

## Assess the Assessment

Michela Nettekoven & Karl Ledermüller,  
23. Mai 2011

- ▶ **Generelle Idee**
- ▶ Generelle technische Umsetzung
  - ▶ Aufbau eines .rnw files
  - ▶ LaTeX
- ▶ zusätzliche Verfahren
  - ▶ Hierarchisches Clustering
  - ▶ Multi Dimensional Scaling
  - ▶ Faktorenanalyse
- ▶ Ergebnisse

Es gibt scheinbar den Wunsch des VRL, Prüfungen (oder besser das Assessmentsystem generell) und die dazupassenden (aufeinander schon - oder nicht abgestimmten) Assessmentmethoden „zu evaluieren“. <sup>1</sup>

Spannend wäre es, ein standardisiertes Reportinginstrumentarium zur Verfügung zu stellen, das (mittels Rasch Modellen) Feedback über die am Prüfungsserver der WU abgewickelten Prüfungen gibt.

---

<sup>1</sup>siehe <http://www.wu.ac.at/academicstaff/info/facts/assess>

- ▶ Generelle Idee
- ▶ **Generelle technische Umsetzung**
  - ▶ Aufbau eines .rnw files
  - ▶ LaTeX
- ▶ zusätzliche Verfahren
  - ▶ Hierarchisches Clustering
  - ▶ Multi Dimensional Scaling
  - ▶ Faktorenanalyse
- ▶ Ergebnisse

Mittels R sweave und  $\text{\LaTeX}$  können automatisiert Berichte erstellt werden. Dafür sind folgende Schritte notwendig:

1. Erstellung eines .rnw files
2. Erstellung des .tex files mittels sweave
3. Erstellung eines .pdf files aus dem .tex files (Bspw. pdflatex)

```
\documentclass[a4paper]{article}
\title{Sweave Example 1}
\author{Friedrich Leisch}
\begin{document}
\maketitle
In this example we embed parts of the examples from the
\texttt{kruskal.test} help page into a \LaTeX{} document:
<<>>=
```

```
data(airquality)
library(ctest)
kruskal.test(Ozone ~ Month, data = airquality)
```

```
@
which shows that the location parameter of the Ozone
distribution varies significantly from month to month. Finally we
include a boxplot of the data:
```

```
\begin{center}
<<fig=TRUE,echo=FALSE>>=
boxplot(Ozone ~ Month, data = airquality)
@
\end{center}
\end{document}
```

2

---

<sup>2</sup>siehe: <http://www.statistik.lmu.de/~leisch/Sweave/example-1.Snw>

„ $\text{\LaTeX}$  is a high quality typesetting system it includes features designed for the production of technical and scientific documentation.  $\text{\LaTeX}$  is the de facto standard for the communication and publication of scientific documents.  $\text{\LaTeX}$  is available as free software.“<sup>3</sup>

Aus der R Console wird mittels Sweave() Befehl das .rnw file aufgerufen und ein .tex ( $\text{\LaTeX}$ ) file erstellt. Dieses .tex file wird in einen .pdf Bericht umgewandelt. (.html wurde vor allem wegen der langen Wartezeit nicht durchgeführt)

Zusätzlich haben wir die Detailauswertungen der Unterfragen mit einem zusätzlichen R file umgesetzt, das außerhalb des Sweave Befehls (Sweave verwendet sink()) erstellt wird.

Diskussionmöglichkeit: Kann man eleganter .tex Code (wie beispielsweise Überschriften) über ein .rnw File erstellen (bspw. eine for Schleife über i Elemente).

---

<sup>3</sup><http://www.latex-project.org/>

- ▶ Generelle Idee
- ▶ Generelle technische Umsetzung
  - ▶ Aufbau eines .rnw files
  - ▶ LaTeX
- ▶ zusätzliche Verfahren
  - ▶ Hierarchisches Clustering
  - ▶ Multi Dimensional Scaling
  - ▶ Faktorenanalyse
- ▶ Ergebnisse



Um mehr über die Dimensionen der Fragen zu erfahren, haben wir folgende Verfahren angewandt:

- ▶ Hierarchisches Clustering
- ▶ Multi Dimensional Scaling
- ▶ Faktorenanalyse

Mittels hierarchical clustering werden aus einer Distanzmatrix (in unserem Fall: 1-cor) Ähnlichkeiten zwischen Items (Prüfungsfragen) dargestellt und in einem Dendrogramm (griechisch dendron=Baum) dargestellt. Dadurch wird gezeigt, wie ähnlich welche gemessenen Konstrukte (Fragen) sind.

Fragen, die eine geringe Distanz aufweisen werden zusammenhängend dargestellt, welche mit hoher Distanz sind nicht im gleichen Cluster.

Mittels Multi Dimensionaler Skalierung (package smacof) werden hochdimensionale statistische Räume auf eine kleinere Anzahl von Dimensionen reduziert.

Auch hier wird eine Distanzmatrix verwendet, um die Ähnlich- bzw. Unähnlichkeiten darzustellen. Objekte die näher zusammen liegen, sind sich ähnlicher, als Objekte, die weiter entfernt sind.

Die Faktorenanalyse versucht, genauso wie die oben genannten Verfahren die Dimensionen des Datensatzes zu reduzieren und auf latente Variablen zurückzuführen. Wir verwenden die Faktorenanalyse explorativ und möchten vor allem herausfinden, wie „multidimensional“ die Daten sind.

Zur Bestimmung der Anzahl der Faktoren wird ein interner Hypothesentest verwendet, der überprüft, ob die Anzahl der Faktoren ausreicht (wir brechen bei 10% ab).

- ▶ Generelle Idee
- ▶ Generelle technische Umsetzung
  - ▶ Aufbau eines .rnw files
  - ▶ LaTeX
- ▶ zusätzliche Verfahren
  - ▶ Hierarchisches Clustering
  - ▶ Multi Dimensional Scaling
  - ▶ Faktorenanalyse
- ▶ Ergebnisse

Die Ergebnisse werden in einem pdf-file zusammengefasst:

- ▶ Kennzahlen des Rasch Modells:
  - ▶  $\beta$ - und  $\eta$ -Parameter
  - ▶ Person-Item Map

- ▶ Partial-Credit Modell
- ▶ mehrstufige Antwortkategorien werden automatisch ermittelt und ergeben sich aus den möglichen Prozentwerten, die auf eine Frage erreicht werden können (Teilpunkteschema der WU)
- ▶ Person-Item Map:
  - ▶ Achtung auf Interpretation der Antwortkategorien
  - ▶ Items mit vertauschter Reihenfolge der Antwortkategorien: bei widersprüchlichen oder offensichtlich falschen Antwortmöglichkeiten werden diese selten angekreuzt, sodass ein bestimmter zu erreichender Prozentwert bei dieser Frage selten auftritt. Dann verschiebt sich die Reihenfolge der Antwortkategorien, der selten auftretende Prozentwert wird als besonders schwere Antwortkategorie interpretiert.

- ▶ Mehrdimensionalität der Test-Items
  - ▶ Martin-Löf Test (Split-Kriterium: Median)
  - ▶ Faktorenanalyse (iterativ)
  - ▶ Clusteranalyse
  - ▶ Multidimensional Scaling
- ▶ Personenhomogenität
  - ▶ untersuchte Gruppen: Geschlecht, Matr.Nr., Erstantritt
  - ▶ Analyse der erreichten Prozentwerte: Boxplot, Regression, ANOVA
  - ▶ Analyse der Test-Items: Andersen LR Test, graphischer Modelltest
  - ▶ Infit/Outfit Item und Person Map
- ▶ Fragenbasierte Auswertung: pro Frage:
  - ▶  $\beta$ -Parameter
  - ▶ Schwellenwerte (thresholds)
  - ▶ Wald Test (Gruppen nach Geschlecht, Matr.Nr., Erstantritt)
  - ▶ Item Characteristic Curve