Introduction to Geostatistics

9. Two-sample T-test: Type-II errors and Power Analysis of Variance

Kristina Helle
(Edzer J. Pebesma)

kristina.helle@uni-muenster.de
Institute for Geoinformatics (**ifgi**)
University of Münster

summer semester 2009,
June 23, 2009

Sample Size for Conficence Intervals

Conficence Interval for mean μ at confidence level α (e.g. $\alpha = 0.05$)

$$\mu \in [\bar{X} - t_{1-\frac{\alpha}{2},df} \cdot \frac{s}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2},df} \cdot \frac{s}{\sqrt{n}}]$$

$\bar{X}$: sample mean; s^2: sample variance; n: sample size; $t_{1-\alpha/2,df}$; $(1-\frac{\alpha}{2})$-quantile of t-distribution

Therefore the confidence interval of width $W = 2 \cdot t_{1-\frac{\alpha}{2},df} \cdot \frac{s}{\sqrt{n}}$ can be obtained by sample size

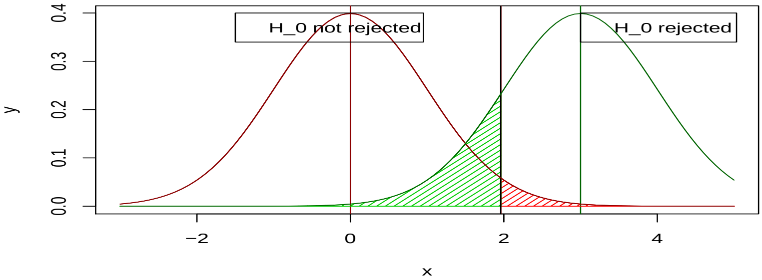
$$n = \left(2 \cdot t_{1-\frac{\alpha}{2},df} \cdot \frac{s}{W}\right)^2$$

Type II Error / Power

For fixed alternative H_0: $\mu = \mu_0$; H_A: $\mu = \mu_A$

Type I error, α: reject H_0 even if it is true

Type II error, β: not reject H_0 even if it is false

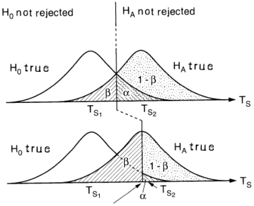


	test result	
truth	H_0 not rejected	H_0 rejected
H_0 true	OK, $(1-\alpha)$	Type I error, α
H_0 false	Type II error, β	OK, $(1-\beta)$, power

Controlling Power (alternative fixed)

Power $(1-\beta)$ can be increased by

increasing α



Sachs & Hedderich, p.315, changed

small $SE = \frac{\sigma}{\sqrt{n}}$ (by big sample size n)

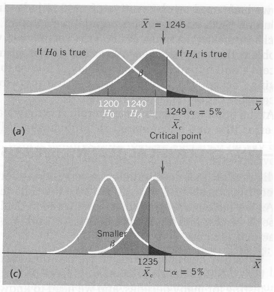
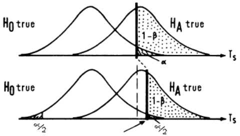


FIGURE 9-7

Wonnacott & Wonnacott, p.307, changed

using one-sided tests

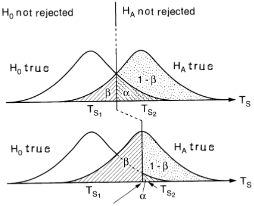


Sachs & Hedderich, p.316, changed

Controlling Power (alternative fixed)

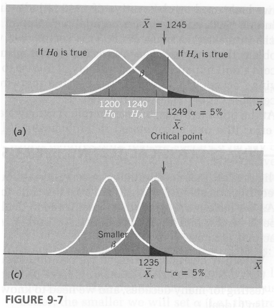
Power $(1-\beta)$ can be increased by

increasing α



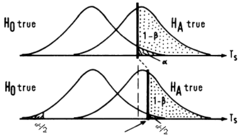
Sachs & Hedderich, p.315, changed

small $SE = \frac{\sigma}{\sqrt{n}}$ (by big sample size n)



Wonnacott & Wonnacott, p.307, changed

using one-sided tests

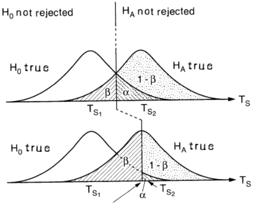


Sachs & Hedderich, p.316, changed

Controlling Power (alternative fixed)

Power $(1-\beta)$ can be increased by

increasing α



Sachs & Hedderich, p.315, changed

small $SE = \frac{\sigma}{\sqrt{n}}$ (by big sample size n)

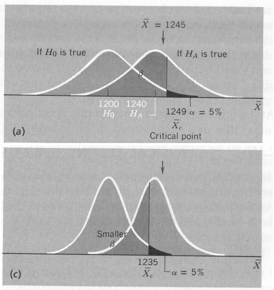
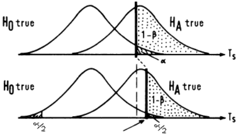


FIGURE 9-7

Wonnacott & Wonnacott, p.307, changed

using one-sided tests

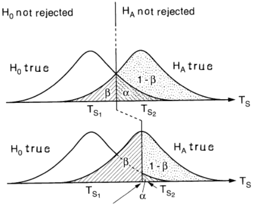


Sachs & Hedderich, p.316, changed

Controlling Power (alternative fixed)

Power $(1-\beta)$ can be increased by

increasing α



Sachs & Hedderich, p.315, changed

small $SE = \frac{\sigma}{\sqrt{n}}$ (by big sample size n)

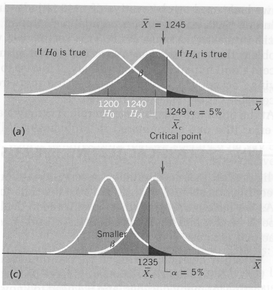
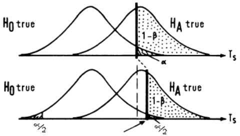


FIGURE 9-7

Wonnacott & Wonnacott, p.307, changed

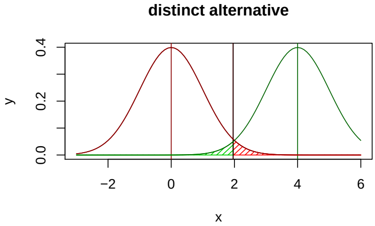
using one-sided tests

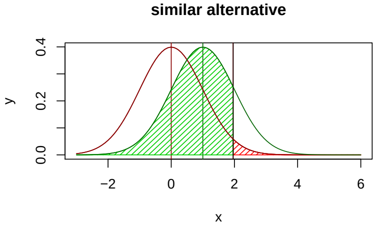


Sachs & Hedderich, p.316, changed

Power dependent on alternative

Very different alternatives are easier to detect than very similar alternatives.

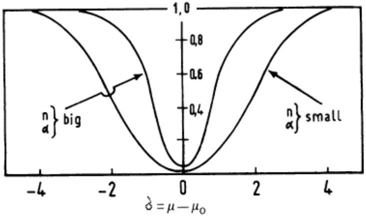




Power $(1-\beta)$ dependent on difference of alternatives $(\delta = \mu_A - \mu_0)$

small: for small sample, small α

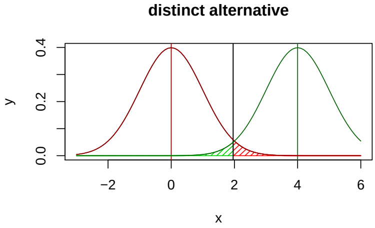
big: for big sample, big α

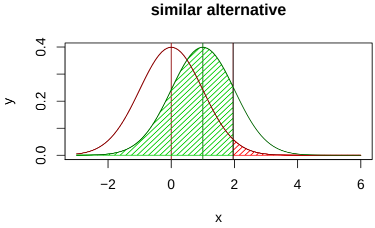


Sachs & Hedderich, p.317, changed

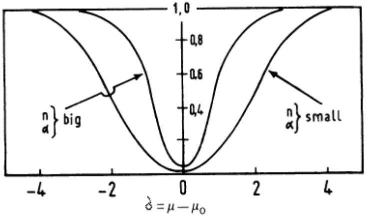
Power dependent on alternative

Very different alternatives are easier to detect than very similar alternatives.





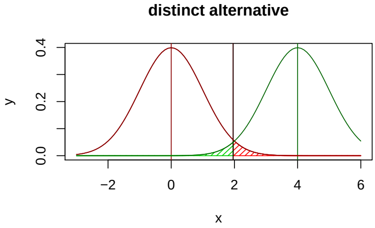
Power $(1 - \beta)$ dependent on difference of alternatives $(\delta = \mu_A - \mu_0)$
small: for small sample, small α
big: for big sample, big α

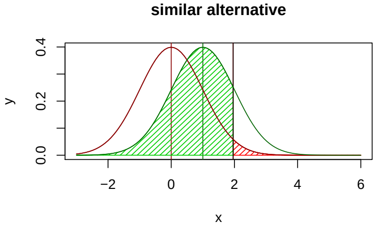


Sachs & Hedderich, p.317, changed

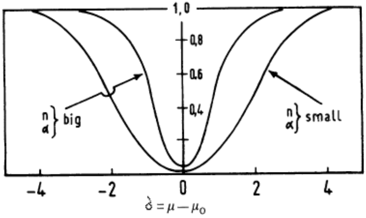
Power dependent on alternative

Very different alternatives are easier to detect than very similar alternatives.





Power $(1-\beta)$ dependent on difference of alternatives $(\delta = \mu_A - \mu_0)$
small: for small sample, small α
big: for big sample, big α



Sachs & Hedderich, p.317, changed

power.t.test()

For given $n = 10$, $\alpha = 0.05$, $\delta = 2$, $s = 1.36$ the power can be computed with

```
> power.t.test(n = 10, sig.level = 0.05, delta = 2, sd = 1.36,
+     power = NULL)

     Two-sample t test power calculation

              n = 10
          delta = 2
             sd = 1.36
      sig.level = 0.05
          power = 0.8746395
    alternative = two.sided

 NOTE: n is number in *each* group
```

setting `n = NULL` computes required sample size . . .

General Concept of Statistical Tests I

a lot of information — strong decision

$\overset{Test}{\Rightarrow}$

little information — weak decision

examples:

information	decision
big sample	narrow confidence intervals
narrow distribution	small p-value (strong rejection of H_0)
known (normal) distribution	
unknown distribution	high p-value (H_0 can not be rejected)
skewed, wide distribution	high significance level (type I error: $\alpha = 10\%$)
...	

General Concept of Statistical Tests I

a lot of information — strong decision

$\overset{Test}{\Rightarrow}$

little information — weak decision

examples:

information	decision
big sample	narrow confidence intervals
narrow distribution	small p-value (strong rejection of H_0)
known (normal) distribution	
unknown distribution	high p-value (H_0 can not be rejected)
skewed, wide distribution	high significance level (type I error: $\alpha = 10\%$)
...	

General Concept of Statistical Tests I

a lot of information — strong decision

$\overset{Test}{\Rightarrow}$

little information — weak decision

examples:

information	decision
big sample	narrow confidence intervals
narrow distribution	small p-value (strong rejection of H_0)
known (normal) distribution	
	high p-value (H_0 can not be rejected)
unknown distribution	
skewed, wide distribution	high significance level (type I error: $\alpha = 10\%$)
...	

paired t-test

If data is given in pairs, we can investigate the mean difference in each pair instead of the difference of the means of each group.
Thus we keep other factors constant and often increase power.

obj	t_1	t_2
1	13.5	12.7
2	15.3	15.1
3	7.5	6.6
4	10.3	8.5
5	8.7	8.0

examples:

- value: N concentration in soil; obj: boreholes at different locations; t1 / t2: before / after use of fertilizer; question: What is the average effect of the fertilizer if location is fixed?
- value: ozone concentration in air; obj: different stations; t1 / t2: 7:00 / 15:00; question: How much does ozone concentration change between 7:00 and 15:00?
- value: ozone concentration in air; obj: different dates / hours; t1 / t2: two different stations; question: How much do the concentrations at the two stations differ?

paired t-test

If data is given in pairs, we can investigate the mean difference in each pair instead of the difference of the means of each group.
Thus we keep other factors constant and often increase power.

obj	t_1	t_2
1	13.5	12.7
2	15.3	15.1
3	7.5	6.6
4	10.3	8.5
5	8.7	8.0

examples:

- value: N concentration in soil; obj: boreholes at different locations; t1 / t2: before / after use of fertilizer; question: What is the average effect of the fertilizer if location is fixed?
- value: ozone concentration in air; obj: different stations; t1 / t2: 7:00 / 15:00; question: How much does ozone concentration change between 7:00 and 15:00?
- value: ozone concentration in air; obj: different dates / hours; t1 / t2: two different stations; question: How much do the concentrations at the two stations differ?

paired t-test

If data is given in pairs, we can investigate the mean difference in each pair instead of the difference of the means of each group.
Thus we keep other factors constant and often increase power.

obj	t_1	t_2
1	13.5	12.7
2	15.3	15.1
3	7.5	6.6
4	10.3	8.5
5	8.7	8.0

examples:

- value: N concentration in soil; obj: boreholes at different locations; t1 / t2: before / after use of fertilizer; question: What is the average effect of the fertilizer if location is fixed?
- value: ozone concentration in air; obj: different stations; t1 / t2: 7:00 / 15:00; question: How much does ozone concentration change between 7:00 and 15:00?
- value: ozone concentration in air; obj: different dates / hours; t1 / t2: two different stations; question: How much do the concentrations at the two stations differ?

paired t-test

If data is given in pairs, we can investigate the mean difference in each pair instead of the difference of the means of each group.
Thus we keep other factors constant and often increase power.

obj	t_1	t_2
1	13.5	12.7
2	15.3	15.1
3	7.5	6.6
4	10.3	8.5
5	8.7	8.0

examples:

- value: N concentration in soil; obj: boreholes at different locations; t1 / t2: before / after use of fertilizer; question: What is the average effect of the fertilizer if location is fixed?
- value: ozone concentration in air; obj: different stations; t1 / t2: 7:00 / 15:00; question: How much does ozone concentration change between 7:00 and 15:00?
- value: ozone concentration in air; obj: different dates / hours; t1 / t2: two different stations; question: How much do the concentrations at the two stations differ?

paired t-test

If data is given in pairs, we can investigate the mean difference in each pair instead of the difference of the means of each group.
Thus we keep other factors constant and often increase power.

obj	t_1	t_2
1	13.5	12.7
2	15.3	15.1
3	7.5	6.6
4	10.3	8.5
5	8.7	8.0

examples:

- value: N concentration in soil; obj: boreholes at different locations; t1 / t2: before / after use of fertilizer; question: What is the average effect of the fertilizer if location is fixed?
- value: ozone concentration in air; obj: different stations; t1 / t2: 7:00 / 15:00; question: How much does ozone concentration change between 7:00 and 15:00?
- value: ozone concentration in air; obj: different dates / hours; t1 / t2: two different stations; question: How much do the concentrations at the two stations differ?

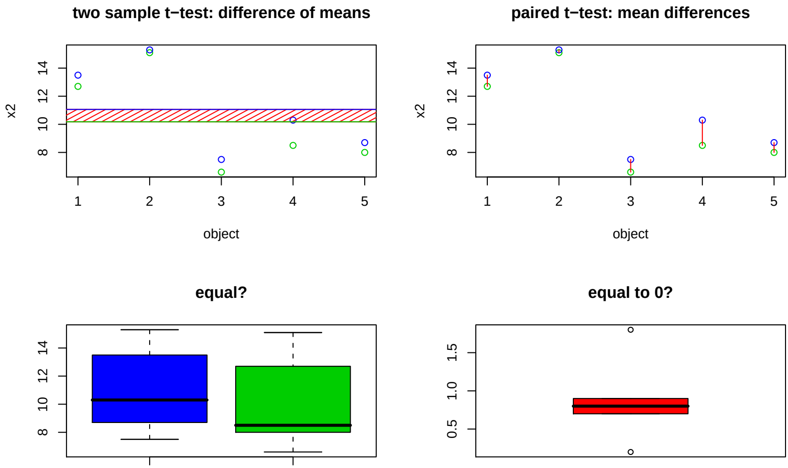
paired t-test

If data is given in pairs, we can investigate the mean difference in each pair instead of the difference of the means of each group.
Thus we keep other factors constant and often increase power.

obj	t_1	t_2
1	13.5	12.7
2	15.3	15.1
3	7.5	6.6
4	10.3	8.5
5	8.7	8.0

examples:

- value: N concentration in soil; obj: boreholes at different locations; t1 / t2: before / after use of fertilizer; question: What is the average effect of the fertilizer if location is fixed?
- value: ozone concentration in air; obj: different stations; t1 / t2: 7:00 / 15:00; question: How much does ozone concentration change between 7:00 and 15:00?
- value: ozone concentration in air; obj: different dates / hours; t1 / t2: two different stations; question: How much do the concentrations at the two stations differ?




```
t.test(x1, x2, var.equal = TRUE)
```

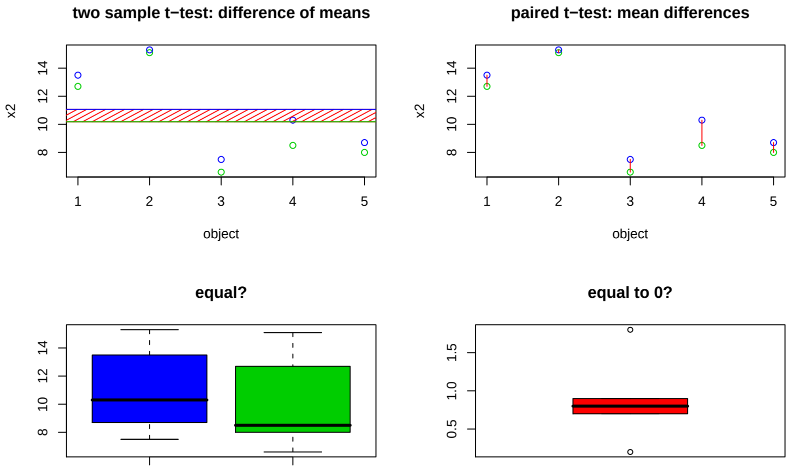
CI: -4.11 5.87

p: 0.70

```
t.test(x1, x2, var.equal = TRUE, paired = TRUE)
```

CI: 0.16 1.60

p: 0.03



`t.test(x1, x2, var.equal = TRUE)`

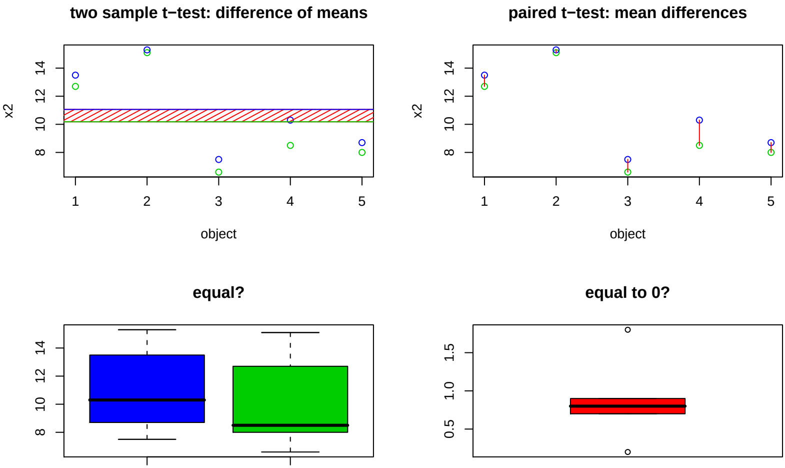
CI: -4.11 5.87

p: 0.70

`t.test(x1, x2, var.equal = TRUE, paired = TRUE)`

CI: 0.16 1.60

p: 0.03



t.test(x1, x2, var.equal = TRUE)

CI: -4.11 5.87

p: 0.70

t.test(x1, x2, var.equal = TRUE, paired = TRUE)

CI: 0.16 1.60

p: 0.03

significant $\neq$ important

significant distinguishable; depends on test (big sample $\rightarrow$ high significance)
Can we be sure that there really is a difference?

important should influence decision; depends on purpose
Is the difference big enough so we should not ignore it?

example

A new fertilizer was tested at 1000 field patches. The wheat yield was with very high significance ($p < 0.005$) increased by 1% (on average 6.43 t/ha instead of 6.37 t/ha).
The fertilizer thus has an significant effect. But the effect is small and may not be important enough to satisfy the unknown side effects, the costs of new production machines...

significant $\neq$ important

significant distinguishable; depends on test (big sample $\rightarrow$ high significance)
Can we be sure that there really is a difference?

important should influence decision; depends on purpose
Is the difference big enough so we should not ignore it?

example
A new fertilizer was tested at 1000 field patches. The wheat yield was with very high significance ($p < 0.005$) increased by 1% (on average 6.43 t/ha instead of 6.37 t/ha).
The fertilizer thus has an significant effect. But the effect is small and may not be important enough to satisfy the unknown side effects, the costs of new production machines...

ANOVA: analysis of variance

The data are nominal values (`Length`), which are separated into several groups by a factor (`I.am.`).
Are the means of the groups equal?

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

μ_1 =`mean(Length[I.am. == "small"])` ...; a: number of groups

How much of the variance of the values is due to the difference between groups compared to the variability within the groups (does the factor explain any of the variability)?

	Length	I.am.
1	165	small
2	176	small
3	158	small
4	174	small
5	180	medium
6	180	medium
7	163	medium
8	187	medium
9	180	tall
10	189	tall
11	173	tall
12	185	tall
13	165	small
14	159	small
15	160	small
16	159	small
17	183	medium
18	165	medium
19	163	medium
20	183	medium
21	182	tall
22	188	tall
23	185	tall
24	188	tall

ANOVA: analysis of variance

The data are nominal values (`Length`), which are separated into several groups by a factor (`I.am.`).
Are the means of the groups equal?

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

μ_1 =`mean(Length[I.am. == "small"])` ...; a: number of groups

How much of the variance of the values is due to the difference between groups compared to the variability within the groups (does the factor explain any of the variability)?

	Length	I.am.
1	165	small
2	176	small
3	158	small
4	174	small
5	180	medium
6	180	medium
7	163	medium
8	187	medium
9	180	tall
10	189	tall
11	173	tall
12	185	tall
13	165	small
14	159	small
15	160	small
16	159	small
17	183	medium
18	165	medium
19	163	medium
20	183	medium
21	182	tall
22	188	tall
23	185	tall
24	188	tall

ANOVA: analysis of variance

The data are nominal values (`Length`), which are separated into several groups by a factor (`I.am.`).
Are the means of the groups equal?

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

μ_1 =`mean(Length[I.am. == "small"])` ...; a: number of groups

How much of the variance of the values is due to the difference between groups compared to the variability within the groups (does the factor explain any of the variability)?

	Length	I.am.
1	165	small
2	176	small
3	158	small
4	174	small
5	180	medium
6	180	medium
7	163	medium
8	187	medium
9	180	tall
10	189	tall
11	173	tall
12	185	tall
13	165	small
14	159	small
15	160	small
16	159	small
17	183	medium
18	165	medium
19	163	medium
20	183	medium
21	182	tall
22	188	tall
23	185	tall
24	188	tall

Why use ANOVA?

Suppose we have three groups.

ANOVA H_0: $\mu_1 = \mu_2 = \mu_3$

alternative H_A: at least one mean is different from the others

t-tests $H_0^a : \mu_1 = \mu_2$, $H_0^b : \mu_2 = \mu_3$, $H_0^c : \mu_1 = \mu_3$.

- It is possible that the first two are rejected but not the third. How to explain?
- Each t-test needs many observations to be powerful; given the same sample size ANOVA has more power. In our example: ANOVA has $df = 21$, each t-test would have $df = 7$.

Why use ANOVA?

Suppose we have three groups.

ANOVA H_0: $\mu_1 = \mu_2 = \mu_3$
alternative H_A: at least one mean is different from the others

t-tests $H_0^a : \mu_1 = \mu_2$, $H_0^b : \mu_2 = \mu_3$, $H_0^c : \mu_1 = \mu_3$.

- It is possible that the first two are rejected but not the third. How to explain?
- Each t-test needs many observations to be powerful, given the same sample size ANOVA has more power. In our example ANOVA has $df = 21$, each t-test would have $df = 7$.

Why use ANOVA?

Suppose we have three groups.

ANOVA H_0: $\mu_1 = \mu_2 = \mu_3$
alternative H_A: at least one mean is different from the others

t-tests $H_0^a : \mu_1 = \mu_2$, $H_0^b : \mu_2 = \mu_3$, $H_0^c : \mu_1 = \mu_3$.

- It is possible that the first two are rejected but not the third. How to explain?
- Each t-test needs many observations to be powerful, given the same sample size ANOVA has more power. In our example ANOVA has $df = 21$, each t-test would have $df = 7$.

Why use ANOVA?

Suppose we have three groups.

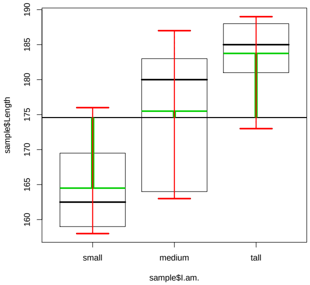
ANOVA H_0: $\mu_1 = \mu_2 = \mu_3$
alternative H_A: at least one mean is different from the others

t-tests $H_0^a : \mu_1 = \mu_2$, $H_0^b : \mu_2 = \mu_3$, $H_0^c : \mu_1 = \mu_3$.

- It is possible that the first two are rejected but not the third. How to explain?
- Each t-test needs many observations to be powerful, given the same sample size ANOVA has more power. In our example ANOVA has $df = 21$, each t-test would have $df = 7$.

test statistic: F value

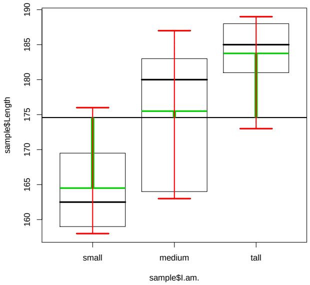
$$\hat{F} = \frac{\text{variance between groups (explained variance)}}{\text{variance within groups (unexplained variance)}} = \frac{n \cdot \text{Var}(\bar{x}_i)}{s_p^2} = \frac{n \cdot \frac{1}{a-1} \sum_{i=1}^{a} (\bar{x}_{i.} - \bar{\bar{x}}_{..})^2}{\frac{1}{a-1} \sum_{i=1}^{a} \left(\frac{1}{n-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_{i.})^2 \right)}$$



x_{ij}: values, i is index for group
a: number of groups
n: number of values within each group (equal)
$\bar{x}_{i.} = \frac{1}{n} \sum_{j=1}^{n} x_{ij}$: mean of group i
$\bar{\bar{x}}_{..} = \sum_{i=1}^{a} x_{i.}$: overall mean
$s_p^2 = \frac{1}{a-1} \sum_{i=1}^{a} s_i^2$: pooled variance

test statistic: F value

$$\hat{F} = \frac{\text{variance between groups (explained variance)}}{\text{variance within groups (unexplained variance)}} = \frac{n \cdot \text{Var}(\bar{x}_i)}{s_p^2} = \frac{n \cdot \frac{1}{a-1} \sum_{i=1}^{a} (\bar{x}_{i.} - \bar{\bar{x}}_{..})^2}{\frac{1}{a-1} \sum_{i=1}^{a} (\frac{1}{n-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_{i.})^2)}$$



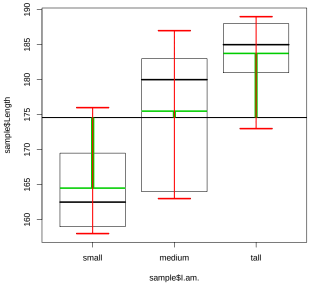
x_{ij}: values, i is index for group
a: number of groups
n: number of values within each group (equal)
$\bar{x}_{i.} = \frac{1}{n} \sum_{j=1}^{n} x_{ij}$: mean of group i
$\bar{\bar{x}}_{..} = \sum_{i=1}^{a} x_{i.}$: overall mean
$s_p^2 = \frac{1}{a-1} \sum_{i=1}^{a} s_i^2$: pooled variance

test statistic: F value

$$\hat{F} = \frac{\text{variance between groups (explained variance)}}{\text{variance within groups (unexplained variance)}}$$

$$= \frac{n \cdot \text{Var}(\bar{x}_i)}{s_p^2}$$

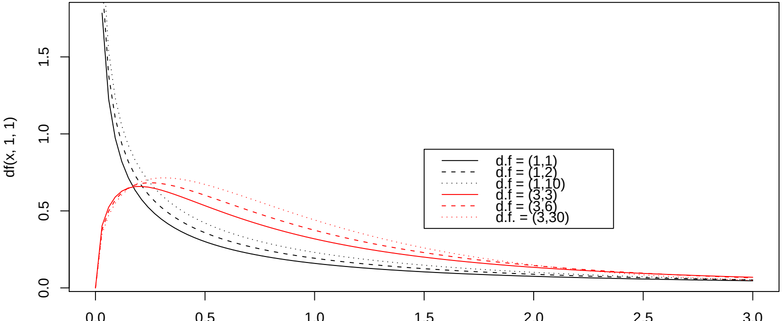
$$= \frac{n \cdot \frac{1}{a-1} \sum_{i=1}^{a} (\bar{x}_{i.} - \bar{\bar{x}}_{..})^2}{\frac{1}{a-1} \sum_{i=1}^{a} (\frac{1}{n-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_{i.})^2)}$$



x_{ij}: values, i is index for group
a: number of groups
n: number of values within each group (equal)
$\bar{x}_{i.} = \frac{1}{n} \sum_{j=1}^{n} x_{ij}$: mean of group i
$\bar{\bar{x}}_{..} = \sum_{i=1}^{a} x_{i.}$: overall mean
$s_p^2 = \frac{1}{a-1} \sum_{i=1}^{a} s_i^2$: pooled variance

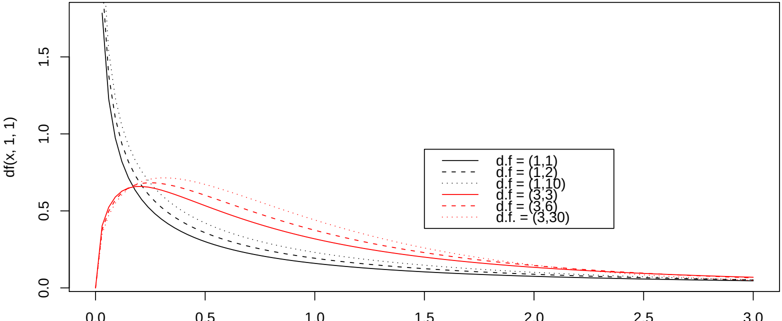
F test

- If the factor explains a lot of the variability, $\hat{F}$ is high.
  A model with many groups is likely to "explain" much of the variability even if there is no real difference.
- If H_0 is true, $\hat{F}$ follows a F-distribution with $a - 1$ (numerator) and $a(n-1)$ (denominator) degrees of freedom
- p value: `p = 1 - pf(`$\hat{F}$`, a - 1, a(n - 1))`
  (probability to have such a high $\hat{F}$ even if all means are equal)



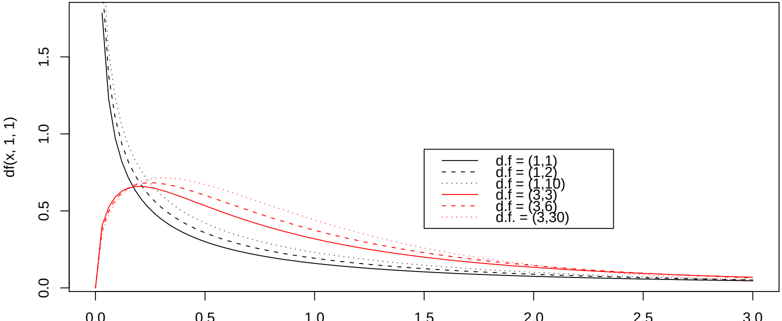
F test

- If the factor explains a lot of the variability, $\hat{F}$ is high.
  A model with many groups is likely to "explain" much of the variability even if there is no real difference.
- If H_0 is true, $\hat{F}$ follows a F-distribution with $a - 1$ (numerator) and $a(n-1)$ (denominator) degrees of freedom
- p value: `p = 1 - pf(`$\hat{F}$`, a - 1, a(n - 1))`
  (probability to have such a high $\hat{F}$ even if all means are equal)



F test

- If the factor explains a lot of the variability, $\hat{F}$ is high.
  A model with many groups is likely to "explain" much of the variability even if there is no real difference.
- If H_0 is true, $\hat{F}$ follows a F-distribution with $a-1$ (numerator) and $a(n-1)$ (denominator) degrees of freedom
- p value: `p = 1 - pf(`$\hat{F}$`, a - 1, a(n - 1))`
  (probability to have such a high $\hat{F}$ even if all means are equal)



ANOVA table

	deg. of freedom	sum squared	mean squared	F	p
factor A	$df_A = a - 1$	$SS_A = n \sum_i (\bar{x}_{i.} - \bar{\bar{x}}_{..})^2$	$MS_A = SS_A / df_A$	$F_A = MS_A / MS_E$	$1 - pt(F_A, df_A, df_E)$
residuals	$df_E = a(n-1)$	$SS_E = \sum_i \sum_j (x_{ij} - \bar{x}_{i.})^2$	$MS_E = SS_E / df_E$		

```
> summary(aov(Length ~ I.am., sample))

            Df  Sum Sq Mean Sq F value    Pr(>F)
I.am.        2 1492.33  746.17  12.501 0.0002656 ***
Residuals   21 1253.50   59.69
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

properties of $\hat{F}$

If there are 2 groups, $\hat{F} = \hat{t}^2$ and the p value is equal to the one of the two-sample t-test.

The total error is the sum of explained and unexplained error

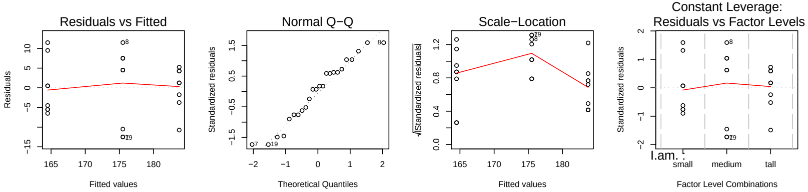
$$\sum_{i=1}^{a} \sum_{j=1}^{n} (x_{ij} - \bar{\bar{x}}_{..})^2 = SS_A + SS_E$$

General Concept of Statistical Tests II

assumptions		results
completely satisfied		exactly true
	$\overset{Test}{\Longrightarrow}$	
only partially satisfied		only approximately true

In reality, assumptions are hardly ever exactly satisfied, this does not mean that results are useless but we should be careful and find out how badly the assumptions are violated.

Assumptions of ANOVA (can be checked by `plot(aov(Length ~ I.am., sample))`)



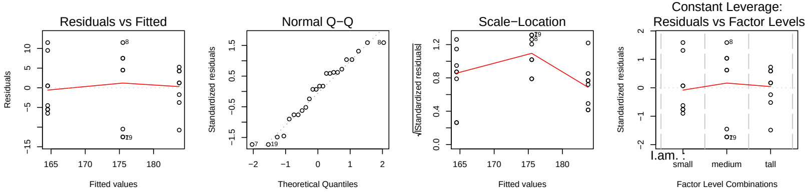
- equal variance in each group (vertical line in "Scale-Location")
- normal distribution of residuals (points on line in "Normal Q-Q")

General Concept of Statistical Tests II

assumptions		results
completely satisfied	$\overset{Test}{\Longrightarrow}$	exactly true
only partially satisfied		only approximately true

In reality, assumptions are hardly ever exactly satisfied, this does not mean that results are useless but we should be careful and find out how badly the assumptions are violated.

Assumptions of ANOVA (can be checked by `plot(aov(Length ~ I.am., sample))`)



- equal variance in each group (vertical line in "Scale-Location")
- normal distribution of residuals (points on line in "Normal Q-Q")

two-way ANOVA

ANOVA can also be used if there are several factors (`I.am.` and `Year`) grouping the data.

	Length	I.am.	Year
1	165	small	8
2	176	small	8
3	158	small	8
4	174	small	8
5	180	medium	8
6	180	medium	8
7	163	medium	8
8	187	medium	8
9	180	tall	8
10	189	tall	8
11	173	tall	8
12	185	tall	8
13	165	small	9
14	159	small	9
15	160	small	9
16	159	small	9
17	183	medium	9
18	165	medium	9
19	163	medium	9
20	183	medium	9
21	182	tall	9
22	188	tall	9
23	185	tall	9
24	188	tall	9

For each factor an analysis of variance is conducted

I.am: H_0: $\mu_{small} = \mu_{medium} = \mu_{tall}$

Year: H_0: $\mu_8 = \mu_9$

The variance is explained by adding the effects of both variables, the residuals are now

$$SS_E = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{\bar{x}}_{...})^2$$

two-way ANOVA

ANOVA can also be used if there are several factors (`I.am.` and `Year`) grouping the data.

	Length	I.am.	Year
1	165	small	8
2	176	small	8
3	158	small	8
4	174	small	8
5	180	medium	8
6	180	medium	8
7	163	medium	8
8	187	medium	8
9	180	tall	8
10	189	tall	8
11	173	tall	8
12	185	tall	8
13	165	small	9
14	159	small	9
15	160	small	9
16	159	small	9
17	183	medium	9
18	165	medium	9
19	163	medium	9
20	183	medium	9
21	182	tall	9
22	188	tall	9
23	185	tall	9
24	188	tall	9

For each factor an analysis of variance is conducted

I.am.: H_0: $\mu_{small} = \mu_{medium} = \mu_{tall}$

Year: H_0: $\mu_8 = \mu_9$

The variance is explained by adding the effects of both variables, the residuals are now

$$SS_E = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{\bar{x}}_{...})^2$$

two-way ANOVA

ANOVA can also be used if there are several factors (`I.am.` and `Year`) grouping the data.

	Length	I.am.	Year
1	165	small	8
2	176	small	8
3	158	small	8
4	174	small	8
5	180	medium	8
6	180	medium	8
7	163	medium	8
8	187	medium	8
9	180	tall	8
10	189	tall	8
11	173	tall	8
12	185	tall	8
13	165	small	9
14	159	small	9
15	160	small	9
16	159	small	9
17	183	medium	9
18	165	medium	9
19	163	medium	9
20	183	medium	9
21	182	tall	9
22	188	tall	9
23	185	tall	9
24	188	tall	9

For each factor an analysis of variance is conducted

I.am. H_0: $\mu_{small} = \mu_{medium} = \mu_{tall}$

Year H_0: $\mu_8 = \mu_9$

The variance is explained by adding the effects of both variables, the residuals are now

$$SS_E = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{\bar{x}}_{...})^2$$

two-way ANOVA

ANOVA can also be used if there are several factors (`I.am.` and `Year`) grouping the data.

	Length	I.am.	Year
1	165	small	8
2	176	small	8
3	158	small	8
4	174	small	8
5	180	medium	8
6	180	medium	8
7	163	medium	8
8	187	medium	8
9	180	tall	8
10	189	tall	8
11	173	tall	8
12	185	tall	8
13	165	small	9
14	159	small	9
15	160	small	9
16	159	small	9
17	183	medium	9
18	165	medium	9
19	163	medium	9
20	183	medium	9
21	182	tall	9
22	188	tall	9
23	185	tall	9
24	188	tall	9

For each factor an analysis of variance is conducted

I.am. H_0: $\mu_{small} = \mu_{medium} = \mu_{tall}$

Year H_0: $\mu_8 = \mu_9$

The variance is explained by adding the effects of both variables, the residuals are now

$$SS_E = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{\bar{x}}_{...})^2$$

two-way ANOVA

ANOVA can also be used if there are several factors (`I.am.` and `Year`) grouping the data.

	Length	I.am.	Year
1	165	small	8
2	176	small	8
3	158	small	8
4	174	small	8
5	180	medium	8
6	180	medium	8
7	163	medium	8
8	187	medium	8
9	180	tall	8
10	189	tall	8
11	173	tall	8
12	185	tall	8
13	165	small	9
14	159	small	9
15	160	small	9
16	159	small	9
17	183	medium	9
18	165	medium	9
19	163	medium	9
20	183	medium	9
21	182	tall	9
22	188	tall	9
23	185	tall	9
24	188	tall	9

For each factor an analysis of variance is conducted

I.am. H_0: $\mu_{small} = \mu_{medium} = \mu_{tall}$

Year H_0: $\mu_8 = \mu_9$

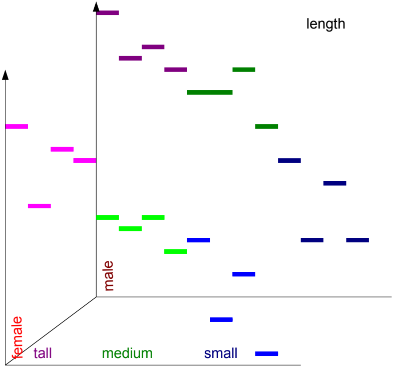
The variance is explained by adding the effects of both variables, the residuals are now

$$SS_E = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{\bar{x}}_{...})^2$$

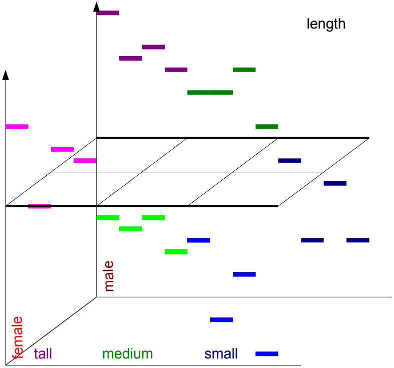
2 way ANOVA in pictures I: data

The pictures shall illustrate the ideas of 2 way ANOVA using artificial values.

data



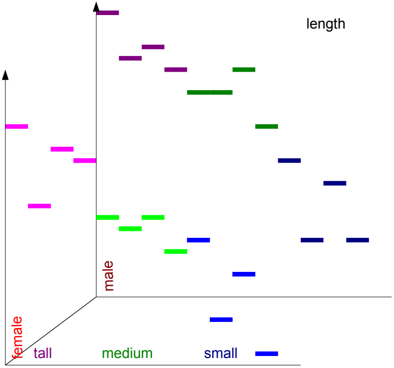
global mean $\bar{\bar{x}}_{\cdot\cdot}$



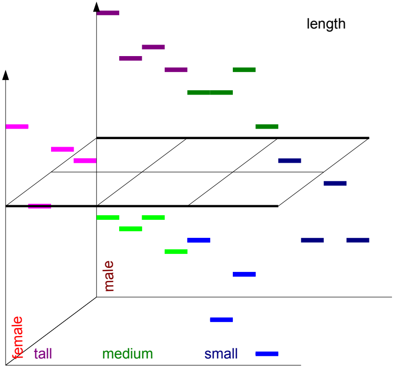
2 way ANOVA in pictures I: data

The pictures shall illustrate the ideas of 2 way ANOVA using artificial values.

data



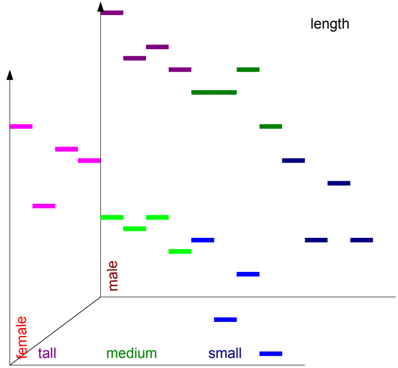
global mean $\bar{\bar{x}}_{\cdot\cdot\cdot}$



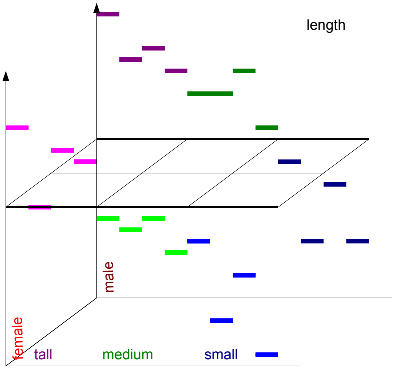
2 way ANOVA in pictures I: data

The pictures shall illustrate the ideas of 2 way ANOVA using artificial values.

data



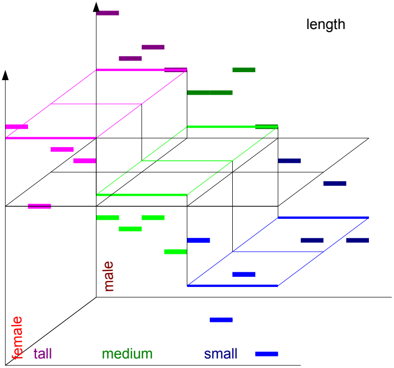
global mean $\bar{\bar{x}}_{\cdot\cdot}$



2 way ANOVA in pictures II: factor effects

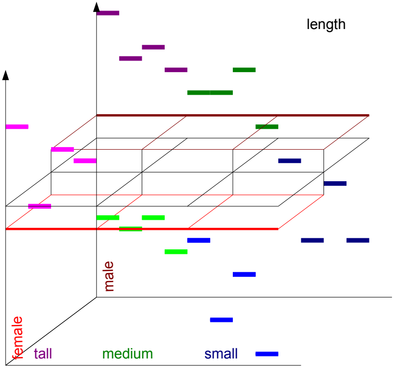
Effect of "l.am."

$\bar{x}_{tall..} > \bar{x}_{medium..} > \bar{x}_{small..}$



Effect of "Gender"

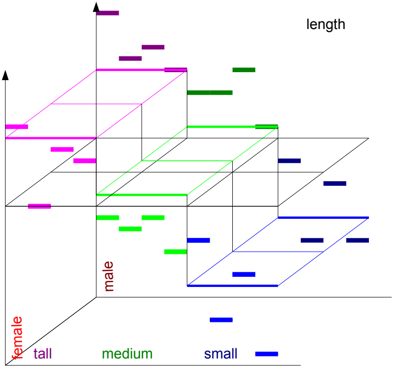
$\bar{x}_{.male.} > \bar{x}_{.female.}$



2 way ANOVA in pictures II: factor effects

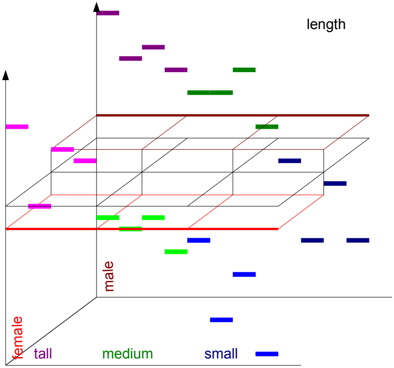
Effect of "l.am."

$\bar{x}_{tall..} > \bar{x}_{medium..} > \bar{x}_{small..}$



Effect of "Gender"

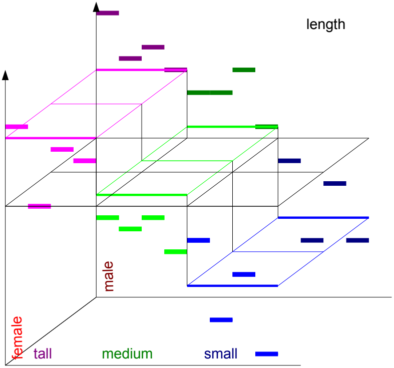
$\bar{x}_{.male.} > \bar{x}_{.female.}$



2 way ANOVA in pictures II: factor effects

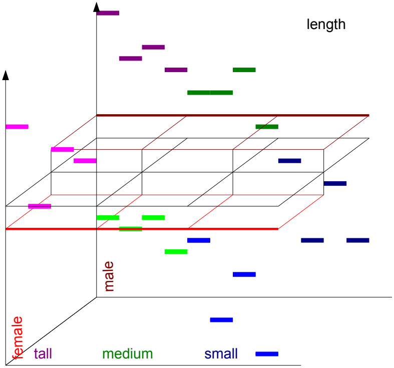
Effect of "I.am."

$\bar{x}_{tall..} > \bar{x}_{medium..} > \bar{x}_{small..}$



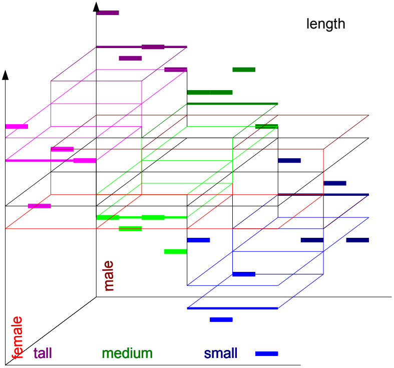
Effect of "Gender"

$\bar{x}_{.male.} > \bar{x}_{.female.}$



2 way ANOVA in pictures III: added effects

Effects added give value for each factor combination



example:

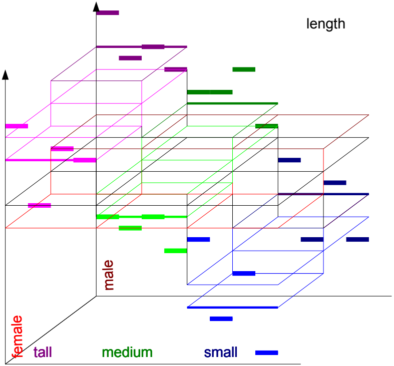
$\bar{x}_{.female.} = \bar{\bar{x}}_{...} - 2$

$\bar{x}_{small..} = \bar{\bar{x}}_{...} - 7$

therefore $\bar{x}_{smallfemale.} = \bar{\bar{x}}_{...} - 9$

2 way ANOVA in pictures III: added effects

Effects added give value for each factor combination



example:

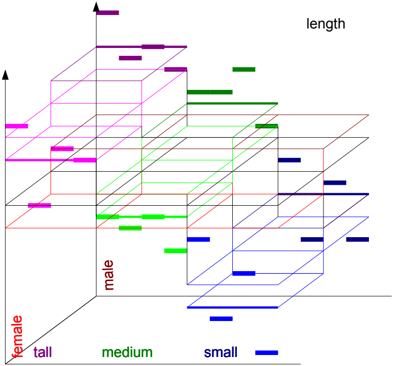
$\bar{x}_{.female.} = \bar{\bar{x}}_{...} - 2$

$\bar{x}_{small..} = \bar{\bar{x}}_{...} - 7$

therefore $\bar{x}_{smallfemale.} = \bar{\bar{x}}_{...} - 9$

2 way ANOVA in pictures III: added effects

Effects added give value for each factor combination



example:

$\bar{x}_{.female.} = \bar{\bar{x}}_{...} - 2$

$\bar{x}_{small..} = \bar{\bar{x}}_{...} - 7$

therefore $\bar{x}_{smallfemale.} = \bar{\bar{x}}_{...} - 9$

more ANOVA

Analysis of variance has more versions which can deal with

- unbalanced samples (unequal group sizes)
- more than 2 factors
- interactions (is the difference between tall male and small male bigger than between tall female and small female?)
- random effects (take 12 samples from year 8 and 9 each, the number of small / medium / tall is not fixed in advance)
- ...

more ANOVA

Analysis of variance has more versions which can deal with

- unbalanced samples (unequal group sizes)
- more than 2 factors
- interactions (is the difference between tall male and small male bigger than between tall female and small female?)
- random effects (take 12 samples from year 8 and 9 each, the number of small / medium / tall is not fixed in advance)
- ...

more ANOVA

Analysis of variance has more versions which can deal with

- unbalanced samples (unequal group sizes)
- more than 2 factors
- interactions (is the difference between tall male and small male bigger than between tall female and small female?)
- random effects (take 12 samples from year 8 and 9 each, the number of small / medium / tall is not fixed in advance)
- ...

more ANOVA

Analysis of variance has more versions which can deal with

- unbalanced samples (unequal group sizes)
- more than 2 factors
- interactions (is the difference between tall male and small male bigger than between tall female and small female?)
- random effects (take 12 samples from year 8 and 9 each, the number of small / medium / tall is not fixed in advance)
- ...

more ANOVA

Analysis of variance has more versions which can deal with

- unbalanced samples (unequal group sizes)
- more than 2 factors
- interactions (is the difference between tall male and small male bigger than between tall female and small female?)
- random effects (take 12 samples from year 8 and 9 each, the number of small / medium / tall is not fixed in advance)
- ...

References

- Sachs, L.; Hedderich, J. (2006): Angewandte Statistik. Methodensammlung mit R. 12.Aufl., Springer.
- Wonnacott, R; Wonnacott, T. (1990): Introductory Statistics. 5.Ed., Wiley.

References

- Sachs, L.; Hedderich, J. (2006): Angewandte Statistik. Methodensammlung mit R. 12.Aufl., Springer.
- Wonnacott, R; Wonnacott, T. (1990): Introductory Statistics. 5.Ed., Wiley.