

Universitäts-Rechenzentrum Trier



AWS.SPSS.5

Trier, den 22.02.2006

B. Baltes-Götz

Conjoint-Analyse mit SPSS

Inhalt

VORWORT	4
1 EINLEITUNG	5
2 ERSTELLUNG EINES ORTHOGONALEN UNTERSUCHUNGSPLANS (ORTHOPLAN)	6
3 ERSTELLUNG DES UNTERSUCHUNGSMATERIALS (PLANCARDS)	10
4 METHODEN DER DATENGEWINNUNG	12
4.1 Ratingdaten	12
4.2 Ordinaldaten	12
5 CONJOINT-ANALYSE DER PRÄFERENZURTEILE	14
5.1 Schätzmethoden in der SPSS-Prozedur CONJOINT	14
5.2 Präferenzmodelle	14
5.2.1 Diskrete Faktoren	15
5.2.2 Lineare Faktoren	16
5.2.3 Quadratische Faktoren	17
5.3 Ergebnisse für eine einzelne Auskunftsperson	17
5.3.1 Das CONJOINT-Kommando	17
5.3.2 Teilnutzenwerte und deren Standardfehler	19
5.3.3 Die Regressionsanalyse zu Fall Nr. 6	22
5.3.4 Relative Bedeutung der Faktoren	24
5.3.5 Fit-Indikatoren	24
5.3.6 Simulationsergebnisse	25
5.3.7 Verletzungen der modellgemäßen Ordnung (Reversals)	25
5.4 Aggregierte Ergebnisse	26
5.4.1 Teilnutzenwerte und Wichtigkeiten	26
5.4.2 Reversals	28
5.4.3 Präferenzen für die Simulationsprodukte	29
5.5 Die Datei mit den Nutzenwerten	30
6 LITERATUR	31
7 STICHWORTREGISTER	32

Herausgeber: Universitäts-Rechenzentrum Trier
 Universitätsring 15
 D-54286 Trier
 Tel.: (0651) 201-3417, Fax.: (0651) 3921

Leiter: Dr. Peter Leinen
Autor: Bernhard Baltes-Götz
Copyright © 2006; URT

Vorwort

In diesem Manuskript wird die metrische, auf der linearen Regression basierende, Variante der Conjoint-Analyse beschrieben, die im Statistikpaket SPSS über das Zusatzmodul Conjoint realisiert ist.

Im Manuskript kommt SPSS 13.0 für Windows zum Einsatz, jedoch können praktisch alle vorgestellten Verfahren auch mit alternativen SPSS-Versionen (ca. ab 8.0) unter Windows, MacOS oder Linux realisiert werden.

Das Manuskript ist als PDF-Dokument zusammen mit den im Kurs benutzten Dateien auf dem Webserver der Universität Trier von der Startseite (<http://www.uni-trier.de/>) ausgehend folgendermaßen zu finden:

[Rechenzentrum](#) > [Studierende](#) > [EDV-Dokumentationen](#) >
[Statistik](#) > [Conjoint-Analyse mit SPSS](#)

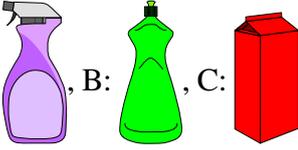
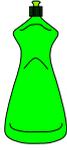
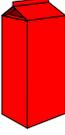
Hinweise auf Unzulänglichkeiten im Manuskript werden mit Dank entgegen genommen

1 Einleitung

Die Conjoint-Analyse wird in der Marketing-Forschung u.a. dazu verwendet, neue Produkte mit hohen Absatzchancen zu entwerfen. Dabei wird versucht, die Attraktivität eines Produktes, das durch seine Ausprägungen bei mehreren Eigenschaften beschrieben ist, auf die **Teilnutzenwerte** der realisierten Merkmalsausprägungen zurückzuführen. Es wird angenommen, dass sich der Gesamtnutzen eines Produktes, von dem die Bevorzugung gegenüber anderen Produkten abhängt, **additiv** aus den Teilnutzenwerten seiner Eigenschaftsausprägungen zusammensetzt. Damit ist der Ansatz dieser Marktforschungsmethode stark verwandt mit der meßtheoretischen Methode *Additive Conjoint Measurement*, was im Titel der Conjoint-Monografie von Green & Wind (1973) zum Ausdruck kommt: *Multiattribute decisions in marketing: A measurement approach*.

Aber die Methode besitzt nicht nur interessante methodologische Wurzeln, sondern hat sich zu einem zentralen Instrument der Marketing-Analyse mit hoher praktischer Relevanz entwickelt (Müller-Hagedorn et al. 1993, S. 123).

Im SPSS-Handbuch zum Zusatzmodul Conjoint (SPSS 1997) wird das aus Green & Wind (1973) stammende Beispiel einer Marketingstudie zur Planung eines neuen Teppichreinigers beschrieben. Hier werden den Untersuchungsteilnehmern mehrere Produktbeschreibungen (Profile) auf je einer Karte vorgelegt, wobei jeder Teppichreiniger charakterisiert wird durch

- | | |
|--|---|
| <ul style="list-style-type: none"> • seine Verpackung • seinen Markennamen • seinen Verkaufspreis • ein Gütesiegel • eine Geld-zurück-Garantie |  <p>A: , B: , C: </p> |
| | <p>K2R, Glory, Bissell</p> <p>1,69 Euro, 1,99 Euro, 2,29 Euro</p> <p>vorhanden oder nicht</p> <p>vorhanden oder nicht</p> |

Die fiktiven Produkte werden dann von jeder Auskunftsperson entweder einzeln auf einer Skala bewertet oder gemeinsam in eine Rangreihe gebracht.

Für jeden einzelnen Teilnehmer wird versucht, aus seinen Gesamtbewertungen bzw. aus seinen ordinalen Präferenzurteilen die metrischen Teilnutzenwerte der Eigenschaftsausprägungen zu erschließen. Dabei geht man von der Annahme aus, dass von zwei Produkten dasjenige mit dem höheren Gesamtnutzen bevorzugt wird, wobei sich der Gesamtnutzen eines Produktes durch Addition der Teilnutzenwerte seiner Eigenschaftsausprägungen ergibt. Der subjektive Gesamtnutzen eines Teppichreinigers mit Verpackung B, der unter dem Namen K2R für 1,69 Euro mit Gütesiegel, aber ohne Geld-zurück-Garantie verkauft wird, für einen individuellen Käufer bzw. Untersuchungsteilnehmer wird also folgendermaßen erklärt:

$$\begin{aligned} \text{Gesamtnutzen} &= \text{Konstante} + \text{Teilnutzen(B)} + \text{Teilnutzen(K2R)} + \text{Teilnutzen(1,69 Euro)} \\ &+ \text{Teilnutzen(Gütesiegel: ja)} + \text{Teilnutzen(Geld-zurück-Garantie: nein)} \end{aligned}$$

In regressionsanalytischer Terminologie haben wir bei Betrachtung der Präferenzdaten *eines individuellen Käufers* bzw. Untersuchungsteilnehmers folgende Situation:

- Die beurteilten Produkte sind die Fälle (Beobachtungseinheiten).
- Die Eigenschaften der Produkte (Verpackung etc.) spielen die Rolle von unabhängigen Variablen.
- Der ordinal oder metrisch gemessene Gesamtnutzen der Produkte fungiert als abhängige Variable.

Durch *Aggregieren der individuellen Nutzenwerte* gewinnt man generalisierbare Aussagen über die Produkte und ihre Eigenschaften, die bei der Planung und Marktplatzierung von Produkten verwendet werden können.

In der Marktforschung ist man vor allem daran interessiert, welche Produkteigenschaften für die Präferenzbildung besonders wichtig sind, weil aus Kostengründen in der Regel nur bei den wichtigsten Eigenschaften die Realisation der attraktivsten Ausprägung lohnend ist.

Weil nur begrenzt viele Eigenschaften und Ausprägungen in eine Untersuchung einbezogen werden können (siehe Abschnitt 2), setzt die Conjoint-Analyse voraus, dass die Produkte anhand relativ weniger Eigenschaften hinreichend charakterisiert werden können.

SPSS unterstützt die Conjoint-Analyse durch die drei Prozeduren ORTHOPLAN, PLANCARDS und CONJOINT, die Bestandteile des Zusatzmoduls Conjoint sind.

2 Erstellung eines orthogonalen Untersuchungsplans (ORTHOPLAN)

Die einbezogenen Eigenschaften müssen im sachlogischen Sinn voneinander **unabhängig** sein, außerdem natürlich **relevant** und bei der Produktgestaltung **realisierbar**. Um den Aufwand für die Teilnehmer in Grenzen zu halten, darf die Zahl der zu untersuchenden Eigenschaften und Ausprägungen nicht zu groß werden. Nach einer Empfehlung von Backhaus et al. (2005, S. 609) sollten maximal 20 Produkte einbezogen werden.

Selbst bei größter Beschränkung auf die wesentlichen Eigenschaften und Ausprägungen wird ein vollständiges faktorielles Design, bei dem jede Kombination von Merkmalsausprägungen realisiert wird, in der Regel unpraktikabel viele Produkte umfassen. Im unserem Beispiel müssten die Untersuchungsteilnehmer ($3 \times 3 \times 3 \times 2 \times 2 = 108$ Produktbeschreibungen in eine Rangordnung bringen oder bewerten. Da die Conjoint-Analyse ohnehin Interaktionen zwischen den Eigenschaften ausschließt, ist ein vollständiges Design auch gar nicht erforderlich. Die Haupteffekte der Eigenschaften lassen sich auch in einem reduzierten Design schätzen, das nur eine Teilmenge der Eigenschaftskombinationen enthält.

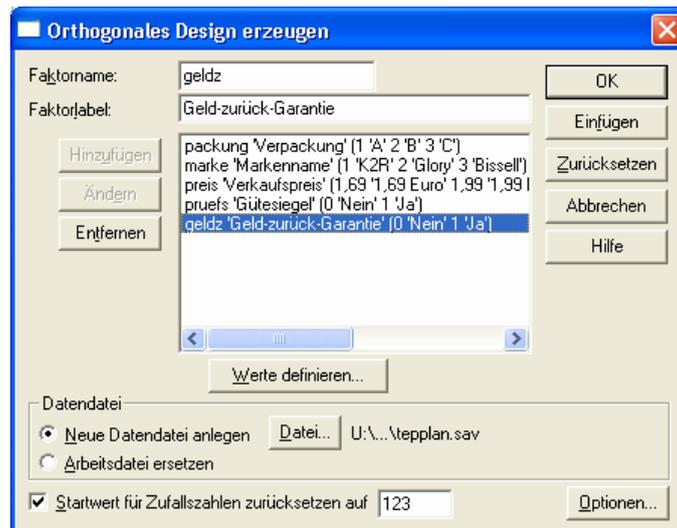
Man kann zeigen, dass die sparsamste Schätzung aller Haupteffekte in einem so genannten **orthogonalen Design** möglich ist. Dabei wird jede Faktorstufe (Eigenschaftsausprägung) mit den Stufen eines anderen Faktors mit identischen oder proportionalen Häufigkeiten kombiniert. Dadurch sind alle Eigenschaften im statistischen Sinne unabhängig voneinander.

SPSS kann für vorgegebene Faktoren und Faktorstufen den kleinstmöglichen orthogonalen Plan ermitteln. Dabei bestimmt SPSS zum vorliegenden Design (bei uns: $3 \times 3 \times 3 \times 2 \times 2$) zunächst ein passendes *symmetrisches Design*, dessen Faktoren gleich viele Stufen haben (bei uns: $3 \times 3 \times 3 \times 3 \times 3$). Dazu kann SPSS ein reduziertes orthogonales Design mit Hilfe vorhandener Lösungsbibliotheken generieren. Schließlich werden die zur symmetrischen Erweiterung hinzugenommenen Merkmalsausprägungen unter Beibehaltung der Orthogonalität durch tatsächlich vorhandene Merkmalsausprägungen ersetzt.

Nach dem Menübefehl:

Daten > Orthogonales Design > Erzeugen

erscheint die folgende Dialogbox:



Hier werden zunächst die **Faktornamen** mit zugehörigem **Label** eingetragen und dann **hinzugefügt**. Nach dem Markieren eines eingetragenen Faktors kann dann mit einem Klick auf **Werte definieren** die folgende Dialogbox angefordert werden:



Hier werden für den gerade bearbeiteten Faktor die einzelnen **Werte** vereinbart, optional ergänzt durch ein **Label**.

Bei metrischen Eigenschaften (z.B. Preis) wird man als Werte die jeweiligen Variablenausprägungen eintragen. Bei kategorialen Eigenschaften kann man als Werte beliebige Zahlen verwenden, wobei man diese im Rahmen **Ausfüllen** auch automatisch generieren lassen kann. SPSS repräsentiert in der zu erzeugenden Plandatei jede Eigenschaft durch eine *numerische* Variable.

In der Subdialogbox **Optionen** kann man SPSS eine minimal zu erzeugende Anzahl von Produktkarten vorgeben, um die automatisch ermittelte Anzahl zu steigern. Dies erhöht die Präzision der Schätzungen, aber natürlich auch den Zeitaufwand der Untersuchungsteilnehmer (s.o.):



Außerdem ist es hier möglich, so genannte **Prüffälle (Holdout-Karten)** anzufordern. Diese werden von den Untersuchungsteilnehmern wie die regulären Produktkarten beurteilt, aber bei der Conjoint-Analyse *nicht* zur Schätzung der Nutzenwerte herangezogen, sondern später zur Kreuzvalidierung verwendet. Sie stammen aus einem alternativen Orthogonalplan. Man kann dafür sorgen, dass die Prüffälle bei der fortlaufenden Nummerierung aller Produktkarten zufällig untergemischt werden.

Schließlich kann man in der Dialogbox **Orthogonales Design erzeugen** noch einen Startwert für den Pseudozufallszahlengenerator festlegen. Dadurch wird das von SPSS unter Zufallsbeteiligung gewählte orthogonale Design reproduzierbar, so dass eine verloren gegangene Plandatei leicht wiederbeschafft werden kann.

Das von SPSS erzeugte Design kann alternativ in eine SPSS-Datendatei oder in die aktuelle Arbeitsdatei geschrieben werden. Die Zeilen bzw. Fälle entsprechen den Produktkarten und die Spalten den untersuchten Eigenschaften, wobei die Eigenschaften grundsätzlich durch numerische Variablen repräsentiert werden. Wählen Sie nach einem Mausklick auf den Schalter **Datei** im Rahmen **Datendatei** der Dialogbox **Orthogonales Design erzeugen** z.B. die folgende Ausgabedatei

U:\Eigene Dateien\SPSS\Conjoint\teplan.sav

und lassen Sie den Plan per **OK**-Schalter generieren.

Wenn Sie die erzeugte Datendatei öffnen und im Menü **Extras** die Werte-Labels anzeigen lassen, erhalten Sie das folgende Design für unser Beispiel:

	packung	marke	preis	pruefs	geldz	STATUS	CARD
1	A	K2R	1,99 Euro	Ja	Nein	Design	1
2	A	Glory	2,29 Euro	Ja	Ja	Design	2
3	B	K2R	2,29 Euro	Nein	Ja	Design	3
4	A	K2R	1,69 Euro	Nein	Nein	Design	4
5	B	Bissell	1,99 Euro	Nein	Nein	Design	5
6	A	Bissell	2,29 Euro	Ja	Nein	Design	6
7	A	Bissell	1,69 Euro	Nein	Ja	Design	7
8	A	K2R	1,69 Euro	Nein	Ja	Design	8
9	C	Bissell	1,69 Euro	Ja	Ja	Design	9
10	C	K2R	1,69 Euro	Ja	Nein	Design	10
11	A	Glory	1,69 Euro	Nein	Nein	Design	11
12	B	Glory	1,69 Euro	Ja	Nein	Design	12
13	C	K2R	2,29 Euro	Nein	Nein	Design	13
14	B	K2R	1,69 Euro	Ja	Ja	Design	14
15	C	Glory	1,99 Euro	Nein	Ja	Design	15
16	A	K2R	1,99 Euro	Ja	Ja	Design	16
17	A	Bissell	1,99 Euro	Ja	Ja	Holdout	17
18	B	Bissell	1,99 Euro	Ja	Ja	Holdout	18
19	A	Glory	2,29 Euro	Nein	Ja	Holdout	19
20	B	Glory	1,99 Euro	Nein	Ja	Holdout	20

Die Variable STATUS_ enthält für jede Produktkarte eine Typangabe:

- Die 16 Karten vom Typ **Design** (numerischer Wert: Null) stellen das eigentliche Design dar. Sie werden zur Schätzung der Nutzenwerte herangezogen und sollen im Folgenden als **Designkarten** bezeichnet werden.
- Die vier angeforderten Prüffälle (Typ: **Holdout**, numerisch: Eins) stehen am Ende des Kartenstapels, da wir keine zufällige Einordnung gewünscht haben.

Die 20 Design- bzw. Prüfkarten werden von den Teilnehmern ohne Unterscheidung gemeinsam bewertet bzw. angeordnet. Daher werden sie in der CARD_ - Spalte der obigen Datenmatrix gemeinsam nummeriert.

Häufig verwendet man neben den Design- bzw. Holdout-Karten auch noch so genannte **Simulationskarten** mit besonders interessanten Produkten bzw. Eigenschaftskombinationen, deren Gesamtnutzen man auf jeden Fall (ohne Einsatz eines Taschenrechners) in Erfahrung bringen möchte. Die Simulationskarten nehmen an der Untersuchung *nicht* teil, so dass Sie ohne Belastung der Probanden beliebig viele Simulationskarten einbeziehen können. SPSS berechnet jedoch aus den empirisch gewonnenen Teilnutzenwerten auch für die Simulationskarten den Gesamtnutzen. Dies demonstriert eine wichtige Eigenschaft der Conjoint-Analyse mit reduziertem Design: Mit Hilfe ihrer Ergebnisse kann man den Gesamtnutzen beliebiger Produkte schätzen. Die Simulationskarten könne z.B. nach dem Generieren des Designs von Hand am Ende der Plandatei angehängt werden, wobei zu beachten ist:

- In der STATUS_ - Spalte muss **Simulation** (numerisch: Zwei) eingetragen werden.
- In der CARD_ - Spalte sollte für die Simulationskarten eine eigenständige Nummerierung begonnen werden.

In unserem Beispiel sollen zwei Simulationskarten ergänzt werden:

	packung	marke	preis	pruefs	geldz	STATUS_	CARD
20	B	Glory	1,99 Euro	Nein	Ja	Holdout	20
21	C	K2R	1,69 Euro	Nein	Nein	Simulation	1
22	B	Glory	1,69 Euro	Ja	Ja	Simulation	2

Weitere Hinweise:

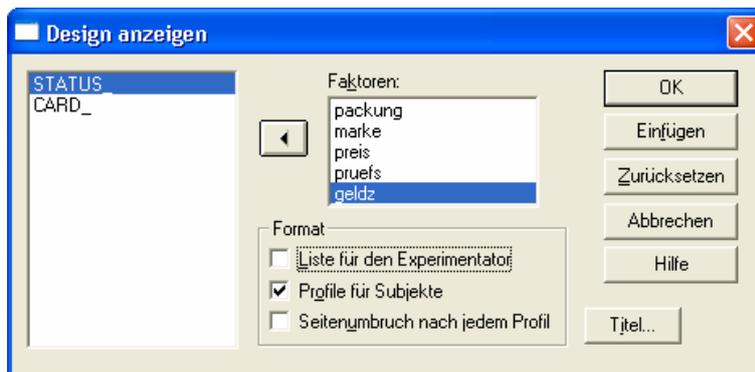
- Gelegentlich erzeugt SPSS identische Designkarten. Wird eine Karte per Hand gelöscht geht die Orthogonalität verloren. Daher ist es sinnvoller, mit einem alternativen Startwert des Pseudozufallszahlengenerators eine neue Plandatei zu erzeugen.
- Das Generieren einer Plandatei setzt im Unterschied zu den meisten anderen SPSS-Prozeduren keine definierte Arbeitsdatei voraus, sondern kann eine solche erzeugen.
- In Abschnitt 5.3.2 wird die ORTHOPLAN-Verwendung per SPSS-Syntax vorgestellt, wobei auch das Erzeugen von Simulationskarten automatisiert werden kann.

3 Erstellung des Untersuchungsmaterials (PLANCARDS)

Nachdem die Design- und Prüfkarten festgelegt sind, müssen diese in präsentable Form gebracht werden. Zu dieser Fleißarbeit kann SPSS durch die Erstellung von „Kartenrohlingen“ einen Beitrag leisten. Laden Sie nötigenfalls die Plandatei **teppplan.sav**, und wählen Sie anschließend den Menübefehl:

Daten > Orthogonales Design > Anzeigen

Wählen Sie in der erscheinenden Dialogbox alle Eigenschaften in der gewünschten Reihenfolge aus, und legen Sie als Ausgabeformat **Profile für Subjekte** fest:



In der **Titel**-Subdialogbox können mehrere Titel- und Fußzeilen festgelegt werden:



Wir beschränken uns auf eine Titelzeile, in der nach der Zeichenfolge "Produkt " eine Kartenummer erscheinen soll. Für die fortlaufende Nummerierung sorgt der Platzhalter

)CARD

Nach dem Anstoßen der Produktion per **OK**-Schalter erhalten wir z.B. für die ersten beiden Karten:

```
Produkt 1

Verpackung A
Markenname K2R
Verkaufspreis 1,99 Euro
Gütesiegel Ja
Geld-zurück-Garantie Nein
```

Produkt 2

Verpackung A
 Markenname Glory
 Verkaufspreis 2,29 Euro
 Gütesiegel Ja
 Geld-zurück-Garantie Ja

Die von SPSS gelieferten Kartenrohlinge kann man mit jedem Textverarbeitungsprogramm importieren und verfeinern, z.B.:

Produkt 1		Produkt 2	
Verpackungsdesign:	A: 	Verpackungsdesign:	A: 
Markenname:	K2R	Markenname:	Glory
Preis:	1,99 €	Preis:	2,29 €
Prüfsiegel?	Ja	Prüfsiegel?	Ja
Geld-zurück-Garantie?	Nein	Geld-zurück-Garantie?	Ja

In der Dialogbox **Design anzeigen** kann neben dem Untersuchungsmaterial auch eine Übersicht für den Untersuchungsleiter angefordert werden, in der technische Angaben enthalten sind (z.B. die Kartentypen).

Man kann sich durchaus vorstellen, dass eine zeitgemäße Marketingstudie an Stelle der traditionellen Karten alternative Produktpräsentationen wählt, z.B. durch multimediale Techniken oder durch reale Produkte (echte Prototypen).

4 Methoden der Datengewinnung

Nachdem das Untersuchungsmaterial vorbereitet worden ist, kann die Datenerhebung stattfinden. Dabei wird man versuchen, eine möglichst große Anzahl von Teilnehmern zu gewinnen, damit generalisierbare Ergebnisse erzielt werden können. Üblicherweise werden Stichproben mit 100 bis 1000 Probanden untersucht.

SPSS-Conjoint kann Rating- und Ordinaldaten verarbeiten:

4.1 Ratingdaten

Ratingdaten entstehen, wenn jede Person alle Produktkarten separat beurteilt, z.B. auf einer Dimension, die durch folgende Pole aufgespannt wird:

1 \cong "Ich würde dieses Produkt auf keinen Fall kaufen"

5 \cong "Ich würde dieses Produkt auf jeden Fall kaufen"

Üblicherweise betrachtet man die resultierenden Daten als intervallskaliert.

Beim Erfassen derartiger Daten mit SPSS definiert man für jede Produktkarte eine Variable und trägt in diese Spalte der Arbeitsdatei die Ratingergebnisse der Teilnehmer ein, z.B.:



	id	prod1	prod2	prod3	prod4	prod5	prod6	prod7	prod8	prod9	prod10	prod11	prod12	prod13
1	1	2	3	3	2	2	3	3	2	5	3	1	2	3
2	2	3	4	4	5	5	5	1	2	2	4	2	3	5

Im CONJOINT-Kommando ist das Datenformat SCORE anzugeben (siehe unten).

4.2 Ordinaldaten

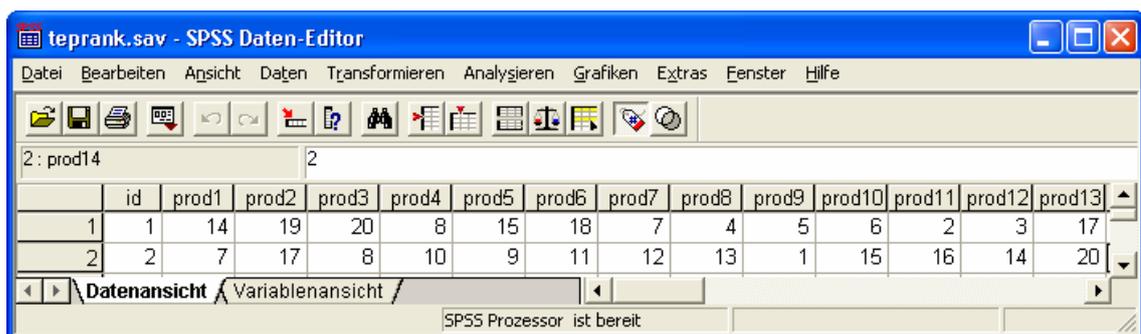
Ordinaldaten resultieren, wenn die Untersuchungsteilnehmer gebeten werden, die Produktkarten in eine persönliche Präferenzordnung zu bringen.

Die Herstellung einer Rangordnung können sich die Probanden eventuell folgendermaßen erleichtern:

- Zunächst Gruppen mit guten, mittleren und schlechten Produkten bilden
- Dann innerhalb der drei Gruppen eine Rangordnung bestimmen

Beim Erfassen von Ordinaldaten mit SPSS sind zwei Wege möglich:

- Man definiert für jede Produktkarte eine Variable und trägt in diese Spalte der Arbeitsdatei die von den Teilnehmern vergebenen Rangplätze ein, z.B.:



	id	prod1	prod2	prod3	prod4	prod5	prod6	prod7	prod8	prod9	prod10	prod11	prod12	prod13
1	1	14	19	20	8	15	18	7	4	5	6	2	3	17
2	2	7	17	8	10	9	11	12	13	1	15	16	14	20

Im CONJOINT-Kommando ist das Datenformat RANK anzugeben (siehe unten).

- Man definiert *für jede Rangposition* eine Variable und trägt in diese Spalte der Arbeitsdatei bei jedem Teilnehmer ein, welche Karte in seinem Stapel diese Rangposition einnimmt. Diese Methode werden wir benutzen und später ausführlich demonstrieren. Im CONJOINT-Kommando ist das Datenformat SEQUENCE anzugeben (siehe unten).

5 Conjoint-Analyse der Präferenzurteile

Zunächst werden die *individuellen* Teilnutzenwerte für jeden einzelnen Probanden ermittelt.

5.1 Schätzmethoden in der SPSS-Prozedur CONJOINT

Unabhängig davon, ob Rating- oder Rangdaten analysiert werden, verwendet SPSS stets einen **metrischen (OLS-)Regressionsalgorithmus** (*Ordinary Least Squares*) zum Schätzen der Teilnutzenwerte. Damit weicht SPSS von der klassischen, nichtmetrischen Regressionsmethodologie der Conjoint-Analyse ab (monotone Varianzanalyse, vgl. z.B. Backhaus et al. 2005).

Die OLS-Methodologie setzt (z.B. bei der Berechnung von Standardfehlern zu geschätzten Parametern) für die (abhängige) Präferenzvariable Intervallskalenqualität und darüber hinaus die varianzhomogene Normalverteilung der Modellresiduen voraus. Diese Voraussetzungen sind bei Verwendung von Rangdaten zumindest fragwürdig.

Außerdem wird bei der metrischen Conjoint-Analyse auf die methodologisch bedeutsame Attraktion der klassischen, nichtmetrischen Methodologie verzichtet, aus **ordinalskalierten** Präferenzurteilen auf messtheoretisch überzeugende Weise **intervallskalierte** Teilnutzenwerte zu gewinnen.

Allerdings hat sich in vielen Methodenstudien erwiesen, dass die OLS-Analyse im Wesentlichen dieselben Ergebnisse liefert wie die nichtmetrischen Techniken. Damit kann die OLS-Methode als praxisbewährte und bequeme Approximation der eigentlich wünschenswerten nichtmetrischen Techniken gelten (Green & Krieger, 1993, S. 478).

Vermutlich machen sich mäßige Fehler in den Schätzungen für die individuellen Teilnutzenwerte bei den letztlich entscheidungsrelevanten aggregierten Daten kaum noch bemerkbar.

5.2 Präferenzmodelle

Wie wir inzwischen wissen, werden in der Conjoint-Analyse Gesamturteile über mehrfaktoriell beschriebene Produkte auf additiv verknüpfte Einzelbeiträge der beteiligten Faktorstufen zurückgeführt. Zunächst soll präzisiert werden, wie nach den Vorstellungen der Conjoint-Analyse bei einer **festen Person** der Gesamtnutzen bzw. die Präferenz p_i des i -ten Produktes bzw. der i -ten Karte erklärt wird:

$$p_i := \beta_0 + \sum_{j=1}^J u_{jk(i)} \quad (1)$$

Wir schreiben in diesem Manuskript generell u_{jk} für den Teilnutzen der k -ten Ausprägung von Merkmal j . In obiger Formel wird durch die Schreibweise $k(i)$ ausgedrückt, dass die bei Produkt i realisierte Ausprägung des j -ten Merkmals gemeint ist. Wie wir gleich sehen werden, kann der Teilnutzen u_{jk} bei einem Merkmal in linearer, quadratischer oder beliebiger Weise von den Ausprägungen abhängen. Die additive Konstante β_0 gibt die Präferenz eines (i. a. hypothetischen) Produktes an, dessen J Teilnutzenwerte sich zu Null aufaddieren.

SPSS schätzt die Präferenz p_i des i -ten Produktes durch seinen **Score**, der bei I realisierten Designkarten folgendermaßen aus dem Rangplatz berechnet wird:

$$\hat{p}_i := \text{Score}(i) := (I + 1) - \text{Rang}(i) \quad (2)$$

Die verrechneten Ränge beruhen ausschließlich auf den I Designkarten.

5.2.1 Diskrete Faktoren

Bei einem diskreten (kategorialen, nominalskalierten) Faktor j mit l_j Ausprägungen versteht SPSS unter dem Teilnutzen u_{jk} der k -ten Ausprägung die Abweichung ihres Präferenzmittelwertes vom *ungewichteten* Präferenzgesamtmittel bzgl. Faktor j .

$$u_{jk} := \bar{p}_{jk} - \frac{1}{l_j} \sum_{v=1}^{l_j} \bar{p}_{jv}$$

Sofern die Kategorien von Faktor j ungleich stark besetzt sind, unterscheidet sich das ungewichtete Präferenzgesamtmittel bzgl. Faktor j von der *gewichteten* Variante (auch als **Grand Mean** bekannt). Auch stimmen bei verschiedenen Faktoren die jeweiligen *ungewichteten* Präferenzgesamtmittel *nicht* überein.

Betrachten wir ein Beispiel, um die verschiedenen Mittelwerte zu erläutern. Die folgende Tabelle enthält für eine konkrete Auskunftsperson¹ zu den drei Ausprägungen des Faktors MARKE die Präferenzmittelwerte und die Häufigkeiten der Ausprägungen, die SPSS im orthogonalen Plan ausgewählt hat:

MARKE	Mittelwert der Scores	N
K2R	8,6250	8
Glory	8,5000	4
Bissel	8,2500	4
Gesamt	8,5000	16

Als *ungewichtetes* Präferenzgesamtmittel bzgl. MARKE erhalten wir:

$$\frac{8,625 + 8,5 + 8,25}{3} = 8,458$$

Als geschätzte Teilnutzenwerte resultieren:

MARKE	\hat{u}_{jk}
K2R	0,1667
Glory	0,0417
Bissel	-0,2083

Das *gewichtete* Präferenzgesamtmittel, also der so genannte **Grand Mean**, ist demgegenüber (siehe **Gesamt** in obiger Tabelle):

$$\frac{8,625 \cdot 8 + 8,5 \cdot 4 + 8,25 \cdot 4}{16} = 8,5$$

und wird bei der Teilnutzenschätzung *nicht* verwendet.

Es ist sinnvoll, in der Definition der Teilnutzenwerte auf die *ungewichteten* Kategorienmittelwerte zurückzugreifen, weil die Teilnutzenwerte sonst von der Verteilung des Faktors abhängen würden, die von der Orthogonalitätsbedingung diktiert wird. Bei wahren Präferenz-Kategorienmittelwerten (2, 6, 10) würden z.B. in Abhängigkeit von den im Orthogonalplan realisierten Häufigkeiten die folgenden gewichteten Mittelwerte und zugehörigen Teilnutzenwerte ergeben:

¹ Es handelt sich um Person Nr. 6, deren Daten später noch genauer analysiert werden.

Häufigkeitsverteilung	Gewichtete Mittelwerte	Kategorienmittel minus gewichtetes Gesamtmittel
(8, 4, 4)	$\frac{(2 \cdot 8 + 6 \cdot 4 + 10 \cdot 4)}{16} = 5$	(-3, 1, 5)
(4, 4, 8)	$\frac{(2 \cdot 4 + 6 \cdot 4 + 10 \cdot 8)}{16} = 7$	(-5, -1, 3)

Auf Grund der in SPSS realisierten Teilnutzendefinition unter Verwendung des ungewichteten Gesamtmittels ($\frac{2+6+8}{3} = 6$) erhalten wir die Teilnutzenwerte (-4, 0, 4).

Bei einem diskreten Faktor j erlaubt die Teilnutzenfunktion u_{jk} in Gleichung (1) für jede Faktorstufe einen eigenständigen Teilnutzenwert, wobei sich alle Werte zu Null addieren müssen. Im Regressionsansatz sind (l_j-1) Kodiervariablen $x_{jk}, k = 1, \dots, l_j-1$, in die Regressionsgleichung aufzunehmen. SPSS verwendet Abweichungskontraste, d.h. die Kodiervariablen werden so gebildet, dass der Regressionskoeffizient $\beta_{jk}, k = 1, \dots, l_j-1$, direkt die Abweichung des Präferenzkategorienmittels vom ungewichteten Präferenzgesamtmittel, also den Teilnutzen der k -ten Faktorstufe, repräsentiert.²

Die geschätzten Teilnutzenwerte zu den l_j Faktorstufen sind also folgendermaßen aus den geschätzten Regressionsgewichten $\hat{\beta}_{jk}, k = 1, \dots, l_j-1$, zu ermitteln:

$$\hat{u}_{jk} := \begin{cases} \hat{\beta}_{jk} & \text{für } k = 1, \dots, l_j - 1 \\ -\sum_{k=1}^{l_j-1} \hat{\beta}_{jk} & \text{für } k = l_j \end{cases} \quad (3)$$

In unserem Teppichreiniger-Beispiel müssen die beiden ersten Faktoren (Verpackung und Markenname) auf diese Weise behandelt werden.

5.2.2 Lineare Faktoren

Bei einem linearen Faktor j wird angenommen, dass die Teilnutzenbeiträge linear mit den Eigenschaftsausprägungen steigen bzw. sinken. Im Regressionsansatz kann ein linearer Faktor also direkt als einzelner Prädiktor x_j mit zugehörigem Regressionsgewicht β_j verwendet werden. Aus dem geschätzten Regressionsgewicht $\hat{\beta}_j$ wird der geschätzte Teilnutzenwert zu der k -ten Eigenschaftsausprägung $x_j^{(k)}$ folgendermaßen von SPSS berechnet:

$$\hat{u}_{jk} := \hat{\beta}_j x_j^{(k)} \quad (4)$$

Es ist also gerade die Veränderung des Gesamtnutzens beim Übergang von $x_j = 0$ zu $x_j = x_j^{(k)}$ (ceteribus paribus).

Damit der Teilnutzenwert der Ausprägung $x_j^{(k)}$ des linearen Faktors j analog zu Abschnitt 5.2.1 als Abweichung der prognostizierten Stufenpräferenz von der prognostizierten Präferenz des ungewichteten Faktormittelwertes

² In Abschnitt 5.3.3 ist die Designmatrix für unser Beispiel mit den Abweichungskodiervariablen zu den Faktoren PACKUNG und MARKE zu sehen.

$$\bar{x}_j := \frac{1}{l_j} \sum_{k=1}^{l_j} x_j^{(k)}$$

interpretiert werden kann, muss der Regressor zum Faktor j **zentriert** werden, d.h. x_j wird ersetzt durch:

$$x_j - \bar{x}_j$$

In unserem Teppichreinigerbeispiel bietet es sich an, Preis als linearen Faktor zu verwenden. Die gerade diskutierte Zentrierung haben wir oben verpasst. Wir werden sie nachholen, wenn uns die „nonzentralen“ Ergebnisse nicht gefallen.

5.2.3 Quadratische Faktoren

Bei einem quadratischen Faktor nimmt der Teilnutzen bei Entfernung von einem idealen bzw. antiidealen Bezugspunkt im quadratischen Sinne ab bzw. zu. Um einen solchen Faktor in Regressionsansatz zu repräsentieren, verwendet SPSS neben dem Prädiktor x_j auch sein Quadrat. Wie die linearen Faktoren sollte man auch die quadratischen zentrieren, so dass SPSS letztlich mit folgenden Prädiktoren arbeitet:

$$(x_j - \bar{x}_j) \text{ und } (x_j - \bar{x}_j)^2$$

Dabei ist \bar{x}_j das ungewichtet gebildete Mittel aller Merkmalsausprägungen von Faktor j . Die beiden zugehörigen Regressionsgewichte sollen mit β_{j1} bzw. β_{j2} bezeichnet werden.

Den Teilnutzenwert zur Eigenschaftsausprägung $x_j^{(k)}$ definiert SPSS wiederum als Abweichung der prognostizierten Stufenpräferenz von der prognostizierten Präferenz des ungewichtet gebildeten Faktormittelwertes:

$$u_{jk} := \beta_{j1} x_j^{(k)} + \beta_{j2} x_j^{(k)2}$$

In unserem Anwendungsbeispiel kommen leider keine quadratischen Faktoren vor.

5.3 Ergebnisse für eine einzelne Auskunftsperson

5.3.1 Das CONJOINT-Kommando

Da wir von jeder Auskunftsperson einen Kartenstapel mit den nach Präferenz geordneten Produktkarten erbeten haben, eignet sich zum Erfassen und Verarbeiten dieser Daten das sequentielle Format am besten (vgl. Abschnitt 4). Wir definieren also 20 Variablen KAP1 bis KAP20 („Karte an Position ...“) und tragen in KAP1 für jeden Fall die Nummer des am stärksten präferierten Produktes ein, in KAP2 die Nummer des Produktes auf dem zweiten Platz usw. So entstand die Datei

tepseq.sav

mit den Daten von 13 Personen:

	id	kurs	kap1	kap2	kap3	kap4	kap5	kap6	kap7
1	1	1195	14	11	12	8	9	10	7
2	2	1195	9	14	16	18	19	20	1
3	3	1195	14	12	1	4	17	16	18
4	4	1195	11	8	4	7	17	16	1
5	5	1195	10	9	8	12	20	18	17

Die eigentliche Conjoint-Analyse ist in der SPSS-Version 13 noch nicht in die grafische Benutzeroberfläche integriert. Wir müssen daher in einem Syntaxfenster (erhältlich mit **Datei > Neu > Syntax**) ein CONJOINT-Kommando erstellen:

```

conjoint
  plan      = 'U:\Eigene Dateien\SPSS\Conjoint\tepplan.sav'
  /data     = *
  /sequence = kap1 to kap20
  /subject  = id
  /factors  = packung marke (discrete) preis (linear less)
             pruefs geldz (discrete more)
  /utility  = 'U:\Eigene Dateien\SPSS\Conjoint\tepnut.sav'
  /plot     = all.
    
```

Im **PLAN-Subkommando** wird mitgeteilt, dass der Studie unser zuvor mit ORTHOPLAN erstelltes und in die Datei **tepplan.sav** gesichertes Design zugrunde liegt. Daraus erfährt SPSS z.B. welche Faktorstufen bei der von unserer ersten Auskunftsperson am stärksten präferierten Karte 14 (s. o.) realisiert wurden.

Im **DATA-Subkommando** wird angegeben, wo sich die Präferenzdaten befinden. In unserem Beispiel gehen wir davon aus, dass sie zuvor in die Arbeitsdatei geladen worden sind, was durch den Stern an Stelle eines Dateinamens ausgedrückt wird. Weil das Lesen aus der aktuellen Arbeitsdatei voreingestellt ist, hätten wir das DATA-Subkommando auch weglassen können.

Im **SEQUENCE-Subkommando** wird die Organisationsform unserer Daten beschrieben (siehe oben und Abschnitt 4).

Das **SUBJECT-Subkommando** benennt die Variable zur Fallidentifikation.

Im **FACTORS-Subkommando** wird für jeden Faktor ein Typ deklariert. Die Bedeutung der Schlüsselwörter DISCRETE und LINEAR dürfte aufgrund der Ausführungen in Abschnitt 5.2 klar sein. Für quadratische Faktoren stehen die Schlüsselwörter IDEAL und ANTIIDEAL zur Verfügung.

Im Beispiel wird bei den Merkmalen PREIS, PRZEFS und GELDZ die Möglichkeit genutzt, für diskrete und lineare Faktoren die postulierte Richtung des Zusammenhangs anzugeben. So ist etwa zu vermuten, dass mit steigendem Preis der Teilnutzen sinkt. Im Ergebnisprotokoll teilt uns SPSS mit, wenn eine Person erwartungswidrig Nutzenverläufe zeigt (Reversal-Analyse).

Durch das **UTILITY-Subkommando** wird SPSS angewiesen, die individuellen Teilnutzenwerte in eine Datei zu sichern, damit sie später z.B. für eine Marktsegmentanalyse zur Verfügung stehen (siehe Abschnitt 5.5).


```

orthoplan factors=
  packung 'Verpackungsdesign' ( 'A' 'B' 'C' )
  marke 'Markenname' ( 'K2R' 'GLORY' 'BISSEL' )
  preis 'Preis' (-0.3 '1,69 Euro' 0 '1.99 Euro' 0.3 '2,29 Euro ' )
  pruefs 'Prüfsiegel?' ( 0 'Nein' 1 'Ja' )
  geldz 'Geld-zurück-Garantie?' ( 0 'Nein' 1 'Ja' )
/holdout=4.

```

Sie finden dieses Kommando mit einigen Ergänzungen in der Datei

tepplan2.sps

und den damit erzeugten Plan in der Datei

tepplan2.sav

Beachten Sie, dass beim Erstellen eines Plans auch (Pseudo)zufall im Spiel ist. In **tepplan2.sps** erhält der Pseudozufallszahlengenerator einen festen Startwert, damit der Plan leicht reproduziert werden kann.

Bei Verwendung des neuen Plans liefert das CONJOINT-Kommando für die Person Nr. 6 ein gut interpretierbares Ergebnis:

```

SUBJECT NAME:      6

Importance  Utility(s.e.)  Factor

      3,24  [ ]          packung  Verpackungsdesign
           ,0000( ,1347)         A
           ,2500( ,1580)         B
           -,2500( ,1580)         C

      2,43  [ ]          marke    Markenname
           ,1667( ,1347)         K2R
           ,0417( ,1580)         Glory
           -,2083( ,1580)         Bissel

      51,88 [ ]          pruefs    Prüfsiegel?
           -4,0000( ,1010) ----  Nein
           4,0000( ,1010) ----  Ja

      25,94 [ ]          geldz     Geld-zurück-Garantie?
           -2,0000( ,1010)  --  Nein
           2,0000( ,1010)  --  Ja

      16,51 [ ]          preis     Preis
           1,2727( ,1219)    -   1,69 Euro
           ,0000( ,0000)    -   1,99 Euro
           -1,2727( ,1219)  -   2,29 Euro
           B = -4,2424( ,4062)

           8,1402( ,1158) CONSTANT

Pearson's R      = ,998              Significance = ,0000

Kendall's tau = 1,000              Significance = ,0000
Kendall's tau = ,667 for 4 holdouts Significance = ,0871

Simulation results:
Card:      1  2
Score:     2,1 14,4

```

Wir stellen fest, dass nunmehr für den günstigsten Preis ein *positiver* Teilnutzenwert resultiert.

Auch die übrigen Teilnutzenwerte sehen plausibel aus:

- Bei PACKUNG und MARKE haben alle Faktorstufen Teilnutzenwerte nahe Null, so dass bei Person 6 eine Variation dieser Faktoren nur wenig Einfluss auf den Gesamtnutzen eines Produktes hat.
- Mit steigendem Preis sinkt erwartungsgemäß der subjektive Gesamtnutzen.
- Die Geld-zurück-Garantie und vor allem das Prüfsiegel sind bei der betrachteten Auskunftsperson wichtige Kaufargumente.

In der SPSS-Ausgabe werden die Teilnutzenwerte nicht nur numerisch sondern auch semigraphisch präsentiert. Später werden attraktivere Grafiken mit Teilnutzenwerten vorgestellt.

Da SPSS bei der CONJOINT-Analyse für jede Auskunftsperson eine OLS-Regressionsanalyse rechnet, können aus den Standardfehlern der Regressionskoeffizienten leicht **Standardfehler der Teilnutzenwerte** berechnet werden. Diese geben Auskunft darüber, wie präzise die Teilnutzenwerte geschätzt werden konnten. Bei einem diskreten Faktor j mit l_j Stufen bildet SPSS $(l_j - 1)$ Kodiervariablen und schätzt dazu die Regressionsgewichte β_{jk} , $k = 1, \dots, l_j - 1$, aus denen gemäß Formel (3) die Teilnutzenwerte berechnet werden. Mit den bekannten Rechenregeln für Varianzen und Kovarianzen kann man dann leicht für die Varianzen der Teilnutzenschätzer nachrechnen:

$$\text{Var}(\hat{u}_{jk}) = \begin{cases} \text{Var}(\hat{\beta}_{jk}) & \text{für } k = 1, \dots, l_j - 1 \\ \sum_{k=1}^{l_j-1} \text{Var}(\hat{\beta}_{jk}) - 2 \sum_{k=1}^{l_j-1} \sum_{v < k} \text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{jv}) & \text{für } k = l_j \end{cases}$$

Den Standardfehler erhält man dann durch Wurzelziehen.

Im Fall von linearen Faktoren erhalten wir nach Formel (4) die Teilnutzenwerte \hat{u}_{jk} sofort aus dem Regressionsgewicht $\hat{\beta}_j$:

$$\hat{u}_{jk} = \hat{\beta}_j x_j^{(k)}$$

Damit gilt für die Varianz des Teilnutzenschätzers:

$$\text{Var}(\hat{u}_{jk}) = x_j^{(k)2} \text{Var}(\hat{\beta}_j)$$

und den zugehörigen Standardfehler erhalten wir durch Wurzelziehen.

Für quadratische Faktoren ergeben sich die Standardfehler der Teilnutzenschätzer analog (siehe SPSS 2004b, S. 6).

Unter den Verteilungsvoraussetzungen der OLS-Regressionsanalyse kann man **Vertrauensintervalle für die Teilnutzenschätzer** bestimmen bzw. **Nullhypothesen über die Teilnutzenwerte** testen. Da wir schon daran gewöhnt sind, mit unserer ordinalen abhängigen Variablen so zu rechnen, als wären die OLS-Voraussetzungen erfüllt, wollen wir noch QAD - Signifikanztests zu den einzelnen Teilnutzenwerten durchführen (QAD = Quick-And-Dirty). Unser Testproblem für den Teilnutzenwert u_{jk} :

$$H_0: u_{jk} = 0 \quad \text{vs.} \quad H_1: u_{jk} \neq 0$$

Wir wollen approximativ auf dem 5%-Niveau testen, und entscheiden uns daher folgendermaßen:

$$|\hat{u}_{jk}| \leq 2\sqrt{\text{Var}(\hat{u}_{jk})} \Rightarrow H_0 \text{ beibehalten}$$

$$|\hat{u}_{jk}| > 2\sqrt{\text{Var}(\hat{u}_{jk})} \Rightarrow H_0 \text{ ablehnen}$$

Bei Person Nr. 6 sind alle Nullhypothesen zu den Teilnutzenwerten der ersten beiden Faktoren (PACKUNG, MARKE) beizubehalten, während die drei restlichen Faktoren signifikante Teilnutzenwerte aufweisen.

Man kann wohl die Zuverlässigkeit bzw. Gewissenhaftigkeit einer Auskunftsperson anhand ihrer Standardfehler beurteilen. Z.B. liefert Person Nr. 2 bei fast allen Teilnutzenwerten erheblich größere Standardfehler ab als Person Nr. 6:

```

SUBJECT NAME:      2

Importance  Utility(s.e.)  Factor
-----
25,09      ,0000(1,1120)      packung      Verpackungsdesign
                2,5000(1,3040)      |--         A
                -2,5000(1,3040)     |--         B
                |                   |         C
36,37      1,5000(1,1120)      marke        Markenname
                -4,3750(1,3040)     ----       K2R
                2,8750(1,3040)     ---        Glory
                |                   |         Bissel
16,31      -1,6250( ,8340)     pruefs       Prüfsiegel?
                1,6250( ,8340)     -         Nein
                |                   |         Ja
11,29      -1,1250( ,8340)     geldz        Geld-zurück-Garantie?
                1,1250( ,8340)     -         Nein
                |                   |         Ja
10,95      1,0909(1,0059)     preis        Preis
                ,0000( ,0000)     |         1,69 Euro
                -1,0909(1,0059)    -         1,99 Euro
                B = -3,6364(3,3529)  -         2,29 Euro
                |                   |
                7,8523( ,9557)     CONSTANT

Pearson's R      = ,859                      Significance = ,0000
Kendall's tau    = ,683                      Significance = ,0001
Kendall's tau    = ,000 for 4 holdouts       Significance = ,5000

Simulation results:
Card:           1  2
Score:         4,1 8,7
    
```

5.3.3 Die Regressionsanalyse zu Fall Nr. 6

Um die von SPSS angewandte OLS-Conjoint-Analyse vollends transparent zu machen, soll nun die komplette Analyse für Person Nr. 6 mit der SPSS-Regressionsprozedur nachgerechnet werden. Zunächst wird die Regressionsgleichung für die Stichprobendaten in Matrixschreibweise angegeben:

Koeffizienten(a)

Modell	Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
	B	Standardfehler	Beta		
1 (Konstante)	8,140	,116		70,303	,000
pack1	1,288E-16	,135	,000	,000	1,000
pack2	,250	,158	,038	1,583	,152
mark1	,167	,135	,030	1,237	,251
mark2	,042	,158	,006	,264	,799
Preis	-4,242	,406	-,229	-10,444	,000
Prüfsiegel?	-4,000	,101	-,868	-39,587	,000
Geld-zurück-Garantie?	-2,000	,101	-,434	-19,794	,000

a Abhängige Variable: scores

5.3.4 Relative Bedeutung der Faktoren

Weisen alle Stufen eines Faktors j nahezu identische Nutzenwerte auf, dann ist dieser Faktor ohne Bedeutung für die Attraktivität der Produkte, weil die Auswahl der Stufe von Faktor j den Gesamtnutzen eines Produktes kaum beeinflusst. Ein gutes Maß für die Bedeutung eines Faktors ist demzufolge der Abstand zwischen dem größten und dem kleinsten Teilnutzenwert der Faktorstufen. Wenn man diesen Wert durch die Summe der MAX-MIN - Abstände aller Faktoren dividiert, erhält man ein Maß für die relative Bedeutung, das von SPSS noch durch Multiplikation mit 100 auf eine Prozentskala gebracht wird:

$$imp_j := 100 \frac{range_j}{\sum_{j=1}^M range_j}, \text{ mit } range_j := \max_k(u_{jk}) - \min_k(u_{jk})$$

Wie die linke Spalte der Ergebnisausgabe zu unserem Fall Nr. 6 zeigt, werden seine Entscheidungen dominiert von den Faktoren PRUEFS, GELD und PREIS mit 51,88, 25,94 und 16,51 Prozent der Gesamtunterschiedlichkeit.

5.3.5 Fit-Indikatoren

Bei der Conjoint-Analyse soll das (meist ordinale) Gesamturteil möglichst gut durch Addition der geschätzten Teilnutzenwerte reproduziert werden. Folglich sollte der rekonstruierte Gesamtnutzen möglichst hoch mit dem beobachteten Gesamturteil korrelieren. SPSS berechnet für diese beiden Größen die Korrelationsmaße nach Pearson (metrisch) und Kendall (ordinal) und gibt jeweils die einseitigen Überschreitungswahrscheinlichkeiten zum Test mit der Alternativhypothese

$$H_0 : \rho > 0$$

an. Die Berechnungen werden getrennt für die Design- und die Holdout-Karten ausgeführt. Während die Korrelationen für die Designkarten aufgrund der meist hohen Anzahl von Regressoren in Relation zur Anzahl von Fällen mit einem positiven Bias belastet sind, also zu hoch liegen, liefern die Holdout-Karten erwartungstreue Schätzungen, die allerdings aufgrund des meist geringen Stichprobenumfangs eine große Fehlerstreuung aufweisen.

Bei unserem Fall Nr. 6 resultiert für die Designkarten eine Kendall-Rangkorrelation von 1,0 und für die Holdout-Karten ein Wert von 0,67, was wohl ein akzeptables Ergebnis darstellt. Bei den Ergebnissen zu den vier

Holdout-Karten ist zu beachten, dass Korrelationen aus einer derart kleinen Stichprobe stark zufallsbelastet sind, weshalb nur sehr hohe Werte signifikant werden.

Für den Fall Nr. 2, der schon durch sehr große Standardfehler aufgefallen war, erhalten wir bei den Designkarten einen Wert von 0,68 und bei den Holdout-Karten einen Wert von 0,0.

Für die 16 Designkarten sind bei Fall Nr. 6 in folgender Tabelle die Beobachtungswerte und die Modellprognosen gegenübergestellt:

PRODUKT	RRANG	SCORES	PREDS6
1,00	7,000	10,00	10,30682
2,00	4,000	13,00	12,90909
3,00	12,000	5,00	5,28409
4,00	13,000	4,00	3,57955
5,00	15,000	2,00	2,18182
6,00	8,000	9,00	8,65909
7,00	10,000	7,00	7,20455
8,00	9,000	8,00	7,57955
9,00	2,000	15,00	14,95455
10,00	6,000	11,00	11,32955
11,00	14,000	3,00	3,45455
12,00	5,000	12,00	11,70455
13,00	16,000	1,00	,78409
14,00	1,000	16,00	15,82955
15,00	11,000	6,00	5,93182
16,00	3,000	14,00	14,30682

In der Spalte RRANG stehen die nach Streichung der Holdout-Karten neu berechneten Ränge, aus denen die SCORES gemäß Formel (2) berechnet werden. In der letzten Spalte befinden sich die vom SPSS-Programm **regs6.sps** ermittelten Modellprognosen.

5.3.6 Simulationsergebnisse

Am Ende der personspezifischen Ausgaben erscheint für die beiden Simulationskarten der Gesamtnutzen aufgrund der geschätzten Regressionsgleichung. Für unseren Fall Nr. 6 ist die zweite Simulationskarte eindeutig attraktiver:

```
Simulation results:
Card:      1      2
Score:    2,1 14,4
```

5.3.7 Verletzungen der modellgemäßen Ordnung (Reversals)

Ist für einen Faktor (linear oder diskret) eine Ordnung definiert (MORE oder LESS bei der Faktor-Definition im CONJOINT-Kommando), dann protokolliert SPSS bei jeder Auskunftsperson jede Verletzung dieser Ordnung. In unserem Datensatz schneidet z.B. bei der Person 12 die *niedrigste* Preiskategorie am schlechtesten ab, was der postulierten Richtung LESS beim Faktor PREIS widerspricht:

Importance	Utility(s.e.)	Factor	** Reversed (1 reversal)
9,84		preis	** Preis
	-,8182(,5828)	-	1,69 Euro
	,0000(,0000)		1,99 Euro
	,8182(,5828)	-	2,29 Euro
	B = 2,7273(1,9426)		

Die erwartungsdiskrepanten Teilnutzenwerte sind nicht signifikant (vgl. Abschnitt 5.3.2). Trotzdem können wir sie vorsichtig als Hinweis auf das Misstrauen mancher Kunden gegenüber allzu preiswerten Angeboten interpretieren.

5.4 Aggregierte Ergebnisse

5.4.1 Teilnutzenwerte und Wichtigkeiten

Natürlich muss sich die Produktplanung an möglichst repräsentativen Ergebnissen orientieren. SPSS bietet eine denkbar einfache Aggregation der individuellen Teilnutzenwerte und Faktorwichtigkeiten an: Es werden einfach die arithmetischen Mittelwerte über alle Auskunftspersonen gebildet. Wir erhalten:³

```

SUBFILE SUMMARY

Averaged
Importance   Utility   Factor
-----
31,80       1,3590   packung
              -,2179   |--
              -1,1410  |--   Verpackungsdesign
                                   A
                                   B
                                   C
11,04       ,1538   marke
              -,4423   |
              ,2885   |   Markenname
                                   K2R
                                   Glory
                                   Bissel
17,89      -1,0769  pruefs
              1,0769  |--   Prüfsiegel?
                                   Nein
                                   Ja
9,26       -,7212   geldz
              ,7212   |   Geld-zurück-Garantie?
                                   Nein
                                   Ja
30,02      2,2867  preis
              ,0000  |----
              -2,2867  |----   Preis
                                   1,69 Euro
                                   1,99 Euro
                                   2,29 Euro
              B = -7,6224
              7,5501  CONSTANT

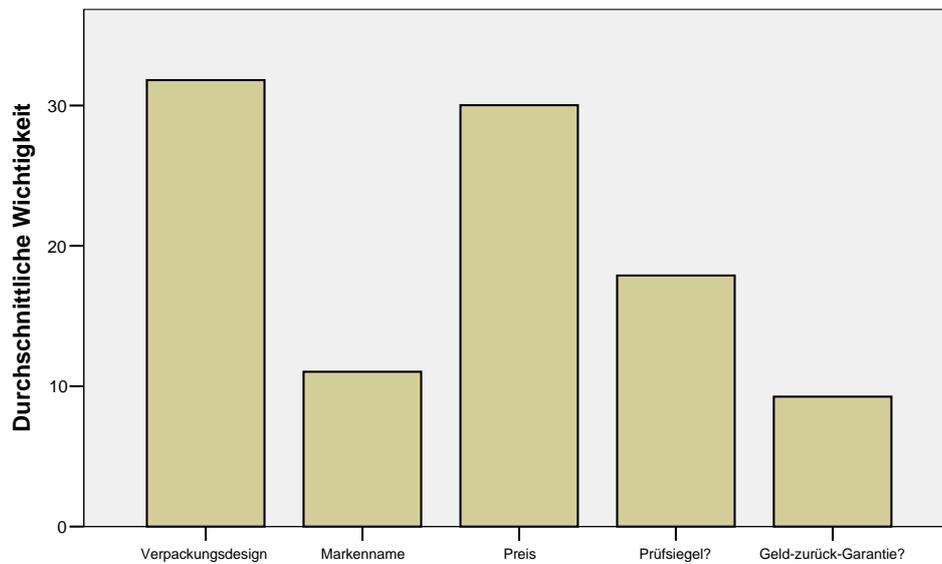
Pearson's R   = ,989           Significance = ,0000
Kendall's tau = ,879           Significance = ,0000
Kendall's tau = 1,000 for 4 holdouts Significance = ,0208

Simulation results:
Card:      1   2
Score:    4,8 8,7
    
```

Die wesentlichen Ansatzpunkte bei der Produktplanung sind in unserem Beispiel also der Preis und das Verpackungsdesign, was in folgender Grafik mit den aggregierten relativen Wichtigkeiten besonders deutlich wird:

³ Der extreme Wert bei Kendall's Tau für die Holdout-Karten (1,0) erscheint unplausibel. Leider ist der Formelsammlung (SPSS 2004b) keine Berechnungsvorschrift für diesen aggregierten Wert zu entnehmen.

Wichtigkeitszusammenfassungen



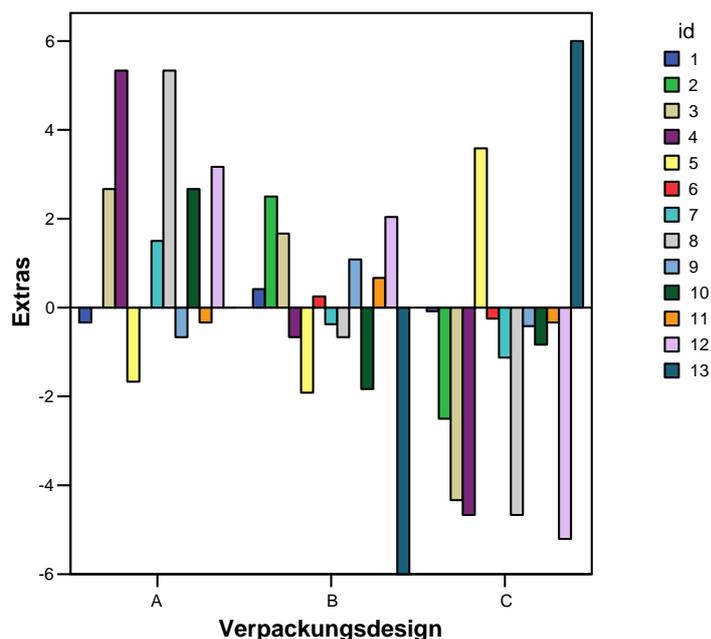
Während der Preisgestaltung enge ökonomische Grenzen gesetzt sind, kann man die Empfehlung bezüglich der Verpackung wohl gut umsetzen. Bei der Inspektion der aggregierten Teilnutzenwerte zeigt sich ein plausibles Ergebnis: Teppichreiniger in Milchkartons (Packungsdesign C) sind recht schwer abzusetzen. Natürlich wird man den aussichtsreichsten Markennamen (*Bissel*) trotz der geringen Wichtigkeit des Faktors verwenden, wenn die Auswahl kostenneutral erfolgen kann. Bei den Faktoren PRUEFS und GELDZ müssen wieder ökonomische Kalkulationen berücksichtigt werden.

Aufgrund des Subkommandos

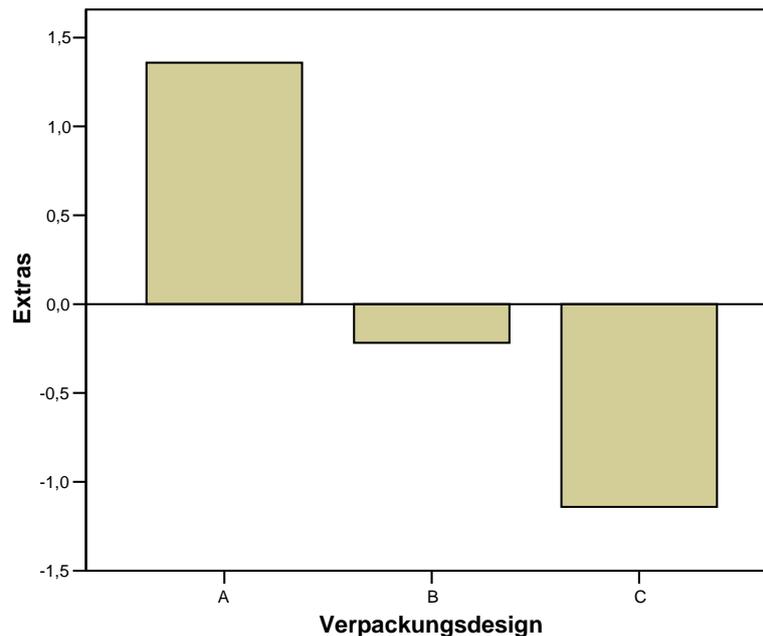
```
/plot = all
```

im CONJOINT-Kommando (vgl. Abschnitt 5.3.1) erhält man nicht nur obige Abbildung mit den gemittelten Wichtigkeiten, sondern außerdem ...

- für jeden Faktor ein Balkendiagramm mit den individuellen Teilnutzenwerten seiner Ausprägungen, z.B. beim Verpackungsdesign:



- für jeden Faktor ein Balkendiagramm mit den mittleren Teilnutzenwerten seiner Ausprägungen, z.B. beim Verpackungsdesign:



Die durch einfache Mittelwertbildung gewonnenen aggregierten Nutzenwerte kann man auch in der so genannten **gemeinsamen Conjoint-Analyse** ermitteln (vgl. Backhaus et al. 2005, S. 582 u. 600). Dabei werden die Regressionsvariablen aller Auskunftspersonen untereinander geschrieben, so dass in unserem Beispiel eine SPSS-Arbeitsdatei mit $20 \times 13 = 260$ Beobachtungseinheiten entsteht, die dann einer OLS-Regressionsanalyse unterzogen werden. Folglich stammen die oben mitgeteilten aggregierten Teilnutzenwerte ebenfalls von den Koeffizienten einer Regressionsanalyse ab. Trotzdem gibt SPSS zu Recht keine Standardfehler aus, weil die 260 Beobachtungseinheiten nicht unabhängig voneinander sind.

5.4.2 Reversals

SPSS bietet eine Übersicht zu erwartungswidrigen Teilnutzenwerten:

```

SUBFILE SUMMARY

Reversal Summary:

    7 subjects had  1 reversals

Reversals by factor:

    pruefs    4
    preis     2
    geldz     1
    marke     0
    packung   0
    
```

Reversal index:

Page	Reversals	Subject
61	0	1
62	0	2
63	1	3
64	1	4
65	0	5
66	0	6
67	1	7
68	1	8
69	0	9
70	1	10
71	0	11
72	1	12
73	1	13

Unter Reversals by factor ist der etwas verblüffende Befund nachzulesen, dass bei immerhin 4 von 13 Teilnehmern ein Prüfsiegel die Attraktivität eines Produktes zu mindern scheint.

5.4.3 Präferenzen für die Simulationsprodukte

Abschließend erhalten wir Ergebnisse zu der Frage, wie die Chancen in einem Wettbewerb zwischen den Simulationsprodukten verteilt sind. SPSS bietet drei Schätzmethode für die Wahrscheinlichkeiten, mit denen die Simulationsprodukte von einer beliebigen Person bevorzugt werden (vgl. SPSS 2004b, S. 7):

- **Max Utility**

Die Wahrscheinlichkeit der i -ten Simulationskarte wird durch die relative Häufigkeit der Auskunftspersonen mit dem höchsten Gesamtnutzen bei der i -ten Simulationskarte geschätzt.

- **BTL**

Bei jeder Auskunftsperson wird die Wahrscheinlichkeit, die i -te Simulationskarte zu bevorzugen, geschätzt durch den Quotienten aus dem Gesamtnutzen der i -ten Simulationskarte und der Summe aller Simulations-Gesamtnutzenwerte:

$$\frac{\hat{p}_i}{\sum_{i=1}^I \hat{p}_i}$$

Die so geschätzten individuellen Wahrscheinlichkeiten werden gemittelt. Dabei wird allerdings Person Nr. 12, die uns in Abschnitt 5.3.7 durch die Bevorzugung *hoher* Preise aufgefallen war, ausgeschlossen, weil ihr Modell für das erste Simulationsprodukt einen negativen Nutzen produziert, was im BTL-Algorithmus zu einer negativen Wahrscheinlichkeitsschätzung führt.

- **Logit**

Diese Schätzung ist analog zum BTL-Wert definiert, wobei jedoch die Nutzenwerte zuvor per Exponentialfunktion transformiert werden. Auch dabei kann Person 12 mit dem negativen Nutzenwert für das erste Simulationsprodukt nicht mitwirken.

In unserem Beispiel hätte wohl das zweite Produkt die besten Marktchancen:

Simulation Summary (13 subjects/ 12 subjects with non-negative scores)

Card	Max Utility	BTL	Logit
1	15,38%	36,66%	22,67%
2	84,62	63,34	77,33

5.5 Die Datei mit den Nutzenwerten

Die im CONJOINT-Kommando zu unserem Beispiel angeforderte Datei **tepnut.sav** enthält für jede Auskunftsperson:

- Alle Teilnutzenwerte (inkl. Regressionskonstante)
- Die Modellprognosen zu den Scores der Design- und Holdout-Karten
- Die Modellprognosen der Gesamtnutzenwerte zu den Simulationskarten

Sofern weitere Daten zu den Auskunftspersonen vorliegen, kann man z.B. bestimmte Teilnutzenwerte als abhängige Variablen in einer Regressionsanalyse näher untersuchen.

6 Literatur

- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2005). *Multivariate Analysemethoden* (11. Aufl.). Berlin: Springer.
- Green, P.E. & Krieger A.M. (1993). Conjoint analysis with product-positioning applications. In: J. Eliashberg & G.L. Lilien (Eds.). *Marketing*. Amsterdam: North-Holland.
- Green, P.E. & Wind Y. (1973). *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, IL: Dryden Press.
- Müller-Hagedorn, L., Sewing, E. & Toporowski, W. (1993). Zur Validität von Conjoint-Analysen. *ZfbF*, 45, 123-148.
- SPSS, Inc. (1997). *SPSS Conjoint 8.0*. Chicago, IL.
- SPSS, Inc. (2004a). *SPSS 13. Statistical Algorithms. ORTHOPLAN*. Chicago, IL.
- SPSS, Inc. (2004b). *SPSS 13. Statistical Algorithms. CONJOINT*. Chicago, IL.

7 Stichwortregister

Abweichungskontraste	16	ORTHOPLAN	6
Additive Conjoint Measurement	5	ORTHOPLAN-Kommando	20
Aggregieren	6	PLANCARDS	10
Aggregierte Ergebnisse	27	PLAN-Subkommando	18
BTL	30	PLOT-Subkommando	19
CONJOINT-Kommando	18, 26	Prüffälle	8
DATA-Subkommando	18	Pseudozufallszahlengenerator	8
Designkarten	9	Ratingdaten	12
Diskrete Faktoren	15	Reversal-Analyse	18, 26, 29
Erhebungsmethoden	12	Scores	14
FACTORS-Subkommando	18	SEQUENCE-Subkommando	18
Fit-Indikatoren	25	Signifikanztest für Teilnutzenwerte	22
Gemeinsame Conjoint-Analyse	29	Simulationskarten	9, 30
Holdout-Karten	8	Standardfehler der Teilnutzenwerte	21
Interaktionen	6	SUBJECT-Subkommando	18
Kendall-Rangkorrelation	26	Symmetrisches Design	6
Kreuzvalidierung	8	Syntaxfenster	18
LESS	26	Teilnutzen	14
Lineare Faktoren	16	Teilnutzenwerte	5
Logit	30	bei diskreten Faktoren	15
Max Utility	30	bei linearen Faktoren	16
Metrische Conjoint-Analyse	14	bei quadratischen Faktoren	17
Monotone Varianzanalyse	14	Utility-Datei	31
MORE	26	UTILITY-Subkommando	19
Nichtmetrische Conjoint-Analyse	14	Vertrauensintervall für Teilnutzenwerte	22
OLS-Conjoint-Analyse	14	Wichtigkeit der Faktoren	25
Ordinaldaten	12	Zufallszahlengenerator	8
Orthogonales Design	6		