											0

program

registration number

name and initials

WAGENINGEN UNIVERSITY
group MAT

Examination Advanced Statistics for Nutritionists (MAT-24306)
Wednesday, January 26, 2011; 14:00 – 17:00 h

Hand in your answers personally to one of the supervisors at the end of the examination.

MAT-24306

Question 1 (27)
In the temperate regions wheat is a main component for pig feed. Attempts are made to give pig feed good antimicrobial properties by fermentation of the wheat before mixing all components together. Fermentation by lactic acid bacteria using the back-slop technique is studied here. (Back-slopping means mixing a portion of previously fermented product in the fresh product to speed up fermentation.) Tanks are filled with 5 kg wheat and 15 kg water and inoculated with 20 ml lactic acid bacteria. Then to a tank either no previously fermented wheat is added (no back-slop) or a certain amount of previously fermented wheat is added (back-slop). The weight fraction of previously fermented wheat in a tank was in this research 0 (no back-slop) or 0.20, 0.33 or 0.42 (back-slop). Altogether 20 tanks are prepared and tested, 5 for each fraction.
After 24h fermentation the pH (next to many other things) is measured in each tank.
The model is $E(\underline{y}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ with
y = pH of tank contents after 24h fermentation,
x_1 = fraction previously fermented wheat (0, 0.20, 0.33 or 0.42)
$x_2 = x_1^2$,
x_3 = 1 in case of no back-slop ($x_1 = 0$), = 0 in case of back-slop ($x_1 > 0$)

Further we trust in validity of the usual additional assumptions.
Data of these 20 tanks:

```
tank fraction   y      x1      x2      x3
-----------------------------------------
 1    0       4.23      0       0       1
 2    0       4.18      0       0       1
 3    0       4.17      0       0       1
 4    0       4.22      0       0       1
 5    0       4.17      0       0       1
 6    0.20    3.73      0.20    0.0400  0
 ---------------- etc -----------------
20    0.42    3.83      0.42    0.1764  0
```

SPSS-output:
Regression

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.813	.140		27.176	.000
	x1	-.652	.981	-.547	-.665	.516
	x2	1.492	1.591	.534	.938	.362
	x3	.381	.141	.877	2.702	.016

a. Dependent Variable: y

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.69002	3	.23001	201.317	.000[a]
	Residual	.01828	16	.00114		
	Total	.70830	19			

a. Predictors: (Constant), x3, x2, x1
b. Dependent Variable: y

Univariate Analysis of Variance with the command **/LMATRIX="difference" x1 0.20 x2 0.12** gives

Contrast Results (K Matrix)[a]

Contrast			Dependent Variable y
L1	Contrast Estimate		.049
	Hypothesized Value		0
	Difference (Estimate - Hypothesized)		.049
	Std. Error		.020
	Sig.		.026
	95% Confidence Interval for Difference	Lower Bound	.007
		Upper Bound	.091

a. Based on the user-specified contrast coefficients (L') matrix: difference *copied lines from spreadsheet after running the Regression:*

```
y       x1     x2      x3      pre_1       sep_1
4.23    .00    .0000   1       4.19400     .01512
4.18    .00    .0000   1       4.19400     .01512
4.17    .00    .0000   1       4.19400     .01512
------------------------ etc ---------------------
3.83    .42    .1764   0       3.80200     .01512
.       .20    .0400   0       3.74200     .01512
```

a (2) What are these usual additional assumptions, referred to above?

▼▲ Observations are independent, normally distributed with the same variance.

b1 (1) Give the estimate for the variance of an observation.

▼▲ s^2 = MS(Residual) = 0.00114

b2 (2) Calculate the proportion of variation explained by the model, adjusted for degrees of freedom.

▼▲ $R^2_{\text{adj}} = 1 - (0.01828/16)/(0.70830/19) = 0.969$

We want to test $H_0 : \beta_3 = 0$ against $H_a : \beta_3 > 0$, at $\alpha = 0.05$.

c1 (3) Carry out the test. Mention

(1) the definition of the test statistic,

▼▲ $t = b_3 / \hat{\text{se}}\,(b_3)$

(2) the outcome of the test statistic,

▼▲ $t = 2.702$

(3) the P-value,

▼▲ P-value = 0.016/2 = 0.008 (since $b_3 > 0$ indeed)

(4) the conclusion.

▼ P-value = 0.008 < α = 0.05, so H_0 is rejected.
▲ It is shown that β_3 is positive.

c2 (2) Explain in ordinary words what $\beta_3 > 0$ means.

▼ When we trust the model for the whole range of x_1-values: no back-slopping is worse (gives higher pH-values) than back-slopping with the slightest amounts of fermented wheat.
[This interpretation implies that we trust the existence of a discontinuity when x_1 tends to 0. If we think that in reality back-slopping with decreasing amounts of fermented wheat should in the limit give the same result as no back-slopping, then the interpretation is that the quadratic curve for $x_1 > 0$ does not
▲ give a good description when x_1 decreases to 0.]

When there is no systematic difference between back-slopping to fraction 0.40 and back-slopping to fraction 0.20, the lower fraction is preferable.

d1 (2) Briefly show that the systematic difference between back-slopping to fraction 0.40 and back-slopping to fraction 0.20 is equal to $0.20\beta_1 + 0.12\beta_2$.

▼ $x_1 = 0.40$: $E(y) = \beta_0 + 0.40\beta_1 + 0.16\beta_2$
$x_1 = 0.20$: $E(y) = \beta_0 + 0.20\beta_1 + 0.04\beta_2$
-------------------------------- –
▲ $0.20\beta_1 + 0.12\beta_2$

d2 (4) Test ($\alpha = 0.05$) the hypothesis that there is no systematic difference between back-slopping to fraction 0.40 and back-slopping to fraction 0.20. Mention
(1) H_0 and H_a,

▼▲ H_0: $0.20\beta_1 + 0.12\beta_2 = 0$ and H_a: $0.20\beta_1 + 0.12\beta_2 \neq 0$

(2) the definition of the test statistic,

▼▲ $t = (0.20b_1 + 0.12b_2)/\ \hat{se}\,(0.20b_1 + 0.12b_2)$

(3) the outcome of the test statistic,

▼▲ $t = 0.049/0.020 = 2.45$

(4) the P-value,

▼▲ P-value = 0.026

(5) the conclusion (also in words).

▼ P-value = 0.026 < α = 0.05, H_0 is rejected.
It is shown that there is a systematic difference between back-slopping to fraction 0.40 and back-
▲ slopping to fraction 0.20

Suppose we know that a pH lower than 4.0 will reduce microbial risks to an acceptable level. We try to show that the expected pH when back-slopping with fraction 0.20 is smaller than 4.0 .

e1 (2) Write H_0 and H_a for this problem in parameters.

▼▲ H_0: $\beta_0 + 0.20\beta_1 + 0.04\beta_2 = 4.0$ and H_a: $\beta_0 + 0.20\beta_1 + 0.04\beta_2 < 4.0$

e2 (3) Carry out the test at $\alpha = 0.05$, using a suitable test statistic. Mention
(1) the definition of the test statistic,

▼▲ $t = (b_0 + 0.20b_1 + 0.04b_2 - 4)/\ \hat{se}\,(\ b_0 + 0.20b_1 + 0.04b_2)$

(2) the outcome of the test statistic,

▼▲ $t = (3.742 - 4.0) / 0.01512 = -0.258/0.01512 = -17.1$

(3) the P-value or the critical region,

▼ critical region is $(-\infty, -1.746]$
▲ (right-sided; see table 2 with $df = 16$ and $\alpha = 0.05$)

(4) the conclusion in words.

▼▲ -17.1 is in $(-\infty, -1.746]$, reject H_0, we showed that the expected pH is indeed smaller than 4.0.

More SPSS-output:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.981[a]	.962	.958	.03957	.962	217.672	2	17	.000
2	.987[b]	.974	.969	.03380	.012	7.299	1	16	.016

a. Predictors: (Constant), x2, x1

b. Predictors: (Constant), x2, x1, x3

and

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.980[a]	.961	.959	.03927	.961	441.269	1	18	.000
2	.987[b]	.974	.969	.03380	.013	4.149	2	16	.035

a. Predictors: (Constant), x3

b. Predictors: (Constant), x3, x2, x1

and

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.797[a]	.636	.615	.11974	.636	31.398	1	18	.000
2	.987[b]	.974	.969	.03380	.339	104.952	2	16	.000

a. Predictors: (Constant), x1

b. Predictors: (Constant), x1, x3, x2

and

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.987[a]	.973	.970	.03324	.973	311.997	2	17	.000
2	.987[b]	.974	.969	.03380	.001	.442	1	16	.516

a. Predictors: (Constant), x3, x2

b. Predictors: (Constant), x3, x2, x1

Also we want to test H_0: "when back-slop is indeed applied, it doesn't matter how much fermented wheat is added to the tank"

f (6) Carry out the test at $\alpha = 0.05$. Mention
(1) the null hypothesis and the alternative hypothesis,

▼▲ H_0: $\beta_1 = \beta_2 = 0$

(2) the outcome of the test statistic,

▼▲ $F = 4.149$

(3) the null distribution of the test statistic,

▼▲ F-distribution with $df_1 = 2$ and $df_2 = 16$

(4) the P-value,

▼▲ P-value = 0.035

(5) the conclusion (also in words).

▼ P-value = 0.035 < α = 0.05, we reject H_0 and showed that when we apply back-slopping, the amount
▲ of fermented wheat is important.

Question 2 (22)
In the development of "light"-products by reduction of fat, care is taken that not only appearance and taste, but also e.g. breaking behaviour, is only mildly influenced. The breaking behaviour of biscuits is quantified by the so-called breaking tension (y). A large breaking tension will be experienced by the consumer as "hard" and is for specific types of biscuits (e.g. "sprits") undesirable.

In an experiment with "sprits" 3 versions for factor Fat are chosen:
F1 (65% margarine), F2 (60% margarine) and F3 (55% margarine).
F1 corresponds with respect to nutritional value to the usual "sprits" on the market; F3 has the highest "light"ness.
Furthermore, two standard methods of preparation of the dough are included, factor Preparation:
P1 (cream method) and P2 (all-in method).
P1 and P2 are not different in composition of the dough, but are different in the order of mixing and stirring times, which might have an effect on breaking behaviour. It is conceivable that one method is more suitable for the preparation of "light"-sprits than the other.

At the difference factor combinations a total of 36 observations are taken. It is assumed that the 36 observations are independent, normally distributed with equal variance σ^2.
The observations are:

		Preparation															
		P1								P2							
Fat	F1	183	144	121	109	173	113	139	189	157	111	93	137	148	172	111	111
	F2	175	222	211	173	145				163	188	179	180	210			
	F3	209	203	212	178	189				259	213	320	293	311			

Part 1
For the analysis of y the interaction model is used:
$y_{ijk} = \mu + \tau_i + \beta_j + \tau\beta_{ij} + \varepsilon_{ijk}$ with $E(\varepsilon_{ijk}) = 0$ and
with obvious meaning of the indices (i for Fat, j for Preparation and k for replication within combination).

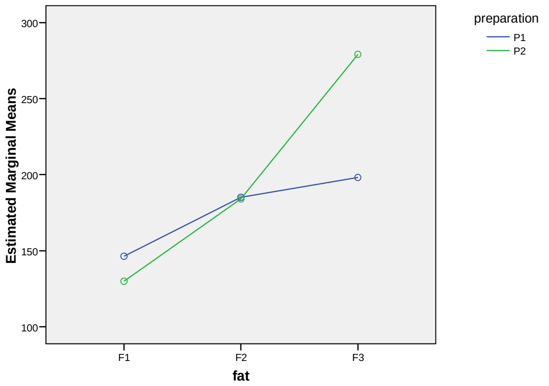
SPSS-output:

Descriptive Statistics

Dependent Variable: breaking tension

fat	preparation	Mean	Std. Deviation	N
F1	P1	146.38	31.807	8
	P2	130.00	27.563	8
	Total	138.19	29.969	16
F2	P1	185.20	31.180	5
	P2	184.00	17.132	5
	Total	184.60	23.726	10
F3	P1	198.20	14.342	5
	P2	279.20	43.752	5
	Total	238.70	52.580	10
Total	P1	171.56	35.435	18
	P2	186.44	69.741	18
	Total	179.00	55.039	36

Estimated Marginal Means of breaking tension



Between-Subjects Factors

		Value Label	N
fat	1	F1	16
	2	F2	10
	3	F3	10
preparation	1	P1	18
	2	P2	18

Tests of Between-Subjects Effects

Dependent Variable: breaking tension

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	80083.725[a]	5	16016.745	18.522	.000
Intercept	1153476.000	1	1153476.000	1333.895	.000
fat	62605.062	2	31302.531	36.199	.000
preparation	1995.111	1	1995.111	2.307	.139
fat * preparation	15483.551	2	7741.776	8.953	.001
Error	25942.275	30	864.743		
Total	1259502.000	36			
Corrected Total	106026.000	35			

a. R Squared = .755 (Adjusted R Squared = .715)

Parameter Estimates

Dependent Variable: breaking tension

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	279.200	13.151	21.230	.000	252.342	306.058
[fat=1]	-149.200	16.764	-8.900	.000	-183.437	-114.963
[fat=2]	-95.200	18.598	-5.119	.000	-133.183	-57.217
[fat=3]	0[a]	.	.	.	.	.
[preparation=1]	-81.000	18.598	-4.355	.000	-118.983	-43.017
[preparation=2]	0[a]	.	.	.	.	.
[fat=1] * [preparation=1]	97.375	23.708	4.107	.000	48.956	145.794
[fat=1] * [preparation=2]	0[a]	.	.	.	.	.
[fat=2] * [preparation=1]	82.200	26.302	3.125	.004	28.484	135.916
[fat=2] * [preparation=2]	0[a]	.	.	.	.	.
[fat=3] * [preparation=1]	0[a]	.	.	.	.	.
[fat=3] * [preparation=2]	0[a]	.	.	.	.	.

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

a1 (2) Is the design orthogonal? Explain briefly.

▼▲ yes, for the number of observations proportional representation: 8:8 = 5:5 = 5:5

a2 (1) Will the design be non-orthogonal if by chance one observation of each factorcombination is lost?

▼▲ no, for the number of observations still proportional representation: 7:7 = 4:4 = 4:4

b (4) Test ($\alpha = 0.05$) whether the Preparation effect is the same at all Fat-levels. Report
(1) the definition of the test statistic,

▼▲ F = MS(fat*preparation)/MS(Error)

(2) the outcome of the test statistic,

▼▲ $F = 8.953$

(3) the P-value,

▼▲ P-value = 0.001

(4) the conclusion (also in words).

▼▲ P-value = 0.001 < $\alpha = 0.05$, H_0 is rejected, it is shown that there is interaction, or it is shown that the Preparation effect is not the same at all Fat-levels.

c1 (2) Estimate the expected breaking tension for a sprits with 65% margarine (F1) and Preparation cream method (P1). Use only the Parameter Estimates output.

▼▲ 279.2 + (–149.2) + (–81.0) + 97.375 = 146.375

c2 (1) Check your answer using other fragments of the output.

▼▲ $\bar{y}_{11}$ = 146.38

c3 (2) Estimate the corresponding standard error.

▼▲ √(864.743/8) = 10.40

Part II
From now on, it is assumed that an additive model is appropriate, if F3 is removed from the analysis (and later on discussed separately).

Then we get the following model: $\underline{y}_{ijk} = \mu + \tau_i + \beta_j + \underline{\varepsilon}_{ijk}$ for the *k*-th observation at Fat *i* and Preparation *j*, with *i* = 1, 2 and *j* = 1, 2 and *k* = 1, 2, . . .

SPSS-output:

Between-Subjects Factors

		Value Label	N
fat	1	F1	16
	2	F2	10
preparation	1	P1	13
	2	P2	13

Descriptive Statistics

Dependent Variable: breaking tension

fat	preparation	Mean	Std. Deviation	N
F1	P1	146.38	31.807	8
	P2	130.00	27.563	8
	Total	138.19	29.969	16
F2	P1	185.20	31.180	5
	P2	184.00	17.132	5
	Total	184.60	23.726	10
Total	P1	161.31	36.066	13
	P2	150.77	35.898	13
	Total	156.04	35.662	26

Tests of Between-Subjects Effects

Dependent Variable: breaking tension

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	13978.009[a]	2	6989.004	9.022	.001
Intercept	633048.038	1	633048.038	817.205	.000
fat	13256.124	1	13256.124	17.112	.000
preparation	721.885	1	721.885	.932	.344
Error	17816.953	23	774.650		
Total	664843.000	26			
Corrected Total	31794.962	25			

a. R Squared = .440 (Adjusted R Squared = .391)

Parameter Estimates

Dependent Variable: breaking tension

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	179.331	10.357	17.316	.000	157.907	200.755
[fat=1]	-46.413	11.220	-4.137	.000	-69.622	-23.203
[fat=2]	0[a]	.	.	.	.	.
[preparation=1]	10.538	10.917	.965	.344	-12.045	33.122
[preparation=2]	0[a]	.	.	.	.	.

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

d (2) Give a 0.95-confidence interval for the expected breaking tension of a sprits with 60% margarine (F2) and method of preparation P2.

▼▲ So for μ: (157.9 , 200.76)

Now we try to show that Preparation 1 tends to give higher breaking tension than Preparation 2. In other words, we want to test H_0: $\beta_1 - \beta_2 = 0$ against a suitable alternative hypothesis.

e1 (1) Estimate $\beta_1 - \beta_2$ using the table with Descriptive Statistics.

▼▲ $\bar{y}_{\text{prep}\,1} - \bar{y}_{\text{prep}\,2} = 161.31 - 150.77 = 10.54$

e2 (2) Estimate the corresponding standard error using the table with Descriptive Statistics.

▼▲ $\hat{\text{se}}\,(\bar{y}_{\text{prep}\,1} - \bar{y}_{\text{prep}\,2}) = \sqrt{}(774.650(1/13 + 1/13)) = 10.917$

e3 (2) Check your answers of the questions e1 and e2 using other fragments of the output.

▼▲ $\hat{\beta}_1 - \hat{\beta}_2 = 10.538 - 0 = 10.538$ and $\hat{\text{se}}\,(\hat{\beta}_1 - \hat{\beta}_2) = \hat{\text{se}}(\hat{\beta}_1) = 10.917$

e4 (3) Carry out the test at $\alpha = 0.05$ using a suitable confidence bound for $\beta_1 - \beta_2$.

▼ H_a: $\beta_1 - \beta_2 > 0$, so we need a confidence lower bound.
$10.54 - 1.714 \times 10.917 = -8.17$ (1.714 follows from a t-distribution with $df = 23$ and $\alpha = 0.05$)
$0 \in (-8.17\ ,\ \infty)$, H_0m is not rejected. It is not shown that Preparation 1 tends to give higher breaking
▲ tension than Preparation 2.

.

Question 3 (12)

Genetic modification (GMO) of apples can result in a more sustainable way of cultivation. Modified apple trees might have a higher resistance against scab, the most common fungal disease in the apple farming industry. The genetically modified apple trees in the research received some additional genetic material from barley. This DNA element creates a substance which protects the barley from invasive fungi. Laboratory research showed that the substance was also effective against the fungus causing apple scab.

In 2004 researchers planted 25 genetically modified apple trees and 25 control trees in a research orchard, where the fungus causing scab was not countered with chemicals. After 4 years the leaf damage by the fungus was measured using a score, averaged over a large number of leaves per tree. A higher score means larger damage. We assume that the 50 observations are independent and normally distributed.

SPSS-output:

T-Test

Group Statistics

	GMO	N	Mean	Std. Deviation	Std. Error Mean
score	yes	25	2.4744	.83382	.16676
	no	25	2.7236	.70858	.14172

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
score	Equal variances assumed	.554	.460	-1.139	48	.260	-.24920	.21885	-.68922	.19082
	Equal variances not assumed			-1.139	46.782	.261	-.24920	.21885	-.68952	.19112

First we want to investigate whether the variance between the GMO apple trees is larger than the variance between the control trees.

a (6) Investigate ($\alpha = 0.05$) whether the variance between the GMO apple trees is larger than the variance between the control trees. Mention:

(1) the null hypothesis and the alternative hypothesis,

▼ H_0: $\sigma_1^2/\sigma_2^2 = 1$ and H_a: $\sigma_1^2/\sigma_2^2 > 1$

▲ with σ_1^2 = variance of the score for a GMO tree and σ_2^2 = variance of the score for a control tree

(2) the definition and the outcome of the test statistic,

▼ definition test statistic: $F = s_1^2/s_2^2$

▲ outcome test statistic: $F = 0.83382^2/0.70858^2 = 1.38$

(3) the P-value or the critical region,

▼ the critical region is right-sided; null distribution test statistic: F-distribution with $df_1 = 24$and $df_2 = 24$

▲ the critical value follows from table 8 with $\alpha = 0.05$: 1.98 and the critical region is: $[1.98\ ,\ \infty]$.

(4) the conclusion (also in words).

▼ $1.38 \notin [1.98\ ,\ \infty]$, so H_0 is not rejected.

It is not shown that the variance between the GMO apple trees is larger than the variance between the

▲ control trees.

From now on we assume that the 50 observations have the same variance (σ^2).

b (2) Estimate σ^2.

▼▲ $s^2 = (24 \times 0.83382^2 + 24 \times 0.70858^2)/\ 48 = 0.5987$

Next we want to know whether the expected score of GMO apple trees is smaller than the expected score of control trees. The conclusion is "it is not shown that expected score of GMO apple trees is smaller than the expected score of control trees". Therefore the researchers have doubts about the power of the test.
They want to have the following power requirements for investigating whether the leaf damage score of GMO apple trees is systematically smaller than the score of control trees: when the expected difference is 0.3 (no vs. yes) then the test with $\alpha = 0.05$ should have a power $(1 - \beta)$ of at least 0.90.

c (4) How many apple trees are needed for this experiment?

▼ H_a is one-sided: $N \approx 2\dfrac{(z_\alpha + z_\beta)^2}{(\Delta/\sigma)^2}$ with $z_\alpha = 1.645$ ($\alpha = 0.05$); $z_\beta = 1.282$ ($\beta = 0.10$); $\Delta = 0.3$; $\sigma \approx 0.60$

▲ $N \approx 2\dfrac{(1.645 + 1.282)^2}{(0.3^2/0.60)} = 114.2$, so one needs about 115 GMO trees and 115 control trees.

Question 4 (8)
A MNH student of our university does the following experiment inspired by a publication. She takes random samples of 12 men and 10 women from the relevant populations. All these persons are willing to receive a special light pain stimulus and immediately after that to drink a cup of coffee. After drinking the coffee each person gives an answer on the question whether he/she feels less pain by now.
We assume that all the observations are independent.

We get the following SPSS output:

gender * less pain Crosstabulation

			less pain		Total
			yes	no	
gender	man	Count	2	10	12
		Expected Count	4.4	7.6	12.0
	woman	Count	6	4	10
		Expected Count	3.6	6.4	10.0
Total		Count	8	14	22
		Expected Count	8.0	14.0	22.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	4.426[b]	1	.035	.074	.048	
Continuity Correction[a]	2.752	1	.097			
Likelihood Ratio	4.567	1	.033	.074	.048	
Fisher's Exact Test				.074	.048	
Linear-by-Linear Association	4.225[c]	1	.040	.074	.048	.043
N of Valid Cases	22					

a. Computed only for a 2x2 table
b. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 3.64.
c. The standardized statistic is -2.055.

a (8) The student wants to investigate ($\alpha = 0.05$) whether the probability for 'yes' (so: 'less pain') is for men smaller than for women. Mention
(1) null hypothesis and alternative hypothesis,

▼ H_0: $\pi_1 = \pi_2$ with π_1 = probability that a man says 'yes' and π_2 = probability that a woman says 'yes',
▲ against H_a: $\pi_1 < \pi_2$

(2) full name of the test (consistent with the sampling design),

▼▲ Fisher's exact test for homogeneity

(3) definition of the test statistic,

▼▲ k = the number of 'yes' for men (given the total number of 'yes')

(4) outcome of the test statistic,

▼▲ $k = 2$ (given the total number of 'yes' is 8)

(5) P-value,

▼▲ P-value = 0.048

(6) conclusion in words.

▼ P-value = 0.048) < $\alpha = 0.05$, so H_0 is rejected. It is shown that the probability that a man says 'yes' is
▲ smaller than the probability that a woman says 'yes'.

Question 5 (6)
Researchers conducted an experiment to study the effects of GLV (green leafy vegetables) supplements on vitamin-C. The diets (4 GLV-based and a non-GLV with extra carotene tablets as control) were as follows (please ignore details!):

Diet 1 (GLV+N): GLV (100 g/day) cooked with normal oil (5 g oil/100 g GLV/day)
Diet 2 (GLV+M): GLV (100 g/day) cooked with more oil (10 g oil/100 g GLV/day)
Diet 3 (GLV+C): GLV (100 g/day) cooked with normal oil+vitamin C (Celin tablet of 100 mg/day)
Diet 4 (GLV+E): GLV (100 g/day) cooked with normal oil+vitamin E (Evion tablet of 100 mg/day)
Diet 5 (non-GLV+β-car): Cereal-based diet with other vegetable but no GLV
+ Parry's β-carotene from spirulina (10 mg) thrice/week

At random 8 persons (young adults) were assigned to each diet and instructed to follow the diet for 3 weeks. For each person, the percentage increase of plasma vitamin-C was measured. The usual assumptions (independence, normal distributions, common variance) are reasonable.
The five expected values for a response variable are denoted by μ_1 to μ_5.

Based on previous investigations it is reasonable to think that if a GLV-diet has effect on vitamin-C as compared to the non-GLV diet, it can only be a positive effect (i.e. leads to higher vitamin-C values than diet 5). The researchers now want to compare each of the GLV-diets with the non-GLV diet, with α = 0.05 simultaneously.

SPSS-output:

Multiple Comparisons

Dependent Variable: vitaminC

	(I) Diet	(J) Diet	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	GLV-N	GLV-M	.5375	1.6575	.997	-4.228	5.303
		GLV-C	-2.7500	1.6575	.471	-7.516	2.016
		GLV-E	-.3000	1.6575	1.000	-5.066	4.466
		non-GLV	2.5125	1.6575	.559	-2.253	7.278
	GLV-M	GLV-N	-.5375	1.6575	.997	-5.303	4.228
		GLV-C	-3.2875	1.6575	.295	-8.053	1.478
		GLV-E	-.8375	1.6575	.986	-5.603	3.928
		non-GLV	1.9750	1.6575	.756	-2.791	6.741
	GLV-C	GLV-N	2.7500	1.6575	.471	-2.016	7.516
		GLV-M	3.2875	1.6575	.295	-1.478	8.053
		GLV-E	2.4500	1.6575	.583	-2.316	7.216
		non-GLV	5.2625*	1.6575	.024	.497	10.028
	GLV-E	GLV-N	.3000	1.6575	1.000	-4.466	5.066
		GLV-M	.8375	1.6575	.986	-3.928	5.603
		GLV-C	-2.4500	1.6575	.583	-7.216	2.316
		non-GLV	2.8125	1.6575	.449	-1.953	7.578
	non-GLV	GLV-N	-2.5125	1.6575	.559	-7.278	2.253
		GLV-M	-1.9750	1.6575	.756	-6.741	2.791
		GLV-C	-5.2625*	1.6575	.024	-10.028	-.497
		GLV-E	-2.8125	1.6575	.449	-7.578	1.953
Dunnett t (>control)[a]	GLV-N	non-GLV	2.5125	1.6575	.188	-1.202	
	GLV-M	non-GLV	1.9750	1.6575	.299	-1.739	
	GLV-C	non-GLV	5.2625*	1.6575	.006	1.548	
	GLV-E	non-GLV	2.8125	1.6575	.140	-.902	
Dunnett t (<control)[a]	GLV-N	non-GLV	2.5125	1.6575	.995		6.227
	GLV-M	non-GLV	1.9750	1.6575	.987		5.689
	GLV-C	non-GLV	5.2625	1.6575	1.000		8.977
	GLV-E	non-GLV	2.8125	1.6575	.997		6.527
Dunnett t (2-sided)[a]	GLV-N	non-GLV	2.5125	1.6575	.373	-1.727	6.752
	GLV-M	non-GLV	1.9750	1.6575	.581	-2.265	6.215
	GLV-C	non-GLV	5.2625*	1.6575	.011	1.023	9.502
	GLV-E	non-GLV	2.8125	1.6575	.279	-1.427	7.052

*. The mean difference is significant at the .05 level.

a. Dunnett t-tests treat one group as a control, and compare all other groups against it.

a1 (1) Write down the *alternative* hypotheses that have to be considered simultaneously.

▼▲ $\mu_1 - \mu_5 > 0$, $\mu_2 - \mu_5 > 0$, $\mu_3 - \mu_5 > 0$, $\mu_4 - \mu_5 > 0$

a2 (5) Carry out the test(s). Defend the choice of testing procedure and mention your conclusion together with the numerical results on which this is based.

▼ Compare all GLVs with "control" non-GLV, so use Dunnett, vs. last, one-sided "larger than control". Only for GLV-C the P-value for the comparison with non-GLV is < 0.05 (namely 0.006). Only GLV-C can be declared to have a positive effect as compared to non-GLV.
▲

Question 6 (15)
In the Western world, increased consumption of carbonated soft drinks combined with a decreasing intake of milk may increase the risk of osteoporosis. In an experiment with 11 healthy young men Diet 1 and Diet 2 were compared. Both diets were the same low-calcium basic diet. During Diet 1 a person had to drink 2.5 liter cola per day but during Diet 2 this was replaced by 2.5 liter semi-skimmed milk per day. Each person was observed during 4 diet periods of 10 days, alternately with diets 1-2-1-2 or 2-1-2-1, with wash-out periods in-between. The diet to start with was chosen at random for each person. During each diet period the increase in serum CTX was recorded (variable y, in µg/l). Note that high CTX values indicate a negative effect of the diet on bone health. Results:

man	1	2	...	11
Diet 1 Cola	0.31 0.40	0.43 –0.06	... ...	0.33 0.00
Diet 2 Milk	0.30 0.21	–0.05 –0.34	... ...	–0.12 0.04

For a statistical analysis we will use the model:
$\underline{y}_{ijk} = \mu + \tau_i + \underline{\beta}_j + \underline{\tau\beta}_{ij} + \underline{\varepsilon}_{ijk}$
with indices i for Diet, j for Man and k for replicate in time.
All random model terms are independent and normally distributed, with expectation 0 and constant variance (resp. σ_β^2, $\sigma_{\tau\beta}^2$ and σ_ε^2). Descriptive and fixed-effects output of SPSS is shown below.

Report

y

Diet	N	Mean	Std. Deviation
Cola	22	.1664	.29167
Milk	22	-.1968	.35041
Total	44	-.0152	.36777

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4.965[a]	21	.236	6.116	.000
Intercept	.010	1	.010	.264	.613
Diet	1.451	1	1.451	37.529	.000
Man	2.579	10	.258	6.670	.000
Diet * Man	.936	10	.094	2.420	.040
Error	.851	22	.039		
Total	5.826	44			
Corrected Total	5.816	43			

a. R Squared = .854 (Adjusted R Squared = .714)

a1 (2) Motivate why Diet is entered as a fixed factor and Man as a random factor.

▼ Diet: only two levels, Diet 1 and Diet 2, so fixed
Man: 11 men are at random chosen from a lot of men, so random. We are not interested in individual
▲ men.

a2 (2) Give the expression for the variance of a measurement y.

▼▲ $\sigma_y^2 = \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma_\varepsilon^2$

First we want to investigate whether there is a positive systematic difference between the Cola and the Milk diets.

b1 (2) Formulate H_0 and H_a in parameters.

▼▲ H_0: $\tau_1 - \tau_2 = 0$ against H_a: $\tau_1 - \tau_2 > 0$

b2 (5) Carry out the test at $\alpha = 0.05$.

(1) Mention the definition of the test statistic.

▼▲ test statistic: $\underline{t} = (\underline{\hat{\tau}}_1 - \underline{\hat{\tau}}_2)/\underline{\hat{se}}$

(2) Calculate the outcome of the test statistic.

▼ $\hat{\tau}_1 - \hat{\tau}_2 = 0.1664 - (-0.1968) = 0.3632$ and $\hat{se} = \sqrt{[(1/22 + 1/22) \times 0.094]} = 0.0924$.

▲ Outcome test statistic: $t = 0.3632/0.0924 = 3.93$

(3) Give the critical region.

▼▲ critical region (see table 2 with $df = 10$ and $\alpha = 0.05$): $[1.81 \ , \ \infty)$

(4) Give the conclusion of the test, also in words.

▼ $3.93 \in$ critical region, so H_0 is rejected. It is shown that Cola diet gives a systematic larger CTX value
▲ than Milk diet.

c1 (2) A mixed model is meant for dependent data. Which measurements are dependent, and which measurements are independent?

▼ dependent: the 4 observations within a man
▲ independent: observations from different men

c2 (2) Which component of variance is associated with the variation between measurements of the same man with the same Diet?

▼▲ σ_ε^2