**Lecture 7**

**Empirical Research Methods**
**IN4304**

**Quantitative Data analysis I – Differences**

TUDelft

---

**Previous lecture**

**Data Entry**
Code data, direction scales, clear name, missing data, level

**Central tendency**
Mean, median, mode, IOR, SD

**Explorative data analysis**
- Boxplot
- Histogram
- Stem and leaf plot

**Normal distribution**
1. Gives a good model of data for some real world data sets (e.g. large populations)
2. Good approximation of results of random outcomes (e.g. throwing a dice many times)
3. Large number of inferential statistics are based on normal distributions

*Check for Normality if you want to use parametric statistics*

TUDelft

---

**Today**

- Hypothesis testing
- Chi-square test
- *t*-test
- ANOVA
- Mann-Whitney U-test
- MANOVA

TUDelft

---

**Learning outcomes of this lecture**

After today's lecture you should be able :
- to explain the idea behind hypothesis testing
- to link conceptual models with specific statistical tests
- to perform a number of statistical tests in SPSS

TUDelft

---

**Idea behind hypothesis testing**



population

samples

- Is the difference between two samples a reflection of the difference of two different populations or simply caused by sampling error?

TUDelft

---

**Hypothesis testing**

- **Null hypothesis**: observed difference in sample is caused by sampling error and when considering the whole population there is no difference
- **Alternative hypothesis**: observed difference in sample is a real effect plus some change variation
- **p-value**: if null hypothesis is true, than what is the probability of observing a difference at least as extreme as the one observed in the sample?
- Social Science: $p < 0.05$ supports rejection of Null hypothesis

TUDelft

---

## Decision

Population

| | H0 = True | H0 = False |
|---|---|---|
| Accept H0 | No Error | Type II Error ß |
| Reject H0 | Type I Error ($\alpha$, p-value) | No Error |

Sample

Power = 1 - ß

**T U** Delft

---

## Interpreting p-value

- **p-value**: if null hypothesis is true, than what is the probability of observing a difference at least as extreme as the one observed in the sample?
- $p. < 0.05$ -> reject H0
- $p. > 0.05$
  1. Inconclusive
  2. Or if there are no methodological problems, argue based on power analysis that the real difference can not be very large

**T U** Delft

---

## Statistical test and Scales of measurement

| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | z-test, One-Sample t-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney U test |
| | | Interval / Ratio | z-test, two independent sample t-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample t-test, repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramér's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

**T U** Delft

---

## Chi-square test – Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

User interface type

- UI A
- UI B
- UI C

Task completed

- NO
- YES

**T U** Delft

---

## Chi-square test - example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo

| Observed | UI A | UI B | UI C | | |
|---|---|---|---|---|---|
| Yes | 30 | 70 | 25 | 125 | 66% |
| No | 35 | 10 | 20 | 65 | 34% |
| | 65 | 80 | 45 | 190 | |
| | 34% | 42% | 24% | | |

| Expected | UI A | UI B | UI C |
|---|---|---|---|
| Yes | 43 | 53 | 30 |
| No | 22 | 27 | 15 |

**T U** Delft

---

## Chi-square test - assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. Nominal data
3. Each observed person (or case) must appear in only one of the frequency cells
4. No more than 20% of the expected frequency cell counts should be less than 5 in large contingency table. For small tables each expected frequency cell count should be greater than 5

**T U** Delft

## Chi-square test – interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- Demo

**Chi-Square Tests**

| | Value | df | Asymp. Sig. (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | 29.983[a] | 2 | .000 |
| Likelihood Ratio | 32.286 | 2 | .000 |
| Linear-by-Linear Association | 2.726 | 1 | .099 |
| N of Valid Cases | 190 | | |

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 15.39.

TUDelft

---

## Chi-square test – SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

File: examples_chi.sav
Var: UI-Type, Task_complete

TUDelft

---

## Statistical test and Scales of measurement

| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | $z$-test, One-Sample $t$-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney $U$ test |
| | | Interval / Ratio | $z$-test, two independent sample $t$-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample $t$-test, repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramér's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

TUDelft

---

## Mann-Whitney U test – Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

Gender

Female

Male

→

Usability File Control in MP3 Player

Textual feedback

TUDelft

---
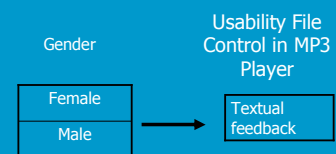
## Mann-Whitney U test – example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo

**(Rater total) rationale for score File Control * Gender (female = 1, male = 2) Crosstabulation**

Count

| | | Gender (female = 1, male = 2) | | Total |
|---|---|---|---|---|
| | | female | male | |
| (Rater total) rationale for score File Control | missing or not relevant | 3 | 4 | 7 |
| | Mainly negative comments, and/or improvement advice | 1 | 5 | 6 |
| | Both negative and positive comments, and/or improvement advi | 2 | 5 | 7 |
| | Mainly positive or no negative comments and no improvement | 8 | 6 | 14 |
| Total | | 14 | 20 | 34 |

TUDelft

---

## Mann-Whitney U test – assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. At least ordinal level

TUDelft

## Slide 19

### Mann-Whitney U test – interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- Demo

**Ranks**

| Gender (female | | N | Mean Rank | Sum of Ranks |
|---|---|---|---|---|
| (Rater total) rationale for score File Control | female | 11 | 17.05 | 187.50 |
| | male | 16 | 11.91 | 190.50 |
| | Total | 27 | | |

**Test Statistics[b]**

| | (Rater total) rationale for score File Control |
|---|---|
| Mann-Whitney U | 54.500 |
| Wilcoxon W | 190.500 |
| Z | -1.811 |
| Asymp. Sig. (2-tailed) | .070 |
| Exact Sig. [2*(1-tailed Sig.)] | .099[a] |

a. Not corrected for ties.
b. Grouping Variable: Gender (female = 1, male = 2)

**T U** Delft

## Slide 20

### Mann-Whitney U test – SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

File: mp3playersurvey.sav
Var: gender, fileRateTotal

**T U** Delft

## Slide 21

### Statistical test and Scales of measurement

| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | $z$-test, One-Sample $t$-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney $U$ test |
| | | Interval / Ratio | $z$-test, two independent sample $t$-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample $t$-test, repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramér's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

**T U** Delft

## Slide 22

### One-sample *t*-test – Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

- Single Condition
- Compare results with a norm or standard

**T U** Delft

## Slide 23

### One-sample *t*-test - example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo

(95% confidence interval)

5.29 Cut off point

Low - High usability rating

Average score Play Control, Average score File Control, Average score Volume Control, Average score Information Control
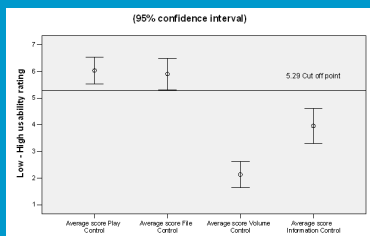
Usability rating of 4 UI components, with Cut of point 5.29

**T U** Delft

## Slide 24

### One-sample *t*-test - assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. Interval (or ratio) data
3. Data are from normally distributed population

**T U** Delft

## One-sample *t*-test — interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- Demo

**One-Sample Test**

| | | | | | |
|---|---|---|---|---|---|
| | | | Test Value = 5.29 | | |
| | | t | df | Sig. (2-tailed) | Mean Difference |
| Average score Play Control | | 3.013 | 33 | .005 | .74922 |
| Average score File Control | | 2.120 | 33 | .042 | .61196 |
| Average score Volume Control | | -13.126 | 33 | .000 | -3.15765 |
| Average score Information Control | | -4.102 | 33 | .000 | -1.33412 |

**T U** Delft

---

## One-sample *t*-test – SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

File: mp3playersurvey.sav
Var: fileAvg, VolumeAvg, InfAvg, PlayAvg

**T U** Delft

---

## Statistical test and Scales of measurement

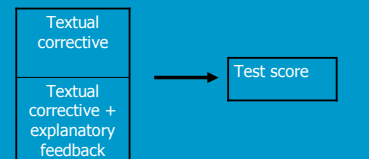| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | $z$-test, One-Sample $t$-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney $U$ test |
| | | Interval / Ratio | $z$-test, two independent sample $t$-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample $t$-test, repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramér's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

**T U** Delft

---

## Independent-sample *t*-test — Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

Feedback type in e-learning module

Score in post test

Textual corrective

Textual corrective + explanatory feedback

→ Test score

**T U** Delft

---

## Independent-sample *t*-test — assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. Interval (or ratio) data
3. Data are from normally distributed population
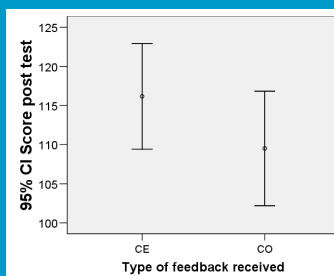4. Homogeneity of variance (variances in populations are roughly equal)*

**T U** Delft

---

## Independent-sample *t*-test — example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo

**T U** Delft

## Slide 31: Independent-sample *t*-test − interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- Demo

| | Levene's Test for Equality of Variances | | t-test for Equality of M | | | |
|---|---|---|---|---|---|---|
| | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference |
| Score post test — Equal variances assumed | .040 | .841 | 1.365 | 59 | .177 | 6.651 |
| Equal variances not assumed | | | 1.367 | 58.741 | .177 | 6.651 |

TUDelft

---

## Slide 32: Independent-sample *t*-test − SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

File: Feedback experiment.sav
Var: feedback (1,2), postscore
Note: Filter feedback <> 3

TUDelft

---

## Slide 33: Statistical test and Scales of measurement

| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | *z*-test, One-Sample *t*-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney *U* test |
| | | Interval / Ratio | *z*-test, two independent sample *t*-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample *t*-test repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramér's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

TUDelft

---

## Slide 34: Paired-sample *t*-test – Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

Time of test          Module test
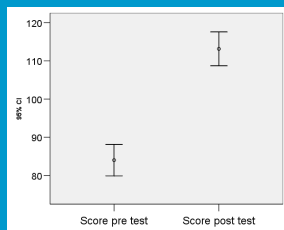
Pre
Post          →     score

TUDelft

---

## Slide 35: Paired-sample *t*-test – example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo



Score pre test    Score post test

TUDelft

---

## Slide 36: Paired-sample *t*-test – assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. Interval (or ratio) data
3. Difference data are from a normally distributed population

TUDelft

## Paired-sample *t*-test – interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- Demo

**Paired Samples Test**

| | | Paired Differences | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 95% Confidence Interval of the Difference | | | | |
| | | Mean | Std. Deviation | Std. Error Mean | Lower | Upper | t | df | Sig. (2-tailed) |
| Pair 1 | Score pre test - Score post test | -29.12088 | 22.02074 | 2.30840 | -33.70692 | -24.53484 | -12.615 | 90 | .000 |

**T**U Delft

---

## Paired-sample *t*-test – SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

File: Feedback experiment.sav
Var: pre_score, post_score
Note: Filter out

**T**U Delft

---

## Statistical test and Scales of measurement

| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | z-test, One-Sample t-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney U test |
| | | Interval / Ratio | z-test, two independent sample t-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample t-test, repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramér's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

**T**U Delft

---

## One-way ANOVA – Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- posthoc
- Demo

Exam year

Student performance

2004
2005
2006
2007

Exam score
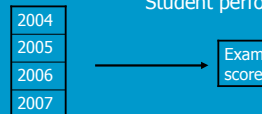
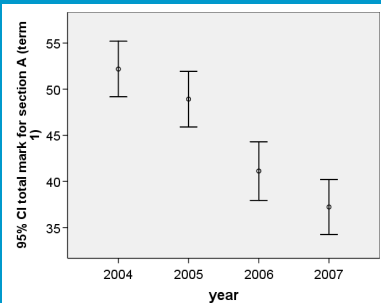**T**U Delft

---

## One-way ANOVA – example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- posthoc
- Demo

**T**U Delft

---

## One-way ANOVA – assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- posthoc
- Demo

1. Independence of observations
2. Interval (or ratio) data
3. data are from a normally distributed population
4. Homogeneity of variance

**T**U Delft

## One-way ANOVA – interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- posthoc
- Demo

**ANOVA**

total mark for section A (term 1)

|  | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 22481.424 | 3 | 7493.808 | 19.933 | .000 |
| Within Groups | 236090.9 | 628 | 375.941 |  |  |
| Total | 258572.3 | 631 |  |  |  |

T U Delft

---

## One-way ANOVA – Post-Hoc

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Post-hoc**
- Demo

**Multiple Comparisons**

Dependent Variable: total mark for section A (term 1)

Tukey HSD

| (I) year | (J) year | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
|  |  |  |  |  | Lower Bound | Upper Bound |
| 2004 | 2005 | 3.266 | 2.181 | .440 | -2.35 | 8.89 |
|  | 2006 | 11.057* | 2.181 | .000 | 5.44 | 16.68 |
|  | 2007 | 14.956* | 2.181 | .000 | 9.34 | 20.57 |
| 2005 | 2004 | -3.266 | 2.181 | .440 | -8.89 | 2.35 |
|  | 2006 | 7.791* | 2.181 | .002 | 2.17 | 13.41 |

- Instead of separate t-tests, post-hoc test control for Type I error
- Keeps family-wise error rate down to 0.05
- Tukey's Honestly Significant Difference (HSD) test

T U Delft

---

## One-way ANOVA – SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- Post-Hoc
- **Demo**

File: exam_marks2004-2007_equal_year_sample.sav
Var: Year, t_term1
Tukey posthoc

T U Delft

---

## Statistical test and Scales of measurement

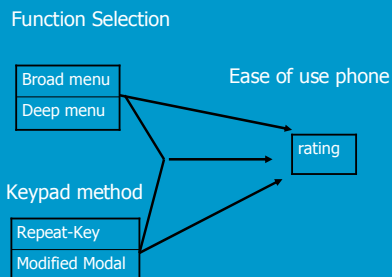| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
|  |  | Ordinal | Wilcoxon signed-ranks test |
|  |  | Interval / Ratio | $z$-test, One-Sample $t$-test |
|  | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
|  |  | Ordinal | Mann-Whitney $U$ test |
|  |  | Interval / Ratio | $z$-test, two independent sample $t$-test, ANOVA, MANOVA |
|  | Dependent | Nominal | McNemar |
|  |  | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
|  |  | Interval / Ratio | Paired-Sample $t$-test, repeated measures ANOVA, MANOVA |
| Find relation |  | Nominal | Cramér's V, phi |
|  |  | Ordinal | Kendall, Spearman correlation |
|  |  | Interval / Ratio | Pearson product moment correlation, regression analysis |

T U Delft

---

## Two-way ANOVA – Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

Function Selection

Broad menu
Deep menu

Ease of use phone

rating

Keypad method

Repeat-Key
Modified Modal

T U Delft

---

## Two-way ANOVA – assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. Interval (or ratio) data
3. data are from a normally distributed population
4. Homogeneity of variance

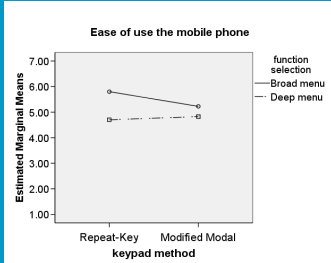File: mobile Phone (Interaction with Computers)
Var: Keypad, function, qamobile

T U Delft

## Slide 49

### Two-way ANOVA – example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo

**Ease of use the mobile phone**

## Slide 50

### Two-way ANOVA – interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- Demo

**Tests of Between-Subjects Effects**

Dependent Variable: Average of questions about ease of use of the mobile phone

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Corrected Model | 14.713[a] | 3 | 4.904 | 5.094 | .003 |
| Intercept | 2111.513 | 1 | 2111.513 | 2193.363 | .000 |
| keypad | 1.013 | 1 | 1.013 | 1.052 | .308 |
| function | 11.250 | 1 | 11.250 | 11.686 | .001 |
| keypad * function | 2.450 | 1 | 2.450 | 2.545 | .115 |
| Error | 73.164 | 76 | .963 | | |
| Total | 2199.389 | 80 | | | |
| Corrected Total | 87.876 | 79 | | | |

a. R Squared = .167 (Adjusted R Squared = .135)

## Slide 51

### Two-way ANOVA – SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

File: mp3playerSurvey.sav
Within factor: Component (4)
(1) playAvg
(2) FileAvg
(3) VolumeAvg
(4) InfAvg

## Slide 52

### Statistical test and Scales of measurement

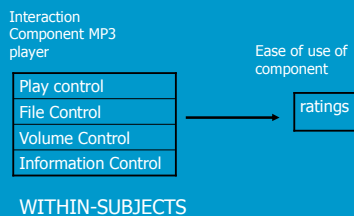| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | $z$-test, One-Sample $t$-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney $U$ test |
| | | Interval / Ratio | $z$-test, two independent sample $t$-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample $t$-test, repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramer's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

## Slide 53

### ANOVA with repeated measures – Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

Interaction Component MP3 player

Play control
File Control
Volume Control
Information Control

WITHIN-SUBJECTS

→ Ease of use of component → ratings

## Slide 54

### ANOVA with repeated measures – assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. Interval (or ratio) data
3. data are from a normally distributed population
4. Sphericity (1. homogeneity of variance, 2. variances of the differences between levels are similar to one another)

9

## Slide 55

**ANOVA with repeated measures - example**

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo


(95% confidence interval)
5.29 Cut off point
Low - High usability rating
Average score Play Control / Average score File Control / Average score Volume Control / Average score Information Control

TUDelft

## Slide 56

**ANOVA with repeated measures - interpretation outcomes**

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**
- Demo

**Mauchly's Test of Sphericity**

Measure: MEASURE_1

| Within Subjects Effect | Mauchly's W | Approx. Chi-Square | df | Sig. | |
|---|---|---|---|---|---|
| component | .836 | 5.673 | 5 | .340 | |

**Tests of Within-Subjects Effects**

Measure: MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| component | Sphericity Assumed | 336.948 | 3 | 112.316 | 67.363 | .000 |
| | Greenhouse-Geisser | 336.948 | 2.727 | 123.553 | 67.363 | .000 |
| | Huynh-Feldt | 336.948 | 2.997 | 112.434 | 67.363 | .000 |
| | Lower-bound | 336.948 | 1.000 | 336.948 | 67.363 | .000 |
| Error(component) | Sphericity Assumed | 165.066 | 99 | 1.667 | | |
| | Greenhouse-Geisser | 165.066 | 90.006 | 1.834 | | |

TUDelft

## Slide 57

**ANOVA with repeated measures - SPSS Demo**

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

File: mobile phone (IwC)
Var: function (IV), DV: tokey,
Tot_time, qamobile

TUDelft

## Slide 58

# Statistical test and Scales of measurement

| Aim | Samples | Data Level | Tests |
|---|---|---|---|
| Find difference | Single sample | Nominal | Binominal test, $\chi^2$ goodness-of-fit |
| | | Ordinal | Wilcoxon signed-ranks test |
| | | Interval / Ratio | $z$-test, One-Sample $t$-test |
| | Independent | Nominal | Fisher-Exact test, $\chi^2$ |
| | | Ordinal | Mann-Whitney $U$ test |
| | | Interval / Ratio | $z$-test, two independent sample $t$-test, ANOVA, MANOVA |
| | Dependent | Nominal | McNemar |
| | | Ordinal | Sign Test, Wilcoxon matched-pairs signed-rank test, Friedman Test |
| | | Interval / Ratio | Paired-Sample $t$-test, repeated measures ANOVA, MANOVA |
| Find relation | | Nominal | Cramér's V, phi |
| | | Ordinal | Kendall, Spearman correlation |
| | | Interval / Ratio | Pearson product moment correlation, regression analysis |

TUDelft

## Slide 59

# MANOVA– Conceptual model

- **Conceptual model**
- Assumptions
- Example
- Interpretation outcome
- Demo

Usability

Function Selection

Broad menu
Deep menu

→ Number of keystrokes
→ Task time
→ Subjective ease of use

TUDelft

## Slide 60

# MANOVA – example

- Conceptual model
- Assumptions
- **Example**
- Interpretation outcome
- Demo


task time
Number of Keystrokes
Ease of Use
function selection: Broad menu / Deep menu

TUDelft

## MANOVA – assumptions

- Conceptual model
- **Assumptions**
- Example
- Interpretation outcome
- Demo

1. Independence of observations
2. Interval (or ratio) data
3. Multivariate normality*
4. Homogeneity of covariance matrices*

*(see Fields, 2005, p.593 for checking assumptions)

**T U**Delft

---

## MANOVA – interpretation outcomes

- Conceptual model
- Assumptions
- Example
- **Interpretation outcome**

**Multivariate Tests[b]**

| Effect | | Value | F | Hypothesis df | Error df | Sig. |
|---|---|---|---|---|---|---|
| Intercept | Pillai's Trace | .981 | 1328.630[a] | 3.000 | 76.000 | .000 |
| | Wilks' Lambda | .019 | 1328.630[a] | 3.000 | 76.000 | .000 |
| | Hotelling's Trace | 52.446 | 1328.630[a] | 3.000 | 76.000 | .000 |
| | Roy's Largest Root | 52.446 | 1328.630[a] | 3.000 | 76.000 | .000 |
| function | Pillai's Trace | .323 | 12.078[a] | 3.000 | 76.000 | .000 |
| | Wilks' Lambda | .677 | 12.078[a] | 3.000 | 76.000 | .000 |
| | Hotelling's Trace | .477 | 12.078[a] | 3.000 | 76.000 | .000 |
| | Roy's Largest Root | .477 | 12.078[a] | 3.000 | 76.000 | .000 |

**T U**Delft

---

## MANOVA – SPSS Demo

- Conceptual model
- Assumptions
- Example
- Interpretation outcome
- **Demo**

**T U**Delft

---

## Summary

**Idea behind hypothesis testing**

- Is the difference between two samples a reflection of the difference of two different populations or simply caused by sampling error?

**Conceptual models with specific statistical tests**

- Different tests depending on: level, number of independent and dependent variables, between or within subjects set up

**Statistical tests in SPSS**

- See the recording of this lecture

**T U**Delft

---

## This week in practicum

Working on project coursework
- Chi-square,
- Error-bars,
- One-sample t-test,
- independent-sample t-test,
- paired-sample t-test,
- Mann-Whitney U test

**T U**Delft

---

## Next time

**Week 4.1 (20/4/2010) Quantitative Data analysis II - Relations**
- **Correlation**
- **Regression**
- **(Robson ch. 13)**

**EWI lecture room B, Tuesday 15:45 - 17:45**

**T U**Delft

## References

- Coolican, H., (2004). *Research methods and statistics in psychology* (4th ed). London, UK: Hodder Arnold.
- Fields, A. (2005). *Discovering statistics using SPSS* (2nd ed). London: UK, Sage.

**T U** Delft