Index of Lecture 7: Multifactorial designs

Practical Information

Today's lecture: multifactorial designs:

- 2-way and 3-way ANOVA: balanced (yes/no) & replication (yes/no),
- main effects and interaction.

Notes on for textbook reading (i.e., the GO text):

- Chapters 8–10 on multifactorial analysis: skip too technical parts[1] while focusing on the models and the ANOVA tables,
  - we are not computing SS-values in ANOVA tables manually (only DF-values),
  - you are not expected to compute contrasts involving more than one factor; furthermore, polynomial contrasts can be replaced by regression modelling.

Other news:

- second home assignment due today,
- third home assignment to be posted soon, due March 3,
- need to think about your project (prepare to discuss in Monday's lab, or set up an appointment next week).

[1] Technical parts: 177_8-179^{15}; 179_9-180_5; $181_{11}-181_4$; $183_{10}-184^{11}$; $184_{16}-185^4$; 192_7-194_1; 205_5-208_6; Figure 8.6; Sections 9.2.3, 9.2.4 and 9.3.

Multi-Factorial Designs (Recap)

Several factors in the same design? — Yes!, in good designs it is possible to separate effects of different factors from each other ⇒

- possible to study combined effect of several factors (the presence of interaction),
- if interaction is absent: cheaper (less experimental units) than in several one-at-a-time experiments,[2]
- increased scope of the experiment/study,

and analysing multi-factorial data by each factor separately: is generally wrong and only gives valid results if at most one factor is of importance.

Design terminology and issues:

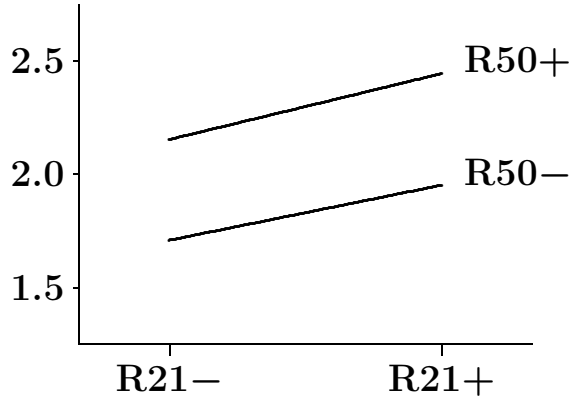
- balancedness: all (combined) groups are equally large, otherwise unbalanced,
- completeness: all (combined) groups are present (no empty cells), otherwise incomplete (should be avoided, or at least analyzed with care),
- replication: some of (combined) groups have $n > 1$, otherwise no replication (all $n = 1$),
- factorial structure can be combined with blocking structures (next lecture).

[2] Simple example: effects of alcohol (A: 0/1) and sleeping pills (B; 0/1). Two "one-at-a-time" studies (effect of A at fixed level of B; effect of B at fixed level of A), each with 20 subjects, give (in the absence of interaction) same precision as a combined study with 5 subjects per alcohol×pill combination.

Decomposing a 2-Way Table of Means I

Example 8.6: Nonstarter bacteria in cheddar cheese:

Total free amino acids		Strain R21 no	Strain R21 yes	Mean
Strain R50	no	1.709	1.952	1.831
	yes	2.153	2.444	2.299
Mean		1.931	2.198	2.065



Different ways to look at the data:

(i) four separate groups,

(ii) two R50 groups for each R21 group, or (iii) two R21 groups for each R50 group,

(iv) (overall level), two R50 groups, two R21 groups, association between R50 and R21 groups.

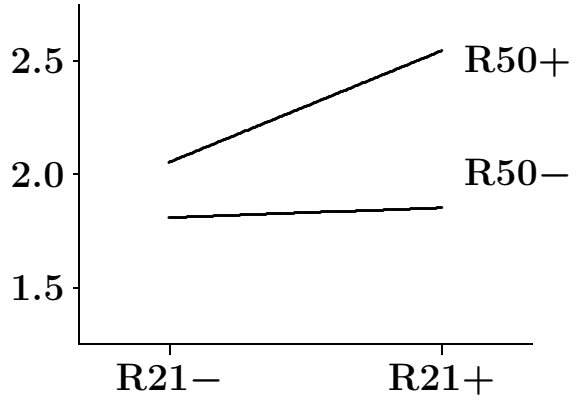
Decomposition of means ∼ (iv):

$\bar{y}_{..}$ overall mean	2.065	2.065	−0.134	0.134	$\bar{y}_{.j} - \bar{y}_{..}$ R21 effect
	2.065	2.065	−0.134	0.134	
$\bar{y}_{i.} - \bar{y}_{..}$ R50 effect	−0.234	−0.234	0.012	−0.012	$\bar{y}_{ij} - \bar{y}_{i.}$ $-\bar{y}_{.j} + \bar{y}_{..}$
	0.234	0.234	−0.012	0.012	

Decomposing a 2-Way Table of Means II

Modified bacteria in cheese data:

Total free amino acids		Strain R21 no	Strain R21 yes	Mean
Strain R50	no	1.809	1.852	1.831
	yes	2.053	2.544	2.299
Mean		1.931	2.198	2.065



Decomposition of means ∼ (iv):

$\bar{y}_{..}$ overall mean	2.065	2.065	−0.134	0.134	$\bar{y}_{.j} - \bar{y}_{..}$ R21 effect
	2.065	2.065	−0.134	0.134	
$\bar{y}_{i.} - \bar{y}_{..}$ R50 effect	−0.234	−0.234	0.112	0.112	$\bar{y}_{ij} - \bar{y}_{i.}$
	0.234	0.234	−0.112	0.112	$-\bar{y}_{.j} + \bar{y}_{..}$

Comparison of two variants of bacteria in cheese data:

- same overall level, same overall (average) effects of both R50 and R21,
- original data: almost parallel lines ⇒ additive effects (same effect of one factor at all levels of other factor),
- modified data: non-parallel lines ⇒ non-additive effects, or interaction between the factors R50 and R21.

Interaction and Additivity (Recap)

Interaction (synergism/antagonism, covariation):

- the combined effect of two factors[3] is not predictable from isolated effects of each of them examined separately,
- the effect of one factor depends on the level of the other factor,
- non-parallel lines in interaction plot (i.e., plot of combined means versus one factor), where parallel lines $\sim$ additive effects (and hence no interaction).
- note: interaction is dependent on scale (of outcome), so affected by transformation.

Dealing with interaction between two factors:

- decomposition of SS for combined factor $A \times B$:

$$\mathrm{SS}_{A\times B} = \mathrm{SS}_A + \mathrm{SS}_B + \mathrm{SS}_{A*B},$$

  and each SS corresponds to treatment contrast(s), as illustrated on next page,
- DF formula: $\mathrm{DF}_{A*B} = \mathrm{DF}_A \cdot \mathrm{DF}_B$ (if $A \times B$ complete),
- in presence of an "important" interaction (significant and strong), the main effects are of no direct interest[4], but contrasts in combined factor may be of interest.

[3] Interaction between 3 factors: the interaction between two of the factors depends on the level of the third factor.

[4] In GO terminology (Section 8.11), the hierarchy is retained by not removing main effects involved in an interaction, so they may not need to be tested at all.

Balanced ANOVA Example: Bacteria in Cheese Data

Example 8.6 (p. 178): Nonstarter bacteria in cheddar cheese, a 2×2-factorial for the bacterial strains R50 and R21:

- SS_{Trt} computed from the four treatments means $(1.709, 1.952, 2.153, 2.444)$,

$$\text{SS}_{\text{Trt}} = 3[(1.709-2.065)^2 + (1.952-2.065)^2 + (2.153-2.065)^2 + (2.444-2.065)^2] = 0.872,$$

- decomposition of SS_{Trt} by orthogonal contrasts:

Contrast $w(\{\mu_i\})$	w_1	w_2	w_3	w_4	$\widehat{w}$	$\text{SS}(\widehat{w})$	$F = t_w^2$
main R50	1	1	−1	−1	−0.935	0.656	7.23
main R21	1	−1	1	−1	−0.535	0.214	2.36
interaction	1	−1	−1	1	0.048	0.002	0.02
						SS_{Trt}	

- ANOVA table:

Source	DF	SS	MS	F	P-value
R50	1	0.656	0.656	7.23	0.028
R21	1	0.214	0.214	2.36	0.16
Interaction	1	0.002	0.002	0.02	0.89
Error	8	0.726	0.091		
Total	11	1.598			

- conclusions:
  - absolutely non-significant interaction between effects of R50 and R21,
  - positive and weakly significant main effect of R50: adding R50 increases TFAA.
  - positive, but non-significant main effect of R21 $\Rightarrow$ no demonstrated R21 effect,
  - the three tests are independent of each other (due to orthogonality of contrasts).

Parametrisation of Interaction (Recap)

Assume a row by columns layout of two factors, and let

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij},$$

where $i \sim$ rows, $j \sim$ columns, and

- μ_{ij} = mean of (i,j)th group,
- μ = overall or baseline mean (or "intercept"),
- α_i = deviation of ith row group from overall mean,
- β_j = deviation of jth column group from overall mean,
- γ_{ij} (or $(\alpha\beta)_{ij}$, GO and commonly used notation) = interaction = deviation of (i,j)th group from additivity.[5]

Technical note: some restrictions on α's, β's and γ's needed (otherwise too many parameters):

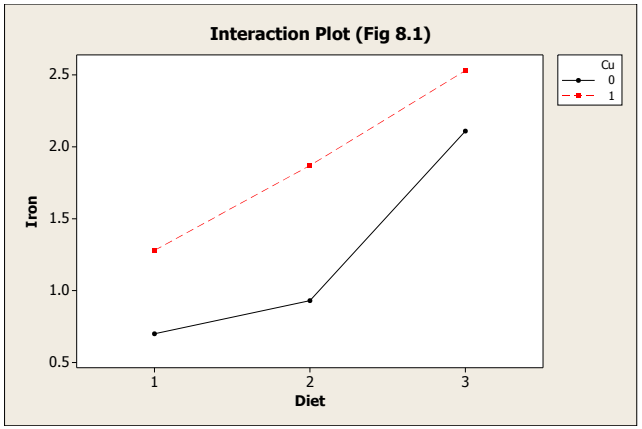
$$\begin{aligned}
\text{GO \& Minitab} &: \textstyle\sum_i \alpha_i = 0,\ \sum_j \beta_j = 0,\ \sum_i \gamma_{ij} = 0,\ \sum_j \gamma_{ij} = 0,\\
\text{Stata} &: \alpha_1 = 0,\ \beta_1 = 0,\ \gamma_{11}, \ldots, \gamma_{1b} = 0,\ \gamma_{11}, \ldots, \gamma_{a1} = 0,\\
\text{SAS} &: \alpha_a = 0,\ \beta_b = 0,\ \gamma_{a1}, \ldots, \gamma_{ab} = 0,\ \gamma_{1b}, \ldots, \gamma_{ab} = 0.
\end{aligned}$$

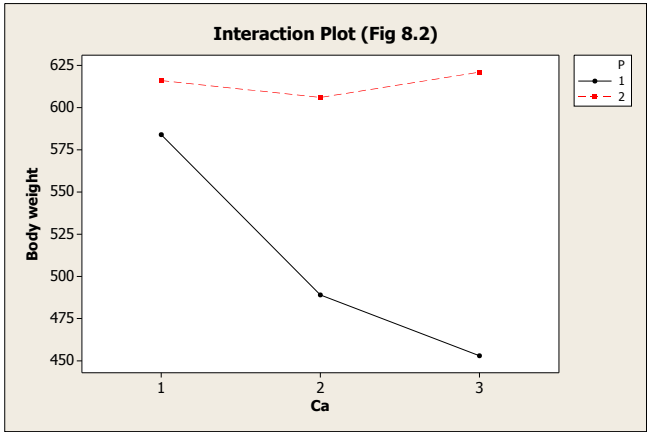
[5] Computer software formalism for factors A and B: Minitab and SAS: $A * B$ = interaction between A and B, and $A \,|\, B = A \;\; B \;\; A * B$ (main effects and interaction); Stata version 12+: # instead of $*$, and ## instead of |; R software: : instead of $*$, and $*$ instead of |.

More Interaction Plot Examples

Rat data (Example 8.2):

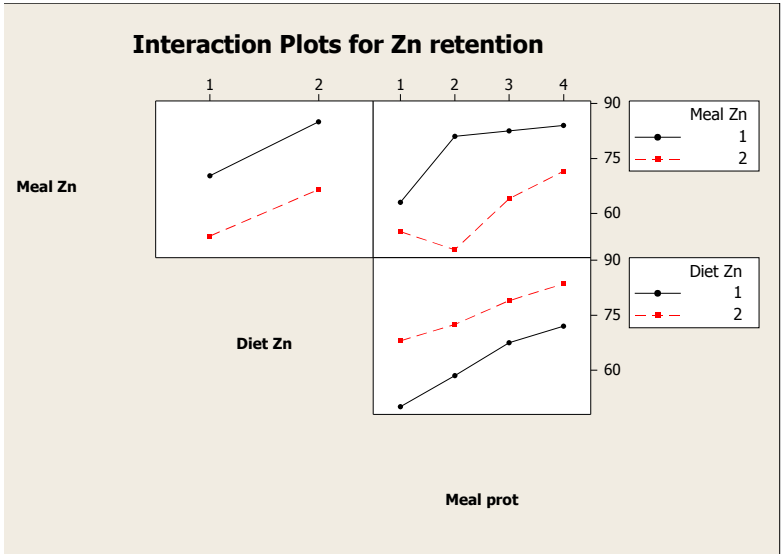
- **outcome**: iron levels in liver tissue,
- **factors**: milk diet (3), copper deficiency (2),
- **interpretation**: close to additive effects.





Chick data (Example 8.4):

- **outcome**: body weights,
- **factors**: Ca supplement (3), P supplement (2),
- **interpretation**: interaction, with Ca effect for low P (P = 1) only.



Rat data II (Example 8.5):

- **outcome**: Zn retention,
- **factors**: Diet Zn (2), final Meal Zn (2), final Meal protein (4),
- **interpretation**: interaction Meal Zn $*$ Meal protein; other effects additive.

Textbook Examples Overview

ANOVA examples of Chapters 8, 9.1-2 and 10.1-2:

- 8.6: non-starter bacteria in cheese: 2×2 factorial with 3 replicates,
- 8.8: page faults (#) in CPU experiment (program algorithms and settings): 2×3×3×3 factorial with no replication (models without replication[6]),
- 8.10, 9.3: amylase activity in maize (analysis and growth temp., variety): 2×2×8 factorial with 3 replicates, both temperatures quantitative (polynomial contrasts[7]),
- 9.2: unspecified outcome and factors: 2^4 $(2 \times 2 \times 2 \times 2)$ factorial with 2 replicates (one-cell interaction[8]),
- 9.4: seed viability (storage conditions): 3×7 factorial with 3 replicates, both factors quantitative (polynomial contrasts[7]),
- 10.1-2: amylase activity (8.10) with one observation omitted (unbalanced data[9]),
- 10.3: unspecified outcome and factors: 2×2 factorial with highly unequal replication (unbalanced data[9]).

[6] Most common method is to omit highest order interaction(s) and thus obtain estimate of residual error: conservative approach if these interactions are non-zero.

[7] For a balanced factor with quantitative (equidistant) values, contrasts ∼ polynomial terms (see also 6L−13).

[8] An interaction may be due to a single mean value deviating from additivity.

[9] Unbalanced data no longer have orthogonal factorial contrasts ⇒ effects depend on other terms in model (as in general regression); two types of sum of squares:

- partial/adjusted SS: remove term while keep all others in;
- sequential SS: remove terms sequentially (↑) so only terms above are kept in.

Three-Way ANOVA with Replication

Amylase activity example (GO Example 8.10, p. 195):

- amylase specific activity of sprouted maize under 32 treatment conditions:

$$\begin{aligned} y_{ijkl} &= \text{activity for maize plant batch } l \text{ of type } (i,j,k) \\ i &= 1,\ldots,8 \sim \text{analysis temperature (10,13,15,20,25,30,35,40°C)}, \\ j &= 1,2 \sim \text{growth temperature (13,25°C)}, \\ k &= 1,2 \sim \text{variety (B73,O43), and } l = 1,2,3 \sim \text{replicate}, \end{aligned}$$

- completely randomized design (if full randomization),
- full statistical model = one-way ANOVA with 32 groups: $y_{ijkl} = \mu_{ijk} + \varepsilon_{ijkl}$, where the errors $\varepsilon_{1111},\ldots,\varepsilon_{8223}$ are i.i.d. and $\sim \mathrm{N}(0,\sigma^2)$.
- decomposition of combined factor levels into main effects and interactions (first order and second order),

$$\mu_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk},$$

  - α_i = main effect of `atemp` (level i),
  - $(\alpha\beta)_{ij}$ = interaction `atemp*gtemp` (levels i and j),
  - $(\alpha\beta\gamma)_{ijk}$ = interaction `atemp*gtemp*var` (levels i, j and k).
- model checks based on full model $\Rightarrow$ same approach as in one-way ANOVA: demonstrating a log-transform to be appropriate.

Amylase Data: Summary of Results

ANOVA table (natural log scale):

Source	DF	Seq SS	Adj SS	Adj MS	F	P
atemp	7	3.01613	3.01613	0.43088	78.86	0.000
gtemp	1	0.00438	0.00438	0.00438	0.80	0.374
var	1	0.58957	0.58957	0.58957	107.91	0.000
atemp*gtemp	7	0.08106	0.08106	0.01158	2.12	0.054
atemp*var	7	0.02758	0.02758	0.00394	0.72	0.654
gtemp*var	1	0.08599	0.08599	0.08599	15.74	0.000
atemp*gtemp*var	7	0.04764	0.04764	0.00681	1.25	0.292
Error	64	0.34967	0.34967	0.00546		
Total	95	4.20202				

- high $R^2 = 1 - 0.3497/4.2020 = 91.7\%$,
- non-significant terms: `atemp*gtemp*var` and `atemp*var`, and `atemp*gtemp` is close to signficant,[10]
- final model: `atemp*gtemp gtemp*var`, but no need to refit model (pool variance terms) because DF_E is large,
- additive effects of `atemp` and `var` for given `gtemp`,
- interaction plots: strong, parabolic-type effect of `atemp` (no obvious interpretation of `gtemp` interaction), and different effects of `gtemp` for the varieties,
- presentation of `gtemp*var` results by means with SE:

var=B73		var=O43		
gtemp=13	gtemp=25	gtemp=13	gtemp=25	SE
5.85	5.92	5.75	5.70	0.015

- `atemp` effects: multiple comparisons or polynomial modelling.

[10] Note that `gtemp` should not be labelled as non-significant, because it is involved in a significant interaction.

Amylase Data: Further Analyses

Polynomial modelling (GO Example 9.3):

- `atemp` polynomial contrasts not attractive due to non-equidistant temperatures,
- for simplicty, refit with (clearly) non-significant `atemp` terms omitted before polynomial modelling,
- quadratic model in `atemp` has lack-of-fit test:
  $$F = [(0.52148-0.42489)/(88-78)]/0.00545 = 1.71 \sim F(10,78) \text{ under } H_0,\ P = 0.094,$$
  − no formal evidence against quadratic model, but higher order terms may improve fit,
- cubic model in `atemp` has lack-of-fit test:
  $$F = [(0.45037-0.42489)/(86-78)]/0.00545 = 0.58 \sim F(8,78) \text{ under } H_0,\ P > 0.5,$$
  − cubic polynomial model seems appropriate,
- interpretation of fitted model(s): plots of predicted curves.

Illustration: effect of unbalancedness (GO Examples 10.1-2): dropping one observation, the first row in dataset:

- sequential and partial/adjusted (SAS type I and III) sum of squares do no longer coincide: SS values for `gtemp` range within 0.00140−0.00330 across different models,
- model building must be sequential (as in regression),
- simple means are no longer same as fitted means ⇒ care needed with margins.

Tools for Factor Effects

Least squares means = tool for interpretation of factor effects:

- in balanced or other "nice" designs: just simple means,
- generally: factor level estimates, while keeping all other predictors at their average:
  - continuous predictors at their sample mean,
  - categorical predictors by averaging across all levels,
- interpretation: estimated level for "average subject" wrt. all other predictors (but often not a "real" subject),
- usually better than simple means because factor levels are compared "all other things being equal"; exception is with strongly correlated predictors (difficult case),
- directly available in Minitab and SAS (in Stata via `margins` command).

The parameter estimates themselves:

- continuous predictors: always look at regression coefficients,
- factors: recall the restrictions (depending on your software program); usually, `lsmeans`/`margins` are sufficient,
- interaction between continuous and factor: separate regression coefficients for different levels of factor (beware of parameter restrictions).

Model Specification

We specify models to document the analysis and to aid the development of good models, by making clear what is taken into account and what is not.

Types of model specifications – illustrated by Amylase example (`atemp`, `gtemp`, `var`):

- fully specified model, multi-index notation:

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \varepsilon_{ijkl},$$

  where $i = 1, \ldots, 8$; $j, k = 1, 2$; $l = 1, 2, 3$,

- fully specified model, single-index notation:

$$y_i = \mu + \alpha_{\texttt{atemp}(i)} + \beta_{\texttt{gtemp}(i)} + \gamma_{\texttt{var}(i)} + (\beta\gamma)_{\texttt{gtemp*var}(i)} + \varepsilon_i,$$

  where $i = 1, \ldots, 96 \sim$ observation number,

- fully specified model, no-index notation:

$$y = \mu + \alpha_{\texttt{atemp}} + \beta_{\texttt{gtemp}} + \gamma_{\texttt{var}} + (\beta\gamma)_{\texttt{gtemp*var}} + \varepsilon,$$

- model formula (software: without +'s and error term):

$$y = \text{atemp} + \text{gtemp} + \text{var} + \text{gtemp*var} + \text{error}.$$

Regression terms / Covariates:

- in specified models: terms like $\beta_1 \cdot \texttt{atemp}_i + \beta_2 \cdot \texttt{atemp}_i^2$,
- in model formulae: special notation (options or boxes in software).