ANOVA and Statistical Models

R. A. Bailey
School of Mathematical Sciences
Queen Mary
University of London
r.a.bailey@qmul.ac.uk

Learning Institute, Queen Mary, University of London,
2 February 2011

Outline

1. Comparing three or more conditions
2. Models with submodels
3. More complicated families of models

An experiment on dairy cows

An experiment to compare the effects of three different diets on milk production used 11 Holstein dairy cows at similar points in their lactation cycles. They were fed the diets for three weeks. During the third week, the average daily milk production was recorded for each cow, in pounds per day.

Do the diets have different effects on milk yield?

Assumptions and vocabulary

Assumption: yield depends on diet.

Model:

$$\text{yield on cow } i = \begin{cases} K_A + \varepsilon_i & \text{if on diet A} \\ K_B + \varepsilon_i & \text{if on diet B} \\ K_C + \varepsilon_i & \text{if on diet C} \end{cases}$$

where $\varepsilon_1, \ldots, \varepsilon_{11}$ are independent (normal) random variables with mean zero and variance σ^2.

Assumptions and vocabulary

Assumption: yield depends on diet.

Model:

$$\text{yield on cow } i = \begin{cases} K_A + \varepsilon_i & \text{if on diet A} \\ K_B + \varepsilon_i & \text{if on diet B} \\ K_C + \varepsilon_i & \text{if on diet C} \end{cases}$$

where $\varepsilon_1, \ldots, \varepsilon_{11}$ are independent (normal) random variables with mean zero and variance σ^2.

Estimate K_A as the average yield of cows on diet A
(and similarly for the other diets).
This gives the fit (of the model) for each cow.

Assumptions and vocabulary

Assumption: yield depends on diet.

Model:

$$\text{yield on cow } i = \begin{cases} K_A + \varepsilon_i & \text{if on diet A} \\ K_B + \varepsilon_i & \text{if on diet B} \\ K_C + \varepsilon_i & \text{if on diet C} \end{cases}$$

where $\varepsilon_1, \ldots, \varepsilon_{11}$ are independent (normal) random variables with mean zero and variance σ^2.

Estimate K_A as the average yield of cows on diet A
(and similarly for the other diets).
This gives the fit (of the model) for each cow.

$$\text{data} = \text{fit} + \text{residual}.$$

Data for milk experiment

Diet	Yield Data	Diet mean Fit	Residual	Data^2	Fit^2	Residual^2
A	60.7	61.2	−0.5	3684.49	3745.44	0.25
A	59.7	61.2	−1.5	3564.09	3745.44	2.25
A	61.9	61.2	0.7	3831.61	3745.44	0.49
A	62.5	61.2	1.3	3906.25	3645.44	1.69
B	55.6	54.3	1.3	3091.36	2948.49	1.69
B	52.9	54.3	−1.4	2798.41	2948.49	1.96
B	52.7	54.3	−1.6	2777.29	2948.49	2.56
B	56.0	54.3	1.7	3136.00	2948.49	2.89
C	62.8	58.2	4.6	3943.84	3387.24	21.16
C	55.8	58.2	−2.4	3113.64	3387.24	5.76
C	56.0	58.2	−2.2	3136.00	3387.24	4.84
				36982.98	36937.44	45.54

Sums of squares: I

Pythagoras' Theorem (in 11 dimensions) shows that

$$\begin{aligned}&\text{sum of the squares of data} = \\ &\qquad \text{sum of squares of fits} + \text{sum of squares of residuals}\end{aligned}$$

Sums of squares: I

Pythagoras' Theorem (in 11 dimensions) shows that

$$\begin{aligned}&\text{sum of the squares of data} =\\ &\qquad \text{sum of squares of fits} + \text{sum of squares of residuals}\end{aligned}$$

$$\begin{aligned}&\text{sum of the squares of fits} =\\ &\qquad (\text{mean for diet A})^2 \times \text{number of cows on diet A}\\ &\qquad + (\text{mean for diet B})^2 \times \text{number of cows on diet B}\\ &\qquad + (\text{mean for diet C})^2 \times \text{number of cows on diet C}\end{aligned}$$

Sums of squares: I

Pythagoras' Theorem (in 11 dimensions) shows that

$$\begin{aligned}&\text{sum of the squares of data} = \\ &\qquad \text{sum of squares of fits} + \text{sum of squares of residuals}\end{aligned}$$

$$\begin{aligned}&\text{sum of the squares of fits} = \\ &\qquad (\text{mean for diet A})^2 \times \text{number of cows on diet A} \\ &\qquad + (\text{mean for diet B})^2 \times \text{number of cows on diet B} \\ &\qquad + (\text{mean for diet C})^2 \times \text{number of cows on diet C}\end{aligned}$$

so the sum of squares of residuals can be calculated simply as

$$\text{sum of squares of data} - \text{sum of squares of fits}.$$

Sums of squares: II

Number of data	11
Number of fitted parameters	3
Difference	8

Theory predicts that the sum of squares of residuals is a random variable with mean $8\sigma^2$.

Sums of squares: II

Number of data	11
Number of fitted parameters	3
Difference	8

Theory predicts that the sum of squares of residuals is a random variable with mean $8\sigma^2$.

So we can estimate σ^2 by

$$\frac{\text{sum of squares of residuals}}{8}$$

This is called the residual mean square.

Null and alternative hypotheses

$$\text{Model: yield on cow } i = \begin{cases} K_A + \varepsilon_i & \text{if on diet A} \\ K_B + \varepsilon_i & \text{if on diet B} \\ K_C + \varepsilon_i & \text{if on diet C} \end{cases}$$

Null and alternative hypotheses

$$\text{Model: yield on cow } i = \begin{cases} K_A + \varepsilon_i & \text{if on diet A} \\ K_B + \varepsilon_i & \text{if on diet B} \\ K_C + \varepsilon_i & \text{if on diet C} \end{cases}$$

$\swarrow$ $\searrow$

$K_A = K_B = K_C$

there is no difference
between the diets

K_A, K_B and K_C are not all the same

Null and alternative hypotheses

$$\text{Model: yield on cow } i = \begin{cases} K_A + \varepsilon_i & \text{if on diet A} \\ K_B + \varepsilon_i & \text{if on diet B} \\ K_C + \varepsilon_i & \text{if on diet C} \end{cases}$$

$\swarrow$ $\searrow$

$K_A = K_B = K_C$
there is no difference
between the diets

K_A, K_B and K_C are not all the same

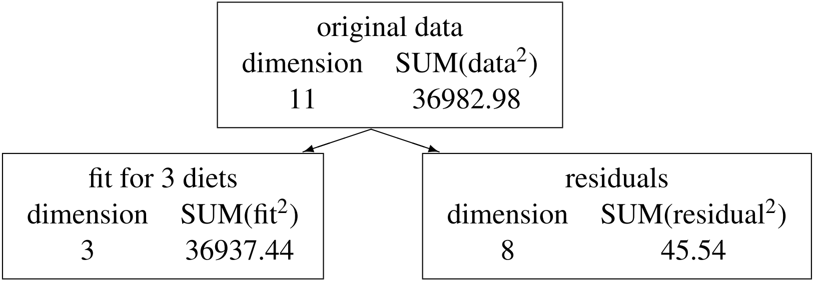
null hypothesis

alternative hypothesis

Keep on splitting it up

original data	
dimension	SUM(data^2)
11	36982.98

Keep on splitting it up



original data

dimension	SUM(data^2)
11	36982.98

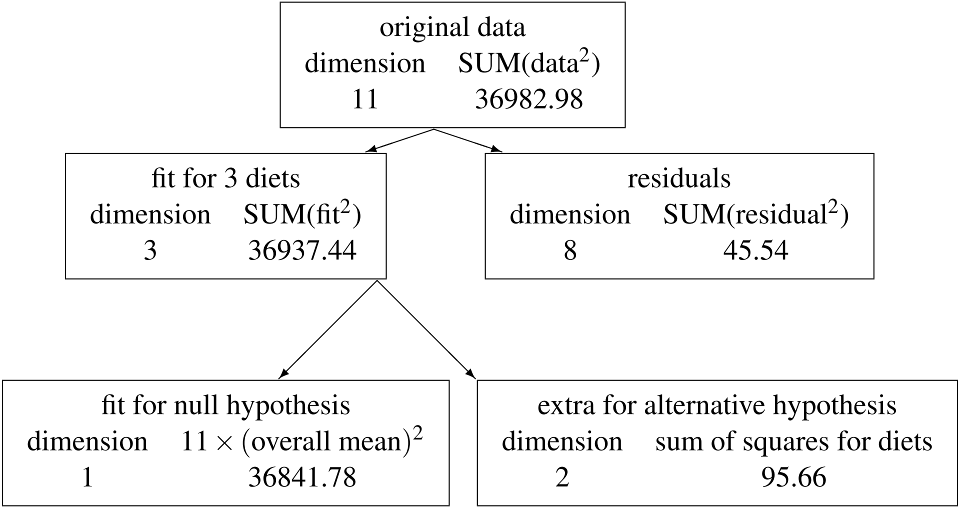
fit for 3 diets

dimension	SUM(fit^2)
3	36937.44

residuals

dimension	SUM(residual^2)
8	45.54

Keep on splitting it up



original data

dimension SUM(data^2)

11 36982.98

fit for 3 diets

dimension SUM(fit^2)

3 36937.44

residuals

dimension SUM(residual^2)

8 45.54

fit for null hypothesis

dimension $11 \times (\text{overall mean})^2$

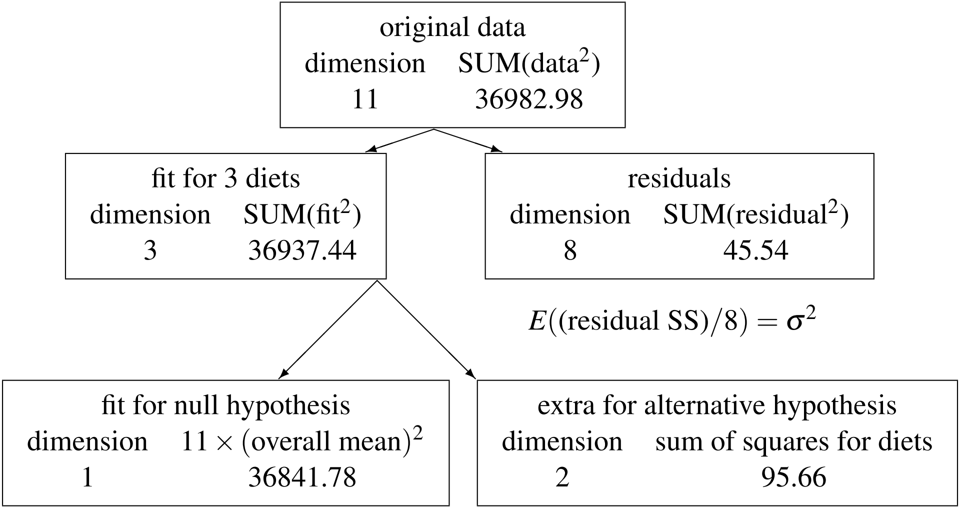
1 36841.78

extra for alternative hypothesis

dimension sum of squares for diets

2 95.66

Keep on splitting it up



original data

dimension	SUM(data^2)
11	36982.98

fit for 3 diets

dimension	SUM(fit^2)
3	36937.44

residuals

dimension	SUM(residual^2)
8	45.54

$$E((\text{residual SS})/8) = \sigma^2$$

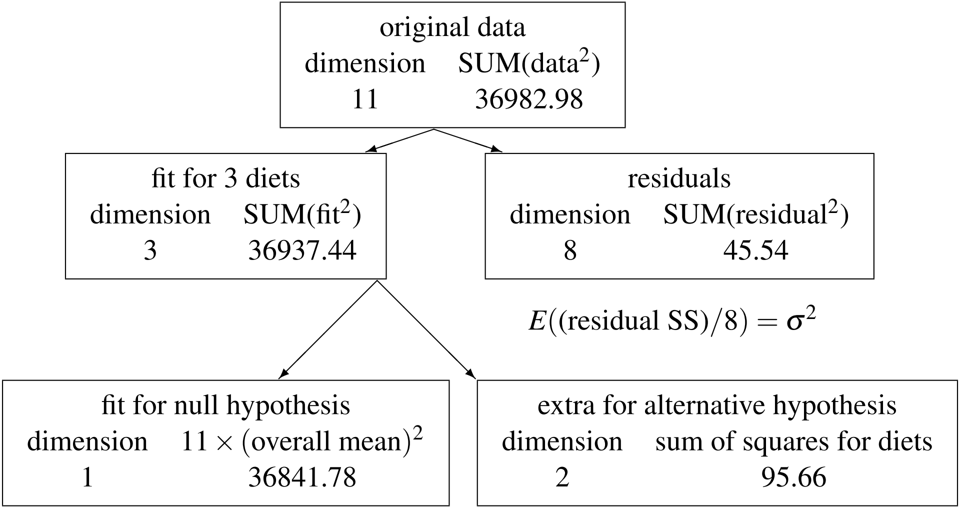
fit for null hypothesis

dimension	$11 \times (\text{overall mean})^2$
1	36841.78

extra for alternative hypothesis

dimension	sum of squares for diets
2	95.66

Keep on splitting it up



original data

dimension 11

SUM(data^2) 36982.98

fit for 3 diets

dimension 3

SUM(fit^2) 36937.44

residuals

dimension 8

SUM(residual^2) 45.54

$E((\text{residual SS})/8) = \sigma^2$

fit for null hypothesis

dimension 1

$11 \times (\text{overall mean})^2$ 36841.78

extra for alternative hypothesis

dimension 2

sum of squares for diets 95.66

If the null hypothesis is true, $E((\text{diets SS})/2) = \sigma^2$

Analysis-of-variance table: I

Source	Dimension	Sum of squares
null	1	36841.78
diets	2	95.66
residual	8	45.54
Total	11	36982.98

Analysis-of-variance table: I

Source	Dimension	Sum of squares	Mean square
null	1	36841.78	
diets	2	95.66	47.83
residual	8	45.54	5.69
Total	11	36982.98	

Analysis-of-variance table: I

Source	Dimension	Sum of squares	Mean square
null	1	36841.78	
diets	2	95.66	47.83
residual	8	45.54	5.69
Total	11	36982.98	

If the null hypothesis is true then the two mean squares should be about the same size. Are they?

Analysis-of-variance table: I

Source	Dimension	Sum of squares	Mean square	Ratio
null	1	36841.78		
diets	2	95.66	47.83	8.40
residual	8	45.54	5.69	
Total	11	36982.98		

If the null hypothesis is true then the two mean squares should be about the same size. Are they?

Find their ratio!

Analysis-of-variance table: I

Source	Dimension	Sum of squares	Mean square	Ratio	P
null	1	36841.78			
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.69		
Total	11	36982.98			

If the null hypothesis is true then the two mean squares should be about the same size. Are they?

Find their ratio!

If the null hypothesis is true and the data are normal then this ratio should have an F distribution with 2 and 8 degrees of freedom (sometimes written F_8^2). Then the probability that the ratio is 8.40 or higher is approximately 0.01 (I looked this up in statistical tables).

Analysis-of-variance table: II

Source	Dimension	Sum of squares	Mean square	Ratio	P
null	1	36841.78			
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	11	36982.98			

Analysis-of-variance table: II

Source	Dimension	Sum of squares	Mean square	Ratio	P
null	1	36841.78			
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	11	36982.98			

Some conventions:

- Write ‘degrees of freedom’ or ‘df’ of ‘d.f.’ instead of ‘dimension’.

Analysis-of-variance table: II

Source	df	Sum of squares	Mean square	Ratio	P
null	1	36841.78			
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	11	36982.98			

Some conventions:

- Write ‘degrees of freedom’ or ‘df’ of ‘d.f.’ instead of ‘dimension’.
- We never use the first row, so omit it:
  then we have adjust the totals to match.

Analysis-of-variance table: II

Source	df	Sum of squares	Mean square	Ratio	P
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	10	141.20			

Some conventions:

- Write ‘degrees of freedom’ or ‘df’ of ‘d.f.’ instead of ‘dimension’.
- We never use the first row, so omit it:
  then we have adjust the totals to match.
- Your software may write ‘variance ratio’ or ‘V.R.’ or ‘F’ instead of ‘Ratio’;
  it may write ‘Prob’ or ‘F.pr.’ instead of ‘P’.

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.
  We conclude that there *are* differences between the conditions.

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.
  We conclude that there *are* differences between the conditions.
  Report the estimates of the effects of the conditions (these are just the averages).

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.
  We conclude that there *are* differences between the conditions.
  Report the estimates of the effects of the conditions (these are just the averages).
  It is good practice to report the standard error of a difference as well.

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.
  We conclude that there *are* differences between the conditions.
  Report the estimates of the effects of the conditions (these are just the averages).
  It is good practice to report the standard error of a difference as well.
- If Ratio > 3 but $P > 0.05$ then we cannot reject the null hypothesis.

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.
  We conclude that there *are* differences between the conditions.
  Report the estimates of the effects of the conditions (these are just the averages).
  It is good practice to report the standard error of a difference as well.
- If Ratio > 3 but $P > 0.05$ then we cannot reject the null hypothesis.
  This does not mean that the null hypothesis is true; it means that you need to gather more data.

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.
  We conclude that there *are* differences between the conditions.
  Report the estimates of the effects of the conditions (these are just the averages).
  It is good practice to report the standard error of a difference as well.
- If Ratio > 3 but $P > 0.05$ then we cannot reject the null hypothesis.
  This does not mean that the null hypothesis is true; it means that you need to gather more data.
- If $1/10 <$ Ratio and $P > 0.05$ then we cannot reject the null hypothesis. Gathering more data may produce a similar conclusion.

Interpreting and reporting the results, in general

- If the Ratio is huge and $P < 0.05$, we can reject the null hypothesis at the 5% significance level.
  We conclude that there *are* differences between the conditions.
  Report the estimates of the effects of the conditions (these are just the averages).
  It is good practice to report the standard error of a difference as well.
- If Ratio > 3 but $P > 0.05$ then we cannot reject the null hypothesis.
  This does not mean that the null hypothesis is true; it means that you need to gather more data.
- If $1/10 <$ Ratio and $P > 0.05$ then we cannot reject the null hypothesis. Gathering more data may produce a similar conclusion.
- If Ratio $< 1/10$ then this is extremely unlikely. There may be systematic errors in your data.

Reporting the results of the diet experiment

Source	df	Sum of squares	Mean square	Ratio	P
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	10	141.20			

Reporting the results of the diet experiment

Source	df	Sum of squares	Mean square	Ratio	P
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	10	141.20			

We can reject the null hypothesis at the 1% level of significance: we conclude that the different diets have different effects on milk yield.

Reporting the results of the diet experiment

Source	df	Sum of squares	Mean square	Ratio	P
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	10	141.20			

We can reject the null hypothesis at the 1% level of significance: we conclude that the different diets have different effects on milk yield.

Table of means

Diet A	Diet B	Diet C
61.2	54.3	58.2

Reporting the results of the diet experiment

Source	df	Sum of squares	Mean square	Ratio	P
diets	2	95.66	47.83	8.40	0.01
residual	8	45.54	5.68		
Total	10	141.20			

We can reject the null hypothesis at the 1% level of significance: we conclude that the different diets have different effects on milk yield.

Table of means

Diet A	Diet B	Diet C
61.2	54.3	58.2

$$\text{Standard error of a difference} = \sqrt{\left(\frac{1}{4}+\frac{1}{4}\right)5.68} = 1.69 \quad \text{for A v B}$$

$$\text{Standard error of a difference} = \sqrt{\left(\frac{1}{4}+\frac{1}{3}\right)5.68} = 1.82 \quad \text{for A v C, B v C}$$

Summary: creating the anova table

Using n data to compare m conditions, with r_i data for condition i.

Summary: creating the anova table

Using n data to compare m conditions, with r_i data for condition i.

$$\begin{aligned}
\text{df for total} &= n-1 \\
\text{df for conditions} &= m-1 \\
\text{df for residual} &= (n-1)-(m-1) = n-m
\end{aligned}$$

Summary: creating the anova table

Using n data to compare m conditions, with r_i data for condition i.

$$\begin{aligned}
\text{df for total} &= n-1 \\
\text{df for conditions} &= m-1 \\
\text{df for residual} &= (n-1)-(m-1) = n-m
\end{aligned}$$

$$\begin{aligned}
\text{SS for total} &= \text{sum of }(\text{data})^2 - n(\text{overall mean})^2 \\
\text{SS for conditions} &= \sum_i r_i(\text{mean for condition } i)^2 - n(\text{overall mean})^2 \\
\text{SS for residual} &= \text{SS for total} - \text{SS for conditions}
\end{aligned}$$

Summary: creating the anova table

Using n data to compare m conditions, with r_i data for condition i.

$$\begin{aligned} \text{df for total} &= n-1 \\ \text{df for conditions} &= m-1 \\ \text{df for residual} &= (n-1)-(m-1) = n-m \end{aligned}$$

$$\begin{aligned} \text{SS for total} &= \text{sum of } (\text{data})^2 - n(\text{overall mean})^2 \\ \text{SS for conditions} &= \sum_i r_i(\text{mean for condition } i)^2 - n(\text{overall mean})^2 \\ \text{SS for residual} &= \text{SS for total} - \text{SS for conditions} \end{aligned}$$

(mean square) $\quad \text{MS} = \text{SS}/\text{df} \quad$ (for both residual and conditions)

Summary: creating the anova table

Using n data to compare m conditions, with r_i data for condition i.

$$\begin{aligned}\text{df for total} &= n-1\\ \text{df for conditions} &= m-1\\ \text{df for residual} &= (n-1)-(m-1)=n-m\end{aligned}$$

$$\begin{aligned}\text{SS for total} &= \text{sum of } (\text{data})^2 - n(\text{overall mean})^2\\ \text{SS for conditions} &= \sum_i r_i(\text{mean for condition } i)^2 - n(\text{overall mean})^2\\ \text{SS for residual} &= \text{SS for total} - \text{SS for conditions}\end{aligned}$$

(mean square) $\qquad \text{MS} = \text{SS}/\text{df} \qquad$ (for both residual and conditions)

$$\text{Ratio} = \frac{\text{MS for conditions}}{\text{MS for residual}}$$

Comments

- It is quite easy to do these calculations by hand,
  or in a spreadsheet:
  then you will need to use tables of the F distribution to get the P-value.

Comments

- It is quite easy to do these calculations by hand,
  or in a spreadsheet:
  then you will need to use tables of the F distribution to get the P-value.
- All statistical software can calculate anova tables and give the P-value,
  but you need to learn how to use it and what its conventions are for labelling the output.

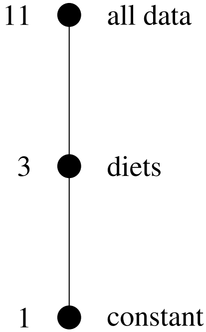
Comments

- It is quite easy to do these calculations by hand,
  or in a spreadsheet:
  then you will need to use tables of the F distribution to get the P-value.
- All statistical software can calculate anova tables and give the P-value,
  but you need to learn how to use it and what its conventions are for labelling the output.
- Some people show the columns in a different order,
  or omit some of them.

Comments

- It is quite easy to do these calculations by hand,
  or in a spreadsheet:
  then you will need to use tables of the F distribution to get the P-value.
- All statistical software can calculate anova tables and give the P-value,
  but you need to learn how to use it and what its conventions are for labelling the output.
- Some people show the columns in a different order,
  or omit some of them.
- If you want to check for normality, you should use the residuals, not the original data.

A diagram

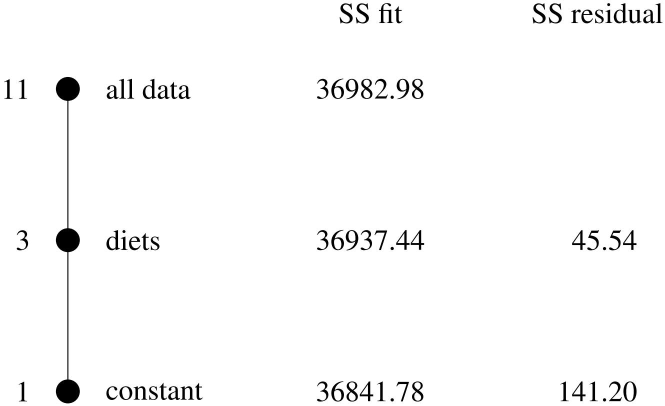


11 all data

3 diets

1 constant

A diagram



		SS fit	SS residual
11	all data	36982.98	
3	diets	36937.44	45.54
1	constant	36841.78	141.20

A diagram

		SS fit	SS residual
11	all data	36982.98	
3	diets	36937.44	45.54
1	constant	36841.78	141.20

df for diets $= 3 - 1 = 2$

SS for diets $= 36937.44 - 36841.78 = 141.20 - 45.54 = 95.66$

End of Part I

Any questions?

An experiment about protecting metal

An experiment was conducted to compare two protective dyes for metal, both with each other and with ‘no dye’. Ten braided metal cords were broken into three pieces. The three pieces of each cord were randomly allocated to the three treatments. After the dyes had been applied, the cords were left to weather for a fixed time, then their strengths were measured, and recorded as a percentage of the nominal strength specification.

Assumptions

Assumption: strength depends additively on dye and cord.

Model:

$$\text{strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$,
where $\varepsilon_{1,1}, \ldots, \varepsilon_{10,3}$ are independent (normal) random variables with mean zero and variance σ^2.

Assumptions

Assumption: strength depends additively on dye and cord.

Model:

$$\text{strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$,
where $\varepsilon_{1,1}, \ldots, \varepsilon_{10,3}$ are independent (normal) random variables with mean zero and variance σ^2.

I want to find out if b_1, b_2 and b_3 are different, and to estimate them.
I suspect that $a_1, \ldots, a_{10}$ are different but I do not really care.

A submodel of the model

$$\text{Model: strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

A submodel of the model

Model: strength of piece of cord i with dye $j = a_i + b_j + \varepsilon_{ij}$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

Submodel: strength of piece of cord i with dye $j = a_i + \varepsilon_{ij}$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

A submodel of the model

$$\text{Model: strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

$$\text{Submodel: strength of piece of cord } i \text{ with dye } j = a_i + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

The submodel is a *special case* of the first model.

A submodel of the model

$$\text{Model: strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

$$\text{Submodel: strength of piece of cord } i \text{ with dye } j = a_i + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

The submodel is a *special case* of the first model.

Null hypothesis: the submodel is true.
Alternative hypothesis: the first model is true but the submodel is not.

Fits and dimensions for both models

$$\text{Model: strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

$$\text{Submodel: strength of piece of cord } i \text{ with dye } j = a_i + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

Fits and dimensions for both models

Model: strength of piece of cord i with dye $j = a_i + b_j + \varepsilon_{ij}$

Submodel: strength of piece of cord i with dye $j = a_i + \varepsilon_{ij}$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

In the submodel, the fit for piece of cord i with dye $j =$ average of cord i.

Dimension of the submodel $= 10$.

Fits and dimensions for both models

$$\text{Model: strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

$$\text{Submodel: strength of piece of cord } i \text{ with dye } j = a_i + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

In the submodel, the fit for piece of cord i with dye $j =$ average of cord i.

$$\text{Dimension of the submodel} = 10.$$

$$\text{In the first model, the fit for piece of cord } i \text{ with dye } j =$$
$$\text{average of cord } i + \text{average of dye } j - \text{overall average}$$

Warning: this simple method does not work unless each dye occurs on each cord equally often.

Fits and dimensions for both models

$$\text{Model: strength of piece of cord } i \text{ with dye } j = a_i + b_j + \varepsilon_{ij}$$

$$\text{Submodel: strength of piece of cord } i \text{ with dye } j = a_i + \varepsilon_{ij}$$

for $i = 1, \ldots, 10$ and $j = 1, \ldots, 3$.

In the submodel, the fit for piece of cord i with dye $j =$ average of cord i.

$$\text{Dimension of the submodel} = 10.$$

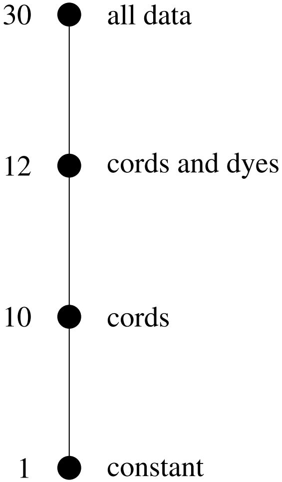
$$\begin{aligned}&\text{In the first model, the fit for piece of cord } i \text{ with dye } j = \\ &\qquad \text{average of cord } i + \text{average of dye } j - \text{overall average}\end{aligned}$$

Warning: this simple method does not work unless each dye occurs on each cord equally often.

$$\text{Dimension of first model} = 10 + 3 - 1 = 12$$

because $a_i + b_j = (a_i + 1) + (b_j - 1)$.

Another diagram



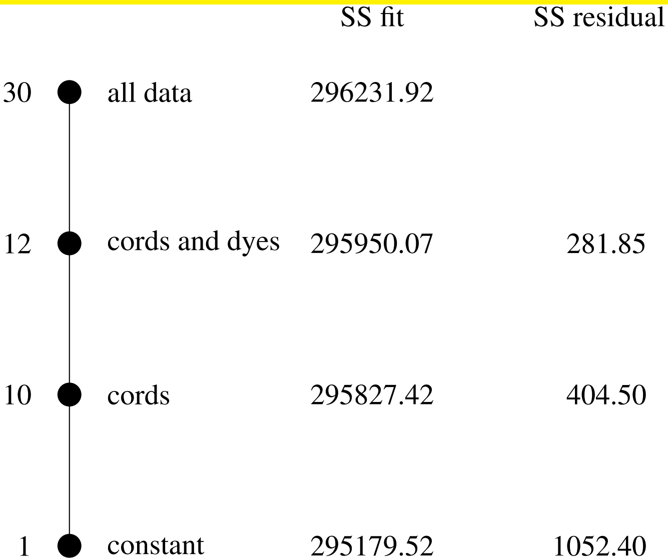
30 all data

12 cords and dyes

10 cords

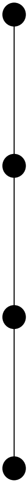
1 constant

Another diagram



		SS fit	SS residual
30	all data	296231.92	
12	cords and dyes	295950.07	281.85
10	cords	295827.42	404.50
1	constant	295179.52	1052.40

Another diagram



		SS fit	SS residual
30	all data	296231.92	
12	cords and dyes	295950.07	281.85
10	cords	295827.42	404.50
1	constant	295179.52	1052.40

Take differences along each edge.

Analysis of variance for dye experiment

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

Analysis of variance for dye experiment

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

We can reject the null hypothesis at the 5% level of significance: we conclude that the different dyes have different effects on metal strength.

Analysis of variance for dye experiment

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

We can reject the null hypothesis at the 5% level of significance: we conclude that the different dyes have different effects on metal strength.

Table of means

No dye	Dye F	Dye G
96.67	99.29	101.62

Analysis of variance for dye experiment

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

We can reject the null hypothesis at the 5% level of significance: we conclude that the different dyes have different effects on metal strength.

Table of means

No dye	Dye F	Dye G
96.67	99.29	101.62

$$\text{Standard error of a difference} = \sqrt{\left(\frac{1}{10}+\frac{1}{10}\right)15.66} = 1.77$$

Warnings about software

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

Warnings about software

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90	72.00	4.60	0.0029
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

- Most software will assume that you are also interested in the differences between cords, and will fill in the row for cords in the anova table.

Warnings about software

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90	72.00	4.60	0.0029
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

- Most software will assume that you are also interested in the differences between cords, and will fill in the row for cords in the anova table.
- If each dye does not occur on each cord equally often,
  the software may produce different sums of squares from those that I recommend here.

Warning about practice

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

Warning about practice

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

The experimenters did well to make sure that each cord had a piece used for each dye, but they were not interested in the cords, so they ignored them in the data analysis.

Warning about practice

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

The experimenters did well to make sure that each cord had a piece used for each dye, but they were not interested in the cords, so they ignored them in the data analysis.

Source	df	Sum of squares	Mean square	Ratio	P
dyes	2	122.65	61.32	1.78	0.1879
residual	27	929.75	34.44		
Total	29	1052.40			

Warning about practice

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dyes	2	122.65	61.32	3.92	0.0386
residual	18	281.85	15.66		
Total	29	1052.40			

The experimenters did well to make sure that each cord had a piece used for each dye, but they were not interested in the cords, so they ignored them in the data analysis.

Source	df	Sum of squares	Mean square	Ratio	P
dyes	2	122.65	61.32	1.78	0.1879
residual	27	929.75	34.44		
Total	29	1052.40			

They were unable to reject the hypothesis of no difference between the dyes.

An intermediate model: to dye or not to dye

Model: strength of piece of cord i with dye $j = a_i + b_j + \varepsilon_{ij}$

Submodel: strength of piece of cord i with dye $j = a_i + \varepsilon_{ij}$

An intermediate model: to dye or not to dye

Model: strength of piece of cord i with dye $j = a_i + b_j + \varepsilon_{ij}$

Intermediate model:

strength of piece of cord i with dye $j = \begin{cases} a_i + \varepsilon_{ij} & \text{if dye } j \text{ is `no dye'} \\ a_i + d + \varepsilon_{ij} & \text{if dye } j \text{ is a real dye} \end{cases}$

Submodel: strength of piece of cord i with dye $j = a_i + \varepsilon_{ij}$

An intermediate model: to dye or not to dye

Model: strength of piece of cord i with dye $j = a_i + b_j + \varepsilon_{ij}$

Intermediate model:

$$\text{strength of piece of cord } i \text{ with dye } j = \begin{cases} a_i + \varepsilon_{ij} & \text{if dye } j \text{ is `no dye'} \\ a_i + d + \varepsilon_{ij} & \text{if dye } j \text{ is a real dye} \end{cases}$$

Submodel: strength of piece of cord i with dye $j = a_i + \varepsilon_{ij}$

The submodel is a special case of the intermediate model,
and the intermediate model is a special case of the first model.

All models on one diagram

		SS fit	SS residual
30	all data	296231.92	
12	cords and dyes	295950.07	281.85
10	cords	295827.42	404.50
1	constant	295179.52	1052.40

All models on one diagram

		SS fit	SS residual
30	all data	296231.92	
12	cords and dyes	295950.07	281.85
11	cords and dye/not		
10	cords	295827.42	404.50
1	constant	295179.52	1052.40

All models on one diagram

		SS fit	SS residual
30	all data	296231.92	
12	cords and dyes	295950.07	281.85
11	cords and dye/not	295922.93	308.99
10	cords	295827.42	404.50
1	constant	295179.52	1052.40

Better analysis of variance for dye experiment: I

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dye/not	1	95.51	95.51	6.10	0.0238
real dyes	1	27.14	27.14	1.73	0.2049
residual	18	281.85	15.66		
Total	29	1052.40			

Better analysis of variance for dye experiment: I

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dye/not	1	95.51	95.51	6.10	0.0238
real dyes	1	27.14	27.14	1.73	0.2049
residual	18	281.85	15.66		
Total	29	1052.40			

First we compare the full model with the intermediate model.
At the 5% level of significance we cannot reject the hypothesis that there is no difference between the two real dyes. So we simplify our assumption to the intermediate model.

Better analysis of variance for dye experiment: I

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dye/not	1	95.51	95.51	6.10	0.0238
real dyes	1	27.14	27.14	1.73	0.2049
residual	18	281.85	15.66		
Total	29	1052.40			

First we compare the full model with the intermediate model.
At the 5% level of significance we cannot reject the hypothesis that there is no difference between the two real dyes. So we simplify our assumption to the intermediate model.

Then we compare the intermediate model to the submodel.
At the 5% level of significance we can reject the hypothesis that ‘no dye’ has the same effect as real dye. We conclude that applying dyes does affect metal strength. We cannot simplify our assumption to the submodel.

Better analysis of variance for dye experiment: II

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dye/not	1	95.51	95.51	6.10	0.0238
real dyes	1	27.14	27.14	1.73	0.2049
residual	18	281.85	15.66		
Total	29	1052.40			

Table of means

No dye	Dye F or Dye G
96.67	100.46

Better analysis of variance for dye experiment: II

Source	df	Sum of squares	Mean square	Ratio	P
cords	9	647.90			
dye/not	1	95.51	95.51	6.10	0.0238
real dyes	1	27.14	27.14	1.73	0.2049
residual	18	281.85	15.66		
Total	29	1052.40			

Table of means

No dye	Dye F or Dye G
96.67	100.46

$$\text{Standard error of difference} = \sqrt{\left(\frac{1}{10} + \frac{1}{20}\right) 15.66} = 1.53$$

An experiment with a quantitative factor

An experiment on forage crops compared five seed mixtures in the presence and absence of nitrogen fertilizer. All ten combinations were grown in plots in five different fields. For each crop mixture in each field, the recorded response is improvement in yield, in tons per acre, if fertilizer is added.

An experiment with a quantitative factor

An experiment on forage crops compared five seed mixtures in the presence and absence of nitrogen fertilizer. All ten combinations were grown in plots in five different fields. For each crop mixture in each field, the recorded response is improvement in yield, in tons per acre, if fertilizer is added.

Fields are like cords: we do not care about their differences.

An experiment with a quantitative factor

An experiment on forage crops compared five seed mixtures in the presence and absence of nitrogen fertilizer. All ten combinations were grown in plots in five different fields. For each crop mixture in each field, the recorded response is improvement in yield, in tons per acre, if fertilizer is added.

Fields are like cords: we do not care about their differences.

Crop mixtures are like diets or dyes: we are interested in their differences.

An experiment with a quantitative factor

An experiment on forage crops compared five seed mixtures in the presence and absence of nitrogen fertilizer. All ten combinations were grown in plots in five different fields. For each crop mixture in each field, the recorded response is improvement in yield, in tons per acre, if fertilizer is added.

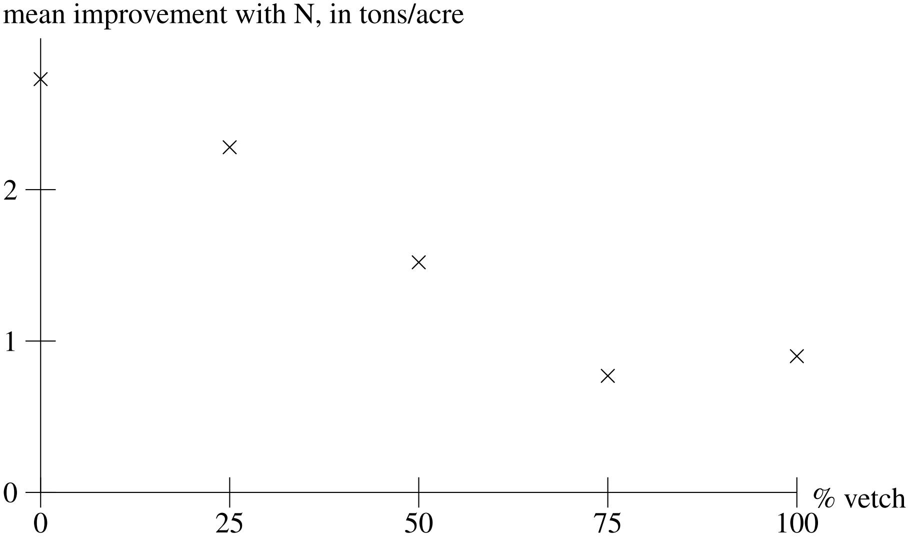
Fields are like cords: we do not care about their differences.

Crop mixtures are like diets or dyes: we are interested in their differences.

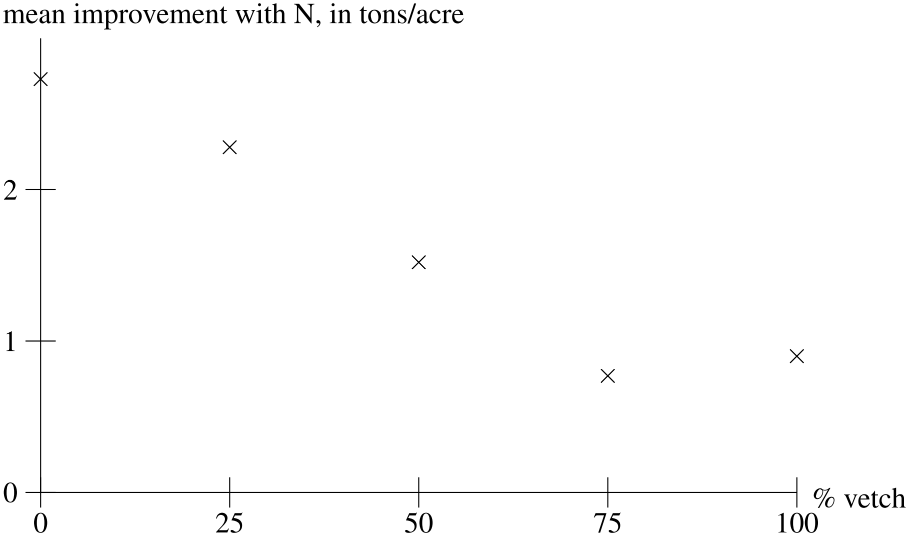
Crop mixtures are not like diets and dyes, because the levels are quantitative:

100% oats	75% oats	50% oats	25% oats	0% oats
0% vetch	25% vetch	50% vetch	75% vetch	100% vetch

An intermediate model: linear in vetch

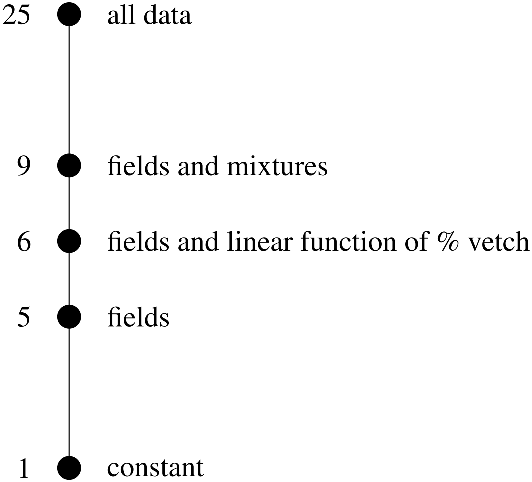


An intermediate model: linear in vetch



Intermediate model: improvement = field parameter + linear(% vetch)

Diagram of models



25	all data
9	fields and mixtures
6	fields and linear function of % vetch
5	fields
1	constant

Another warning about practice

In the forage crops example, almost all the differences between the mixtures can be explained as a linear function of the percentage of vetch.

Source	df	Sum of squares	Mean square	Ratio
fields				
linear in vetch	1	0.2663	0.2663	6.62
mixtures	3	0.0254	0.0085	0.21
residual			0.0402	

Another warning about practice

In the forage crops example, almost all the differences between the mixtures can be explained as a linear function of the percentage of vetch.

Source	df	Sum of squares	Mean square	Ratio
fields				
linear in vetch	1	0.2663	0.2663	6.62
mixtures	3	0.0254	0.0085	0.21
residual			0.0402	

The first statistician who analysed these data did it in the simple way:

Source	df	Sum of squares	Mean square	Ratio
fields				
mixtures	4	0.2917	0.0729	1.81
residual			0.0402	

and concluded that there was not enough evidence that the mixtures had different effects.

Polynomial models

Suppose that we measure quantities x_i and y_i on item i,
for $i = 1, \ldots, n$.
We want a model that predicts y as a function of x.
We might try a cubic polynomial:

$$y_i = a + bx_i + cx_i^2 + dx_i^3 + \varepsilon_i.$$

Polynomial models

Suppose that we measure quantities x_i and y_i on item i, for $i = 1, \ldots, n$.
We want a model that predicts y as a function of x.
We might try a cubic polynomial:

$$y_i = a + bx_i + cx_i^2 + dx_i^3 + \varepsilon_i.$$

A special case is the quadratic polynomial:

$$y_i = e + fx_i + gx_i^2 + \varepsilon_i,$$

Polynomial models

Suppose that we measure quantities x_i and y_i on item i,
for $i = 1, \ldots, n$.
We want a model that predicts y as a function of x.
We might try a cubic polynomial:

$$y_i = a + bx_i + cx_i^2 + dx_i^3 + \varepsilon_i.$$

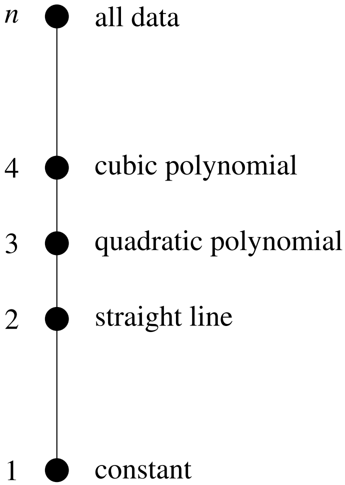
A special case is the quadratic polynomial:

$$y_i = e + fx_i + gx_i^2 + \varepsilon_i,$$

and a special case of that is the straight line:

$$y_i = h + kx_i + \varepsilon_i.$$

Diagram of polynomial models



n all data

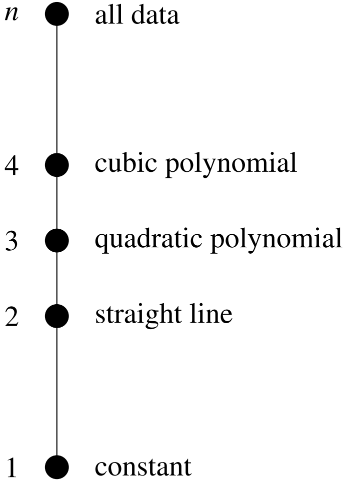
4 cubic polynomial

3 quadratic polynomial

2 straight line

1 constant

Diagram of polynomial models



n all data

4 cubic polynomial

3 quadratic polynomial

2 straight line

1 constant

Warning: the best-fitting quadratic polynomial is not usually obtained by taking the best-fitting cubic polynomial and removing the term in x^3

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.
3. For each model, fit it, and find either its residual sum of squares or the sum of the squares of the fits.

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.
3. For each model, fit it, and find either its residual sum of squares or the sum of the squares of the fits.
4. Each line of the diagram corresponds to a difference between models. For each line, calculate

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.
3. For each model, fit it, and find either its residual sum of squares or the sum of the squares of the fits.
4. Each line of the diagram corresponds to a difference between models. For each line, calculate
   - 4.1 df = difference in dimensions

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.
3. For each model, fit it, and find either its residual sum of squares or the sum of the squares of the fits.
4. Each line of the diagram corresponds to a difference between models. For each line, calculate
    4.1 df = difference in dimensions
    4.2 SS = difference in residual sums of squares

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.
3. For each model, fit it, and find either its residual sum of squares or the sum of the squares of the fits.
4. Each line of the diagram corresponds to a difference between models. For each line, calculate
    4.1 df = difference in dimensions
    4.2 SS = difference in residual sums of squares
    4.3 MS = SS/df.

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.
3. For each model, fit it, and find either its residual sum of squares or the sum of the squares of the fits.
4. Each line of the diagram corresponds to a difference between models. For each line, calculate
    4.1 df = difference in dimensions
    4.2 SS = difference in residual sums of squares
    4.3 MS = SS/df.
5. Each row of the anova table corresponds to a line on the diagram, but the top-to-bottom order is reversed.

Creating the anova table for a chain of models

Your software *may* be able to do this.

1. Draw the the diagram showing all the models.
2. Show the dimension of each model.
3. For each model, fit it, and find either its residual sum of squares or the sum of the squares of the fits.
4. Each line of the diagram corresponds to a difference between models. For each line, calculate
    4.1 df = difference in dimensions
    4.2 SS = difference in residual sums of squares
    4.3 MS = SS/df.
5. Each row of the anova table corresponds to a line on the diagram, but the top-to-bottom order is reversed.
6. The bottom row gives the residual MS; divide every other MS by this.

Interpreting the anova table for a chain of models

Start at the bottom.
The bottom row gives the residual.
The next row up compares the largest model with the second-largest.

Interpreting the anova table for a chain of models

Start at the bottom.
The bottom row gives the residual.
The next row up compares the largest model with the second-largest.
If the P-value is “small”, then we cannot simplify the model: report that we need the largest model and estimate its parameters.

Interpreting the anova table for a chain of models

Start at the bottom.
The bottom row gives the residual.
The next row up compares the largest model with the second-largest.
If the P-value is “small”, then we cannot simplify the model: report that we need the largest model and estimate its parameters.
If the P-value is “large”, then we can simplify the model, so move up to the next row.

Interpreting the anova table for a chain of models

Start at the bottom.
The bottom row gives the residual.
The next row up compares the largest model with the second-largest.
If the P-value is “small”, then we cannot simplify the model: report that we need the largest model and estimate its parameters.
If the P-value is “large”, then we can simplify the model, so move up to the next row.
Continue like this until the model cannot be simplified any further: at this stage, report this model and estimate its parameters.

End of Part II

Any questions?

An experiment to investigate combinations of two factors

Eight newly-hatched chicks took part in a feeding experiment. Four different feeds (*A*, *B*, *C* and *D*) were made available to two chicks each. The protein in feeds *A* and *B* was groundnuts, while the protein in feeds *C* and *D* was soya bean. Moreover, feeds *B* and *D* contained added fishmeal.

The chicks were weighed (in gm) at the end of six weeks.

An experiment to investigate combinations of two factors

Eight newly-hatched chicks took part in a feeding experiment. Four different feeds (A, B, C and D) were made available to two chicks each. The protein in feeds A and B was groundnuts, while the protein in feeds C and D was soya bean. Moreover, feeds B and D contained added fishmeal.

The chicks were weighed (in gm) at the end of six weeks.

protein	groundnuts		soya bean	
no added fishmeal	*A*		*C*	
	410	393	443	441
added fishmeal	*B*		*D*	
	442	424	500	479

Some models

protein	groundnuts	soya bean
no added fishmeal	A K_A	C K_C
added fishmeal	B K_B	D K_D

One model: weight depends on diet (dimension 4).

Some models

protein	groundnuts	soya bean
no added fishmeal	A K_A a_1	C K_C a_2
added fishmeal	B K_B a_1	D K_D a_2

One model: weight depends on diet (dimension 4).
Special case (protein): $K_A = K_B = a_1$ and $K_C = K_D = a_2$ (dim 2).

Some models

protein	groundnuts	soya bean
no added fishmeal	A K_A a_1 b_1	C K_C a_2 b_1
added fishmeal	B K_B a_1 b_2	D K_D a_2 b_2

One model: weight depends on diet (dimension 4).
Special case (protein): $K_A = K_B = a_1$ and $K_C = K_D = a_2$ (dim 2).
Special case (fishmeal): $K_A = K_C = b_1$ and $K_B = K_D = b_2$ (dim 2).

Some models

protein	groundnuts	soya bean
no added fishmeal	A K_A a_1 b_1	C K_C a_2 b_1
added fishmeal	B K_B a_1 b_2	D K_D a_2 b_2

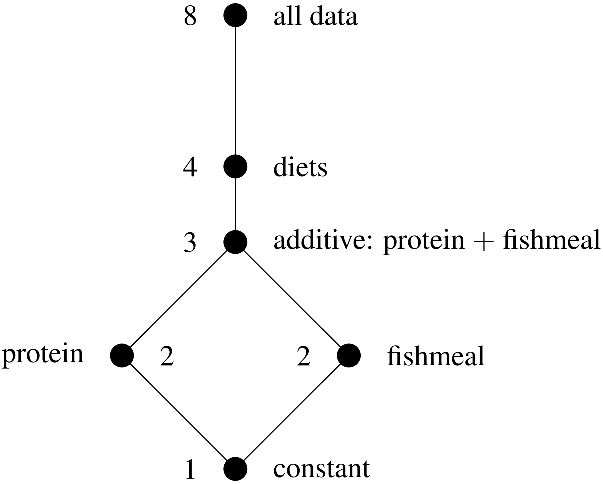
One model: weight depends on diet (dimension 4).
Special case (protein): $K_A = K_B = a_1$ and $K_C = K_D = a_2$ (dim 2).
Special case (fishmeal): $K_A = K_C = b_1$ and $K_B = K_D = b_2$ (dim 2).
Very special case: $K_A = K_B = K_C = K_D$ (dimension 1).

Some models

protein	groundnuts	soya bean
no added fishmeal	A K_A a_1 b_1 $a_1 + b_1$	C K_C a_2 b_1 $a_2 + b_1$
added fishmeal	B K_B a_1 b_2 $a_1 + b_2$	D K_D a_2 b_2 $a_2 + b_2$

One model: weight depends on diet (dimension 4).
Special case (protein): $K_A = K_B = a_1$ and $K_C = K_D = a_2$ (dim 2).
Special case (fishmeal): $K_A = K_C = b_1$ and $K_B = K_D = b_2$ (dim 2).
Very special case: $K_A = K_B = K_C = K_D$ (dimension 1).
The additive model (dimension $2 + 2 - 1 = 3$) is intermediate between both special cases and the full model.

Diagram of models



Anova for chick-feeding experiment: I

Source	df	Sum of squares	Mean square	Ratio	P
protein	1	4704.5	4704.50	35.57	0.004
fishmeal	1	3120.5	3120.50	23.60	0.0083
diets	1	128.0	128.00	0.97	0.3804
residual	4	529.0	132.25		
Total	7	8482.0			

Anova for chick-feeding experiment: I

Source	df	Sum of squares	Mean square	Ratio	P
protein	1	4704.5	4704.50	35.57	0.004
fishmeal	1	3120.5	3120.50	23.60	0.0083
diets	1	128.0	128.00	0.97	0.3804
residual	4	529.0	132.25		
Total	7	8482.0			

First we compare the full model with the additive model. At the 5% level of significance we cannot reject the hypothesis that the additive model is true.
This means that the difference caused by adding fishmeal is the same for both proteins. We say that "fishmeal does not interact with protein" or "there is no interaction between the two factors".
So we simplify our assumption to the additive model.

Anova for chick-feeding experiment: II

Source	df	Sum of squares	Mean square	Ratio	P
protein	1	4704.5	4704.50	35.57	0.004
fishmeal	1	3120.5	3120.50	23.60	0.0083
interaction	1	128.0	128.00	0.97	0.3804
residual	4	529.0	132.25		
Total	7	8482.0			

Anova for chick-feeding experiment: II

Source	df	Sum of squares	Mean square	Ratio	P
protein	1	4704.5	4704.50	35.57	0.004
fishmeal	1	3120.5	3120.50	23.60	0.0083
interaction	1	128.0	128.00	0.97	0.3804
residual	4	529.0	132.25		
Total	7	8482.0			

Then we compare the additive model to both submodels.
At the 5% level of significance we can reject the hypothesis that there is no difference between the proteins, and we can reject the hypothesis that there is no difference between adding fishmeal and not.

Anova for chick-feeding experiment: II

Source	df	Sum of squares	Mean square	Ratio	P
protein	1	4704.5	4704.50	35.57	0.004
fishmeal	1	3120.5	3120.50	23.60	0.0083
interaction	1	128.0	128.00	0.97	0.3804
residual	4	529.0	132.25		
Total	7	8482.0			

Then we compare the additive model to both submodels.
At the 5% level of significance we can reject the hypothesis that there is no difference between the proteins, and we can reject the hypothesis that there is no difference between adding fishmeal and not.

Table of means

groundnuts	soya bean
417.25	465.75

Table of means

no added fishmeal	added fishmeal
421.75	461.25

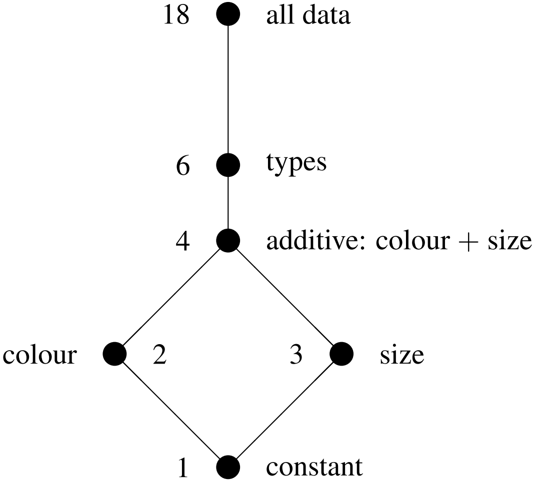
Colour and size

Six types of seed (two colours and three sizes) for a certain plant were compared. For each type, a known number were planted in each of three pots, and the percentage which germinated was recorded.

Colour		Size	
	small	medium	large
brown	73.0	89.0	84.7
red	22.0	26.3	48.7

(Data are average of three pots.)

Diagram of models



Anova for germination experiment

Source	df	Sum of squares	Mean square	Ratio	P
colour	1	11200.06	11200.06	165.06	< 0.0001
size	2	1103.44	551.72	8.13	0.0059
colour × size	2	536.12	268.06	3.95	0.0481
residual	12	814.23	67.85		
Total	17	13653.85			

Anova for germination experiment

Source	df	Sum of squares	Mean square	Ratio	P
colour	1	11200.06	11200.06	165.06	< 0.0001
size	2	1103.44	551.72	8.13	0.0059
colour × size	2	536.12	268.06	3.95	0.0481
residual	12	814.23	67.85		
Total	17	13653.85			

First we compare the full model with the additive model. This is often called 'testing for interaction'. At the 5% level of significance we reject the hypothesis that the additive model is true. This means that the difference between the sizes is not the same for both colours.

Anova for germination experiment

Source	df	Sum of squares	Mean square	Ratio	P
colour	1	11200.06	11200.06	165.06	< 0.0001
size	2	1103.44	551.72	8.13	0.0059
colour × size	2	536.12	268.06	3.95	0.0481
residual	12	814.23	67.85		
Total	17	13653.85			

First we compare the full model with the additive model. This is often called ‘testing for interaction’. At the 5% level of significance we reject the hypothesis that the additive model is true. This means that the difference between the sizes is not the same for both colours.

Present the table of means for the six types, and stop there, because we cannot simplify the model.

Anova for germination experiment

Source	df	Sum of squares	Mean square	Ratio	P
colour	1	11200.06	11200.06	165.06	< 0.0001
size	2	1103.44	551.72	8.13	0.0059
colour × size	2	536.12	268.06	3.95	0.0481
residual	12	814.23	67.85		
Total	17	13653.85			

First we compare the full model with the additive model. This is often called ‘testing for interaction’. At the 5% level of significance we reject the hypothesis that the additive model is true. This means that the difference between the sizes is not the same for both colours.

Present the table of means for the six types, and stop there, because we cannot simplify the model.

Is it useful to present the tables of means for colours and for sizes?

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages
- sums of squares are obtained from simple averages

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages
- sums of squares are obtained from simple averages
- all software presents the same anova table

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages
- sums of squares are obtained from simple averages
- all software presents the same anova table
- different routes from the top of the model diagram to the bottom all give the same conclusions.

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages
- sums of squares are obtained from simple averages
- all software presents the same anova table
- different routes from the top of the model diagram to the bottom all give the same conclusions.

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages
- sums of squares are obtained from simple averages
- all software presents the same anova table
- different routes from the top of the model diagram to the bottom all give the same conclusions.

If they don't occur equally often then

- you need software to calculate the fits and the sums of squares

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages
- sums of squares are obtained from simple averages
- all software presents the same anova table
- different routes from the top of the model diagram to the bottom all give the same conclusions.

If they don't occur equally often then

- you need software to calculate the fits and the sums of squares
- different software presents different sums of squares, and you have to know how to interpret it

Warning about unequal replication

If every combination of levels of the two factors occurs equally often then

- fits are obtained from simple averages
- sums of squares are obtained from simple averages
- all software presents the same anova table
- different routes from the top of the model diagram to the bottom all give the same conclusions.

If they don't occur equally often then

- you need software to calculate the fits and the sums of squares
- different software presents different sums of squares, and you have to know how to interpret it
- different routes from the top of the model diagram can sometimes give contradictory conclusions.

End of Part III

Any questions?