Multivariate Statistics in Ecology and Quantitative Genetics
1. ANalysis Of VAriance (ANOVA)

Dirk Metzler

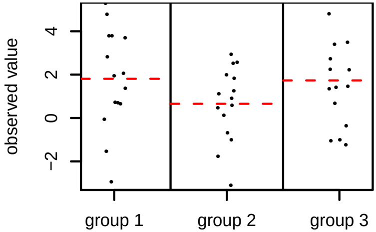
http://evol.bio.lmu.de/_statgen.html

26. Juni 2016

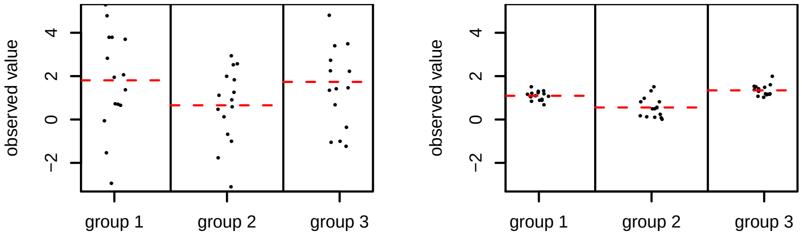
ANOVA and F-Test

Kruskal-Walis Test

Are the group means significantly different?

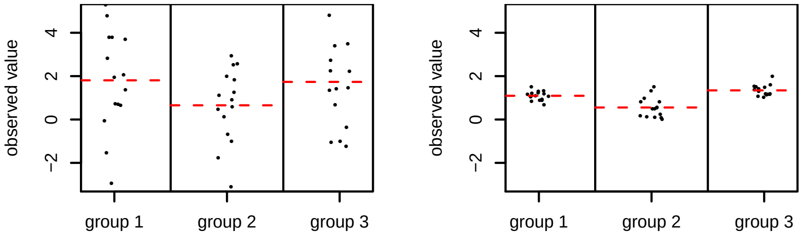


Are the group means significantly different?



Or does this look like random deviations?

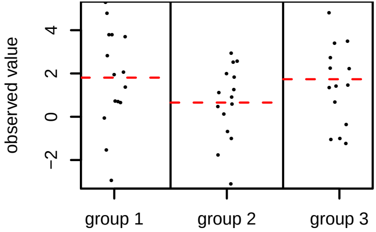
Are the group means significantly different?



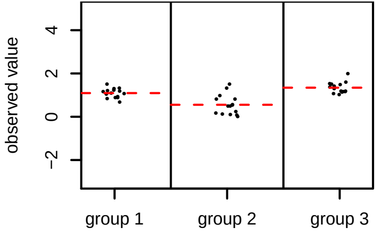
Or does this look like random deviations?

This depends on the ratio of the varibility of group means and the variability within the groups.

Are the group means significantly different?



Large variability
within groups

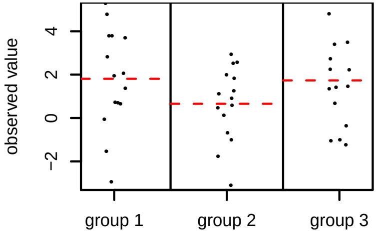


Small variability
with groups

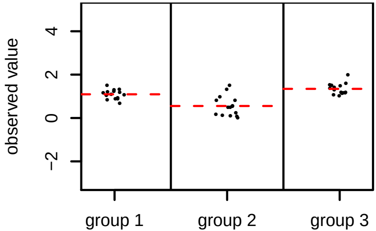
Or does this look like random deviations?

This depends on the ratio of the varibility of group means and the variability within the groups.

Are the group means significantly different?



Large variability
within groups



Small variability
with groups

Or does this look like random deviations?

This depends on the ratio of the varibility of group means and the variability within the groups.
The analysis of variance (ANOVA) quantifies this ratio and its significance.

Example

Blood-clotting time in rats under 4 different treatments

group	observation							
1	62	60	63	59				
2	63	67	71	64	65	66		
3	68	66	71	67	68	68		
4	56	62	60	61	63	64	63	59

Example

Blood-clotting time in rats under 4 different treatments

group	observation							
1	62	60	63	59				
2	63	67	71	64	65	66		
3	68	66	71	67	68	68		
4	56	62	60	61	63	64	63	59

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

Example

Blood-clotting times in rats under 4 different treatments

gr.	$\overline{x}_{i\cdot}$	observations							
1	61	62 $(62-61)^2$	60 $(60-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$				
2	66	63 $(63-66)^2$	67 $(67-66)^2$	71 $(71-66)^2$	64 $(64-66)^2$	65 $(65-66)^2$	66 $(66-66)^2$		
3	68	68 $(68-68)^2$	66 $(66-68)^2$	71 $(71-68)^2$	67 $(67-68)^2$	68 $(68-68)^2$	68 $(68-68)^2$		
4	61	56 $(56-61)^2$	62 $(62-61)^2$	60 $(60-61)^2$	61 $(61-61)^2$	63 $(63-61)^2$	64 $(64-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

Example

Blood-clotting times in rats under 4 different treatments

gr.	$\overline{x}_{i\cdot}$	observations							
1	61	62 $(62-61)^2$	60 $(60-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$				
2	66	63 $(63-66)^2$	67 $(67-66)^2$	71 $(71-66)^2$	64 $(64-66)^2$	65 $(65-66)^2$	66 $(66-66)^2$		
3	68	68 $(68-68)^2$	66 $(66-68)^2$	71 $(71-68)^2$	67 $(67-68)^2$	68 $(68-68)^2$	68 $(68-68)^2$		
4	61	56 $(56-61)^2$	62 $(62-61)^2$	60 $(60-61)^2$	61 $(61-61)^2$	63 $(63-61)^2$	64 $(64-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

The red Differences (unsquared) are the *residuals*: they are the residual variability which is not explained by the model.

Example

Blood-clotting times in rats under 4 different treatments

gr.	$\overline{x}_{i\cdot}$	observations							
1	61	62 $(62-61)^2$	60 $(60-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$				
2	66	63 $(63-66)^2$	67 $(67-66)^2$	71 $(71-66)^2$	64 $(64-66)^2$	65 $(65-66)^2$	66 $(66-66)^2$		
3	68	68 $(68-68)^2$	66 $(66-68)^2$	71 $(71-68)^2$	67 $(67-68)^2$	68 $(68-68)^2$	68 $(68-68)^2$		
4	61	56 $(56-61)^2$	62 $(62-61)^2$	60 $(60-61)^2$	61 $(61-61)^2$	63 $(63-61)^2$	64 $(64-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

The red Differences (unsquared) are the *residuals*: they are the residual variability which is not explained by the model.

Sums of squares within groups:

$ss_{\mathrm{within}} = 112$,

Example

Blood-clotting times in rats under 4 different treatments

gr.	$\overline{x}_{i\cdot}$	observations							
1	61	62 $(62-61)^2$	60 $(60-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$				
2	66	63 $(63-66)^2$	67 $(67-66)^2$	71 $(71-66)^2$	64 $(64-66)^2$	65 $(65-66)^2$	66 $(66-66)^2$		
3	68	68 $(68-68)^2$	66 $(66-68)^2$	71 $(71-68)^2$	67 $(67-68)^2$	68 $(68-68)^2$	68 $(68-68)^2$		
4	61	56 $(56-61)^2$	62 $(62-61)^2$	60 $(60-61)^2$	61 $(61-61)^2$	63 $(63-61)^2$	64 $(64-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

The red Differences (unsquared) are the *residuals*: they are the residual variability which is not explained by the model.

Sums of squares within groups:
$ss_{\text{within}} = 112$, 20 degrees of freedom (df)

Example

Blood-clotting times in rats under 4 different treatments

gr.	$\overline{x}_{i\cdot}$	observations							
1	61	62 $(62-61)^2$	60 $(60-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$				
2	66	63 $(63-66)^2$	67 $(67-66)^2$	71 $(71-66)^2$	64 $(64-66)^2$	65 $(65-66)^2$	66 $(66-66)^2$		
3	68	68 $(68-68)^2$	66 $(66-68)^2$	71 $(71-68)^2$	67 $(67-68)^2$	68 $(68-68)^2$	68 $(68-68)^2$		
4	61	56 $(56-61)^2$	62 $(62-61)^2$	60 $(60-61)^2$	61 $(61-61)^2$	63 $(63-61)^2$	64 $(64-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

The red Differences (unsquared) are the *residuals*: they are the residual variability which is not explained by the model.

Sums of squares within groups:
$ss_{\mathrm{within}} = 112$, 20 degrees of freedom (df)
Sums of squares between groups:
$ss_{\mathrm{betw}} = 4 \cdot (61-64)^2 + 6 \cdot (66-64)^2 + 6 \cdot (68-64)^2 + 8 \cdot (61-64)^2 = 228$,

Example

Blood-clotting times in rats under 4 different treatments

gr.	$\overline{x}_{i\cdot}$	observations							
1	61	62	60	63	59				
		$(62-61)^2$	$(60-61)^2$	$(63-61)^2$	$(59-61)^2$				
2	66	63	67	71	64	65	66		
		$(63-66)^2$	$(67-66)^2$	$(71-66)^2$	$(64-66)^2$	$(65-66)^2$	$(66-66)^2$		
3	68	68	66	71	67	68	68		
		$(68-68)^2$	$(66-68)^2$	$(71-68)^2$	$(67-68)^2$	$(68-68)^2$	$(68-68)^2$		
4	61	56	62	60	61	63	64	63	59
		$(56-61)^2$	$(62-61)^2$	$(60-61)^2$	$(61-61)^2$	$(63-61)^2$	$(64-61)^2$	$(63-61)^2$	$(59-61)^2$

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

The red Differences (unsquared) are the *residuals*: they are the residual variability which is not explained by the model.

Sums of squares within groups:
$ss_{\text{within}} = 112$, 20 degrees of freedom (df)
Sums of squares between groups:
$ss_{\text{betw}} = 4 \cdot (61-64)^2 + 6 \cdot (66-64)^2 + 6 \cdot (68-64)^2 + 8 \cdot (61-64)^2 = 228$,
3 degrees of freedom (df)

Example

Blood-clotting times in rats under 4 different treatments

gr.	$\overline{x}_{i\cdot}$	observations							
1	61	62 $(62-61)^2$	60 $(60-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$				
2	66	63 $(63-66)^2$	67 $(67-66)^2$	71 $(71-66)^2$	64 $(64-66)^2$	65 $(65-66)^2$	66 $(66-66)^2$		
3	68	68 $(68-68)^2$	66 $(66-68)^2$	71 $(71-68)^2$	67 $(67-68)^2$	68 $(68-68)^2$	68 $(68-68)^2$		
4	61	56 $(56-61)^2$	62 $(62-61)^2$	60 $(60-61)^2$	61 $(61-61)^2$	63 $(63-61)^2$	64 $(64-61)^2$	63 $(63-61)^2$	59 $(59-61)^2$

global mean $\overline{x}_{\cdot\cdot} = 64$,
group means $\overline{x}_{1\cdot} = 61$, $\overline{x}_{2\cdot} = 66$, $\overline{x}_{3\cdot} = 68$, $\overline{x}_{4\cdot} = 61$.

The red Differences (unsquared) are the *residuals*: they are the residual variability which is not explained by the model.

Sums of squares within groups:
$ss_{\text{within}} = 112$, 20 degrees of freedom (df)
Sums of squares between groups:
$ss_{\text{betw}} = 4 \cdot (61-64)^2 + 6 \cdot (66-64)^2 + 6 \cdot (68-64)^2 + 8 \cdot (61-64)^2 = 228$,
3 degrees of freedom (df)

$$F = \frac{ss_{\text{betw}}/3}{ss_{\text{within}}/20} = \frac{76}{5.6} = 13.57$$

Example: Blood-clotting times in rats under 4 different treatments.

ANOVA table („ANalysis Of VAriance“)

	df	sum of squares (ss)	mean sum of squares (ss/df)	*F* value
groups	3	228	76	13.57
residuals	20	112	5.6	

Example: Blood-clotting times in rats under 4 different treatments.

ANOVA table („ANalysis Of VAriance“)

	df	sum of squares (ss)	mean sum of squares (ss/df)	F value
groups	3	228	76	13.57
residuals	20	112	5.6	

Under the Null-Hypothesis H_0 “the group means are equal”
(and assuming independent, normally distributed observations)
is F Fisher-distributed with 3 and 20 degrees of freedom, and
$p = \mathrm{Fisher}_{3,20}([13.57, \infty)) \leq 5 \cdot 10^{-5}$.

Example: Blood-clotting times in rats under 4 different treatments.

ANOVA table („ANalysis Of VAriance“)

	df	sum of squares (ss)	mean sum of squares (ss/df)	F value
groups	3	228	76	13.57
residuals	20	112	5.6	

Under the Null-Hypothesis H_0 “the group means are equal”
(and assuming independent, normally distributed observations)
is F Fisher-distributed with 3 and 20 degrees of freedom, and
$p = \mathrm{Fisher}_{3,20}([13.57, \infty)) \leq 5 \cdot 10^{-5}$.

Thus, we can reject H_0.

Sir Ronald Aylmer Fisher,
1890–1962

F-Test

$n = n_1 + n_2 + \cdots + n_I$ obersvations in I groups,
$X_{ij} = j$-th observation in i-th group, $j = 1, \ldots, n_i$.

F-Test

$n = n_1 + n_2 + \cdots + n_I$ obersvations in I groups,
$X_{ij} = j$-th observation in i-th group, $j = 1, \ldots, n_i$.

Model assumption: $X_{ij} = \mu_i + \varepsilon_{ij}$,
with independent, normally distributed ε_{ij}, $\mathbb{E}[\varepsilon_{ij}] = 0$, $\mathrm{Var}[\varepsilon_{ij}] = \sigma^2$

F-Test

$n = n_1 + n_2 + \cdots + n_I$ obersvations in I groups,
X_{ij} = j-th observation in i-th group, $j = 1, \ldots, n_i$.

Model assumption: $X_{ij} = \mu_i + \varepsilon_{ij}$,
with independent, normally distributed ε_{ij}, $\mathbb{E}[\varepsilon_{ij}] = 0$, $\mathrm{Var}[\varepsilon_{ij}] = \sigma^2$

(μ_i is the "true" mean within group i.)

F-Test

$n = n_1 + n_2 + \cdots + n_I$ obersvations in I groups,
$X_{ij} = j$-th observation in i-th group, $j = 1, \ldots, n_i$.

Model assumption: $X_{ij} = \mu_i + \varepsilon_{ij}$,
with independent, normally distributed ε_{ij}, $\mathbb{E}[\varepsilon_{ij}] = 0$, $\mathrm{Var}[\varepsilon_{ij}] = \sigma^2$

(μ_i is the “true” mean within group i.)

$\overline{X}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{I} \sum_{j=1}^{n_i} X_{ij}$ (empirical) “global mean”
$\overline{X}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$ (empirical) mean of group i

F-Test

$n = n_1 + n_2 + \cdots + n_I$ obersvations in I groups,
X_{ij} = j-th observation in i-th group, $j = 1, \ldots, n_i$.

Model assumption: $X_{ij} = \mu_i + \varepsilon_{ij}$,
with independent, normally distributed ε_{ij}, $\mathbb{E}[\varepsilon_{ij}] = 0$, $\mathrm{Var}[\varepsilon_{ij}] = \sigma^2$

(μ_i is the "true" mean within group i.)

$\overline{X}_{\cdot\cdot} = \frac{1}{n}\sum_{i=1}^{I}\sum_{j=1}^{n_i} X_{ij}$ (empirical) "global mean"
$\overline{X}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$ (empirical) mean of group i

$$SS_{\text{within}} = \sum_{i=1}^{I}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_{i\cdot})^2$$

sum of squares within the groups,
$n - I$ degrees of freedom

$$SS_{\text{betw}} = \sum_{i=1}^{I} n_i(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2$$

sum of squares between the groups,
$I - 1$ degrees of freedom

F-Test

$n = n_1 + n_2 + \cdots + n_I$ obersvations in I groups,
$X_{ij} = j$-th observation in i-th group, $j = 1, \ldots, n_i$.

Model assumption: $X_{ij} = \mu_i + \varepsilon_{ij}$,
with independent, normally distributed ε_{ij}, $\mathbb{E}[\varepsilon_{ij}] = 0$, $\mathrm{Var}[\varepsilon_{ij}] = \sigma^2$

(μ_i is the "true" mean within group i.)

$\overline{X}_{\cdot\cdot} = \frac{1}{n}\sum_{i=1}^{I}\sum_{j=1}^{n_i} X_{ij}$ (empirical) "global mean"
$\overline{X}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$ (empirical) mean of group i

$$SS_{\mathrm{within}} = \sum_{i=1}^{I}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_{i\cdot})^2$$

sum of squares within the groups,
$n - I$ degrees of freedom

$$SS_{\mathrm{betw}} = \sum_{i=1}^{I} n_i(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2$$

sum of squares between the groups,
$I - 1$ degrees of freedom

$$F = \frac{SS_{\mathrm{betw}}/(I-1)}{SS_{\mathrm{within}}/(n-I)}$$

F-Test

X_{ij} = j-th observation i-th group, $j = 1, \dots, n_i$,

Model assumption: $X_{ij} = \mu_i + \varepsilon_{ij}$. $\mathbb{E}[\varepsilon_{ij}] = 0$, $\mathrm{Var}[\varepsilon_{ij}] = \sigma^2$

$$SS_{\text{within}} = \sum_{i=1}^{I}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_{i\cdot})^2$$
sum of squares within groups, $n - I$ degrees of feedom

$$SS_{\text{betw}} = \sum_{i=1}^{I} n_i(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2$$
sum of squares between groups, $I - 1$ degrees of freedom

$$F = \frac{SS_{\text{betw}}/(I-1)}{SS_{\text{within}}/(n-I)}$$

Under the hypothesis $H_0 : \mu_1 = \cdots = \mu_I$ ("all μ_i are equal")
F is Fisher-distributed with $I - 1$ and $n - I$ degrees of freedom
(no matter what the true joint value of μ_i is).

F-Test: We reject H_0 on the level of significance α if $F \geq q_\alpha$,
whereas q_α is the $(1-\alpha)$-quantile of the Fisher-distribution with
$I - 1$ and $n - I$ degrees of freedom.

The statistic of the Kruskal-Wallis test

$$\left[\frac{12}{\sum^{a} n_i \left(\sum^{a} n_i + 1\right)} \sum^{a} \frac{\left(\sum^{n_i} R\right)_i^2}{n_i}\right] - 3 \cdot \left(\sum^{a} n_i + 1\right)$$

The statistic of the Kruskal-Wallis test

$$\left[\frac{12}{\sum^a n_i \left(\sum^a n_i + 1\right)} \sum^a \frac{\left(\sum^{n_i} R\right)_i^2}{n_i}\right] - 3 \cdot \left(\sum^a n_i + 1\right)$$

Explanations are given verbally in the lecture and in Sokal and Rohlf (1995) *Biometry*, 3rd ed., on pp. 423–426.