## Overview of Lecture

- Advantages and disadvantages of within subjects designs
- One-way within subjects ANOVA
- Two-way within subjects ANOVA
- The sphericity assumption

- NB repeated measures is synonymous with within subjects



## Advantages of within subjects design

- The main advantage of the within subjects design is that it controls for individual differences between participants.
  - In between groups designs some fluctuation in the scores of the groups that is due to different participants providing scores
  - To control this unwanted variability participants provide scores for each of the treatment levels
  - The variability due to the participants is assumed not to vary across the treatment levels



## Within Subjects or Repeated Measures Designs

- So far the examples given have only examined the between groups situation
  - Different groups of participants randomly allocated to different treatment levels
- Analysis of variance can also handle within subject (or repeated measures of designs)
  - A groups of participants all completing each level of the treatment variable



## Disadvantages of within subjects designs

- Practice Effects
  - Participants may improve simply through the effect of practice on providing scores.
  - Participants may become tired or bored and their performance may deteriorate as the provide the scores.
- Differential Carry-Over Effects
  - The provision of a single score at one treatment level may positively influence a score at a second treatment level and simultaneously negatively influence a score at a third treatment level
- Data not completely independent (assumption of ANOVA)
- Sphericity assumption (more later)
- Not always possible (e.g. comparing men vs women)



## Partitioning the variability

- We can partition the basic deviation between the individual score and the grand mean of the experiment into two components

$$AS - \bar{T} = (AS - \bar{A}) + (\bar{A} - \bar{T})$$

Basic Deviation ($AS - \bar{T}$); Within Treatment Deviation ($AS - \bar{A}$); Between Treatment Deviation ($\bar{A} - \bar{T}$)

- Between Treatment Component - measures effect plus error
- Within Treatment Component - measures error alone



## Partitioning the variability

- The Within Treatment Component

$$AS - \bar{T}$$

- estimates the error
- At least some of that error is individual differences error, i.e., at least some of that error can be explained by the subject variability
- In a repeated measures design we have a measure of subject variability

$$\bar{S} - \bar{T}$$

## Partitioning the variability

- If we subtract the effect of subject variability away from the within treatment component

$$(AS - \bar{T}) - (\bar{S} - \bar{T})$$

- We are left with a more representative measure of experimental error
- This error is known as the residual
- The residual error is an interaction between
  - The Treatment Variable
  - The Subject Variable



## Calculating mean squares

- Mean square estimates of variability are obtained by dividing the sums of squares by their respective degrees of freedom
  - Main Effect $MS_A = SS_A / df_A = SS_A / (a - 1)$
  - Subject $MS_S = SS_S / df_S = SS_S / (s - 1)$
  - Error (Residual) $MS_{AxS} = SS_{AxS} / df_{AxS} = SS_{AxS} / (a - 1)(s - 1)$



## Calculating F-ratios

- We can calculate F-ratios for both the main effect and the subject variables

$$F_A = MS_A / MS_{AxS}$$

$$F_S = MS_S / MS_{AxS}$$



## Example one-way within subjects design

- An experimenter is interested in finding out if the time taken to walk to the Coates building is influenced by practice

| | n=1 | n=2 | n=3 | n=4 |
|---|---|---|---|---|
| s1 | 40 | 20 | 10 | 10 |
| s2 | 30 | 25 | 15 | 10 |
| s3 | 25 | 20 | 10 | 5 |
| s4 | 25 | 20 | 15 | 10 |
| s5 | 20 | 15 | 10 | 5 |



## Experstat output - Anova summary table

```
Within Subjects Design (alias Randomized Blocks)

Source of    Sum of      df   Mean      F       p
Variation    Squares          Squares

Subjects       170.000    4    42.500
A (Practice)  1180.000    3   393.333  27.765  0.0000
(Error AxS)    170.000   12    14.167
```



## Analytical Comparisons

- As with a one-way between groups analysis of variance a significant main effect means
  - There is a significant difference between at least one pair of means
- A significant main effect doesn't say where that difference lies
- We can use planned and unplanned (post hoc) comparisons to identify where the differences are

## Experstat output - tukey tests

```
Comparisons Between Means for Selected Factor(s)

* = p<0.05 ** = p<0.01 *** = p<0.001 **** = p<0.0001

Tukey test

Comparison between levels of Practice

n=1          vs n=2          q =    4.75     *
n=1          vs n=3          q =    9.51     ***
n=1          vs n=4          q =   11.88     ***
n=2          vs n=3          q =    4.75     *
n=2          vs n=4          q =    7.13     **
n=3          vs n=4          q =    2.38
```



## Reporting the results

- Give a table showing the means and standard errors (don't forget to label the table)

| n=1 | n=2 | n=3 | n=4 |
|---|---|---|---|
| 28 (3.391) | 20 (1.581) | 12 (1.225) | 8 (1.225) |

Table 1 shows the means (and standard errors) of the number of minutes taken to arrive at the Coates building for the first four attempts.

- Write a short summary
- "There was a significant main effect of practice ($F_{3,12}$=27.765, Mse=14.167, p<0.0001). Post hoc tukey tests (p≤0.05) showed that all four levels of practice were significantly different with the exception of levels n=3 and n=4 which did not differ significantly."



## A main effect of the subject variable

- A significant main effect of the subject variable is common and usually is not a problem
- A significant main effect of the subject variable is a problem
  - when specific predictions are made about performance
  - when there is a hidden aptitude treatment interaction



## An example aptitude treatment interaction

- A simple repeated measures experiment. Participants completed three graded crosswords and the time taken was measured.

| Easy | Medium | Difficult |
|---|---|---|
| 4.5 | 6.8 | 9.5 |

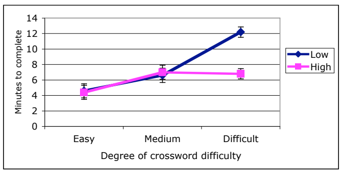
```
Within Subjects Design (alias Randomized Blocks)
Source of      Sum of      df    Mean       F        p
Variation      Squares           Squares

Subjects        25.867      9     2.874
A (Crossword)  125.267      2    62.633    15.500   0.0001
(Error AxS)     72.733     18     4.041
```

- Fcritical with 9,18 df=2.456



## An example aptitude x treatment interaction

- Divide participants into two groups
  - A group with little practice at crosswords and a group with a lot of practice.






## Partitioning a two-way within subjects design

- There are two possible ways to partition the variability in a two-way repeated measures design
  - Construct an overall error term for all the effects of interest
    - We may overestimate the error for some individual effects
    - Ignore an additional assumption made in repeated measures designs
      - Homogeneity of Difference Variances Assumption
  - Construct an error term for each of the effects of interest
    - We will never overestimate the error
    - We can temporarily ignore the homogeneity of difference variance assumption

## Error terms in a two-way within subjects design

- In the two-way repeated measures design
  - The error terms for the main effects are the residual for each main effects.
  - The error term for the interaction is based on the interaction between the two independent variables and the subject variable
- Each effect has a different error term in a within subjects design



## Testing the main effects and the interaction effect

- As in all other ANOVAs the effects are tested by constructing F-ratios

$$F_A = {MS_A}/{MS_{AxS}} \quad \text{with } (a-1) \;\&\; (a-1)(s-1)\; df$$

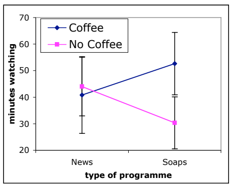
$$F_B = {MS_B}/{MS_{BxS}} \quad \text{with } (b-1) \;\&\; (b-1)(s-1)\; df$$

$$F_{AB} = {MS_{AB}}/{MS_{ABxS}} \quad \text{with } (a-1)(b-1) \;\&\; (a-1)(b-1)(s-1)\; df$$



## An example two-way repeated measures design

| | Coffee | | No Coffee | |
|---|---|---|---|---|
| | News | Soaps | News | Soaps |
| s1 | 42 | 56 | 38 | 28 |
| s2 | 18 | 32 | 26 | 20 |
| s3 | 30 | 46 | 44 | 24 |
| s4 | 47 | 58 | 46 | 26 |
| s5 | 57 | 60 | 56 | 46 |
| s6 | 51 | 64 | 54 | 38 |
| Mean | 40.83 | 52.67 | 44.00 | 30.33 |






## Results of analysis

Within Subjects Design (alias Randomized Blocks)

| Source of Variation | Sum of Squares | df | Mean Squares | F | p |
|---|---|---|---|---|---|
| A (Coffee) | 551.042 | 1 | 551.042 | 14.836 | 0.0120 |
| (Error AxS) | 185.708 | 5 | 37.142 | | |
| B (Programme) | 5.042 | 1 | 5.042 | 0.469 | 0.5237 |
| (Error BxS) | 53.708 | 5 | 10.742 | | |
| AB | 975.375 | 1 | 975.375 | 57.123 | 0.0006 |
| (Error ABxS) | 85.375 | 5 | 17.075 | | |



## Analytical Comparisons

- Planned comparisons can be conducted on main effects and interactions.
- Significant main effects can be further analysed using the appropriate post hoc tests
- When analysing significant interactions simple main effects analysis can be conducted
  - If there is a significant simple main effect with more than two levels then the appropriate post hoc tests can be used to further analyse these data



## Assumptions underlying a within subjects ANOVA

- ANOVA makes several assumptions
  - Data from interval or ratio scale (continuous)
  - Normal distributions
  - Independence
  - Homogeneity of variance
- Within subjects ANOVA adds another assumption
  - 'Sphericity': homogeneity of treatment difference variances
    - Sphericity is a special case of 'compound symmetry', so some people use this term
  - There is no need to test for sphericity if each IV has only two levels

## Testing the Sphericity Assumption

- SPSS provides a test of sphericity called Mauchly's test of sphericity

| | Mauchly's W | Approx. Chi-Square | df | Sig. |
|---|---|---|---|---|
| FACTOR1 | .557 | 2.176 | 5 | .830 |

- If it is not significant then we assume homogeneity of difference variances
- If it is significant then we cannot assume homogeneity of difference variances
  - If we do not correct for violations, ANOVA becomes too liberal
  - We will increase our rate of type 1 errors



## Testing the Sphericity Assumption

- SPSS provides alternative tests when sphericity assumption has not been met

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| FACTOR1 | Sphericity Assumed | 1531.458 | 3 | 510.486 | 23.576 | .000 |
| | Greenhouse-Geisser | 1531.458 | 2.182 | 701.834 | 23.576 | .000 |
| | Huynh-Feldt | 1531.458 | 3.000 | 510.486 | 23.576 | .000 |
| | Lower-bound | 1531.458 | 1.000 | 1531.458 | 23.576 | .005 |

- they adjust DFs (same SS for effect and error)
- G-G is conservative, and H-F liberal