BCN Statistics Course

Repeated measures ANOVA

Jorge Tendeiro

25 June 2014

university of groningen

Overview

1. Usual ANOVA

2. Repeated measures (M)ANOVA
   - Main ideas
   - Testing effects in RM-ANOVA: Univariate test
   - Testing effects in RM-MANOVA: Multivariate test
   - Assumptions of RM-(M)ANOVA
   - Use univariate or multivariate?
   - Decomposing the within-factor effect
   - Post hoc procedures
   - Extending RM-ANOVA

'Usual' ANOVA

Generalization of independent samples t-test to more than 2 groups.

One-way ANOVA (Stats II)

- One DV Y, one categorical IV A
- Compares the means of Y among all groups defined by A

Two-way ANOVA (Stats II)

- One DV Y, two categorical IVs A, B (factors)
- Compares the means of Y among all groups jointly defined by A, B
- It analyzes the main effect of A (across all levels of B), main effect of B (across all levels of A), and interaction effect $A \times B$ ('differences between differences')

'Usual' ANOVA: Assumptions

Assumptions in ANOVA:

- Independence of observations.
- Groups defined by factors are normally distributed.
- Groups defined by factors have equal population variances (homogeneity of variance).

Obs: One-way and two-way ANOVA are just special regression models, where the factors are dummy-coded into the model (Stats II & Stats III).

'Usual' ANOVA: Limitation

- So far, we only allowed each subject to contribute with one score (one measurement) on the DV:
  "Different subjects contribute to different means."
- What if each subject is measured in ALL levels of the treatment factor?

'Usual' ANOVA: Limitation

- So far, we only allowed each subject to contribute with one score (one measurement) on the DV:
  "Different subjects contribute to different means."
- What if each subject is measured in ALL levels of the treatment factor?

Repeated measures (M)ANOVA

Repeated measures (M)ANOVA

Generalization of paired samples t-test to more than 2 groups.

Repeated measures (M)ANOVA

Generalization of paired samples t-test to more than 2 groups.
Imagine the following situations:

Situation 1 A group of subjects is measured under various treatment conditions (or at different points in time).

Situation 2 Subjects are grouped according to their gender and measured under various treatment conditions (or at different points in time).

Situation 3 Two groups of subjects are administered two types of drugs at each of three doses.

Situation 4 Subjects are grouped according to their age (2 groups) and diet program (3 groups), and their weight loss is measured in three different moments (2,4,6 months after beginning of diet).

Repeated measures (M)ANOVA

	Treatments			
	1	2	···	k
Subjects 1				
2				
⋮				
n				

Situation 1

	Treatments			
	1	2	···	k
Males				
Females				

Situation 2

	Drug 1			Drug 2		
	1	2	3	1	2	3
Group1						
Group2						

Situation 3

		Weight loss		
		1	2	3
Control	Age 1			
	Age 2			
Diet 1	Age 1			
	Age 2			
Diet 2	Age 1			
	Age 2			

Situation 4

Repeated measures (M)ANOVA

These cases do not fit the usual ANOVA framework.

- Subjects are now measured more than once (for different treatments/at different points in time): Repeated measures.

The ANOVA model needs to be adjusted in order to fit this type of data.

Repeated measures (M)ANOVA

These cases do not fit the usual ANOVA framework.

- Subjects are now measured more than once (for different treatments/at different points in time): Repeated measures.

The ANOVA model needs to be adjusted in order to fit this type of data.

There are 2 types of factors (categorical IVs) in repeated measures ANOVA designs:

- Between-factors: Variables grouping subjects.
- Within-factors: Repeated measures variables.

Repeated measures (M)ANOVA

These cases do not fit the usual ANOVA framework.

- Subjects are now measured more than once (for different treatments/at different points in time): Repeated measures.

The ANOVA model needs to be adjusted in order to fit this type of data.

There are 2 types of factors (categorical IVs) in repeated measures ANOVA designs:

- Between-factors: Variables grouping subjects.
- Within-factors: Repeated measures variables.

Situation	Between factor(s)	Within factor(s)	Design
Situation 1	—	Treatments	one within
Situation 2	Gender	Treatments	one between, one within
Situation 3	Group	Drug, Dosage	one between, two within
Situation 4	Age, Diet program	Weight loss	two between, one within

RM-(M)ANOVA: Testing effects

Main goal in RM-(M)ANOVA: Test the significance of the effect of the within-factor on the DV (the treatment effect).

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

RM-(M)ANOVA: Testing effects

Main goal in RM-(M)ANOVA: Test the significance of the effect of the within-factor on the DV (the treatment effect).

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

- With usual ANOVA we perform an F test

$$F = \frac{SS_B/df_B}{SS_W/df_W} = \frac{MS_B}{MS_W} \sim F(k-1, n-k)$$

RM-(M)ANOVA: Testing effects

Main goal in RM-(M)ANOVA: Test the significance of the effect of the within-factor on the DV (the treatment effect).

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

- With usual ANOVA we perform an F test

$$F = \frac{SS_B/df_B}{SS_W/df_W} = \frac{MS_B}{MS_W} \sim F(k-1, n-k)$$

- With repeated measures ANOVA we will now have two options (both using F tests):
  - Run a univariate test.
  - Run a multivariate test.

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?
  A:

$$\text{ANOVA:} \quad SS_T = SS_B + \quad SS_W$$

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?
  A:

$$\text{ANOVA:} \quad SS_T = SS_B + \overbrace{SS_W}$$

$$\text{repeated measures ANOVA:} \quad SS_T = SS_B + \overbrace{SS_{B_1} + SS_E}^{},$$

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?
  A:

$$\text{ANOVA:} \quad SS_T = SS_B + SS_W$$

$$\text{repeated measures ANOVA:} \quad SS_T = SS_B + \overbrace{SS_{B_1} + SS_E}^{SS_W},$$

where:
SS_T = total SS

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?
  A:

$$\text{ANOVA:} \quad SS_T = SS_B + SS_W$$

$$\text{repeated measures ANOVA:} \quad SS_T = SS_B + \overbrace{SS_{B_1} + SS_E}^{SS_W},$$

where:
SS_T = total SS
SS_B = between treatments SS (the model SS)

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?
  A:

$$\text{ANOVA:} \quad SS_T = SS_B + SS_W$$

$$\text{repeated measures ANOVA:} \quad SS_T = SS_B + \overbrace{SS_{B_1} + SS_E}^{SS_W},$$

where:
SS_T = total SS
SS_B = between treatments SS (the model SS)
SS_W = within treatments SS

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?
  A:

$$\text{ANOVA:} \quad SS_T = SS_B + SS_W$$

$$\text{repeated measures ANOVA:} \quad SS_T = SS_B + \overbrace{SS_{B_1} + SS_E}^{SS_W},$$

where:
SS_T = total SS
SS_B = between treatments SS (the model SS)
SS_W = within treatments SS
SS_{B_1} = between subjects SS (New! Consider subjects as a **random** factor!)

RM-ANOVA: Univariate test

- Partition the variance of the scores for the DV, just as with regular ANOVA
- Q: How is this done?
  A:

$$\text{ANOVA:} \quad SS_T = SS_B + SS_W$$

$$\text{repeated measures ANOVA:} \quad SS_T = SS_B + \overbrace{SS_{B_1} + SS_E}^{SS_W},$$

where:
SS_T = total SS
SS_B = between treatments SS (the model SS)
SS_W = within treatments SS
SS_{B_1} = between subjects SS (New! Consider subjects as a **random** factor!)
SS_E = residual SS

RM-ANOVA: Univariate test

So, writing the SS in RM-ANOVA in the 'usual-ANOVA' way:

$$\begin{aligned} SS_T &= \mathbf{SS_{between\ treatments}} + SS_{\text{within treatments}} \\ &= \mathbf{SS_{between\ treatments}} + (SS_{\text{between subjects}} + SS_E) \end{aligned}$$

Individual differences between subjects are removed from error term SS_W

RM-ANOVA: Univariate test

So, writing the SS in RM-ANOVA in the 'usual-ANOVA' way:

$$\begin{aligned} SS_T &= \mathbf{SS_{between\ treatments}} + SS_{\text{within treatments}} \\ &= \mathbf{SS_{between\ treatments}} + (SS_{\text{between subjects}} + SS_E) \end{aligned}$$

Individual differences between subjects are removed from error term SS_W

Equivalently, and looking at the subjects blocking:

$$\begin{aligned} SS_T &= SS_{\text{between subjects}} + SS_{\text{within subjects}} \\ &= SS_{\text{between subjects}} + (\mathbf{SS_{between\ treatments}} + SS_E) \end{aligned}$$

The treatment effect is part of the within-subjects variance

RM-ANOVA: Univariate test

Advantages:

- Reduction of within-group variance (error variance) by removing individual differences between subjects (by blocking on subjects).

greater power to test within-subjects effects

- Less subjects are needed, since more information is collected per subject.

RM-ANOVA: Univariate test

Advantages:

- Reduction of within-group variance (error variance) by removing individual differences between subjects (by blocking on subjects).

greater power to test within-subjects effects

- Less subjects are needed, since more information is collected per subject.

Disadvantages:

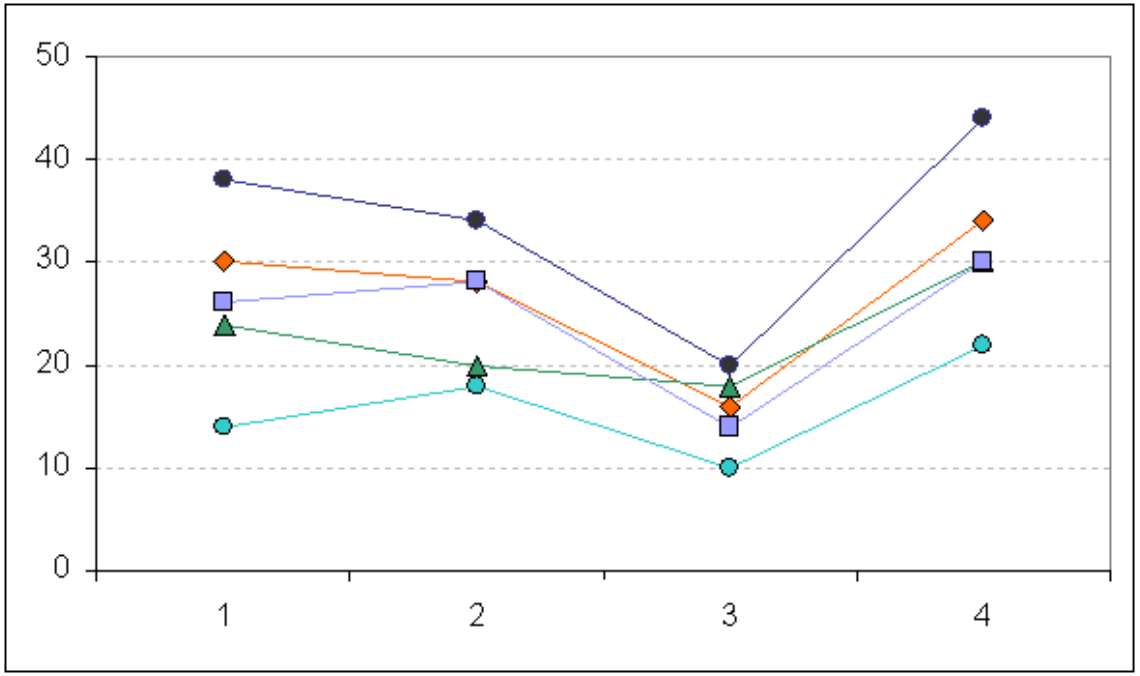
- Order in which conditions are experienced may affect performance (practice effect, fatigue effect, carry-over effect).
- Model assumptions are harder to meet.
- More complicated model (of course...).

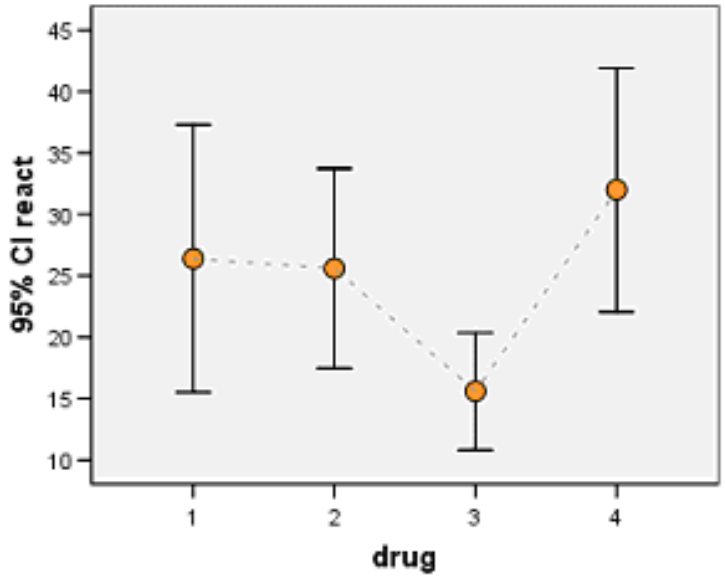
Example (Stevens): Univariate test

Goal: Study the effect of four drugs on reaction time.
Each subject is tested using each drug.

Subjects	Drugs 1	2	3	4
1	30	28	16	34
2	14	18	10	22
3	24	20	18	30
4	38	34	20	44
5	26	28	14	30

Example (Stevens): Univariate test





Is there an effect of drug on reaction times?

Example (Stevens): Univariate test

You can compute each SS with the General Linear Model (`Analyze>General Linear Model>Univariate`)

Subject	**Drug**	Reaction
1	1	30
1	2	28
1	3	16
1	4	34
⋮	⋮	⋮
5	1	26
5	2	28
5	3	14
5	4	30

Factors: Drug (fixed), Subject (random)
DV: Reaction
Model: Full factorial

Example (Stevens): Univariate test

Tests of Between-Subjects Effects

Dependent Variable:reaction

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	12400,200	1	12400,200	72,857	,001
	Error	680,800	4	170,200[a]		
drug	Hypothesis	SS_B 698,200	3	232,733	24,759	,000
	Error	112,800	12	9,400[b]		
subject	Hypothesis	SS_{B_1} 680,800	4	170,200	18,106	,000
	Error	112,800	12	9,400[b]		
drug * subject	Hypothesis	SS_E 112,800	12	9,400	.	.
	Error	,000	0	.[c]		

a. MS(subject)

b. MS(drug * subject)

c. MS(Error)

Example (Stevens): Univariate test

Tests of Between-Subjects Effects

Dependent Variable:reaction

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	12400,200	1	12400,200	72,857	,001
	Error	680,800	4	170,200[a]		
drug	Hypothesis	SS_B 698,200	3	232,733	24,759	,000
	Error	112,800	12	9,400[b]		
subject	Hypothesis	SS_{B_1} 680,800	4	170,200	18,106	,000
	Error	112,800	12	9,400[b]		
drug * subject	Hypothesis	SS_E 112,800	12	9,400	.	.
	Error	,000	0	.[c]		

a. MS(subject)

b. MS(drug * subject)

c. MS(Error)

Notes:

- $SS_W = SS_{B_1} + SS_E = 793.600$
- $SS_{\text{Model}} = SS_B + SS_{B_1} = 1379.000$
- $SS_T = SS_{\text{Model}} + SS_E = 1491.800$
- Drug has a significant effect on reaction time ($F(3, 12) = 24.76$, $p < .001$)

Example (Stevens): Univariate test

Alternative way to do the univariate test:

`Analyze>General Linear Model>Repeated Measures.`

Focus on the 'Sphericity Assumed' part (to be explained later)

Example (Stevens): Univariate test

Alternative way to do the univariate test:

`Analyze>General Linear Model>Repeated Measures.`

Focus on the 'Sphericity Assumed' part (to be explained later)

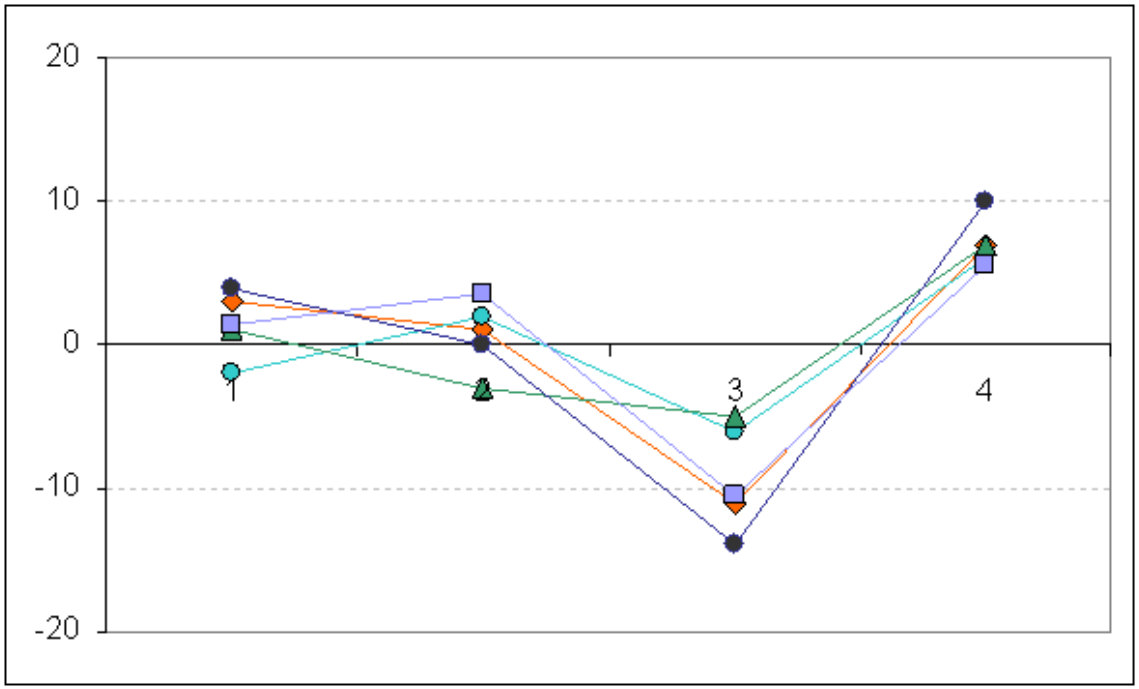
Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Drug	Sphericity Assumed	698.200	3	232.733	24.759	.000
	Greenhouse-Geisser	698.200	1.815	384.763	24.759	.001
	Huynh-Feldt	698.200	3.000	232.733	24.759	.000
	Lower-bound	698.200	1.000	698.200	24.759	.008
Error(Drug)	Sphericity Assumed	112.800	12	9.400		
	Greenhouse-Geisser	112.800	7.258	15.540		
	Huynh-Feldt	112.800	12.000	9.400		
	Lower-bound	112.800	4.000	28.200		

Example (Stevens): Univariate test

"RM-ANOVA removes individual differences between subjects from error SS_W"



Compare with the setting disregarding subjects blocking: ▸ ...

RM-ANOVA: Univariate test

In general,

Source	**SS**	**df**	**MS**	**F**
Within-factor	SS_B	$k-1$	SS_B/df_B	MS_B/MS_E
Between-factor	SS_{B_1}	$n-1$	SS_{B_1}/df_{B_1}	MS_{B_1}/MS_E
Error	SS_E	$(k-1)(n-1)$	SS_E/df_E	
Total	SS_T	$nk-1$		

RM-ANOVA: Univariate test

In general,

Source	SS	df	MS	F
Within-factor	SS_B	$k-1$	SS_B/df_B	MS_B/MS_E
Between-factor	SS_{B_1}	$n-1$	SS_{B_1}/df_{B_1}	MS_{B_1}/MS_E
Error	SS_E	$(k-1)(n-1)$	SS_E/df_E	
Total	SS_T	$nk-1$		

- Subject SS (SS_{B_1}): Variance between subjects.

RM-ANOVA: Univariate test

In general,

Source	SS	df	MS	F
Within-factor	SS_B	$k-1$	SS_B/df_B	MS_B/MS_E
Between-factor	SS_{B_1}	$n-1$	SS_{B_1}/df_{B_1}	MS_{B_1}/MS_E
Error	SS_E	$(k-1)(n-1)$	SS_E/df_E	
Total	SS_T	$nk-1$		

- Subject SS (SS_{B_1}): Variance between subjects.
- Treatment SS (SS_B): Part of variance within subjects due to differences beetween treatments.
- Error SS (SS_E): Interaction part (to which extent subjects respond differently to treatments).

RM-ANOVA: Univariate test

In general,

Source	SS	df	MS	F
Within-factor	SS_B	$k-1$	SS_B/df_B	MS_B/MS_E
Between-factor	SS_{B_1}	$n-1$	SS_{B_1}/df_{B_1}	MS_{B_1}/MS_E
Error	SS_E	$(k-1)(n-1)$	SS_E/df_E	
Total	SS_T	$nk-1$		

- Subject SS (SS_{B_1}): Variance between subjects.
- Treatment SS (SS_B): Part of variance within subjects due to differences beetween treatments.
- Error SS (SS_E): Interaction part (to which extent subjects respond differently to treatments).

The F tests in the table are the univariate tests for repeated measures analysis.

RM-MANOVA: Multivariate test

New approach: Regard each treatment level (or timepoint) as a DV.

In this way the repeated measures of one subject are a multivariate observation of the dependent variables.

For our running example...

Subject	Drug1	Drug2	Drug3	Drug4
1	30	28	16	34
2	14	18	10	22
3	24	20	18	30
4	38	34	20	44
5	26	28	14	30

multivariate observation for subject 2

The main goal remains the same: Test

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k.$$

RM-MANOVA: Multivariate test

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

RM-MANOVA: Multivariate test

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

Testing H_0 is similar to testing

$$H_0 : \begin{pmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \\ \vdots \\ \mu_{k-1} - \mu_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

RM-MANOVA: Multivariate test

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

Testing H_0 is similar to testing

$$H_0 : \begin{pmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \\ \vdots \\ \mu_{k-1} - \mu_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This test is multivariate because several (transformed) DVs are being compared in simultaneous. This is the generalization of the matched-pairs t-test

RM-MANOVA: Multivariate test

- There are several multivariate tests available; SPSS routinely reports four of them:
  - Wilk's Λ (use this test 'by default').
  - Pillai-Bartlett trace.
  - Hotelling-Lawley trace.
  - Roy's largest root.

RM-MANOVA: Multivariate test

- There are several multivariate tests available; SPSS routinely reports four of them:
  - Wilk's Λ (use this test 'by default').
  - Pillai-Bartlett trace.
  - Hotelling-Lawley trace.
  - Roy's largest root.

For our drug example:

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
Drug	Pillai's Trace	.977	28.412[a]	3.000	2.000	.034
	Wilks' Lambda	.023	28.412[a]	3.000	2.000	.034
	Hotelling's Trace	42.618	28.412[a]	3.000	2.000	.034
	Roy's Largest Root	42.618	28.412[a]	3.000	2.000	.034

a. Exact statistic

b. Design: Intercept
Within Subjects Design: Drug

Any of the tests rejects H_0, i.e., drug has an effect on reaction times.

RM-(M)ANOVA: Assumptions

Assumptions needed to perform the univariate/multivariate tests:
For both tests:

1. **Independence** of observations (between subjects).
2. Multivariate **normality** (i.e., each subgroup is normally distributed).

Extra assumption for univariate F-test:

3. Compound symmetry:
   - Variance is homogeneous across treatment levels (homogeneity of variances)

     $$\sigma^2_{T_1} = \sigma^2_{T_2} = \cdots = \sigma^2_{T_k},$$

     and
   - Covariances are homogeneous across treatment levels

     $$\sigma_{T_1,T_2} = \sigma_{T_1,T_3} = \cdots = \sigma_{T_{k-1},T_k}.$$

RM-(M)ANOVA: Assumptions

- **Compound symmetry** is a highly restrictive condition which is not realistic in practice.
- It was shown that compound symmetry is sufficient but not necessary for the F-tests to be valid (i.e., compound symmetry is too strong).
- A more relaxed condition, under which the F-tests are valid, is sphericity:

  All variances of all pairwise differences between pairs of repeated measures are equal:

$$\sigma^2_{T_1-T_2} = \sigma^2_{T_1-T_3} = \cdots = \sigma^2_{T_{k-1}-T_k}.$$

RM-ANOVA: Sphericity

SPSS tests the sphericity assumption with Mauchly's test.
The goal is to NOT reject H_0.
(This test is sensitive to departures from normality.)

RM-ANOVA: Sphericity

SPSS tests the sphericity assumption with Mauchly's test.
The goal is to NOT reject H_0.
(This test is sensitive to departures from normality.)

For our drug example:

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Drug	.186	4.572	5	.495	.605	1.000	.333

The null is not rejected, so sphericity holds

RM-ANOVA: Sphericity

Q: What if sphericity is violated?

RM-ANOVA: Sphericity

Q: What if sphericity is violated?

A: Then the univariate F-test is biased (Type I error is increased).

RM-ANOVA: Sphericity

Q: What if sphericity is violated?
A: Then the univariate F-test is biased (Type I error is increased).

Q: How to avoid this problem?

RM-ANOVA: Sphericity

Q: What if sphericity is violated?
A: Then the univariate F-test is biased (Type I error is increased).

Q: How to avoid this problem?
A: There are two possibilities:

- Introduce a correction in the univariate test, or
- Use the multivariate test.

RM-ANOVA: Sphericity

Q: What if sphericity is violated?
A: Then the univariate F-test is biased (Type I error is increased).

Q: How to avoid this problem?
A: There are two possibilities:

- Introduce a correction in the univariate test, or
- Use the multivariate test.

Q: How to correct lack of sphericity in the univariate test?

RM-ANOVA: Sphericity

Q: What if sphericity is violated?
A: Then the univariate F-test is biased (Type I error is increased).

Q: How to avoid this problem?
A: There are two possibilities:

- Introduce a correction in the univariate test, or
- Use the multivariate test.

Q: How to correct lack of sphericity in the univariate test?
A: Adjust the degrees of freedom of the F statistic before doing the tests (epsilon corrections).

In SPSS: Look at

- Greenhouse-Geisser correction.
- Huynh-Feldt correction.

RM-ANOVA: Sphericity

Univariate test of $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ with epsilon corrections for the drug example:

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Drug	.186	4.572	5	.495	.605	1.000	.333

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Drug	Sphericity Assumed	698.200	3	232.733	24.759	.000
	Greenhouse-Geisser	698.200	1.815	384.763	24.759	.001
	Huynh-Feldt	698.200	3.000	232.733	24.759	.000
	Lower-bound	698.200	1.000	698.200	24.759	.008
Error(Drug)	Sphericity Assumed	112.800	12	9.400		
	Greenhouse-Geisser	112.800	7.258	15.540		
	Huynh-Feldt	112.800	12.000	9.400		
	Lower-bound	112.800	4.000	28.200		

RM-ANOVA: Sphericity

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

In general:

- Greenhouse-Geisser correction is too conservative: Overreject H_0.
- Huynh-Feldt correction too liberal: Underreject H_0.

RM-ANOVA: Sphericity

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

In general:

- Greenhouse-Geisser correction is too conservative: Overreject H_0.
- Huynh-Feldt correction too liberal: Underreject H_0.

A simplified rule which is half-way between Greenhouse-Geisser and Huynh-Feldt corrections is the following:

1. Compute the average of both epsilons (G-G and H-F), say $\bar{\varepsilon}$
2. If...
   - $\bar{\varepsilon} > .7$: Use Huynh-Feldt correction.
   - $\bar{\varepsilon} < .7$ or nothing is known about sphericity: Use Greenhouse-Geisser correction.

Use univariate or multivariate?

Both methods are good for controlling for type I errors.

Use univariate or multivariate?

Both methods are good for controlling for type I errors.

'Sphericity' criterion

- **If sphericity is violated ($\varepsilon < .7$) and sample size is larger than $k + 10$:** Use multivariate test statistics.
  (Because these statistics do not depend on this assumption, and they are more powerful under these conditions.)

Use univariate or multivariate?

Both methods are good for controlling for type I errors.

'Sphericity' criterion

- **If sphericity is violated ($\varepsilon < .7$) and sample size is larger than $k + 10$:** Use multivariate test statistics.
  (Because these statistics do not depend on this assumption, and they are more powerful under these conditions.)
- **If sphericity holds ($\varepsilon > .7$) or sample sizes are small:** Use univariate test statistics
  (Because it is more powerful, i.e., presence of effects is better detected.)

Use univariate or multivariate?

Both methods are good for controlling for type I errors.

'Sphericity' criterion

- **If sphericity is violated ($\varepsilon < .7$) and sample size is larger than $k + 10$:** Use multivariate test statistics.
  (Because these statistics do not depend on this assumption, and they are more powerful under these conditions.)
- **If sphericity holds ($\varepsilon > .7$) or sample sizes are small:** Use univariate test statistics
  (Because it is more powerful, i.e., presence of effects is better detected.)

'Missing values' criterion

Use univariate tests if there are missing values.

Decomposing the within-factor effect

Both the univariate and multivariate tests are omnibus tests:
They only detect overall effects.

We can understand better the within-factor effect by decomposing it into parts.

Decomposing the within-factor effect

Both the univariate and multivariate tests are omnibus tests:
They only detect overall effects.

We can understand better the within-factor effect by decomposing it into parts.

Q: How to do it?

Decomposing the within-factor effect

Both the univariate and multivariate tests are omnibus tests:
They only detect overall effects.

We can understand better the within-factor effect by decomposing it into parts.

Q: How to do it?
A: Using contrasts.

SPSS has several contrasts available (Deviation, Helmert, Polynomial...).

Polynomial contrasts are interesting since it additively decomposes (and tests) the within-factor effect in linear, quadratic, cubic,..., effects: Trend analysis

('Additively' due to the orthogonality of polynomial contrasts.)

Decomposing the within-factor effect

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Drug	Sphericity Assumed	698.200	3	232.733	24.759	.000
	Greenhouse-Geisser	698.200	1.815	384.763	24.759	.001
	Huynh-Feldt	698.200	3.000	232.733	24.759	.000
	Lower-bound	698.200	1.000	698.200	24.759	.008
Error(Drug)	Sphericity Assumed	112.800	12	9.400		
	Greenhouse-Geisser	112.800	7.258	15.540		
	Huynh-Feldt	112.800	12.000	9.400		
	Lower-bound	112.800	4.000	28.200		

Drug has a significant effect. But what else can we say?

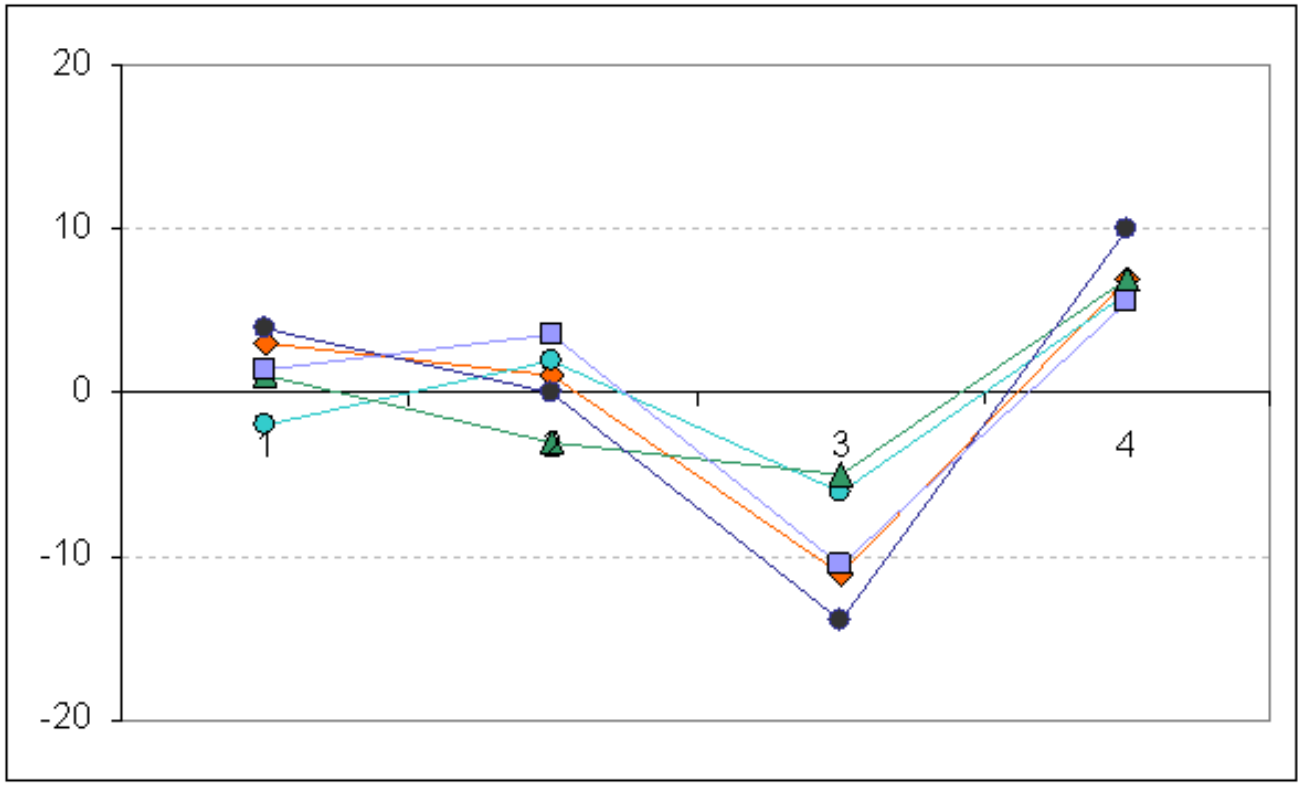
Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	Drug	Type III Sum of Squares	df	Mean Square	F	Sig.
Drug	Linear	11.560	1	11.560	3.074	.154
	Quadratic	369.800	1	369.800	26.797	.007
	Cubic	316.840	1	316.840	29.778	.005
Error(Drug)	Linear	15.040	4	3.760		
	Quadratic	55.200	4	13.800		
	Cubic	42.560	4	10.640		

The effect of drug on reaction times is not significantly linear (but quadratic and cubic trends are sig.)

Decomposing the within-factor effect



Post hoc procedures

In RM-(M)ANOVA we can also use post hoc procedures.

These procedures allow doing multiple comparisons while preventing capitalization by chance (inflation of overall type I error).

In general:

- **If sphericity is violated ($\varepsilon < .7$):** Use Bonferroni procedure.
- **If sphericity holds ($\varepsilon > .7$):** Use Tukey procedure.

Note: Bonferroni is directly available in SPSS in 'Options' button of GLM's `Repeated Measures` menu; Tukey is not. Use of syntax files is needed (downloadable from SPSS' website, for example).

Post hoc procedures

For our drug example:

Pairwise Comparisons

Measure:MEASURE_1

(I) Drug	(J) Drug	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	.800	1.625	1.000	-7.082	8.682
	3	10.800	2.577	.083	-1.700	23.300
	4	-5.600*	.748	.010	-9.230	-1.970
2	1	-.800	1.625	1.000	-8.682	7.082
	3	10.000	2.280	.071	-1.062	21.062
	4	-6.400	1.600	.097	-14.162	1.362
3	1	-10.800	2.577	.083	-23.300	1.700
	2	-10.000	2.280	.071	-21.062	1.062
	4	-16.400*	2.227	.011	-27.204	-5.596
4	1	5.600*	.748	.010	1.970	9.230
	2	6.400	1.600	.097	-1.362	14.162
	3	16.400*	2.227	.011	5.596	27.204

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Extending RM-ANOVA

- The RM-ANOVA situation covered here is the simplest: One within-factor (drug) only.
- The experimental design can be more complicated:
  - One between, one within design (Situation 2).
  - One between, two within design (Situation 3).
  - Two between, one within design (Situation 4).
  - ...

RM-ANOVA can fit any of these extensions.