NTNU
Norwegian University of
Science and Technology

TMA4267 Linear Statistical Models V2014 (20)

Design of experiments (note): significant effects (p 7-12) and blocking (p 15-20)

Mette Langaas

To be lectured: March 11, 2014
wiki.math.ntnu.no/emner/tma4267/2014v/start/

Plan DOE today

- The full 2^k experiment: significance of effects (p 7-12)
- Blocking in full 2^k experiments (p 15-20)

Lima beans example

Experiment from Box, Hunter, Hunter, Statistics for Experimenters, page 321.

- A: depth of planting (0.5 inch or 1.5 inch)
- B: watering daily (once or twice)
- C: type of lima bean (baby or large)
- Y: yield

A	B	C	AB	AC	BC	ABC	Level code	Response
-	-	-	+	+	+	-	1	6
+	-	-	-	-	+	+	a	4
-	+	-	-	+	-	+	b	10
+	+	-	+	-	-	-	ab	7
-	-	+	+	-	-	+	c	4
+	-	+	-	+	-	-	ac	3
-	+	+	-	-	+	-	bc	8
+	+	+	+	+	+	+	abc	5
x_1	x_2	x_3	x_{12}	x_{13}	x_{23}	x_{123}		y

DOE 2^k full factorial : Significant effect ?

$$Y = X\beta + \varepsilon \qquad \varepsilon \sim N_n(0, \sigma^2 I)$$

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 & + \text{interaction} \\ 1 & 1 & -1 & -1 & \\ 1 & -1 & 1 & -1 & \\ 1 & 1 & 1 & -1 & \\ \vdots & -1 & -1 & 1 & \\ \vdots & 1 & -1 & 1 & \end{bmatrix}_{8 \times 8}$$

$$\text{Effect}_j = 2 \cdot \beta_j$$

$$\widehat{\text{Effect}}_j = 2 \hat{\beta}_j$$

$$\widehat{\text{Effect}}_j = 2\hat{\beta}_j = \frac{2}{n} \sum_{i=1}^{n} x_{ij} y_i$$

$$\text{Var}(\widehat{\text{Effect}}_j) = \frac{4}{n} \sigma^2 \equiv \sigma^2_{\text{effect}}$$

$$\widehat{\text{Effect}}_j \sim N(\text{Effect}_j, \sigma^2_{\text{effect}})$$

If we have s^2_{effect} as estimator for σ^2_{effect}

$H_0: \text{Effect}_j = 0$ vs $H_1: \text{Effect}_j \neq 0$

$$T = \frac{\widehat{\text{Effect}}_j - 0}{s_{\text{effect}}} \sim t_\nu \leftarrow \text{df from estimator } s^2_{\text{effect}}$$

Reject H_0 when $|t| > t_{\frac{\alpha}{2},\nu} \iff |\widehat{\text{Effect}}_j| > t_{\frac{\alpha}{2},\nu} \cdot s_{\text{effect}}$

Significant effect?

$$H_0 : Effect_j = 0 \text{ vs } Effect_j \neq 0$$

or equivalently

$$H_0 : \beta_j = 0 \text{ vs } \beta_j \neq 0$$

If σ_{effect} is estimated by s_{effect}, respectively, then we reject H_0 and say that $Effect_j$ is significant if

$$|\widehat{Effect}_j| > t_{\alpha/2,\nu} s_{effect}$$

where ν is the number of degrees of freedom connected to the estimates of σ_{effect} that are used.

R: DOE set-up

```
> summary(lm3)

Call:
lm.default(formula = y ~ (.)^3, data = plan)

Residuals:
ALL 8 residuals are 0: no residual degrees of freedom!

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.875         NA      NA       NA
A1            -1.125         NA      NA       NA
B1             1.625         NA      NA       NA
C1            -0.875         NA      NA       NA
A1:B1         -0.375         NA      NA       NA
A1:C1          0.125         NA      NA       NA
B1:C1         -0.125         NA      NA       NA
A1:B1:C1      -0.125         NA      NA       NA

Residual standard error: NaN on 0 degrees of freedom
Multiple R-squared:     1,Adjusted R-squared:   NaN
F-statistic:   NaN on 7 and 0 DF,  p-value: NA
```

Estimation of σ^2

1. Lenth's Pseudo Standard Error (PSE).
2. Assuming specified higher order interactions are zero (changing the MLR model).
3. Perform replicates, estimate the full model and use s^2 from MLR.

Lenth's PSE

Let $C_1, C_2, \ldots, C_m$ be estimated effects, e.g. $\hat{A}, \hat{B}, \widehat{AB}$, etc.

1. Order absolute values $|C_j|$ in increasing order.
2. Find the median of the $|C_j|$ and compute preliminary estimate

$$s_0 = 1.5 \cdot \text{median}_j |C_j|$$

3. Take out the effects C_j with $|C_j| \geq 2.5 \cdot s_0$ and find the median of the rest of the $|C_j|$. Then PSE is this median multiplied by 1.5, i.e.

$$\text{PSE} = 1.5 \cdot \text{median}\{|C_j| : |C_j| < 2.5 s_0\}$$

and this is Lenth's estimate of σ_{effect}.

4. Lenth has suggested empirically that the degrees of freedom to be used with PSE is $m/3$ where m is the initial number of effects in the algorithm (intercept not included). Thus we claim as significant the effects for which $|C_j| > t_{\alpha/2, m/3} \cdot PSE$.

Methods for estimating σ^2_{effect}

1) Lenth's method : PSE is a conservative (too large) estimate for σ_{effect}.

Idea: many factors have zero or near zero effect $\Rightarrow$
use a function of the median of the absolute value of trimmed effects to estimate σ_{effect}.

Lima beans : $PSE = 0.75$, $m = 7$, $\nu = \frac{m}{3} = \frac{7}{3}$, $\alpha = 0.05$

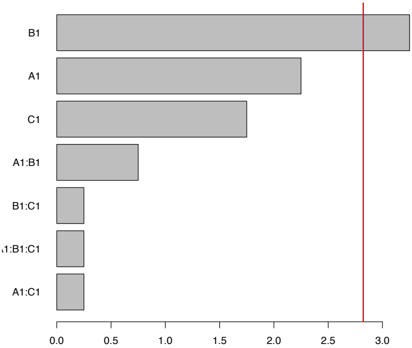
$$t_{0.025, 7/3} = 3.76$$

$$t_{0.025, 7/3} \cdot PSE = \underline{2.823}$$

Only $|\hat{B}| > 2.823 \Rightarrow$ significant with Lenth's method.

Pareto plot : ordered histogram of $|\widehat{Effect_j}|$.

R: Pareto plot for Lima beans



Pareto plot: ordered histogram of absolute value of estimated effects, Length sign line added.

Assuming specified higher order interactions are zero

- In general

$$\widehat{Effect}_j \sim N(Effect_j, \sigma^2_{effect})$$

- If we assume that the effect is zero ($\beta_j = 0$), then $\mathrm{E}(Effect_j) = 0$ and

$$\mathrm{E}(\widehat{Effect}_j^2) = \sigma^2_{effect}$$

- Thus $\widehat{Effect}_j^2$ is an unbiased estimator of σ^2_{effect} if $\beta_j = 0$.
- If several effects are assumed to be 0, we use the average of the $\widehat{Effect}_j^2$ to estimate σ^2_{effect}.

2) Assuming higher order interactions are zero ⇒ changing the MLR model

We may assume that interactions are zero, and fit a simpler model.

Here (Lima beans) we reduce our model to only contain main effects.

MLR: just fit a model with $Y_i = \beta_0 + \underset{A}{\beta_1 x_{i1}} + \underset{B}{\beta_2 x_{i2}} + \underset{C}{\beta_3 x_{i3}} + \varepsilon$

and automatically we have s^2 and significant values

For interpretation: DOE " s^2_{effect} is estimated as the mean of $(\widehat{Effect}_j)^2$ that we think are zero". Why?

Remember: $\widehat{Effect}_j \sim N(Effect_j, \sigma^2_{effect} = \frac{4}{n}\sigma^2)$

We assume that $Effect_j = 0$, then

$$\sigma^2_{effect} = Var(\widehat{Effect}_j) = E(\widehat{Effect}_j^2) - \underbrace{E(\widehat{Effect}_j)^2}_{0} = E(\widehat{Effect}_j^2)$$

$s^2_{effect} = mean(\widehat{Effect}_j^2)$ for the effects assumed to be zero.

Lima beans:

1) lm with only main effects: printout show

Std. error: 0.2165 for coeffs, so $S_{effect} = 2 \cdot 0.2165 = 0.433$

2) $S_{effect}^2 = \frac{1}{4}(\underset{-0.75}{\widehat{AB}^2} + \underset{0.25}{\widehat{AC}^2} + \underset{0.25}{\widehat{BC}^2} + \underset{-0.25}{\widehat{ABC}^2}) = \dots = 0.1875$

$S_{effect} = \sqrt{0.1875} = 0.433$

To do inference: ν = # effects in an mean$(\widehat{\text{Effect}}_j)$

Limabeano: $\nu = 4$

$t_{0.025,4} = 2.78$

Cut-off: $2.78 \cdot \underbrace{0.432}_{s_{effect}} = 1.2$

If $|\widehat{\text{Effect}}_j| > 1.2 \Rightarrow$ assume significant

Conclusion: A, B, C significant.

Warning: it may be very confusing to look at software output:

MSE, s, s_{effect}, Std. Err $= SD(\hat{\beta}_j)$

Lima beans estimated effects: full model

```
Estimated effects (2*coeff):
(Intercept)  A1    B1     C1    A1:B1   A1:C1   B1:C1  A1:B1:C1
  11.75    -2.25  3.25  -1.75   -0.75   0.25    -0.25   -0.25

Analysis of Variance Table
          Df Sum Sq Mean Sq F value Pr(>F)
A          1 10.125  10.125
B          1 21.125  21.125
C          1  6.125   6.125
A:B        1  1.125   1.125
A:C        1  0.125   0.125
B:C        1  0.125   0.125
A:B:C      1  0.125   0.125
Residuals  0  0.000
```

Lima beans: only main effects

```
> lm1 <- lm(y~.,data=plan)
> summary(lm1)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.8750     0.2165  27.135  1.1e-05 ***
A1           -1.1250     0.2165  -5.196  0.00653 **
B1            1.6250     0.2165   7.506  0.00169 **
C1           -0.8750     0.2165  -4.041  0.01559 *
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

Residual standard error: 0.6124 on 4 degrees of freedom
Multiple R-squared: 0.9614,Adjusted R-squared: 0.9325
F-statistic: 33.22 on 3 and 4 DF,  p-value: 0.002755
> anova(lm1)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value   Pr(>F)
A          1 10.125  10.125  27.000 0.006533 **
B          1 21.125  21.125  56.333 0.001686 **
C          1  6.125   6.125  16.333 0.015585 *
Residuals  4  1.500   0.375
```

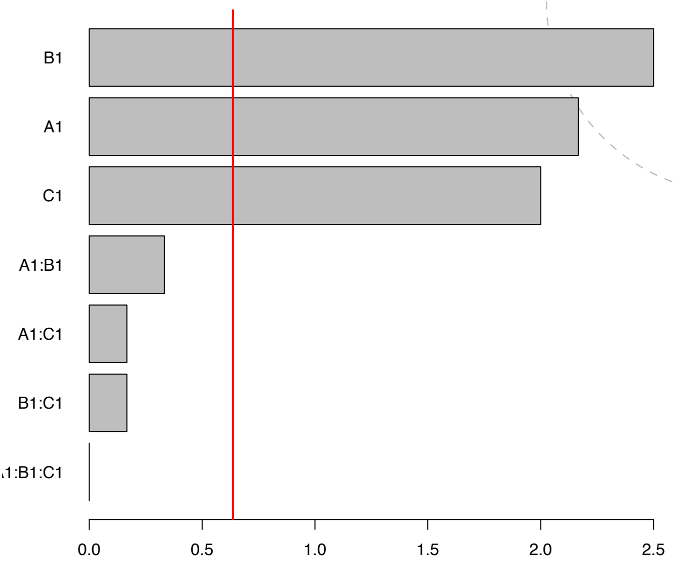
Three factors in three full replicates

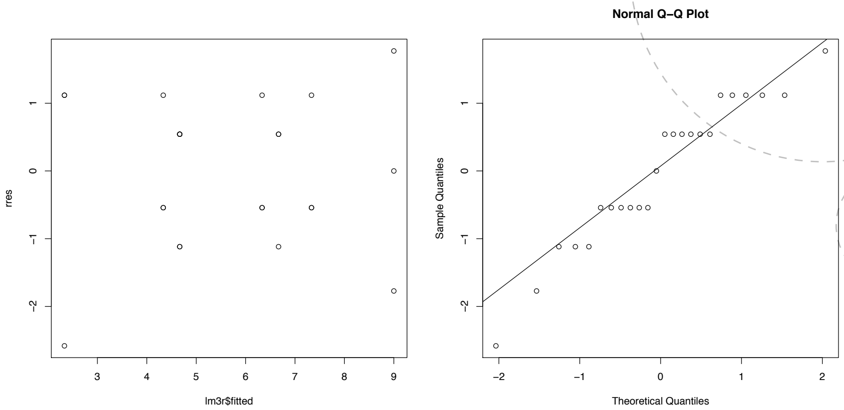
— Lima beans experiment from Box, Hunter, Hunter page 321.
  - A: depth of planting (0.5 inch or 1.5 inch)
  - B: watering daily (once or twice)
  - C: type of limabean (baby or large)
  - Y: yield

— $r = 3$: Performed in three full replicate experiments, i.e. three measurements for each combination of A, B and C.

— We then have $(r-1)2^3 = 2 \cdot 8 = 16$ degrees of freedom for estimating the error variance.

— Estimates follow automatically. Perform this for yourself. Data from course www-page with title “limabeans.r”.


B1
A1
C1
A1:B1
A1:C1
B1:C1
1:B1:C1
0.0
0.5
1.0
1.5
2.0
2.5


Normal Q−Q Plot
rres
lm3r$fitted
Sample Quantiles
Theoretical Quantiles

ANOVA output: R

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
A          1 28.167  28.167 52.0000 2.075e-06 ***
B          1 37.500  37.500 69.2308 3.319e-07 ***
C          1 24.000  24.000 44.3077 5.517e-06 ***
A:B        1  0.667   0.667  1.2308    0.2837
A:C        1  0.167   0.167  0.3077    0.5868
B:C        1  0.167   0.167  0.3077    0.5868
A:B:C      1  0.000   0.000  0.0000    1.0000
Residuals 16  8.667   0.542
```

Which ν?

From the previous slide, connection between ν and your chosen estimation method for σ and σ_{effect}.

1. When Lenth's PSE is used, the degrees of freedom is

$$\nu = \frac{2^k - 1}{3}$$

   where $2^k - 1$ is the number of effects in the model, while the 3 in the denominator has been found empirically by Lenth.

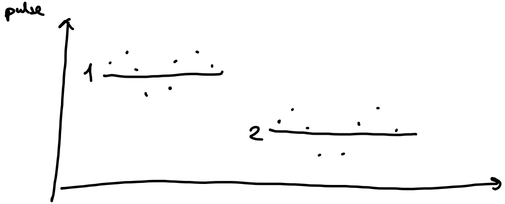
2. If m effects (preferrable higher order interactions) are assumed to be zero, then $\nu = m$.
3. If you have performed the 2^k experiment r times, then $\nu = (r-1)2^k$.

3) Perform replicates: ordinary MLR → no fuzz!

But beware!

Jens and Erna want to perform a 2^3 experiment together, and to get 16 observations they will both conduct the same 2^3 experiment.

Should then a covariate ($x=-1$ Erna, $x=+1$ Jens) telling who did the experiment be added to the regression model?



If the above figure gives a correct picture of the experiment NOT including a person covariate will make SSE very large, and therefore you may get the result that no factor is significant.

Randomize: B 71%
D 29%

Q: Randomization

Why do you need to randomize the order in which you perform the experiments?

To make the experiments

- A: random.
- B: robust to external factors.
- C: have constant variance.
- D: independent.

Vote at clicker.math.ntnu.no, TMA4267 classroom.

Pilot plant: A, B and C

A=Temperature, B=Concentration, C=Catalyst, Y=yield.

A	B	C	AB	AC	BC	ABC	Level code	Response
-	-	-	+	+	+	-	1	60
+	-	-	-	-	+	+	a	72
-	+	-	-	+	-	+	b	54
+	+	-	+	-	-	-	ab	68
-	-	+	+	-	-	+	c	52
+	-	+	-	+	-	-	ac	83
-	+	+	-	-	+	-	bc	45
+	+	+	+	+	+	+	abc	80
x_1	x_2	x_3	x_{12}	x_{13}	x_{23}	x_{123}		y

DOE workflow

1. Set up full factorial design with *k* factors in R, and
2. randomize the runs.
3. Perform experiments, and enter data into R.
4. Fit a full model (all interactions).
5. If you do not have replications, look at Pareto plots and estimate variability with Lenths metod, use this to suggest at reduced model (if possible). Refit the reduced model.
6. Assess model fit (residual plots, need transformations?).
7. Assess significance.
8. Interpret you results (main and interaction plots).

Genuine run replicates

"When genuine run replicates are made under a given set of experimental conditions, the variation between the associated observations may be used to estimate the standard deviation of the effects. By *genuine* run replicated we mean that variation between runs made at the same experimental conditions is a reflection of the total variability afflicting runs made at different experimental conditions. This point requires careful consideration."
From Box, Hunter, Hunter (1978, 2005): "Statistics for Experimenters", Ch.10.6.

Genuine run replicates

Randomization of run order usually ensures that replicates are genuine. Pilot plant example: each run consists of

1. cleaning the reactor
2. inserting the appropriate catalyst carge
3. running the apparatus at at given temperature and a given feed concentration for 3 hrs to allow the process to settle down at the chosen experimental conditions, and
4. combining chemical analyses made on these samples.

A genuine run replicate must involve the taking of all these steps again. In particular, several chemical analyses from a single run would provide only an estimate of *analytical* variance, usually only a small part of the run-to-run variance.

From Box, Hunter, Hunter (1978, 2005): “Statistics for Experimenters”, Ch.10.6.

Performing a full 2^k factorial experiment

Two important concepts:

1) Each experiment is a genuine run replicate, that is, reflects the total variability of the experiment.

2) The run order is random, so that potential external factors are not confused (confounded) with experimental factors.

Blocking

Pilot plant, 2^3 in two blocks: We will perform a 2^3 experiment, but must use two batches of raw material.

Solution: use the ABC column to define the blocks.
ABC is the block generator.

What could you consider a block in your DOE experiment?

2^4 in two days: use ABCD as block generator

Read yourself: more than two blocks!

Blocking on ABC

Block 1 consists of experiments with ABC=-1.
Block 2 consists of experiments with ABC=1.

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
StdOrder	RunOrder	CenterPt	Blocks	A	B	C	ABC		Y	block effect
1	1	1	1	-1	-1	-1	-1	1	60	60
4	4	1	1	-1	1	1	-1	7	45	45
3	3	1	1	1	-1	1	-1	6	83	83
2	2	1	1	1	1	-1	-1	4	68	68
7	7	1	2	-1	-1	1	1	5	52	62
6	6	1	2	-1	1	-1	1	3	54	64
5	5	1	2	1	-1	-1	1	2	72	82
8	8	1	2	1	1	1	1	8	80	90

Blocking on ABC

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
StdOrder	RunOrder	CenterPt	Blocks	A	B	C	ABC		Y	block effect
1	1	1	1	-1	-1	-1	-1	1	60	60
4	4	1	1	-1	1	1	-1	7	45	45
3	3	1	1	1	-1	1	-1	6	83	83
2	2	1	1	1	1	-1	-1	4	68	68
7	7	1	2	-1	-1	1	1	5	52	62
6	6	1	2	-1	1	-1	1	3	54	64
5	5	1	2	1	-1	-1	1	2	72	82
8	8	1	2	1	1	1	1	8	80	90

— ABC is counfunded with the block effect. We can not separate these two effects from eachother.

— Suppose all values in block 2 is increased by 10 units.

- Then the estimated effect of ABC will increase by 10.
- But all other estimated effects remain unchanged - and these are the most important to estimate.

Original data

```
Factorial Fit:
Y versus
Block A B C
Term       Effect     Coef

Constant             64,250
Block                -0,250
A          23,000    11,500
B          -5,000    -2,500
C           1,500     0,750
A*B         1,500     0,750
A*C        10,000     5,000
B*C         0,000     0,000
```

Added 10 to all obs in Block 2.

```
Factorial Fit:
"block effect" versus
Block A B C

Term       Effect     Coef
Constant             69,250
Block                -5,250
A          23,000    11,500
B          -5,000    -2,500
C           1,500     0,750
A*B         1,500     0,750
A*C        10,000     5,000
B*C         0,000     0,000
```

2^3 with four blocks

We need two generators (columns) to define four blocks: the optimal choice is AB and AC

- Block 1: AB=AC=-1 (- -)
- Block 2: AB=-1, AC=1 (- +)
- Block 3: AB=1, AC=-1 (+ -)
- Block 4: AB=AC=1 (+ +)

Std order	A	B	C	AB	AC	BC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

2^3 with AB and AC as generators

Std order	A	B	C	AB	AC	BC	ABC	Block
2	+	-	-	-	-	+	+	1
7	-	+	+	-	-	+	-	1
3	-	+	-	-	+	-	+	2
6	+	-	+	-	+	-	-	2
4	+	+	-	+	-	-	-	3
5	-	-	+	+	-	-	+	3
1	-	-	-	+	+	+	-	4
8	+	+	+	+	+	+	+	4

2^3 with AB and AC as generators

— Interaction effects AB and AC are confounded with the block effect, since they are the generators.

— Their product, $AB * AC = A^2BC = BC$, is alco confounded with the block effect (see that BC is constant within each block).

— Adding h_2 to block 2, h_3 to block 3 and h_4 to block 4 does not change the estimated main effects A, B, or C, and not the interaction effect ABC.

— However, AB will change with $2 \cdot h_3 + 2 \cdot h_4 - 2 \cdot h_2$, and we will NOT be able to separate the true AB effect from the block effect.

How to choose which blocks to be used for blocking?

- Idea: try to leave estimates for main effects and low order interaction unchanged by the blocking.
- Note: I=AA=BB=CC, where I is a column of 1's.
- How NOT to do this:
  - Find the blocks for a 2^3 experiment using generators ABC and AC.
  - The interaction between ABC and AC is ABC*AC=B.
  - This means chosing ABC and AC is not a good idea since then we can not trust our estimate of B.

Questions

Should you use a blocking factor in your compulsory project?
Do you understand the difference between blocking and repetition?