WAGENINGEN UNIVERSITY **Biometris, ASN 1202**

Examination Statistics (MAT-24306)
Wednesday, February 1, 2012; 14:00 – 17:00 h

											0

program registration number

Name and first name

You can bring the book by Ott & Longnecker, the Lecture Notes, one A4 (two-sided) with a hand written summary, and a calculator, but keep e.g. printed PowerPoint sheets in your bag. There are five questions with credits 22, 14, 20, 16 and 18 (with 10 bonus credits this makes a total of 100). Below, credits (per part of an exercise) are in parentheses.
You are free to answer the questions in English or in Dutch (or both).

Put your answers in the empty spaces indicated in this exam. If you use extra paper, put your name on the extra pages, and refer to these extra pages in this exam.

Question 1 (22) Predicting the lean meat percentage of a pig carcass

In slaughterhouses within the European community the percentage of lean meat of each pig carcass is predicted in the slaughter line. In this study, with a probe, at a particular position on a carcass, the thickness of the fat layer (variable x_1 in mm) and the thickness of the muscle underneath (variable x_2 in mm) are determined. Variables x_1 and x_2 will be used to predict the percentage lean meat (variable y in %) in the carcass.

To determine a prediction formula, a random sample of 120 carcasses is collected from a national pig population. For each carcass, variables x_1 and x_2 are measured with the probe. Subsequently, each carcass is dissected, and the percentage lean meat y is determined. Part of the data is reproduced on the right.

Carcass nr.	y	x_1	x_2
1	57.93	17.61	76.32
2	61.79	13.24	70.26
3	60.16	15.38	80.59
4	60.20	13.22	67.76
5	62.12	13.03	69.35
6	62.47	13.69	65.77
7	63.01	16.83	65.14
8	62.11	15.01	71.17
9	65.89	11.56	63.30
…	…	…	…
116	59.29	13.30	60.05
117	53.70	15.82	62.16
118	57.69	13.72	57.24
119	58.95	13.80	68.64
120	60.66	13.41	61.16

The researcher is going to use linear regression:

$$\underline{y}_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \underline{\epsilon}_i,$$

where the residual (error) terms $\underline{\epsilon}_i$ are assumed to be independently Normally distributed around 0 with constant variance σ^2 (i = 1…120).

a1(2) Explain why the usual interpretation of β_0 is not practically meaningful in this study.

Explanation:

According to the usual interpretation, β_0 is the expected percentage lean for a pig carcass with a fat and muscle thickness of 0 mm. From a practical point of view, this makes little sense.

a2(2) Give the practical interpretation of coefficient β_1 in this study.

Interpretation:

When we keep the muscle thickness x_2 at the same value, and we increase the fat thickness x_1 by 1 mm, the expected percentage lean increases by β_1 %.

[*1 point off when keeping muscle thickness at the saem value is not mentioned*]

For the next two questions, you can use the SPSS output below.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	64,931	1,813		35,811	,000
	x1	-,552	,106	-,525	-5,231	,000
	x2	,061	,032	,193	1,921	,057

a. Dependent Variable: y

b1(1) Give the expression for the prediction formula that can be used in a slaughterhouse.

Expression: $\hat{y} = 64.931 - 0.552x_1 + 0.061x_2$.

b2(1) Derive the predicted percentage lean meat ($\hat{y}$) for a carcass with a fat thickness of 14 mm and a muscle thickness of 70 mm.

Predicted percentage: 64.931 - 0.552 * 14 + 0.061 * 70 = 61.5 %.

Use the output below to answer the next questions.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,445[a]	,198	,185	2,203

a. Predictors: (Constant), x2, x1

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	140,482	2	70,241	14,477	,000[a]
	Residual	567,684	117	4,852		
	Total	708,167	119			

a. Predictors: (Constant), x2, x1

b. Dependent Variable: y

b3(2) Show the calculation of $R^2 = 0.198$ from the ANOVA table.

Calculation:

$R^2 = SSR / TSS$ = 140.482 / 708.167 = 0.198.

[*1 point for adjusted mean square*]

b4(1) Give an estimate for the residual standard deviation σ.

Estimate of σ = 2.203 (from 1st table) or $\sqrt{4.852}$ = 2.203 (from 2nd table).

b5(1) Use the estimate from part b4 (take value 2, if you could not find the answer) to indicate (roughly) how large the difference can be, due to variation between carcasses, between the true percentage (y) of a carcass, with a fat thickness of 14 mm and a muscle thickness of 70 mm, and the corresponding prediction ($\hat{y}$) from part b2.

Rough indication of difference:

Roughly 1.96 (or rounded 2) times the residual standard deviation σ, *so 2 * 2.203 = 4.41 %.*

c1(4) Test, with a single test statistic, whether fat and / or muscle thickness (one of the two or both) have any predictive value for the percentage lean. Give:
(1) the null and alternative hypothesis in terms of the model parameters,
(2) the value of the test statistic,
(3) the distribution of the test statistic under the null hypothesis,
(4) the P-value, and your conclusion (α = 0.05) in words.

Null and alternative hypothesis:

$$H_0: \beta_1 = \beta_2 = 0, \;\; H_a: \beta_1 \neq 0 \text{ or } \beta_2 \neq 0 \text{ (or both)}$$

Value of test statistic: F = 14.477 (2^{nd} *table*)

*Distribution: F distribution, with degrees of freedom df*1 = 2, *df*2 = 117.

P-value: 0.000

Conclusion:

P-value = 0.000 < 0.05, *so the null hypothesis is rejected, and the alternative hypothesis is proven. We have shown that either fat or muscle thickness or both have predictive value for the percentage lean meat in a carcass.*

c2 (3) Negative correlation between the lean meat percentage and muscle thickness of carcasses is unlikely. With this in mind, test (α = 0.05) whether muscle thickness adds significantly to the model (has any predictive value for the percentage lean). Give:
(1) the null and alternative hypothesis in terms of the model parameter(s),
(2) the distribution of the test statistic under the null hypothesis,
(3) the P-value, and your conclusion in words.

Null and alternative hypothesis: $H_0: \beta_2 = 0, \;\; H_a: \beta_2 > 0.$

[*1 point off for two-sided alternative*]

Distribution: t-distribution, df = 117.

P-value: 0.057 / 2 = 0.028 (from table on page 2) .

Conclusion:

P-value = 0.028 < 0.05, *so the null hypothesis is rejected, and the alternative hypothesis is proven. We have shown that muscle thickness adds significantly to the model, i.e. has predictive value for the percentage lean meat in a carcass.*

Carcasses are of female and male pigs. The researcher wonders if gender matters for prediction. He introduces three extra explanatory variables in the model: dummy variable x_3 that takes value 1 for females and 0 for males, and products $x_4 = x_1 * x_3$ and $x_5 = x_2 * x_3$.

d1 (2) Explain, in words, what changes this new model brings about with respect to the intercept and the coefficients of x_1 and x_2.

Answer: *Either of the next two is enough for two points.*

With the extra variables, the model now has a separate intercept and separate coefficients for fat an muscle thickness for the two sexes.

The coefficients of x_1 and x_2 are now the *coefficients for fat and muscle thickness for males.*

Additional SPSS output with the model with all five variables $x_1 \ldots x_5$ is provided below.

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	302,651	5	60,530	17,016	,000[a]
	Residual	405,516	114	3,557		
	Total	708,167	119			

a. Predictors: (Constant), x5, x1, x2, x4, x3

b. Dependent Variable: y

d2 (3) The researcher wants to test for any gender effect at all, with a single test. Give:
(1) the outcome of the test statistic (you need to do some calculations by hand, provide these calculations as well),
(2) the distribution of the test statistic when there is no effect of gender (including degrees of freedom).

Outcome of the test statistic (with calculations):

Use the comparison between the full ($x_1 \ldots x_5$) and reduced (x_1 and x_2) model:

$$F = \frac{(567.684 - 405.516)/(117 - 114)}{3.557} = 15.2.$$

Distribution (and degrees of freedom):

F distribution with $df1 = 3$, $df2 = 114$.

Question 2 (14) Salt in bread

This exercise is loosely based on a real experiment. For a month, 116 students had their breakfast in the Restaurant of the Future at Wageningen University campus. Unknown to them, they were randomly divided into three groups receiving either normal bread, or bread with gradually lowered salt, or bread with gradually lowered salt and more food flavourings. At the end of the month, students were given to taste low salt bread and normal bread, without being told about the difference, and were asked which bread they preferred. The (fictitious) data are shown in the table on the right.

Type of breakfast \ Preference at the end	Low salt bread	Normal bread	Total
Normal	16	24	40
Low salt	22	16	38
Low salt & flavourings	25	13	38
Total	63	53	116

First, the researcher want to know whether the three types of breakfast (rows), all have the same effect upon the preference at the end (columns) or not.

a (1) Will this be a test for homogeneity or independence?

Answer:

Since the row totals are fixed beforehand (students are divided into three groups) this is a test for homogeneity.

Let π_1, π_2 and π_3 be the probabilities for a (random) student to prefer low salt bread at the end after the normal, low salt, or low salt & flavourings breakfast, respectively.

b1(2) Formulate the relevant null and alternative hypotheses H_0 and H_a, in terms of π_1, π_2, π_3.

Null hypothesis: $H_0: \pi_1 = \pi_2 = \pi_3.$

Alternative hypothesis: $H_a: \pi_i \neq \pi_j, \textit{for at least one pair } i \neq j.$

b2(2) Give an estimate for π_1 and the corresponding standard error.

Estimate for π_1: 16 / 40 = 0.4.

Corresponding standard error: $\sqrt{0.4 * 0.6 / 40}$ =0.077.

b3(2) For group 1, calculate the expected number of students E_{11} that prefer low salt bread at the end and the expected number of students E_{12} that prefer normal bread at the end, assuming that the null hypothesis is true.

Expected number E_{11}: 63 * 40 / 116 = 21.72.

Expected number E_{12}: 53 * 40 / 116 = 18.28.

For the next questions, you can use this SPSS output:

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5.516[a]	?	.063
Likelihood Ratio	5.554	?	.062
Linear-by-Linear Association	5.213	?	.022
N of Valid Cases	116		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 17.36.

c1(1) Give the name of the test that you would use, and the outcome of the test statistic.

Name of test: Pearson's chi-square test (for homogeneity).

Outcome of test statistic: 5.516.

c2(1) Give the approximate (or asymptotic) distribution of the test statistic under H_0.

Distribution (include degrees of freedom if any):

Chi-square distribution with df = (3 - 1) * (2 - 1) = 2.

c4(2) What is the P-value, and what is your conclusion ($\alpha = 0.05$) in words?

P-value: 0.063.

Conclusion:

The null hypothesis is not rejected. We cannot show that there are differences among the three types of breakfast in the way they affect the final choice of students.
[*1 point off for conclusion that is has been shown that that probabilities are the same… etc.*]
Second, the researchers focus upon the two types of breakfast with low salt, to see whether adding more food flavourings makes a difference. They want to perform an **exact** test to compare probabilities π_2 and π_3 only.

d1(2) Fill in the 2 x 2 table they should look at (numbers in the cells, marginal totals, and an explanatory text for the rows and columns of the table):

	Text: … low salt bread	*Text: … normal bread*	*Row totals*
Text: …low salt	*…22*	*…16*	*…38*
Text: …low salt & flavourings	*…25*	*…13*	*…38*
Column totals	*…47*	*…29*	*…76*

d2(1) What is the name of the exact test that you would advise the researchers to use?

Name of test: Fisher's exact test.

Question 3 (20) Carrots and sowing rate

Below you see a field that is divided into 24 plots. An experiment is going to be conducted involving the effects of four sowing rates (1.5, 2, 2.5 or 3 lb per acre) on yield of carrots for two stocks (batch 1 or batch 2). The field is along the slope of a hill, and the downward direction is indicated by the arrow below the scheme: plots on the left are located higher upon the hill than plots on the right. Height on the hill makes a difference for the yield of a plot.

a(4) Complete the scheme for a randomized complete block design (RCBD) with three blocks. Indicate the blocks, by different shading or otherwise. Put letters A, B, C, D, E, F, G, H, representing the eight combinations of sowing rates and stocks, in the plots. Explain your choice of blocks, and how A … H should be assigned to the plots.

A	B	D
D	D	B
E	G	H
B	C	A
C	F	G
G	E	C
F	H	F
H	A	E

high ⟶ low

Brief explanation of the choice of blocks and way treatments are assigned to plots:

Blocks are shaded and involve plots at the same height. Treatments (A … H) are randomly assigned to plots within each block.

[*2 points for scheme, 2 points for explanation*]

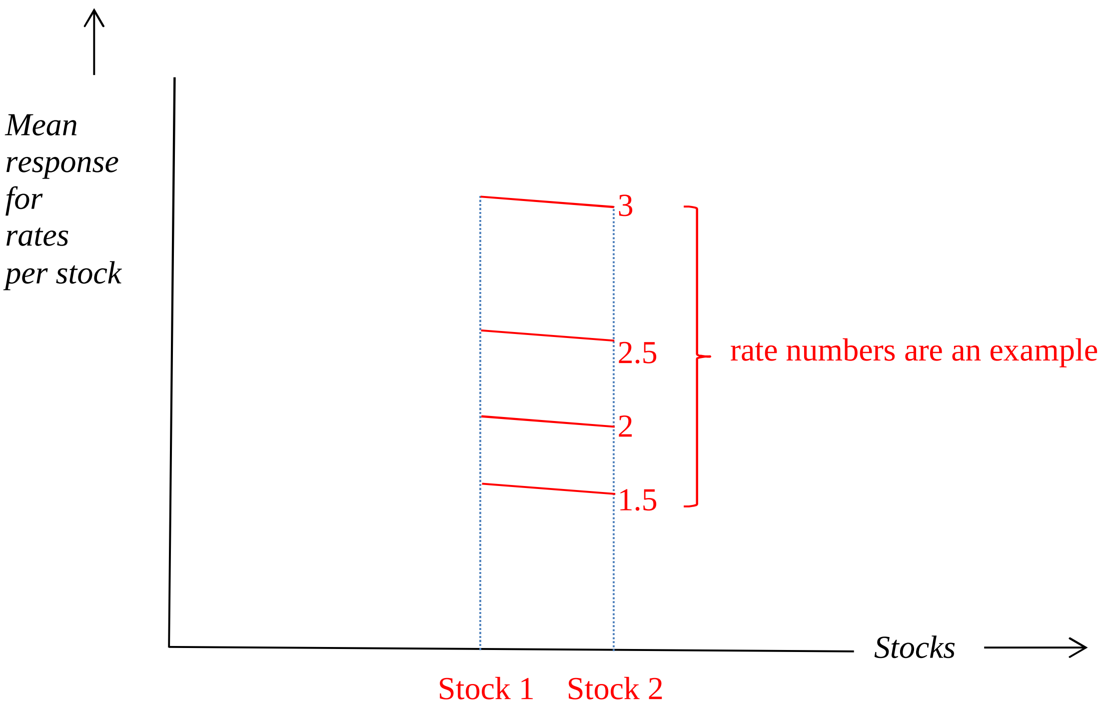
The linear model that is used for the analysis comprises main effects for blocks, sowing rates and stocks, and interaction between sowing rates and stocks.

b1(2) Explain what interaction between sowing rates and stocks would imply, in terms of this experiment, for the differences among sowing rates.

Explanation interaction:

It would imply that the difference in expected yield between at least two (maybe more) of the sowing rates is not the same for the two stocks.

b2(2) Draw a profile plot that would suggest that there is no interaction. Do not worry about the actual data yet. Just draw a typical plot and explain the feature(s) you wanted to bring out in the plot.



Explanation of the plot:

The four lines should be parallel when there is no interaction, i.e. distance between expected yield for any two sowing rates is the same for each stock.

[*Although this was not the intention: also full points for sowing rates along the horizontal and means for stocks along the vertical axis, with the explanation that lines should be parallel.*]

Below you find SPSS output of an analysis of variance of data of the experiment and some descriptive statistics.

Tests of Between-Subjects Effects

Dependent Variable:yield

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	23.988[a]	9	2.665	7.779	.000
Intercept	451.360	1	451.360	1317.373	.000
sowing_rate	13.750	3	4.583	13.377	.000
block	.357	2	.179	.522	.605
stock	1.242	1	1.242	3.625	.078
sowing_rate * stock	8.639	3	2.880	8.405	.002
Error	4.797	14	.343		
Total	480.145	24			
Corrected Total	28.785	23			

a. R Squared = .833 (Adjusted R Squared = .726)

Descriptive Statistics

sowing_rate	stock		N	Minimum	Maximum	Mean	Std. Deviation
1.5	1	yield	3	4.20	4.94	4.5300	.37643
		Valid N (listwise)	3				
	2	yield	3	2.82	3.80	3.2533	.49973
		Valid N (listwise)	3				
2.0	1	yield	3	3.50	4.36	4.0100	.45177
		Valid N (listwise)	3				
	2	yield	3	1.92	3.43	2.6967	.75593
		Valid N (listwise)	3				
2.5	1	yield	3	4.55	5.75	5.2333	.61712
		Valid N (listwise)	3				
	2	yield	3	3.90	4.82	4.4067	.46705
		Valid N (listwise)	3				
3.0	1	yield	3	3.90	5.15	4.4833	.62915
		Valid N (listwise)	3				
	2	yield	3	5.35	6.57	6.0800	.64444
		Valid N (listwise)	3				

c (4) Test whether there is interaction between sowing rates and stocks ($\alpha = 0.05$). Give:
(1) the expression of the test statistic in terms of mean squares,
(2) the outcome of the test statistic,
(3) the distribution of the test statistic under the null hypothesis of no interaction,
(4) the P-value and your conclusion ($\alpha = 0.05$) in words.

Expression test statistic: $F = MS_{sowing\ rate} / MSE$.

Outcome test statistic: $F = 8.405$.

Distribution: *F-distribution,* $df1 = 3$, $df2 = 14$.

P-value: 0.002.

Conclusion:

P-value = 0.002 < 0.05, so the null hypothesis is rejected and the alternative hypothesis is proven. We have shown that there is interaction between sowing rates and stocks with respect to expected yield of carrots.

Suppose that the test from part c was significant. The next step would be to make pairwise comparisons among the eight treatments. SPSS output on the next page shows the results with Fisher's LSD method. To save space, redundant information is deleted (indicated by ...), e.g. only A versus B is included, and B versus A is deleted.

Multiple Comparisons

Yield
LSD

(I) sowingrate+stock	(J) sowingrate+stock	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
A[1]	B	1.2767*	.47793	.018	.2516	2.3017
	C	.5200	.47793	.295	-.5051	1.5451
	D	1.8333*	.47793	.002	.8083	2.8584
	E	-.7033	.47793	.163	-1.7284	.3217
	F	.1233	.47793	.800	-.9017	1.1484
	G	.0467	.47793	.924	-.9784	1.0717
	H	-1.5500*	.47793	.006	-2.5751	-.5249
B	...	...	...	...	...	...
	C	-.7567	.47793	.136	-1.7817	.2684
	D	.5567	.47793	.264	-.4684	1.5817
	E	-1.9800*	.47793	.001	-3.0051	-.9549
	F	-1.1533*	.47793	.030	-2.1784	-.1283
	G	-1.2300*	.47793	.022	-2.2551	-.2049
	H	-2.8267*	.47793	.000	-3.8517	-1.8016
C	...	...	...	...	...	...
	D	1.3133*	.47793	.016	.2883	2.3384
	E	-1.2233*	.47793	.023	-2.2484	-.1983
	F	-.3967	.47793	.420	-1.4217	.6284
	G	-.4733	.47793	.339	-1.4984	.5517
	H	-2.0700*	.47793	.001	-3.0951	-1.0449
D	...	...	...	...	...	...
	E	-2.5367*	.47793	.000	-3.5617	-1.5116
	F	-1.7100*	.47793	.003	-2.7351	-.6849
	G	-1.7867*	.47793	.002	-2.8117	-.7616
	H	-3.3833*	.47793	.000	-4.4084	-2.3583
E	...	...	...	...	...	...
	F	.8267	.47793	.106	-.1984	1.8517
	G	.7500	.47793	.139	-.2751	1.7751
	H	-.8467	.47793	.098	-1.8717	.1784
F	...	...	...	....	...	...
	G	-.0767	.47793	.875	-1.1017	.9484
	H	-1.6733*	.47793	.004	-2.6984	-.6483
G	...	...	...	...	...	...
	H	-1.5967*	.47793	.005	-2.6217	-.5716

Based on observed means.
The error term is Mean Square(Error) = .343.
*. The mean difference is significant at the 0.05 level.

[1] Explanation of the letters:

Combination sowing rate+stock	Sowing rate	Stock
A	1.5	1
B	1.5	2
C	2.0	1
D	2.0	2
E	2.5	1
F	2.5	2
G	3.0	1
H	3.0	2

d (4) Below treatment means (from the descriptive statistics) are presented from low to high. Summarize the pairwise comparisons with the underline notation (giving treatments that do not significantly differ a common underline).

D	B	C	F	G	A	E	H
2.6967	3.2533	4.0100	4.4067	4.4833	4.5300	5.2333	6.0800

The researcher is anxious that the overall experiment wise error rate of the comparisons among the sowing rates is not above 0.05. She has been given to understand that Fisher's protected LSD is not a full proof method in this respect, so she wants to use another method. Moreover, she is not interested in comparing treatment means that correspond to different stocks. She only wants to compare the four sowing rates within each stock of seed.

e1(2) How many pairwise comparisons does she have to perform?

Number of pairwise comparisons: (4 * 3 / 2) + (4 * 3 / 2) = 6 + 6 = 12

e2(2) Suggest one of the methods for performing multiple comparisons that we discussed in class, and explain why you think this would be an appropriate choice.

Method for pairwise comparison (including explanation):

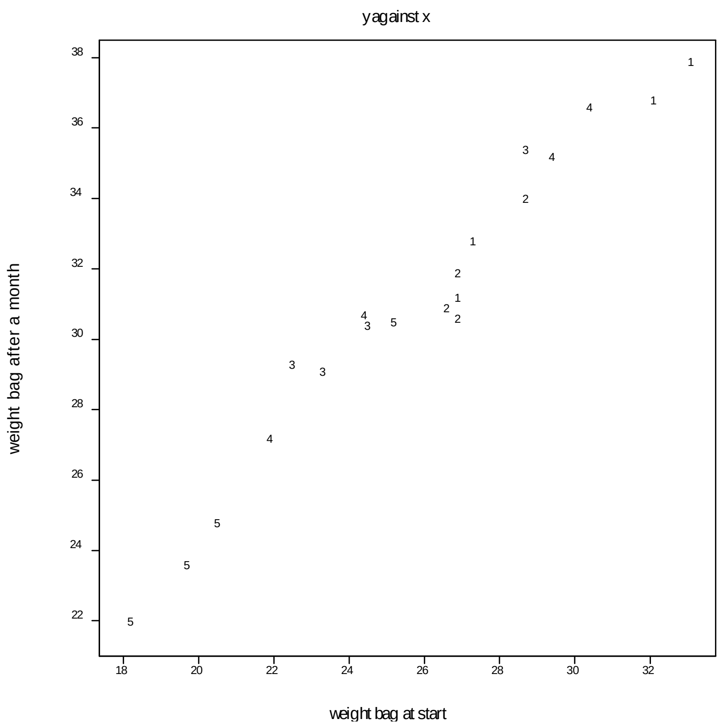
Bonferoni is the most likely choice.

Tukey is made for all pairwise comparisons (28 of them) and will be too severe for 12 comparisons only.

[*However, if you argued that you cannot be sure that Bonferroni is better, because it uses an inequality, and therefore you suggested Tukey instead, you get full points. No points for Dunnett, since interest is not restricted to comparisons with a control.*]

Question 4 (16) Growth of oysters

A researcher wants to determine whether growth of oysters is affected by exposure to artificially heated water and also by the depth of the oysters in the water. Twenty bags of 10 oysters each are randomly assigned to five locations within the cooling water runoff of a power plant. These locations represent five treatments: location 1 = cool / bottom, 2 = cool / surface, 3 = hot / bottom, 4 = hot / surface, 5 = control (mid temperature / mid depth). The oysters in the bags are cleaned and weighted at the beginning (variable *x*) and, a month later, at the end (variable *y*) of the experiment. Below, weights *x* and *y* are shown for each of the twenty bags, and plotted, with symbols 1…5 for the locations.



Location	x	y
1	27.2	32.6
1	32.0	36.6
1	33.0	37.7
1	26.8	31.0
2	28.6	33.8
2	26.8	31.7
2	26.5	30.7
2	26.8	30.4
3	28.6	35.2
3	22.4	29.1
3	23.2	28.9
3	24.4	30.2
4	29.3	35.0
4	21.8	27.0
4	30.3	36.4
4	24.3	30.5
5	20.4	24.6
5	19.6	23.4
5	25.1	30.3
5	18.1	21.8

Initially, a linear model was fitted, with response y, that comprised factor Location at five levels, covariate x, and interaction between Location and x:

$$\underline{y}_{ij} = \beta_0 + \tau_i + \beta_1 x_{ij} + \tau\beta_{i1} x_{ij} + \underline{\varepsilon}_{ij}.$$

The τ_i's are the main effects for the locations, and the $\tau\beta_{i1}$'s are the interactions, $i = 1\ldots5$. Bags at the same location are numbered $j = 1\ldots4$. The fifth location (the control) is the reference: $\tau_5 = 0$ and $\tau\beta_{51} = 0$. The residual (error) terms $\underline{\varepsilon}_{ij}$ are assumed to be independently normally distributed around 0 with constant variance σ^2.

The systematic part of the model, inspired by the plot, consists of five different lines, a separate line for each location, for the relationship between the expected value of y and x.

a (2) What are the intercept and slope (coefficient of *x*), in terms of the parameters of the model, for the lines corresponding to locations 2 and 5?

Intercept location 2: $\beta_0 + \tau_2$.

Slope location 2: $\beta_1 + \tau\beta_{21}$.

Intercept location 5: $\beta_0 + \tau_5 = \beta_0$.

Slope location 5: $\beta_1 + \tau\beta_{51} = \beta_1$.

[also full points when $\tau_5 = \tau\beta_{51} = 0$ is not used]

The researcher primarily wants to compare the five locations with respect to their effect on growth of the oysters. She is very keen on all lines having the same slope, i.e. being parallel lines, and would like to simplify the model accordingly.

b1(2) Explain why parallel lines simplify the comparison between the locations.

Explanation:

The distance between any two lines in the vertical direction represents the expected difference in weight after a month between the corresponding locations. The distance, and therefore the expected difference will be the same, irrespective of the value of *x*, when the lines are parallel. This simplifies matters enormously.

b2(1) Formulate the null hypothesis, in terms of the parameters of the model, that would correspond to the five lines being parallel.

Null hypothesis (*in terms of the model parameters*):

$H_0: \tau\beta_{i1} = 0 \; for \; all \; i = 1, 2, 3, 4, 5.$

Some output of SPSS is shown on the next page.

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	355.835[a]	9	39.537	139.510	.000
Intercept	1.719	1	1.719	6.065	.034
Location	1.696	4	.424	1.496	.275
x	156.040	1	156.040	550.599	.000
Location * x	1.388	4	.347	1.225	.360
Error	2.834	10	.283		
Total	19386.950	20			
Corrected Total	358.670	19			

a. R Squared = .992 (Adjusted R Squared = .985)

c (3) With the output, test (α = 0.05) the hypothesis that the five lines are parallel. Give:
(1) the outcome of the test statistic,
(2) the distribution of the test statistic under the hypothesis that the lines are parallel,
(3) the P-value of the test, and your conclusion in words (α = 0.05).

Outcome of test statistic: F = 1.225 (F test for interaction Location * x).

Distribution of test statistic: F distribution with $df1 = 4$ *and* $df2 = 10$.

P-value: 0.360.

Conclusion:

P-value = 0.360 > 0.05, the null hypothesis is not rejected, *we have not been able to show that the slopes of the five lines for the five locations are different.*

The research worker now proceeds with the simplified model with parallel lines:

$$\underline{y}_{ij} = \beta_0 + \tau_i + \beta_1 x_{ij} + \underline{\varepsilon}_{ij},$$

with the same assumptions about the $\underline{\varepsilon}_{ij}$'s as before. This is an example of analysis of covariance (ANCOVA).

The output on the next page includes the adjusted treatment means ("estimated marginal means" in SPSS) for the locations after analysis of covariance. The output also includes the sample means of the weight of the bags at the start (x) and after a month (y) for the locations.

Take care: there is also output for the model without variable x!

Descriptive Statistics

Location		N	Mean	Std. Deviation
1	x	4	29.7500	3.20572
	y	4	34.4750	3.18891
2	x	4	27.1750	.96047
	y	4	31.6500	1.53731
3	x	4	24.6500	2.75862
	y	4	30.8500	2.95579
4	x	4	26.4250	4.04918
	y	4	32.2250	4.29758
5	x	4	20.8000	3.02104
	y	4	25.0250	3.69899

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	354.447[a]	5	70.889	235.049	.000
Intercept	6.466	1	6.466	21.438	.000
Location	12.089	4	3.022	10.021	.000
x	156.040	1	156.040	517.384	.000
Error	4.222	14	.302		
Total	19386.950	20			
Corrected Total	358.670	19			

a. R Squared = .988 (Adjusted R Squared = .984)

Parameter Estimates

Dependent Variable: y

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	2.495	1.028	2.427	.029	.290	4.699
[Location=1]	-.244	.577	-.424	.678	-1.481	.992
[Location=2]	-.280	.493	-.569	.579	-1.337	.777
[Location=3]	1.655	.429	3.853	.002	.734	2.576
[Location=4]	1.107	.472	2.347	.034	.095	2.119
[Location=5]	0[a]	.	.	.	.	.
x	1.083	.048	22.746	.000	.981	1.185

a. This parameter is set to zero because it is redundant.

Estimated Marginal Means

Location

Dependent Variable: y

Location	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	30.153[a]	.334	29.437	30.869
2	30.117[a]	.283	29.511	30.724
3	32.052[a]	.280	31.453	32.652
4	31.505[a]	.276	30.912	32.098
5	30.398[a]	.362	29.621	31.174

a. Covariates appearing in the model are evaluated at the following values: x = 25.7600.

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	198.407[a]	4	49.602	4.643	.012
Intercept	19028.280	1	19028.280	1780.979	.000
Location	198.407	4	49.602	4.643	.012
Error	160.263	15	10.684		
Total	19386.950	20			
Corrected Total	358.670	19			

a. R Squared = .553 (Adjusted R Squared = .434) d1 (1) Why is the adjusted mean (estimated marginal mean in SPSS) lower than the sample mean for the weight after a month for location 1?

Explanation:

Because the mean starting weight of the oysters for location 1 was 29.75, which is higher than the average starting weight 25.76 (below table with expected marginal means).
So, location 1 started with an advantage, therefore the adjusted mean will be smaller than the sample mean for weight y after a month.

d2 (2) Provide evidence from a test that suggests that it is worthwhile to include x as a covariate, take into consideration whether the alternative hypothesis (and the P-value) should be one- or two-sided.

Answer:

This is the test for the coefficient of variable x. We do not expect a negative coefficient. So, we need a one-sided P-value. Since the estimated slope (1.083) is positive, we can take half of the P-value from SPSS: 0.000 / 2 = 0.000. This P-value is below 0.05, so we have shown that x is worthwhile in the sense that it has predictive value for final weight and therefore is able to reduce the residual (error) variance.
[*1 point for two-sided P-value*]

e1 (2) Test ($\alpha = 0.05$) whether there are differences among the locations. Give:
(1) the outcome of the test statistic,
(2) the P-value of the test, and your conclusion ($\alpha = 0.05$) in words.

Outcome test statistic: $F = 10.021$.

P-value: 0.000.

Conclusion:

The P-value = 0.000 < 0.05, so the null hypothesis is rejected and the alternative hypothesis is proven. We have shown that at least two of the locations (maybe more) differ in their effect upon the expected weight of oysters after a month.

e2(1) Give a 0.95 confidence interval for the difference between locations 3 and 5 in expected final weight of bags of oysters, when the starting weight of the bags is 28.

Interval:

Starting weight is immaterial, since we need the distance between two parallel lines. So, you can look at the difference between intercepts, and read the interval directly from the table with parameter estimates, because location 5 *is the reference:* (0.734, 2.576).

f (2) Give a rough and ready estimate for the number of bags per location that would have been needed with a one-way analysis of variance (ANOVA), with factor Location, but without covariate x in the model, for a completely randomized design (CRD), in order to achieve the same accuracy for comparison of locations as the aforementioned ANCOVA (explain your answer).

Answer (and explanation):

The efficiency is 10.684 / 0.302 = 35.38. *So, you need about* 35.38 * 4 ≈142 *bags per location to achieve the same accuracy in a one-way ANOVA for a CRD.*
[*Extra comment: this quite extreme result is due to the very unfortunate distribution of weight x at start over the locations, see the plot of the data.*]

Question 5 (18) Floral scent

In this (rather eccentric) study it is investigated whether pleasant aromas, here a floral scent, can affect a person's mental performance. 21 people work through different pencil and paper mazes (these are puzzles they have to finish) six times, three times while wearing a floral-scented mask (scented = Sc) and three times while wearing an unscented mask (not scented = NSc). The two masks and 126 mazes are offered in a random order to each person. The response (y) is the time (in seconds) it takes a person, wearing a mask, to complete a maze. Part of the 126 observations is reproduced on the right.

Person	Mask	y
1	NSc	38.40
2	NSc	46.20
3	NSc	72.50
4	NSc	38.00
5	NSc	82.80
6	NSc	33.90
7	NSc	50.40
…	…	…
15	Sc	33.70
16	Sc	42.60
17	Sc	54.90
18	Sc	64.50
19	Sc	43.10
20	Sc	52.80
21	Sc	44.30

The data are analysed with a mixed analysis of variance model, comprising fixed effects for the two types of masks, random main effects for the 21 persons, random interactions between persons and types of mask, and random residual error terms. The model is as follows:

$$\underline{y}_{ijk} = \mu + \tau_i + \underline{\beta}_j + \underline{\tau\beta}_{ij} + \underline{\varepsilon}_{ijk},$$

where τ_i is the fixed effect of a mask (i = 1, 2 for Sc and NSc), $\underline{\beta}_j$ is the random effect of a person (j = 1…21), $\underline{\tau\beta}_{ij}$ is random interaction, and $\underline{\varepsilon}_{ijk}$ is the residual error term (k = 1, 2, 3). We assume that $\underline{\beta}_j \sim N(0, \sigma_\beta^2)$, $\underline{\tau\beta}_{ij} \sim N(0, \sigma_{\tau\beta}^2)$ and $\underline{\varepsilon}_{ijk} \sim N(0, \sigma_\varepsilon^2)$, with all random effects mutually independent.

a1 (2) Explain why it is reasonable to assume that effects of persons are random and effects of types of mask are fixed.

Explanation:

Should you repeat the experiment, you would use the same two masks. It makes little sense to think as if they were sampled from a population of masks. So, effects of mask are fixed effects. However, in a repeat of the experiment, we would use other people, and it makes sense to regard the people being sampled from a population of people. So, effects for people are random effects.

a2 (2) Explain why observations $\underline{y}_{111}$ and $\underline{y}_{121}$ are independent, and observations $\underline{y}_{111}$ and $\underline{y}_{211}$ are dependent.

Explanation:

The first pair of observations is from different persons, but the second pair is from the same person (person 1). Consequently, observations from the last pair are dependent [in the model they share the random effect β_1 for person 1].

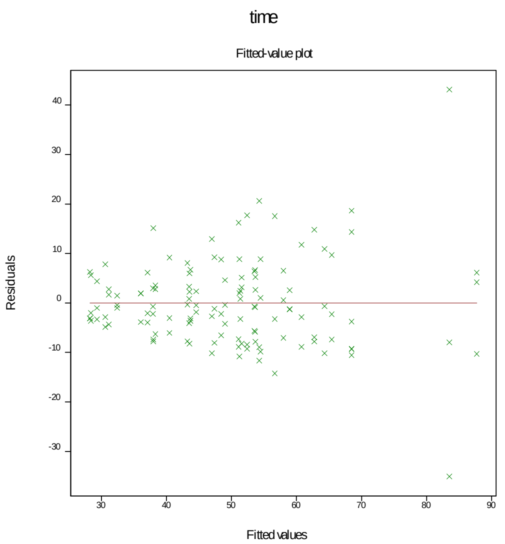
a3 (1) Give the two expressions for the expected value of $\underline{y}$, for a random person, completing a random maze, with each type of mask, i.e. the two population means, in terms of the parameters of the model.

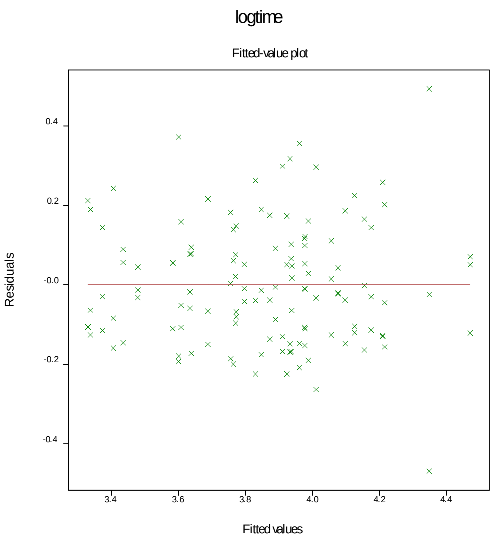
Expressions: $\mu + \tau_1$ *and* $\mu + \tau_2$ *(or just* μ *for the last mean, when the second mask is taken as a reference).*

a4 (1) Give the two expressions for the variance of $\underline{y}$ for a random person, completing a random maze, for each type of mask, i.e. the two population variances, in terms of the parameters of the model.

Expressions: *the two variances are the same, i.e.* $\sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma_\epsilon^2$.

Below, on the left, you see a plot of the residuals (e, estimated values for the ε's) against predicted (fitted) values ($\hat{y}$) made after the analysis. It was decided to analyse the data again, using log transformed times ($\ln(\underline{y})$) as a response. The plot of the residuals after analysis of $\ln(\underline{y})$ is shown on the right.





b (1) Explain, from the plots, which model assumption was the reason to change to an analysis of log transformed time.

Explanation:

In the plot on the left residuals increase with fitted values. So, the variance does not seem to be constant. The plot on the right, after log transformation, looks better: no obvious pattern left.

c (1) When individuals can be easily distracted, the same individual sometimes taking a lot more and sometimes taking a lot less time to finish mazes that are essentially of the same complexity, what component or components of variance would you expect to be relatively large?

Answer: the residual variance σ_{ϵ}^2.

You can use the SPSS output below to answer the next questions d1 and d2.

Tests of Between-Subjects Effects

Dependent Variable:ln_time

Source		Type II Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	1871.238	1	1871.238	4866.233	.000
	Error	7.691	20	.385[a]		
mask	Hypothesis	.024	1	.024	.314	.582
	Error	1.512	20	.076[b]		
person	Hypothesis	7.691	20	.385	5.086	.000
	Error	1.512	20	.076[b]		
mask * person	Hypothesis	1.512	20	.076	2.159	.008
	Error	2.942	84	.035[c]		

a. MS(person)
b. MS(mask * person)
c. MS(Error)

d1(4) Test whether the component of variance for random interaction between persons and masks differs from 0. Give:
(4) the null hypothesis and the alternative hypothesis, in the notation of the model,
(5) the expression of the test statistic (in terms of the relevant mean squares),
(6) the distribution of the test statistic under the null hypothesis,
(7) the P-value, and your conclusion (α = 0.05) in words.

Null and alternative hypothesis: $H_0: \sigma_{\tau\beta}^2 = 0,\ H_a: \sigma_{\tau\beta}^2 > 0.$

[*1 point off for* $H_a: \sigma_{\tau\beta}^2 \neq 0$]

Expression test statistic: $F = MSmask * person/MSE.$

*Distribution: F-distribution with df*1 = 20 *and df*2 = 84.

P-value: 0.008.

Conclusion:

P-value = 0.008 < 0.05, the null hypothesis is rejected, the alternative hypothesis is proven. So, we have shown that there is (random) interaction between persons and masks, the associated component of variance is shown to be positive.

d2(3) Test whether there is a difference in the expected time to finish a maze wearing a mask that is a floral scented or not. Give:
(1) the null hypothesis and the alternative hypothesis, in the notation of the model,
(2) the outcome of the test statistic,
(3) the P-value, and your conclusion (α = 0.05) in words.

Null and alternative hypothesis:

$$H_0: \tau_1 = \tau_2 \ (= 0, when\ e.g.\ mask\ 2\ is\ the\ reference) \ \ and \ \ H_a: \tau_1 \neq \tau_2\ .$$

Outcome test statistic: F = 0.314.

P-value: *0.582.*

Conclusion:

The P-value = 0.582 > 0.05, so the null hypothesis is not rejected. We cannot show that there is a difference between the masks in the way they affect the time to finish a maze.

[*1 point off for: "H_0 is shown/proven", or "it is shown/proven that masks do not differ", etc.*]

Expected Mean Squares

Source	Variance Component			
	Var(person)	Var(person * mask)	Var(Error)	Quadratic Term
Intercept	6.000	3.000	1.000	Intercept, mask
mask	.000	3.000	1.000	mask
person	6.000	3.000	1.000	
person * mask	.000	3.000	1.000	
Error	.000	.000	1.000	

Variance Estimates

Component	Estimate
Var(person)	.051
Var(person * mask)	.014
Var(Error)	.035

Dependent Variable: ln_time

e (3) How are the estimates derived for the three components of variance σ_β^2, $\sigma_{\tau\beta}^2$ and σ_ε^2 that you see in the output above (show the calculations)?

Answer for σ_ε^2: $\hat{\sigma}_\epsilon^2 = MSE = 0.035.$

Answer for $\sigma_{\tau\beta}^2$: $\hat{\sigma}_{\tau\beta}^2 = \frac{MSperson*mask - MSE}{3} = \frac{0.076 - 0.035}{3} = 0.014.$

Answer for σ_β^2: $\hat{\sigma}_\beta^2 = \frac{MSperson - MSperson*mask}{6} = \frac{0.385 - 0.076}{6} = 0.051.$

[*Extra: MS table page 20, coefficients 3 and 6 table Expected Mean Squares on this page.*]