können aus den Faktorladungen die Faktoren interpretieren

Es gibt keine neuen LEF, weil wir eine Übungssitzung haben. Es gibt also am Ende nochmal die LEF 8.

Übung 3



0.1 Laden Sie die Daten unserer Befragung

(Für den Fall, dass es Probleme gibt, laden Sie die Daten [hier](#) herunter. Wenn das auch zu Problemen führt, melden Sie sich im Forum.)

```
1 # Wenn wir die Dateien einmal geladen und aufbereitet haben (mit der Extradatei "Aufbereitung.qmd"), dann ist es besser,
2
3 # Prüfe, ob es in dem Ordner in der die Uebung_1_ab.qmd gespeichert ist, einen Unterordner "files" gibt und, wenn nicht,
4 if(dir.exists("data")){} else {dir.create("data")}
5 ## NULL
6
7 download.file("https://stat.ikmz.uzh.ch/Aufbau/Folien/Sitzung_04/data/Stat_Aufbau_Befragung.RDS", "data/Stat_Aufbau_Befr
8
9 DATEN <- readRDS("data/Stat_Aufbau_Befragung.RDS") |>
10   haven::zap_formats()
11
12 # DATEN |> sjlabelled::get_label()
13
14 DATEN |>
15   saveRDS("data/Stat_Aufbau_Befragung.RDS")
```

0.2 Führen Sie eine Faktorenanalyse für unsere RAQ aus mit Hilfe des Paketes **psych**

Ziehen Sie dazu die Hilfe dieser Website heran: <https://md.psych.bio.uni-goettingen.de/mv/unit/fa/efa.html>

A scenic landscape featuring rolling hills and a valley. In the foreground, a herd of black and white cows is grazing in a field of tall, dry grass. The middle ground shows a valley with green fields and scattered trees. The background consists of distant, hazy mountains under a cloudy sky. The text "Take Home – Ausblick – Vokabeln" is overlaid on the image.

Take Home – Ausblick – Vokabeln

Take Home

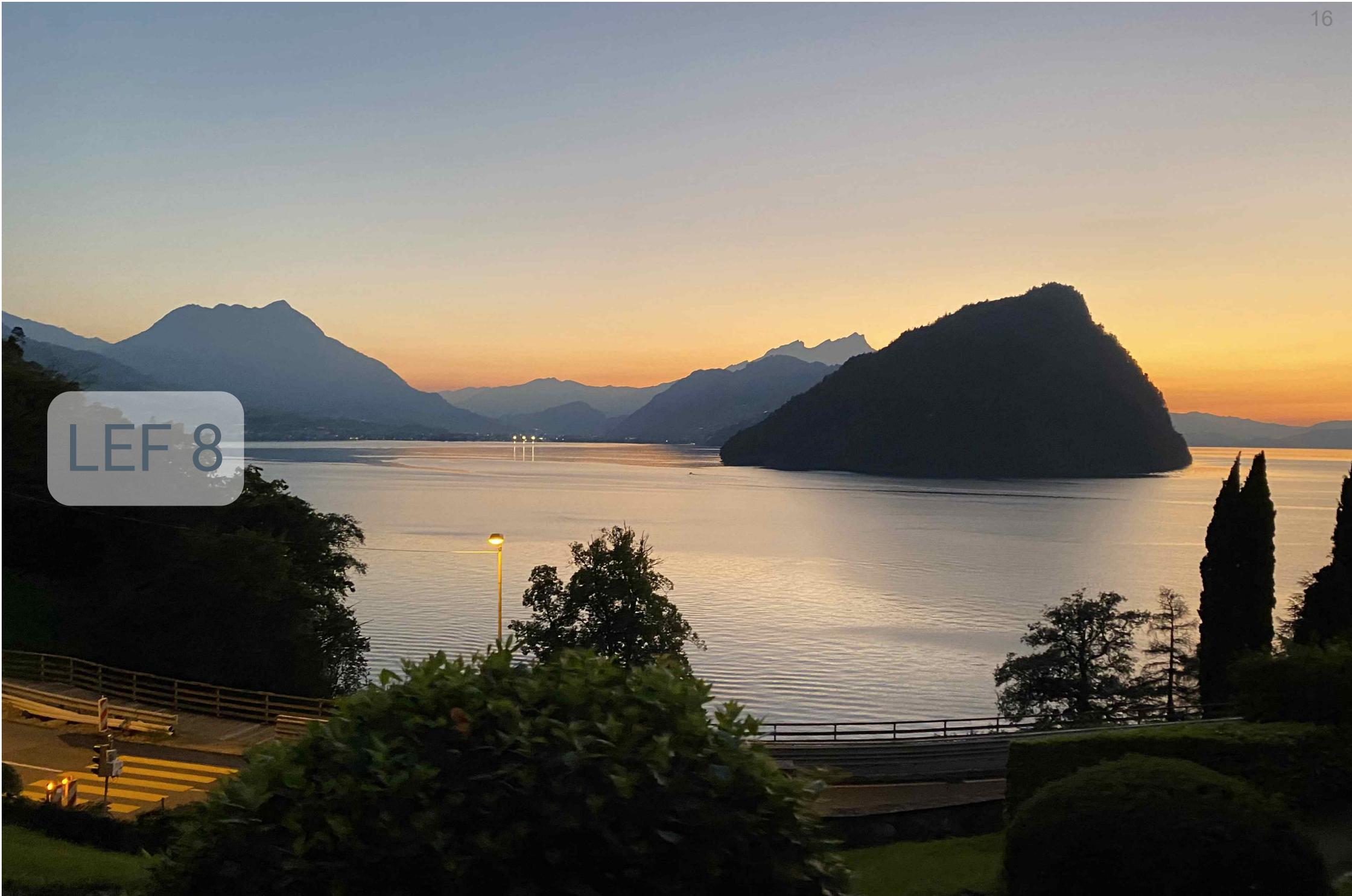
Interaktionen

- Man findet für alle gewünschten Analysen im Netz Anleitungen
- Mit Anleitung ist eine Faktorenanalyse in R kein Problem, wenn man die Vorlage an den richtigen Stellen anpasst.
- Faktorenanalyse können Sie interpretieren
- Sie können die Faktorenanalysen in Publikationen

Ausblick

- Wir schauen uns Regressionsmodelle für den Fall an, dass die AV eine Dummy ist.
- Wir beschäftigen uns in dem Zusammenhang mit Maschine Learning (ML)

LEF 8



Essayfragen 8

E8.1 Bitte schauen Sie sich die folgenden drei (mistigen) Tabellen an, die aus fiktiven Publikationen stammen. Jedes mal werden die Koeffizienten als "beta" bezeichnet, aber es könnten nur die unstandardisierten Regressionskoeffizienten b sein oder die standardisierten Regressionskoeffizienten $\text{std. } b$. Was ist wohl was? a) Begründen Sie Ihre Entscheidung! b) Erklären Sie, warum es nicht β s sein können.

| | A | | |
|-------|-------|------|-------|
| | beta | t | p |
| Var 1 | 0.02 | 1.4 | 0.168 |
| Var 2 | 0.35 | 2.9 | 0.005 |
| Var 3 | 0.43 | 3.1 | 0.003 |
| Var 4 | -0.21 | -2.2 | 0.032 |
| Var 5 | 0.03 | 1.8 | 0.078 |

| | B | | |
|--------|-------|------|-------|
| | beta | t | p |
| pred 1 | 3.45 | 3.25 | 0.002 |
| pred 2 | 1.23 | 2.34 | 0.023 |
| pred 3 | 0.001 | 4.1 | 0.000 |
| pred 4 | 2.5 | -3.2 | 0.002 |
| pred 5 | 0.3 | 1.8 | 0.078 |

| | C | | |
|------|-------|------|-------|
| | beta | t | p |
| UV 1 | 0.3 | 3.6 | 0.001 |
| UV 2 | 0.5 | 1.28 | 0.206 |
| UV 3 | -0.2 | -3.9 | 0.000 |
| UV 4 | -0.08 | -4.2 | 0.000 |
| UV 5 | 0.04 | 1.8 | 0.078 |

Essayfragen 8.2-8.3

E8.2 Schreiben Sie zu jedem der folgenden Schlagworte in ganzen Sätzen auf, zu welchen Zwecken Faktorenanalysen alles eingesetzt werden können: a) latente Variablen entdecken b) Multikollinearität c) Indices d) Messung latenter Konstrukte

E8.3 a) Erklären Sie, was Faktorladungen sind. b) Wie gehen Sie vor, wenn Sie Faktoren interpretieren wollen?

Essayfragen 8.4-8.10

Lesen Sie Song et al. (2004) und beantworten Sie dann folgende Fragen und die anschliessenden MCs.

E8.4 Was wird mit Cronbachs α beschrieben?

E8.5 Was für eine Faktorenanalyse haben Song et al durchgeführt?

E8.6 Mit welcher Methoden wurden die Faktoren rotiert?

E8.7 Wie viele initiale Faktoren hat die Faktorenanalyse herausgegeben?

E8.8 a) Nach welchen Kriterien wurden die Faktoren ausgewählt? b) Wie bewerten Sie das Vorgehen?

E8.9 a) Wie viele Items (Fragen) wurden mit der Faktorenanalyse ausgewertet? b) Auf wie viele Faktoren wurden die Items reduziert?

E8.10 Wie viel Varianz erklären die Faktoren, die von Song et al ausgewählt wurden?

MC-Fragen 8

MC 8.1.

MC 8.1: Was sagen die Zahlen im Text von Song et al. (2004)?

| richtig | falsch | Aussagen |
|-----------------------|-----------------------|---|
| <input type="radio"/> | <input type="radio"/> | Die Variablen erklären 67.4 Prozent der extrahierten Faktoren. |
| <input type="radio"/> | <input type="radio"/> | Die Faktoren wurden mit Hilfe der Varimax-Methode rotiert. |
| <input type="radio"/> | <input type="radio"/> | Alle Faktoren zeigen ein hohes Cronbachs α . |
| <input type="radio"/> | <input type="radio"/> | Der "Factor 5" erklärt etwa das Vierfache einer einfachen Variable. |

Punkte: 0

MC 8.2.

MC 8.2: Was sagen die Zahlen im Text von Song et al. (2004)?

- | richtig | falsch | Aussagen |
|-----------------------|-----------------------|---|
| <input type="radio"/> | <input type="radio"/> | Der "Factor 1" erklärt die meiste Varianz der ursprünglichen Variablen. |
| <input type="radio"/> | <input type="radio"/> | Da auf "Factor 3" und "Factor 4" gleich viele Variablen laden, haben Sie dieselbe Varianzaufklärung. |
| <input type="radio"/> | <input type="radio"/> | Wenn der Eigenwert eines Faktors für einen Faktor A grösser ist als bei Faktor B, muss bei A Cronbach α auch grösser sein als bei B. |
| <input type="radio"/> | <input type="radio"/> | In der Korrelationstabelle sind die Faktoren abgebildet, die teilweise stark korrelieren. |

Punkte: 0

MC 8.3.

MC 8.3: Was sagen die Zahlen im Text von Song et al. (2004)?

- | richtig | falsch | Aussagen |
|-----------------------|-----------------------|---|
| <input type="radio"/> | <input type="radio"/> | In Tabelle 3 sind die β s der GG angegeben, die mit den b's in der Stichprobe geschätzt wurden. |
| <input type="radio"/> | <input type="radio"/> | Die standardisierten BETAS sind mit den Korrelationswerten der Tabelle 2 vergleichbar. |
| <input type="radio"/> | <input type="radio"/> | Die Varianzaufklärung der Regression beträgt 54 Prozent. |
| <input type="radio"/> | <input type="radio"/> | Der p-Wert von "Virutal community" ist genau 0. |

Punkte: 0

MC 8.4.

MC 8.4: Was sagen die Zahlen im Text von Song et al. (2004)?

- | richtig | falsch | Aussagen |
|-----------------------|-----------------------|--|
| <input type="radio"/> | <input type="radio"/> | Je stärker Studierende versuchen sich in eine «Virtuelle community» zu integrieren, desto eher werden sie internetsüchtig. |
| <input type="radio"/> | <input type="radio"/> | Multikollinearität in den UVs kann ausgeschlossen werden. |
| <input type="radio"/> | <input type="radio"/> | Wer das Internet nutzt, um Informationen zu suchen, gerät eher nicht in eine Internetsucht. |
| <input type="radio"/> | <input type="radio"/> | Wer das Internet nutzt, um seinen «Personal status» zu verbessern, neigt stark zu Internetsucht. |

Punkte: 0

MC 8.5.

MC 8.5: Was sagen die Zahlen im Text von Song et al. (2004)?

- | richtig | falsch | Aussagen |
|-----------------------|-----------------------|--|
| <input type="radio"/> | <input type="radio"/> | Die signifikante Korrelation zwischen «Personal status» und «Internetaddiction» ist eher durch Drittvariablen zustande gekommen. |
| <input type="radio"/> | <input type="radio"/> | Das Konfidenzintervall des BETA von «Monetary compensation» schliesst die 0 mit ein. |
| <input type="radio"/> | <input type="radio"/> | Da «Personal status» einen p-Wert $> .05$ hat, ist β in der GG 0 und daher deutlich kleiner als das in der Stichprobe. |
| <input type="radio"/> | <input type="radio"/> | Da in der Tabelle standardisierte BETAs angezeigt werden, sind die t-Werte eher redundant. |

Punkte: 0

Insgesamt (nur MCs) 0 von 10 Punkten, was 0% und etwa einer 1 entspricht.

Vokabeln [↗](#)

 Search:

| Nr | Sitzung | Inhalt | Deutsch | Englisch | Erläuterung |
|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|--|
| <input type="text" value="All"/> |
| 87 | 7 | EFA, CFA, Lesekompetenz | Cronbachs α | Cronbachs α | Cronbachs α ist für die innere Ko einer Skala. Grur die durchschnitt Korrelation aller Items. Cronbach bis max. 1 gehen währe dann perf |
| 71 | 6 | Faktorenanalysen | Dimensionsreduktion | dimensional reduction | Dimensionsredu bedeutet, dass c der Merkmalsdir (Variablen) redu. |
| 72 | 6 | Faktorenanalysen | Eigenwerte | eigenvalues | Varianz, die ein f sich vereint. |

| Nr | Sitzung | Inhalt | Deutsch | Englisch | Erläuterung |
|----|---------|------------------|-----------------|-----------------|---|
| 73 | 6 | Faktorenanalysen | Faktorenanalyse | factor analysis | Verfahren zur Dimensionsreduktion. Mehrere Variablen werden in weniger Faktoren zusammengefasst. |
| 76 | 6 | Faktorenanalysen | Faktorladung | factor loadings | Die Korrelationen der Variablen mit den Faktoren. |
| 77 | 6 | Faktorenanalysen | Faktorlösung | Solution | Die Faktorlösung, bei der die Anzahl der Faktoren optimal ist für die Darstellung der Varianz. |
| 81 | 6 | Faktorenanalysen | Faktorrotation | factor rotation | Die Faktorrotation, bei der die Faktoren so an die Variablen angepasst werden, so dass die Varianz besser erklärt wird. |

| Nr | Sitzung | Inhalt | Deutsch | Englisch | Erläuterung |
|----|---------|-------------------------|-------------------------|------------------------------------|--|
| 78 | 6 | Faktorenanalysen | Hauptkomponentenanalyse | Principle Component Analysis (PCA) | Die Hauptkomponentenanalyse fasst Variablen zusammen. Das latente Konzept interessiert dabei. |
| 79 | 6 | Faktorenanalysen | Kommunalitäten | communalities | Kommunalitäten sind der Varianzanteil einer Variable, der durch die Faktorenlösung erklärt ist. Unter .4 ist ein |
| 84 | 6 | Faktorenanalysen | Scree Plot | Scree Plot | Eine graphische Darstellung der Eigenwerte, die zeigt, welche Faktorenlösung |
| 88 | 7 | EFA, CFA, Lesekompetenz | Skalenreliabilität | scale reliability | Skalenreliabilität bezeichnet die interne Konsistenz einer Skala, also wie stark die Items eines Fragebogens |

| Nr | Sitzung | Inhalt | Deutsch | Englisch | Erläuterung |
|----|---------|------------------|-----------------------------------|---|---|
| | | | | | zusammengehö Cronbachs α) |
| 85 | 6 | Faktorenanalysen | Strukturgleichungsmodell (SEM) | structural equation model (SEM) | Strukturgleichun sind Modelle über vielfältige Zusammenhäng mehrerer Variab Konfirmatorisch Faktorenanalyse der Regel teil von |
| 86 | 6 | Faktorenanalysen | Uniqueness | Uniqueness | Uniqueness ist 1 Kommunalität. V Variable viel Var die sie nicht mit Faktoren teilt, ist einzigartig. |
| 74 | 6 | Faktorenanalysen | explorative Faktorenanalyse | explorative factor analysis (EFA) | Es werden Faktore Fülle von Variablen (exploriert). |

| Nr | Sitzung | Inhalt | Deutsch | Englisch | Erläuterung |
|----|---------|------------------|----------------------------------|------------------------------------|---|
| 75 | 6 | Faktorenanalysen | konfirmatorische Faktorenanalyse | confirmatory factor analysis (CFA) | Faktoren werden vorgegeben und ob die Daten sie (confirm). |
| 80 | 6 | Faktorenanalysen | latente Variablen | latent variable | Latente Variable solche, die nicht sondern von gemessenen Variablen gemessen werden |
| 82 | 6 | Faktorenanalysen | oblique Rotation | oblique rotation | Die Faktoren können korrelieren, stehen nicht ganz senkrecht zueinander. |
| 83 | 6 | Faktorenanalysen | orthogonale Rotation | orthogonal rotation | Die Faktoren stehen bei der Rotation senkrecht (orthogonal) zueinander. Die Korrelation zwischen den Faktoren ist |

Literatur

Song, Indeok, Robert Larose, Matthew S. Eastin, and Carolyn A. Lin. 2004. "Internet Gratifications and Internet Addiction: On the Uses and Abuses of New Media." *CyberPsychology & Behavior* 7 (4): 384–94. <https://doi.org/10.1089/cpb.2004.7.384>.