TENTAMEN: Experimental design (MSA250/TMS031)
Onsdagen den 16 mars 2016 kl 8:30 - 12:30
Lärare och jour: Kerstin Wiklander, tel 772 5355
Hjälpmedel: Valfri miniräknare (med tömt minne) och utdelade papper med formler och tabeller. Minicalculator (with emptied memory) and papers with formulas and tables (handed out).

1. (13 p)
a) Suppose you have a 2^3-design with the factors A, B and C and no replicates. Show how the effect from the AC interaction is computed with the data $y_1, ..., y_8$. The indices follow the standard order.
b) The formula $(-y_1 - y_2 + y_3 + y_4 - y_5 - y_6 + y_7 + y_8)/4$ is used to estimate the main effect from one of the factors in the 2^3-design. Give an alternative expression of this formula and an explanation which makes it more intuitive of how to estimate the effect.
c) A 2^{6-2} design has the generating relations E=ABC and F=BCD. Give the resolution of this design and the total alias pattern.
d) An experiment with four treatments (A, B, C and D) will be performed with the Latin Square design below. How can it be randomized?

		Block I			
		1	*2*	*3*	*4*
	1	A	B	C	D
Block	*2*	B	C	D	A
II	*3*	C	D	A	B
	4	D	A	B	C

e) Two factors, each at two levels have been studied. The data was then analyzed in a one-way Anova where the "groups" were defined as the four level settings. The result is presented in the Anova table:

Source	*SS*	*df*	*MS*	*F*	$p - value$
Between groups	332.7	3	110.9	8.599	0.032
Within groups	51.59	4	12.90		
Total	384.5	7			

Which hypothesis is tested here and what conclusion can you draw? Comment also on the choice of statistical method.

2. (18 p)

A variables which can be used to analyze blood samples is the weight of hemoglobin per red blood cell. Participants in a clinical trial were treated with a combination therapy of three different kind of drugs (A, B and C). The aim of this study was to get the variable as close as possible to a target value.

A 2^3 fully randomized factorial design with two runs for each level setting was used in this trial, with the outcome:

A	B	C	y_1	y_2	sample variance
low	low	low	38	34	8
high	low	low	38	40	2
low	high	low	20	21	0.5
high	high	low	29	26	4.5
low	low	high	35	32	4.5
high	low	high	43	46	4.5
low	high	high	24	27	4.5
high	high	high	46	43	4.5

a) Estimate all the effects.

b) Test the effect from factor B and the effect from the AC interaction. State the hypotheses in both cases. Do also an illustration (interaction plot) of the BC interaction.

c) The responses were round to integers. That makes the assumptions for the test method doubtful. Explain why.

d) What settings do you recommend if you have a target value of 27? State this for three different models; the full model, the model without the highest order interaction and the model with only main effects. Explain how you arrived at your recommendations.

3. (6 p)

Five specimen were used to test for a difference between two treatments. The outcome of some response variable is given in the table below.

Treatment A	Treatment B
1	7
4	8
	9

a) Compute the p-value from a randomization test.

b) What assumptions are needed for this test?

c) Is it possible to get a significant result if the significance level (α) is 5%? Is it possible if the alternative hypothesis is one-sided, again with α at 5%? Motivate your answers.

4. (7 p)

There is an interest in the precison for instruments made by two different manufacturer. They are tested using measurements on some variable Y. The assumption is that Y_i from manufacturer $i = 1, 2$ are independent and follow the normal distributions $N(\mu_i, \sigma_i^2)$.

A summary of the descriptive statistics:

	$\bar{y}$	s	s^2	n
Manufacturer 1:	21.1	2.72	7.40	13
Manufacturer 2:	13.8	3.72	13.84	11

a) Calculate a 90% confidence interval for the ratio of the two theoretical variances.

b) Use information only from the confidence interval calculated in a) to test if the two theoretical variances are equal. What is the significance level in this test?

5. (6 p)

Consider a study of two factors of interest; Preparation method (P) and Temperature (T). Since a complete randomization was not possible to do, a split-plot design was used. A batch produced by one of the preparation methods was divided into the same number of samples as the levels of the factor Temperature. The samples were then cooked with the different temperatures, one for each sample. This continued during the day until batches from all preparation methods were made. Thus, in this design, one replicate of the whole experiment was run each day. The factor Day (D) was therefore represented as a block factor. Randomization was only done for the factor Temperature (within the factor Preparation method).

The Anova table from the whole experiment (where d, p and t are the number of factor levels for the factors D, P and T):

$Source$	SS	df	$MS = SS/df$	F
$D : Day$	SS_D	$d-1$	MS_D	
$P : Preparation$	SS_P	$p-1$	MS_P	
DP	SS_{DP}	$(d-1)(p-1)$	MS_{DP}	
$T : Temperature$	SS_T	$t-1$	MS_T	
DT	SS_{DT}	$(d-1)(t-1)$	MS_{DT}	
PT	SS_{PT}	$(p-1)(t-1)$	MS_{PT}	
DPT	SS_{DPT}	$(d-1)(p-1)(t-1)$	MS_{DPT}	

Give expressions for the test statistic (in column F above) used to test effects from the two relevant factors. Give also the distribution for these test statistic.

TENTAMEN: Experimental design, 16 mars 2016
Short solutions

1. a) $(y_1 - y_2 + y_3 - y_4 - y_5 + y_6 - y_7 + y_8)/4$

b) The formula $(-y_1 - y_2 + y_3 + y_4 - y_5 - y_6 + y_7 + y_8)/4$ is used to estimate the effect from factor B. Alternative form: $\frac{y_3+y_4+y_7+y_8}{4} - \frac{y_1+y_2+y_5+y_6}{4}$. This is the average with factor B on its high level minus the average with B on its low lewel. This difference is a measure of the effect from that factor. You can also compute the difference between the values of high level and low level with the other factors fix, for example, $y_3 - y_1$ and then take the mean of all those.

c) I_1=ABCE and I_2=BCDF gives I_3=ADEF. The resolution is IV.
Comment: The question should have been formulated as: "Give the resolution of this design and the total alias pattern *for the main effects*."
l_A: A+BCE+DEF+ABCDF l_B: B+ACE+CDF+ABDEF
l_C: C+ABE+BDF+ACDEF l_D: D+BCF+AEF+ABCDE
l_E: E+ABC+ADF+BCDEF l_F: F+BCD+ADE+ABCEF

d) Randomize the meaning of the symbols A, B, C and D for the treatments.

e) $H_0 : \mu_1 = ... = \mu_4$, where μ_i is the expected value in group no. i. The conclusion is that there is at least one expected value that differs from the others. Thus, there is no information about which one, nor any separate information about the factors. Use instead a model for a 2^2-design, where parameters for the influence from both factors are included. Then, the information would be decomposed with opportunity to draw conscusions for each factor and interaction separately.

2. a) $l_A = 10, l_B = -8.75, l_C = 6.25, l_{AB} = 3, l_{AC} = 5, l_{BC} = 4.75, l_{ABC} = 1$

b) Test of effect from B: $H_0 : \tau_B = 0, H_1 : \tau_B \neq 0$, where τ_B is the effect from factor B. $s_{effect} = 1.0155$ Test statistic for factor B: $\frac{l_B - 0}{s_{effect}} = -8.616$ and it is t-distributed with 8 degrees of freedom. Reject H_0 if the test statistic is $\geq t_{8,0.025} = 2.306$ or $\leq -t_{8,0.025} - 2.306$. Test of effect from the AC-interaction: $H_0 : \tau_{AC} = 0, H_1 : \tau_{AC} \neq 0$, where τ_{AC} is the effect from the AC-interaction. Test statistic: $\frac{l_{AC} - 0}{s_{effect}} = 4.92$.
Reject the null hypothesis in both cases. We have found a significat effect from factor B and a significat effect from the interaction between the factors A and C.

Interaction plot:

c) The assumption of Normal distribution is not so suitable (the responses are

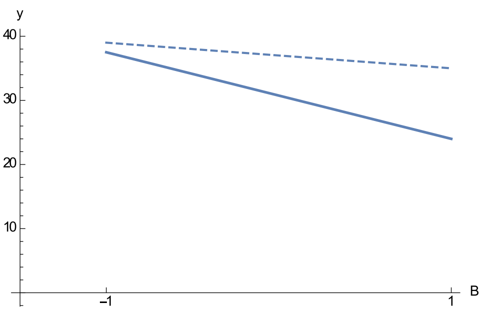

Figure 1: Solid line for C-. Dashed line for C+.

integers) nor is the assumption of constant variances (check the s_i^2).

d) Calculate the predicted values for the three models ($l_M = 33.875$). The best settings for A, B and C are high, high and low for the full model, the same setting for the model without the highest order interaction and low, high, high for the model with only main effects.

3. a) The observed value $\bar{y}_A - \bar{y}_B = -5.5$. Calculate the difference between the means from two A's and three B's from all $\binom{5}{2}$ permutations of the data set. Check the number of cases this difference is $\geq |-5.5|$. This relative frequency is the p-value of the test with $H_0 : E[X_A] = E[X_B]$ and $H_1 : E[X_A] \neq E[X_B]$ (where $E[X]$ is the expected value).

b) Randomization of the treatments.

c) No, there are $\binom{5}{2} = 10$ permutations. So the smallest possible p-value is 1/10.

4. a) The 90% confidence interval for σ_1^2/σ_2^2 is (0.68 , 5.45). The 90% confidence interval for σ_2^2/σ_1^2 is (0.18 , 1.48). (Check the book, page 104.)

b) Since the value one (the value according to $H_0 : \sigma_1^2/\sigma_2^2 = 1$) is included in the confidence interval, the null hypothesis of equal variance can not be rejected. The significance level in this test is 10%.

5. Hint: The whole plot consists of the three first rows in the Anova table. The subplot - the four last rows in the Anova table. Use the highest order interaction within each plot as an error term.

The statistic for P: MS_P/MS_{DP} with a $F_{(p-1),(d-1)(p-1)}$-distribution
The statistic for T: MS_T/MS_{DPT} with a $F_{(t-1),(d-1)(p-1)(t-1)}$-distribution
The statistic for P*T: MS_{PT}/MS_{DPT} with a $F_{(p-1)(t-1),(d-1)(p-1)(t-1)}$-distribution

TENTAMEN: Experimental design (MSA250/TMS031)
Torsdagen den 21 april 2016 kl 8:30 - 12:30
Lärare och jour: Kerstin Wiklander, tel 772 5355
Hjälpmedel: Valfri miniräknare (med tömt minne) och utdelade papper med formler och tabeller. Minicalculator (with emptied memory) and papers with formulas and tables (handed out).

1. (15 p)
a) Data from a replicated factorial design with two factors is given below. Calculate the predicted values and the residuals from the full model.

A	B	response
-1	-1	37
1	-1	38
-1	1	21
1	1	35
-1	-1	33
1	-1	43
-1	1	26
1	1	27

b) Suppose you have done an unreplicated two-level full factorial experiment with four factors. You assume a model where main effects and interaction of order two are active. Show how you can estimate the theoretical variance.

c) You have a 2^{6-3}-design with generators D=AB, E=AC and F=BC. Give the resolution of this design and the total alias pattern for factor E.

d) Suppose you want to do a test where you have a relevant external reference set to your data. What is the minimal p-value that you can get in such a test? Motivate your answer.

e) Suppose you do a permutation test with your data set. What is the minimal p-value that you can get in this test? Motivate your answer.

f) After estimation of the parameters in a model that you assume is suitable, you can calculate the residuals. Give four ways to use these before you continue to work with the statistical inference.

2. (4 p) A replicated 2^3 full factorial design with the three categorical factors A, B and C was analyzed. The total mean was 34.85 and some more results are summarized below:

Effect from	A	B	C	AB	AC	BC	ABC
Estimate	3.350	-10.925	10.325	-10.075	1.025	0.900	0.900
p-value	0.01	< 0.001	< 0.001	< 0.001	0.33	0.39	0.39

The aim was to find a level setting for the factors to minimize the response. The signs of the estimates of the main effects suggest that $-, +, -$ is the optimal choice. Explain why this setting is not the best one. Give also another recommendation with an explanation of your choice.

3. (12 p) A study was conducted of how halibuts are reacting on different temperatures and different pH in the water. After a time in tanks, the relative heart mass was measured (relative to the body mass) aiming to see if the factors had any effect on the heart. The temperatures 6, 12 and 18 degrees Celsius were chosen in combination with acid and normal pH in the water. Five animals per combinations were used in the fully randomized experiment.

The mean values for the combinations:

	6°	12°	18°
Low pH	0.092	0.093	0.071
Normal pH	0.101	0.084	0.075

The incomplete Anova table:

Source of Variation	SS	df	MS	F
Temperature (T)				
pH (P)	1.94×10^{-5}			
Interaction (TxP)	4.54×10^{-4}	2		
Within groups (error)	0.0042			
Total	0.00756	29		

a) Fill in the missing parts in the Anova table. State also the assumptions for using Anova as a test method.
b) Is there evidence of effects from the factors? State all hypotheses and motivate your conclusions.
c) Do an interaction plot.

4. (4 p) Calculation of the sample size (per group) in a two-sample case with a two-sided alternative hypothesis is done with the formula

$$n = 2\frac{(z_{\alpha/2} + z_{\beta})^2\sigma^2}{\Delta^2}$$

Here we assume two independent normally distributed variables with the theoretical variances equal to two. There is an interest to find a difference $\Delta = |\mu_1 - \mu_2|$ of at least 2.5.

a) What sample sizes do you need to find a difference between the expected values that is whortwile. The choice of the power is $1-\beta = 0.95$ and significance level 0.05?

b) If you have resources to take 12 measurements in total, what is the power of the test given that the true difference between the expected values is $\Delta = 2.5$. Select your own value for the significance level.

5. (5 p) The least squares estimators of the parameters in a linear model $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ is expressed as $\mathbf{b} = (X^T X)^{-1} X^T \mathbf{y}$ using matrix notation. To use this formula you need to express X according to your model and design set-up.

a) Give an example of a design matrix X in a regression model with two independent explanatory variables ("regressors") and their interaction.

b) Give an example of a design matrix for a one-way Anova where the grouping variable has three levels.

6. (10 p) In eight randomly selected areas, measurements on a random variable was registered before and after an action for the environment had taken place. The aim was to investigate if the action had any effect, that is, to study if there had been some kind of change.

Area	Response before action	Response after action	Change
1	x_1	y_1	+
2	x_2	y_2	+
3	x_3	y_3	−
4	x_4	y_4	+
5	x_5	y_5	+
5	x_6	y_6	+
7	x_7	y_7	+
8	x_8	y_8	+

With a sign test, you only consider the signs of the differences between the responses, here in the column Change. Then you apply the binomial distribution $\text{bin}(n, p)$ for a variable connected to that information. The probability mass function for such a variable Y is: $P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$ where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ and $n! = n \cdot (n-1) \cdot \ldots \cdot 1$ and $0! = 1$.

a) Do a sign test on these data. State also your null and alternative hypothesis. (Hint: use the parameter in the binomial distribution to express these.) Do not forget to explain all notations you use.

b) Explain short the principles of how a randomization test can be done using the data x_i, y_i.

TENTAMEN: Experimental design, 21 april 2016

Short solutions

1. a) Calculate predicted values as $\hat{y} = l_M \pm l_A/2 \pm l_B/2 \pm l_{AB}/2$ and residuals $y_i - \hat{y}_i, i = 1, \ldots 8$.
b) Use higher order interactions (assumed to be inactive). An estimator of σ^2 is $(l^2_{ABC} + l^2_{ABD} + l^2_{ACD} + l^2_{BCD} + l^2_{ABCD})/5$.
c) $I_1 = ABD, I_2 = ACE, I_3 = BCF, I_4 = I_1I_2 = BCDE, I_5 = ACDF, I_6 = ABEF, I_7 = DEF$. Thus, the resolution is III. Alias for E: l_E estimates the effect $E + AC + DF + BCD + ABF + ABDE + BCEF + ACDEF$.
d) 0 if no value in the reference data is equal to or more extreme than the observed value.
e) 1/N since the observed value is one of the outcomes (N=the number of permutations).
f) For estimation of σ^2. Also, information if the model fits the data: check the form of the model and the distribution, check for constant variance. Illustrate residuals against time order to check for time effect.

2. In the main effect model, you only need to check the signs of the estimate. For minimum: choose minus, plus, minus for A, B and C respectively. Caluculate the predicted values in the model including the significant interaction from AB. The minimum is then found for the setting plus, plus, minus. The impact from the AB-interaction is so large that it explains why the setting for a recommended minimum is influenced by this term. (The same recommendation holds also in models where more or where all interactions are present.)

3. a)

Source of Variation	SS	df	MS	F
Temperature (T)	0.0029	2	0.00145	8.29
pH (P)	1.94 $\times 10^{-5}$	1	1.94 $\times 10^{-5}$	0.111
Interaction (TxP)	4.54 $\times 10^{-4}$	2	2.27 $\times 10^{-4}$	1.297
Within groups (error)	0.0042	24	0.000175	
Total	0.00756	29		

The assumptions are: Independent variables following the normal distribution and with constant variance.

b) Model: $Y_{ijk} = \mu + \tau_i + \pi_j + \omega_{ij} + \epsilon_{ijk}$. For the factor T: $H_0 : \tau_1 = \tau_2 = \tau_3 = 0$, where τ_i is the effect from Temperature. For the factor P: $H_0 : \pi_1 = \pi_2 = 0$, where π_i is the effect from pH. For the interactions TxP: $H_0 : \omega_{ij} = 0$ for all i and j, where ω_{ij} is the interaction effect. The alternative hypothesis in all three cases is that at least one effect is not equal to zero. When the significance level is 5%, the decision rule is: Reject the H_0 for T and for TxP if the test statistic

(F) is $\geq F_{2,24,0.05} = 3.40$ and for P, if the test statistic (F) is $\geq F_{1,24,0.05} = 4.26$. Here, the H_0 for T can be rejected.
c) Interaction plot:

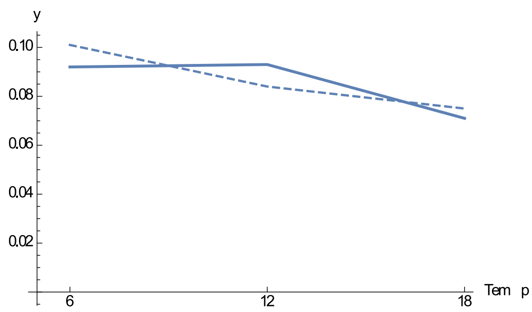

Figure 2: Solid line for low pH. Dashed line for normal pH.

4. a) Calculations give n=8.3. Round to nearest larger integer. You need 9 per group to have a power of at least 95% (when the theoretical variance is known, as here).
b) Calculations give $z_\beta = 1.11$ which gives a power of approximately 86.6%.

5. a)

$$\begin{pmatrix} 1 & x_{11} & x_{21} & x_{11}x_{21} \\ 1 & x_{12} & x_{22} & x_{12}x_{22} \\ 1 & x_{13} & x_{23} & x_{13}x_{23} \\ 1 & x_{14} & x_{24} & x_{14}x_{24} \end{pmatrix} \text{ A numerical example: } \begin{pmatrix} 1 & 2 & 10 & 20 \\ 1 & 3 & 10 & 30 \\ 1 & 2 & 40 & 80 \\ 1 & 3 & 40 & 120 \end{pmatrix}$$

b) Use "dummy variables". Representations with sample size two for each "group" in the model $Y_{ij} = \mu_i + \epsilon_{ij}$ (to the left) and $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$ (to the right):

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

6. a) Let Y=no. of minus (no. of negative change). Then Y is bin(8,p). No change means that there is equal probability of getting a $+$ as getting a $-$, which means that p=1/2. Then, $H_0 : p = 1/2, H_1 : p \neq 1/2$. Calculate half the p-value (since the alternative hypothesis is two-sided) by $P(Y \leq 1) = P(Y = 0) + P(Y = 1) = \binom{8}{0}0.5^0 0.5^8 + \binom{8}{1}0.5^1 0.5^7 = [\binom{8}{0} + \binom{8}{1}]0.5^8 = 9/256 = 0.0352$. The p-value is then 0.07 and a change can not be stated at the 5% significance level.

b) Keep the pairs together. Possible values for the difference in Area i (pair no. i) is then $+d_i$ and $-d_i$. Calculate all 2^8 combinations of them and take the mean of the differences for each combination.

Calculate the p-value; the relative frequency of occurencies in this set of mean differences which are at least as extreme as the outcome in the experiment.