# Paired Comparison Preference Models

## Practicals and Home work IV

Regina Dittrich & Reinhold Hatzinger
Department of Statistics and Mathematics, WU Vienna





**Exercise 1:** Data file: `muspc.Rdata`

This is the same data set as `music.dat` (practicals day II) without NAs

| object | | object | |
|---|---|---|---|
| BIGB | bigband music | LATI | latin music |
| BLUG | bluegrass music | MOOD | easy listening music |
| COUN | country western music | NEWA | new age music |
| BLUE | blues music | OPER | opera |
| MUSI | broadway musicals | RAP | rap music |
| CLAS | classical music | REGG | reggae music |
| FOLK | folk music | CONR | contemporary rock music |
| GOSP | gospel music | OLDI | oldies rock music |
| JAZZ | jazz | HVYM | heavy metal music |

Subject variables:

```
MARITAL (marital status)     SEX (respondents sex)       RACE (race of respondent)
     1  MARRIED                  1  MALE                     1  WHITE
     2  WIDOWED                  2  FEMALE                   2  BLACK
     3  DIVORCED                                            3  OTHER
     4  SEPARATED
     5  NEVER MARRIED
```





```
XNORCSIZ (size of living location)              AGE (in years)
    1  CITY GT 250000      6  UNINC,MED CITY
    2  CITY,50-250000      7  CITY,10-49999      EDUC (highest year of school completed)
    3  SUBURB, LRG CITY    8  TOWN GT 2500
    4  SUBURB, MED CITY    9  SMALLER AREAS
    5  UNINC,LRG CITY     10  OPEN COUNTRY
```

**Tasks:**

**1** select some objects (4 – 5) and the corresponding comparisons
**2** select one or two subject covariates (and recode them)
**3** from 1 and 2 generate an new data file
**4** calculate a common undecided term and add it to the design data frame
**5** find a (minimal) fitting PATTERN model using subject covariates
**6** plot the worth for this model
**7** define one or two object covariates (e.g. easy listening / not easy listening)
**8** fit a pattern model with object specific covariates only
**9** find a model for both subject and object specific covariates
**10** fit the same model using the LLBT (llbt.design, gnm)
**11** fit interaction parameters to the PATTERN model
**11** compare the results, draw a plot for models with and without interactions

*Hints*

**1** see solutions day II
**4** sum up all undecided covariates
**11** different plots are possible:
(i) side-by-side worth plots (using `option(mfrow=c(1,2))`
(ii) using `barplot()`
(iii) scatterplot with 45° line





**Exercise 2:**

generate some PC data, subject and object variables, e.g.,

```
> set.seed(12345)                  # for replicability
> ex<-simPC(3,100,c(3,2,1))        # PC data
> s1<-rnorm(100)                   # numeric subject variable
> s2<-rnorm(100)
> c1<-factor(sample(1:2,100,r=T)) # categorical subject variable
> c2<-factor(sample(1:2,100,r=T))
> exf<-data.frame(ex,s1,s2,c1,c2) # data frame
> O1<-c(1,1,0)                     # object specific variable
> O2<-c(0,1,1)
> objs<-data.frame(O1=O1,O2=O2)    # object data frame
```

**Tasks:**
for each of the follwing compute a PATTERN model and an LLBT, calculate worth, and plot the results (use `design` functions and `gnm`) – you have to fit at least 22 models

**1** simple model without covariates
**2** model with 1 categorical subject covariate, interaction with one/all object(s)
**3** as 2, but 2 categorical subject covariates
**4** as 2, but 1/2 numerical subject covariates
**5** as 2 – 4, but introduce some interactions between subject covariates
**6** as 1, but 1/2 object specific covariates
**7** combination of 4 and 5

## For nerds

**Morning Session:**
**Show that the LLBT and the independence PATTERN model are equivalent**

**Tasks:**

1. generate a data file for 3 objects using `simPC()`
2. fit an LLBT and an independence PATTERN model (estimates are the same)
3. show that $P(Y_{ij})$ is the same in both models:
   calculate, e.g., $P(Y_{12})$, from the LLBT and from the PATTERN model
   both (i) analytically and (ii) using R
   hint: use fitted values, sum up corresponding pattern probabilities
   (e.g. calculate $P(Y_{13} = 1, Y_{23} = 1)$, $P(Y_{13} = 1, Y_{23} = -1)$, etc. and use this)

**Afternoon Session:**
**Investigate intransitive patterns**

1. use 4 items from the `teacher7` data set (day III) and remove NAs
2. generate a data frame with two variables:
   – observed count for all patterns (y in a design data frame)
   – is pattern transitive (yes/no)
   the rows are the PC patterns:
   – generate a character vector with patterns and use it as row names
3. investigate the most frequent intransitive patterns, can you interpret the intransitivities?
4. remove the intransitive observations
5. fit LLBTs for both data sets (with/without intransitive observations)
6. is the goodness-of-fit different? why?





## Exercise 3a:

**Collect ranking data:**

if everybody (during lunch time, by phone, on the street, ...) collects ranking data plus 2 or 3 subject variable from 3 - 5 persons we obtain a nice data set for Exercise 3b (afternoon session)

**Proposal for question:**

Was ist Ihnen/Dir bei einem Restaurantbesuch am wichtigsten, zweitwichtigsten, etc.

a) Qualität der Speisen und Getränke
b) Freundlichkeit des Personals
c) günstige Preise
d) Schnelligkeit des Services

**Proposal for subject variables:**

`SEX`

`AGE` (if possible in years)

`STUDENT` (yes/no)

?





## Exercise 3b:

**Analyse the collected data**

**Tasks:**

1. find a suitable model using the LLBT
2. find a suitable model using the PATTERN approach
3. compare the results

for 1) you have to transform the rankings into PCs, e.g.,

```
> dat <- read.table()
> nobj <- 4
> resdat <- NULL
> for (j in 2:nobj) for (i in 1:(j - 1)) {
+     v <- ifelse(dat[i] < dat[j], 1, ifelse(dat[i] == dat[j],
+         0, -1))
+     resdat <- cbind(resdat, v)
+ }
```





## Exercise 4: (Data file: `princ.Rdata`)

– TALIS (Teaching And Learning International Survey) – OECD 2009
– 24 countries, over 4 000 schools and over 70 000 teachers
– data contains evaluation of school principals on importance of certain tasks

**Responses: Importance of Tasks** (Response Format: Piling)

Six tasks were printed on cards. The cards had to be sorted into piles according to their importance. The number of piles was not predetermined. For each task the pile number was recorded. Higher values correspond to higher importance.

`MINSTRY`
: An important part of my job is to ensure ministry- approved instructional approaches are explained to new teachers, and that more experienced teachers are using these approaches.

`GOALS`
: An important part of my job is to ensure that teachers are held accountable for the attainment of the school's goals.

`PARENTS`
: An important part of my job is to present new ideas to the parents in a convincing way.

`RULES`
: It is important for the school that I see to it that everyone sticks to the rules.

`ADMIN`
: It is important for the school that I check for mistakes and errors in administrative procedures and reports.

`TIMTAB`
: An important part of my job is to resolve problems with the timetable and/or lesson planning.

Subject variables:

SEX

| | |
|---|---|
| 1 | Female |
| 2 | Male |

AGE

| | |
|---|---|
| 1 | Under 40 |
| 2 | 40-49 |
| 3 | 50-59 |
| 4 | 60+ |

SCHOOL

| | |
|---|---|
| 1 | A public school |
| 2 | A private school |

LOCSIZ

| | |
|---|---|
| 1 | A village or rural area (< 3000 people) |
| 2 | A small town (3 000 - about 15 000 people) |
| 3 | A town (15 000 - about 100 000 people) |
| 4 | A city (100 000 - about 1 000 000 people) |
| 5 | A large city with over 1 000 000 people |

EXP

| | |
|---|---|
| 1 | This is my first year |
| 2 | 1-2 years |
| 3 | 3-5 years |
| 4 | 6-10 years |
| 5 | 11-15 years |
| 6 | 16-20 years |
| 7 | More than 20 year |

CNTRY

| | | | | | |
|---|---|---|---|---|---|
| 1 | Australia | 9 | Iceland | 17 | Netherlands |
| 2 | Austria | 10 | Ireland | 18 | Norway |
| 3 | Belgium (Flemish) | 11 | Italy | 19 | Poland |
| 4 | Brazil | 12 | Korea | 20 | Portugal |
| 5 | Bulgaria | 13 | Lithuania | 21 | Slovak Republic |
| 6 | Denmark | 14 | Malaysia | 22 | Slovenia |
| 7 | Estonia | 15 | Malta | 23 | Spain |
| 8 | Hungary | 16 | Mexico | 24 | Turkey |





**Tasks:**

1. decide on the method to analyse the data (if necessary, transform the data accordingly)
2. select some subject covariates (and recode them)
3. decide if you want to compare countries, if yes choose some
4. from 1 and 3 generate an new data file
5. perform an analysis (proceed as in previous exercises)