Statistics for the Sciences

Repeated Measures - Part II

Xuemao Zhang
East Stroudsburg University

January 18, 2025

Outline

- Factorial designs with repeated measures
  - Data format
  - Model description
  - Model fit
- Lab
- Repeated measures on both factors

Factorial designs with repeated measures

- Example (`garcia.csv`): Garcia et al. (2015) studied the effect of the appetite-regulating hormone leptin on appetite and mating preferences in the spadefoot toad Spea bombifrons. Eighteen female toads collected from the wild were allocated to a treatment group ($n = 9$) which received a subcutaneous injection of leptin once per day for six days, and a control group ($n = 9$), which received saline injections with the same frequency. One hour after the day 6 injections, each toad was presented with approximately 50 crickets. The response variable was the cumulative number of attacks by each toad over three-minute intervals for 15 minutes. Treatment (leptin versus control was the fixed between-subject factor and toads were the subjects. The within-subjects fixed factor was time with five groups representing 3, 6, 9, 12 and 15 minutes after the introduction of crickets.
  - Response variable: `cumattack` cumulative number of attacks
  - between-subject factor: `treatment`
  - within-subjects factor: `timefac`
  - Subjects `toad`
    - Repeated measures on one factor only `treatment`

Factorial designs with repeated measures

```
##    toad treatment time numattack cumattack timefac
## 1     1    Leptin    3         7         7       3
## 2     1    Leptin    6         2         9       6
## 3     1    Leptin    9         7        16       9
## 4     1    Leptin   12         4        20      12
## 5     1    Leptin   15         0        20      15
## 6     2    Leptin    3         1         1       3
## 7     2    Leptin    6         2         3       6
## 8     2    Leptin    9         4         7       9
## 9     2    Leptin   12         4        11      12
## 10    2    Leptin   15         3        14      15
## 11    3    Saline    3        12        12       3
## 12    3    Saline    6         8        20       6
## 13    3    Saline    9        11        31       9
## 14    3    Saline   12        11        42      12
## 15    3    Saline   15         5        47      15
## 16    4    Saline    3        14        14       3
## 17    4    Saline    6         6        20       6
## 18    4    Saline    9         8        28       9
## 19    4    Saline   12         5        33      12
## 20    4    Saline   15         1        34      15
## 21    5    Saline    3         5         5       3
## 22    5    Saline    6         6        11       6
## 23    5    Saline    9         6        17       9
## 24    5    Saline   12         9        26      12
## 25    5    Saline   15         6        32      15
## 26    6    Leptin    3        11        11       3
## 27    6    Leptin    6         3        14       6
## 28    6    Leptin    9         3        17       9
## 29    6    Leptin   12         0        17      12
## 30    6    Leptin   15         4        21      15
## 31    7    Saline    3         8         8       3
## 32    7    Saline    6         3        11       6
## 33    7    Saline    9         7        18       9
## 34    7    Saline   12         4        22      12
## 35    7    Saline   15        11        33      15
## 36    8    Leptin    3         9         9       3
## 37    8    Leptin    6         5        14       6
## 38    8    Leptin    9         5        19       9
## 39    8    Leptin   12         2        21      12
```

Factorial designs with repeated measures

- Levels of `treatment`

```
## [1] "Leptin" "Saline"
```

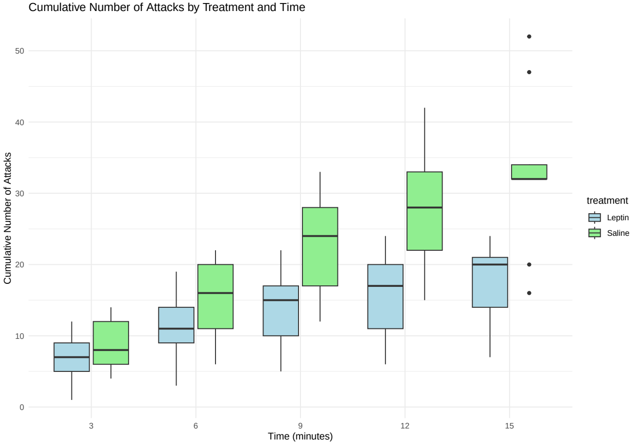
- Levels of `timefac`

```
## [1] "3"  "6"  "9"  "12" "15"
```

- Levels of `toad`

```
##  [1] "1"  "2"  "3"  "4"  "5"  "6"  "7"  "8"  "9"  "10" "11" "12"
## [16] "16" "17" "18"
```

Factorial designs with repeated measures



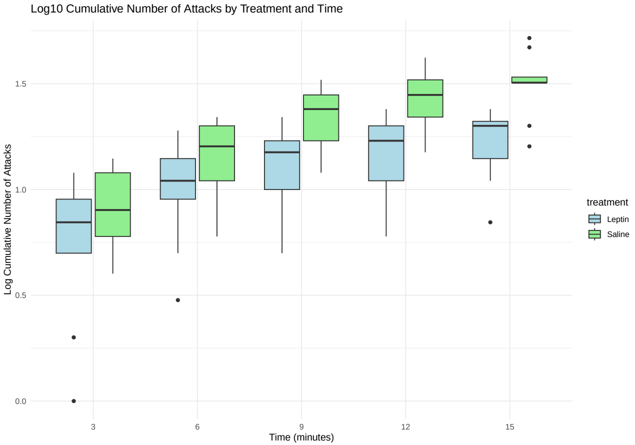
Factorial designs with repeated measures

- Check homogeneity of within-group variances
  - very different variances

```
##    treatment time N cumattack        sd       se       ci
## 1     Leptin    3 9  6.888889  3.789606 1.263202 2.912949
## 2     Leptin    6 9 10.777778  4.918785 1.639595 3.780913
## 3     Leptin    9 9 13.666667  5.612486 1.870829 4.314139
## 4     Leptin   12 9 15.888889  5.883121 1.961040 4.522167
## 5     Leptin   15 9 17.444444  5.811865 1.937288 4.467395
## 6     Saline    3 9  8.888889  3.855011 1.285004 2.963224
## 7     Saline    6 9 14.777778  5.674015 1.891338 4.361434
## 8     Saline    9 9 22.444444  7.731609 2.577203 5.943041
## 9     Saline   12 9 28.222222  9.351173 3.117058 7.187948
## 10    Saline   15 9 33.111111 11.285438 3.761813 8.674756
```

Factorial designs with repeated measures

- So consider log10 transformation of the response



Factorial designs with repeated measures

- Check homogeneity of within-group variances again

```
##    treatment time N logcumattack        sd         se        ci
## 1     Leptin    3 9    0.7391454 0.3626558 0.12088527 0.2787619
## 2     Leptin    6 9    0.9789913 0.2498358 0.08327859 0.1920408
## 3     Leptin    9 9    1.0950095 0.2122436 0.07074786 0.1631449
## 4     Leptin   12 9    1.1675887 0.1933196 0.06443987 0.1485986
## 5     Leptin   15 9    1.2137940 0.1771414 0.05904712 0.1361629
## 6     Saline    3 9    0.9087618 0.2027180 0.06757267 0.1558229
## 7     Saline    6 9    1.1354505 0.1920079 0.06400263 0.1475903
## 8     Saline    9 9    1.3256738 0.1620029 0.05400096 0.1245264
## 9     Saline   12 9    1.4280959 0.1511606 0.05038687 0.1161923
## 10    Saline   15 9    1.4954104 0.1599347 0.05331156 0.1229367
```

Factorial designs with repeated measures

- Description of Design (with **repeated measures on one factor only**):
- Suppose there are two factors: A and B
  - A has two levels: 1, 2
  - B has two levels: 1, 2
  - Repeated measures on factor B
  - Again, each subject is a block

A	Subject	B: 1	B: 2
	1	A_1B_1	A_1B_2
A_1	⋮	⋮	
	s	A_1B_1	A_1B_2
	$s+1$	A_2B_2	A_2B_1
A_2	⋮	⋮	
	$2s$	A_2B_1	A_2B_2

Factorial designs with repeated measures

- Assumption: there are no interactions between treatments and subjects
- Statistical model:
  - subject effect is random; it is nested within factor A

$$Y_{ijk} = \mu + \rho_{i(j)} + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk}, i = 1, \ldots, s; j = 1, \ldots, a; k = 1, \ldots, b$$

  - $\rho_{i(j)} \sim N(0, \sigma_\rho^2)$ are independent
  - $\sum \alpha_j = 0$
  - $\sum \beta_k = 0$
  - $(\alpha\beta)_{jk}$ are constants subject to $\sum_j (\alpha\beta)_{jk} = 0$ for all k and $\sum_k (\alpha\beta)_{jk} = 0$ for all j
  - $\varepsilon_{ijk} \sim N(0, \sigma^2)$ are independent
  - $\rho_{i(j)}$ and ε_{ij} are independent

Factorial designs with repeated measures

Table 1: ANOVA Table for 2-factor with repeated measures on factor B

Source	df	SS	MS	F
Factor A	$a-1$	SS_A	$SS_A/(a-1)$	$MS_A/MS_{S(A)}$
Factor B	$b-1$	SS_B	$SS_B/(b-1)$	MS_B/MS_E
AB interactions	$(a-1)(b-1)$	SS_{AB}	$SS_{AB}/[(a-1)(b-1)]$	SS_{AB}/MS_E
Subjects (within A)	$a(s-1)$	$SS_{S(A)}$	$SS_{S(A)}/[a(s-1)]$	$SS_{S(A)}/MS_E$
Error	$a(s-1)(b-1)$	SSE	$SSE/[a(s-1)(b-1)]$	
Total	$abs-1$	SS_{total}		

- $E(MS_A) = \sigma^2 + b\sigma_\rho^2 + bs\dfrac{\sum \alpha_j^2}{a-1}$
- $E(MS_B) = \sigma^2 + as\dfrac{\sum \beta_k^2}{b-1}$
- $E(MS_{AB}) = \sigma^2 + s\dfrac{\sum\sum(\alpha\beta)_{jk}^2}{(a-1)(b-1)}$
- $E(MS_{S(A)}) = \sigma^2 + b\sigma_\rho^2$
- $E(MS_E) = \sigma^2$

Factorial designs with repeated measures

- Testing interaction H_0 : all $(\alpha\beta)_{jk} = 0$
- Testing factor A main effects H_0 : all $\alpha_j = 0$
- Testing factor B main effects H_0 : all $\beta_k = 0$

Example

- Model fit

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's metho
## lmerModLmerTest]
## Formula: logcumattack ~ treatment * timefac + (1 | toad)
##    Data: garcia
##
## REML criterion at convergence: -70.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.2504 -0.4142  0.0109  0.4192  2.5557
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  toad     (Intercept) 0.036304 0.19053
##  Residual             0.009734 0.09866
## Number of obs: 90, groups:  toad, 18
##
## Fixed effects:
```

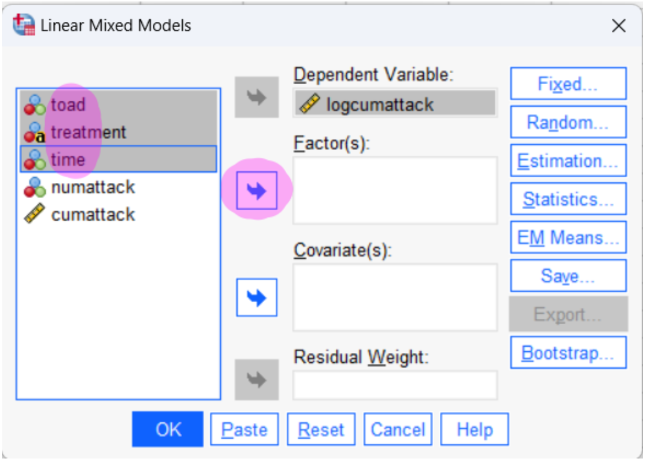
Example

- ANOVA table

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                   Sum Sq Mean Sq NumDF DenDF F value  Pr(>F)
## treatment         0.0553 0.05531     1    16  5.6823 0.02987 *
## timefac           3.2808 0.82020     4    64 84.2656 < 2e-16 ***
## treatment:timefac 0.0546 0.01364     4    64  1.4016 0.24353
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

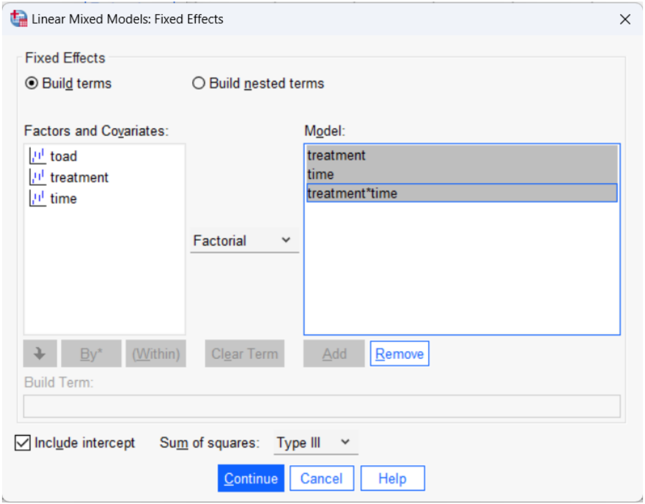
Lab

- After importing data `garcia.csv`, we add a variable of log10 transformation of the response `cumattack`
- In the `Linear Mixed Models` dialog, click on `Continue`
- Then add the dependent variable and factors



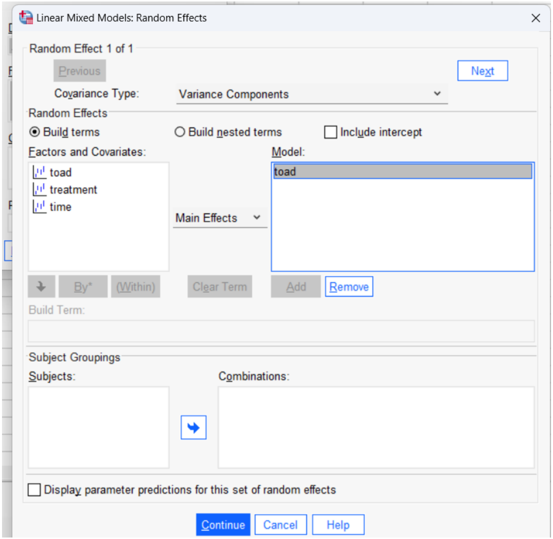
Lab

- Add the fixed effects

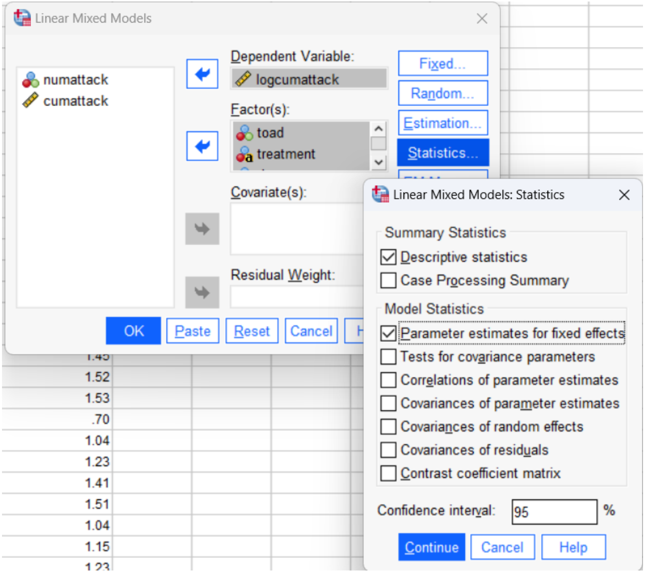


Lab

- Add the random effect



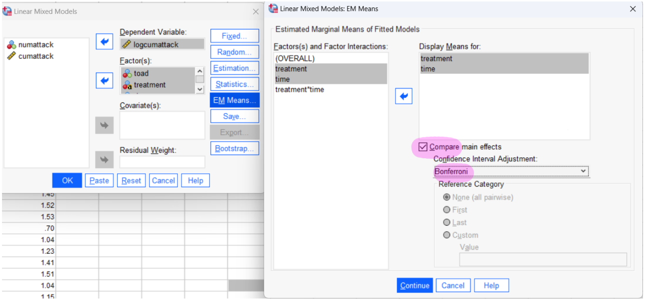
Lab



Linear Mixed Models
numattack
cumattack
Dependent Variable:
logcumattack
Factor(s):
toad
treatment
Covariate(s):
Residual Weight:
Fixed...
Random...
Estimation...
Statistics...
OK
Paste
Reset
Cancel
1.52
1.53
.70
1.04
1.23
1.41
1.51
1.04
1.15
Linear Mixed Models: Statistics
Summary Statistics
Descriptive statistics
Case Processing Summary
Model Statistics
Parameter estimates for fixed effects
Tests for covariance parameters
Correlations of parameter estimates
Covariances of parameter estimates
Covariances of random effects
Covariances of residuals
Contrast coefficient matrix
Confidence interval:
95
%
Continue
Cancel
Help

Lab

- We may want to compare the means of the fixed effects



Lab

Model Dimension[a]

		Number of Levels	Covariance Structure	Number of Parameters
Fixed Effects	Intercept	1		1
	treatment	2		1
	time	5		4
	treatment * time	10		4
Random Effects	toad	18	Variance Components	1
Residual				1
Total		36		12

a. Dependent Variable: logcumattack.

Lab

Fixed Effects

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16	621.042	<.001
treatment	1	16	5.682	.030
time	4	64.000	84.266	<.001
treatment * time	4	64.000	1.402	.244

a. Dependent Variable: logcumattack.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	1.495	.072	22.940	20.909	<.001	1.347	1.643
[treatment=Leptin]	-.282	.101	22.940	-2.784	.011	-.491	-.072
[treatment=Saline]	0[b]	0	.	.	.	.	.
[time=3]	-.587	.047	64.000	-12.614	<.001	-.680	-.494
[time=6]	-.360	.047	64.000	-7.740	<.001	-.453	-.267
[time=9]	-.170	.047	64.000	-3.650	<.001	-.263	-.077
[time=12]	-.067	.047	64.000	-1.447	.153	-.160	.026
[time=15]	0[b]	0	.	.	.	.	.
[treatment=Leptin] * [time=3]	.112	.066	64.000	1.703	.093	-.019	.243
[treatment=Leptin] * [time=6]	.125	.066	64.000	1.903	.062	-.006	.257
[treatment=Leptin] * [time=9]	.051	.066	64.000	.775	.441	-.080	.182
[treatment=Leptin] * [time=12]	.021	.066	64.000	.321	.749	-.110	.153
[treatment=Leptin] * [time=15]	0[b]	0	.	.	.	.	.

Lab

Estimated Marginal Means

1. treatment

Estimates[a]

treatment	Mean	Std. Error	df	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Leptin	1.039	.065	16	.901	1.177
Saline	1.259	.065	16	1.120	1.397

a. Dependent Variable: logcumattack.

Pairwise Comparisons[a]

(I) treatment	(J) treatment	Mean Difference (I-J)	Std. Error	df	Sig.[c]	95% Confidence Interval for Difference[c] Lower Bound	95% Confidence Interval for Difference[c] Upper Bound
Leptin	Saline	-.220*	.092	16	.030	-.415	-.024
Saline	Leptin	.220*	.092	16	.030	.024	.415

Based on estimated marginal means

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

a. Dependent Variable: logcumattack.

Lab

2. time

Estimates[a]

time	Mean	Std. Error	df	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
3	.824	.051	22.940	.719	.929
6	1.057	.051	22.940	.953	1.162
9	1.210	.051	22.940	1.106	1.315
12	1.298	.051	22.940	1.193	1.402
15	1.355	.051	22.940	1.250	1.459

a. Dependent Variable: logcumattack.

Pairwise Comparisons[a]

(I) time	(J) time	Mean Difference (I-J)	Std. Error	df	Sig.[c]	95% Confidence Interval for Difference[c] Lower Bound	95% Confidence Interval for Difference[c] Upper Bound
3	6	-.233*	.033	64.000	<.001	-.329	-.138
	9	-.386*	.033	64.000	<.001	-.482	-.291
	12	-.474*	.033	64.000	<.001	-.570	-.378
	15	-.531*	.033	64.000	<.001	-.626	-.435
6	3	.233*	.033	64.000	<.001	.138	.329
	9	-.153*	.033	64.000	<.001	-.249	-.058
	12	-.241*	.033	64.000	<.001	-.336	-.145
	15	-.297*	.033	64.000	<.001	-.393	-.202
9	3	.386*	.033	64.000	<.001	.291	.482
	6	.153*	.033	64.000	<.001	.058	.249
	12	-.088	.033	64.000	.098	-.183	.008
	15	-.144*	.033	64.000	<.001	-.240	-.049
12	3	.474*	.033	64.000	<.001	.378	.570

Repeated measures on both factors

Description of Design:

- Suppose there are two factors: A and B
  - A has two levels: 1, 2
  - B has two levels: 1, 2
  - Repeated measures on both factors: A and B
  - Again, each subject is a block

	Treatment Order 1	2	3	4
Subject 1	A_1B_2	A_2B_2	A_1B_1	A_2B_1
2	A_2B_1	A_1B_2	A_2B_2	A_1B_1
3	A_2B_2	A_1B_1	A_2B_1	A_1B_2
4	A_1B_1	A_2B_1	A_1B_2	A_2B_2

Repeated measures on both factors

- Statistical model:

  $Y_{ijk} = \mu + \rho_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + (\rho\alpha)_{ij} + (\rho\beta)_{ik} + \varepsilon_{ijk}, i = 1, \ldots, s; j = 1, \ldots, a; k = 1, \ldots, b$

  - $\rho_i \sim N(0, \sigma_\rho^2)$ are independent
  - $\sum \alpha_j = 0$
  - $\sum \beta_k = 0$
  - $(\alpha\beta)_{jk}$ are constants subject to $\sum_j (\alpha\beta)_{jk} = 0$ for all k and $\sum_k (\alpha\beta)_{jk} = 0$ for all j
  - $(\rho\alpha)_{ij}$ and $(\rho\beta)_{ik}$ are mixed effects: $\cdots\cdots$
  - $\varepsilon_{ijk} \sim N(0, \sigma^2)$ are independent
  - ρ_i, $(\rho\alpha)_{ij}$ and $(\rho\beta)_{ik}$ are pairwise independent
  - ε_{ij} are independent of ρ_i, $(\rho\alpha)_{ij}$ and $(\rho\beta)_{ik}$

Repeated measures on both factors

Table 2: ANOVA Table for 2-factor with repeated measures on both factors

Source	df	SS	MS	F
Subjects(S)	$s-1$	SS_S	$SS_S/(s-1)$	SS_S/MS_E
Factor A	$a-1$	SS_A	$SS_A/(a-1)$	MS_A/MS_{AS}
Factor B	$b-1$	SS_B	$SS_B/(b-1)$	MS_B/MS_{BS}
AB interactions	$(a-1)(b-1)$	SS_{AB}	$SS_{AB}/[(a-1)(b-1)]$	SS_{AB}/MS_E
AS interactions	$(a-1)(s-1)$	SS_{AS}	$SS_{AS}/[(a-1)(s-1)]$	SS_{AS}/MS_E
BS interactions	$(b-1)(s-1)$	SS_{BS}	$SS_{BS}/[(b-1)(s-1)]$	SS_{BS}/MS_E
Error	$(a-1)(b-1)(s-1)$	SSE	$SSE/[(a-1)(b-1)(s-1)]$	
Total	$abs-1$	SS_{total}		

- $E(MS_S) = \sigma^2 + ab\sigma_\rho^2$
- $E(MS_A) = \sigma^2 + b\sigma_{\rho\alpha}^2 + bs\dfrac{\sum \alpha_j^2}{a-1}$
- $E(MS_B) = \sigma^2 + a\sigma_{\rho\beta}^2 + as\dfrac{\sum \alpha_k^2}{b-1}$
- $E(MS_{AB}) = \sigma^2 + s\dfrac{\sum\sum(\alpha\beta)_{jk}^2}{(a-1)(b-1)}$
- $E(MS_{AS}) = \sigma^2 + b\sigma_{\rho\alpha}^2$
- $E(MS_{BS}) = \sigma^2 + a\sigma_{\rho\beta}^2$
- $E(MS_E) = \sigma^2$

License