

## Final exam, 19 April 2018

All aids are allowed, except a computer-like device (including tablets and smartphones) and personal assistance. The exam consists of three questions that should all be answered. The weights for each of the three questions and also for each subquestion within a question are indicated; these weights total *50 points*. Note that questions, and often also subquestions, can be answered independently of each other. The duration of the exam is 3 hours.

Generally, **statistical models and methods should be specified**, and every statistical analysis should be summarized in a conclusion. Throughout, if you realize that you need more information than is provided to carry out the analysis, specify what information you need, how you would obtain it using Minitab or Stata, and how you would use it in the analysis.

### Question 1. (*15 points*)

Fifty (50) pills were randomly selected from the production of a certain kind of headache pill, and weighed. The table below gives descriptive statistics for the weights measured in grams (*g*), as computed by Minitab.

statistic	value	statistic	value	statistic	value
mean	0.657	standard dev.	0.045	minimum	0.589
1st quartile	0.624	median	0.653	3rd quartile	0.678
maximum	0.817	skewness	1.43	kurtosis	3.15

**a)** (*5 points*)

Based on the information provided, carry out a descriptive analysis for the distribution of weights. As part hereof, discuss how well these data could be described by a normal distribution; describe any deviations from a normal distribution noted.

**b)** (*3 points*)

Compute a 90% confidence interval for the mean weight of a single pill. In view of **a)**, should this interval be interpreted as exact or approximate? — motivate your answer.

**c)** (*3 points*)

The prescribed weight of each pill is 0.65 *g*. Give a statistical assessment of whether the production is systematically off its target value. In view of **a)**, are you confident about your conclusion? — motivate your answer.

**d)** (*4 points*)

The pills are packaged in boxes with 100 pills, according to the specification on the boxes. However, in the production it is easier to weigh the pills than to count them. Assume that the following method is used to fill boxes with pills: the box is put on a scale and pills are added until their total weight exceeds 66 grams. Compute (estimate) the probability that the box contains at least 100 pills. (*Hint*: it may be helpful to express the event in question in terms of the total (or average) weight of 100 pills.)

**Question 2.** (15 points)

An experiment with 12 animals was carried out to investigate the impact of a vitamin on blood coagulation. Three fodders were used: C, V1, and V2, where C is a standard feed, and V1 and V2 denote the standard feed with vitamin additives (from two different manufacturers). The animals were randomly assigned to the three fodder groups, as shown in the table below, and kept on their designated fodder for a suitable period of time, after which the time (in seconds) until blood coagulation was measured for a sample from each animal.

Fodder					
C		V1		V2	
Animal	Time	Animal	Time	Animal	Time
1	70	5	57	4	53
2	61	7	61	8	53
3	60	9	56	12	59
6	65				
10	82				
11	66				

Animal 10 fell sick during the experiment, for reasons considered to be unrelated to the presence or absence of the vitamin.

Use the description of the study and the Minitab listings on the following pages to answer the parts **a)-e)**.

**a)** (1 point)

Characterize (briefly) the experimental/statistical design, using standard descriptors.

**b)** (7 points)

Formulate an appropriate statistical model for the data. Pay particular attention to whether Animal 10 should be included in the analysis or not. Explain your reasoning regarding this issue carefully, including both information about the study and information extracted from the data, as needed.

**c)** (2 points)

Carry out a statistical analysis to determine whether the data show evidence of a difference between the three fodders with respect to blood coagulation.

**d)** (2 points)

The researchers behind the study did not expect any major difference between the two vitamin products (with respect to blood coagulation). Use statistical inference to quantify the evidence in the data about any difference between the vitamin products.

**e)** (3 points)

The main purpose of the experiment was to establish whether there was any effect of vitamin supplementation, and to quantify it. In continuation of the previous analyses, compute both suitable interval(s) and test(s) (one or several) to quantify such an effect and assess its statistical significance.

*Minitab listing for Question 2):*

(Note: in the listing, some Minitab commands have been abbreviated.)

```
MTB > name c1 'animal' c2 'fodder' c3 'time'
MTB > Copy 'time' c4;
SUBC>  Varnames.
MTB > let 'time_1'(5)='*'
MTB > Describe 'time' 'time_1';
SUBC>  By 'fodder';
...
Descriptive Statistics: time, time_1
```

Variable	fodder	N	N*	Mean	StDev	Minimum	Median	Maximum
time	c	6	0	67.33	8.04	60.00	65.50	82.00
	v1	3	0	58.00	2.65	56.00	57.00	61.00
	v2	3	0	55.00	3.46	53.00	53.00	59.00
time_1	c	5	1	64.40	4.04	60.00	65.00	70.00
	v1	3	0	58.00	2.65	56.00	57.00	61.00
	v2	3	0	55.00	3.46	53.00	53.00	59.00

```
MTB > Name C5 "SRES".
MTB > GLM;
SUBC>  Response 'time';
SUBC>  Categorical 'fodder';
...
SUBC>  SResiduals 'SRES'.
General Linear Model: time versus fodder
```

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
fodder	2	365.6	182.79	4.55	0.043
Error	9	361.3	40.15		
Total	11	726.9			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
6.33626	50.29%	39.25%	24.19%

Fits and Diagnostics for Unusual Observations

Obs	time	Fit	Resid	Resid	Std
5	82.00	67.33	14.67	2.54	R

R Large residual

Means

Term	Fitted	
	Mean	SE Mean
fodder		
c	67.33	2.59
v1	58.00	3.66
v2	55.00	3.66

MTB > Kruskal-Wallis 'time' 'fodder'.  
Kruskal-Wallis Test: time versus fodder

Kruskal-Wallis Test on time

fodder	N	Median	Ave	
			Rank	Z
c	6	65.50	9.3	2.64
v1	3	57.00	4.8	-0.92
v2	3	53.00	2.7	-2.13
Overall	12		6.5	

H = 7.52 DF = 2 P = 0.023  
H = 7.58 DF = 2 P = 0.023 (adjusted for ties)

\* NOTE \* One or more small samples

MTB > Name C6 "SRES\_1".  
MTB > GLM;  
SUBC> Response 'time\_1';  
SUBC> Categorical 'fodder';  
...  
SUBC> SResiduals 'SRES\_1'.  
General Linear Model: time\_1 versus fodder

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
fodder	2	183.7	91.85	7.12	0.017
Error	8	103.2	12.90		
Total	10	286.9			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
3.59166	64.03%	55.04%	34.69%

Means

Term	Fitted	
	Mean	SE Mean
fodder		
c	64.40	1.61
v1	58.00	2.07
v2	55.00	2.07

MTB > Kruskal-Wallis 'time\_1' 'fodder'.  
Kruskal-Wallis Test: time\_1 versus fodder

11 cases were used  
1 cases contained missing values

Kruskal-Wallis Test on time\_1

fodder	N	Median	Ave	
			Rank	Z
c	5	65.00	8.7	2.46
v1	3	57.00	4.8	-0.71
v2	3	53.00	2.7	-2.04
Overall	11		6.0	

H = 6.72 DF = 2 P = 0.035  
H = 6.78 DF = 2 P = 0.034 (adjusted for ties)

\* NOTE \* One or more small samples

MTB > Print 'animal'-'SRES\_1'  
Data Display

Row	animal	fodder	time	time_1	SRES	SRES_1
1	1	c	70	70	0.46103	1.74320
2	2	c	61	61	-1.09494	-1.05837
3	3	c	60	60	-1.26782	-1.36966
4	6	c	65	65	-0.40340	0.18677
5	10	c	82	*	2.53565	*
6	11	c	66	66	-0.23051	0.49806
7	5	v1	57	57	-0.19329	-0.34100
8	7	v1	61	61	0.57987	1.02299
9	9	v1	56	56	-0.38658	-0.68199
10	4	v2	53	53	-0.38658	-0.68199
11	8	v2	53	53	-0.38658	-0.68199
12	12	v2	59	59	0.77317	1.36399

**Question 3.** (20 points)

For measurement of the fat content in milk, the Gerber method is an established standard procedure. A total of 44 human milk samples were tested by both the Gerber method and a “new” enzymatic method. The purpose of the study was to compare the measurements obtained by the two methods. Use the enclosed Minitab listings to answer to the questions; note that it includes a listing of all values recorded by the two methods, in  $g$  per  $100\text{ ml}$ .

**a)** (5 points)

Formulate a statistical model to give a quantitative description of the measurements obtained by the new method as a function of the standard method. Estimate the model’s parameters, and use them to predict the value for the enzymatic method corresponding to a measured fat content of  $4.0\text{ g}/100\text{ ml}$  by the Gerber method. Give also assessments of the model’s predictive ability and of the data’s compliance with model assumptions.

**b)** (5 points)

A simple way of assessing the agreement between the measurements of the two methods is to note how close the estimated relation is to the “identity” relation:  $y = x$  (where the two methods take the same values). Compute 95% confidence intervals for the parameters of the estimated relation, and use either these intervals or corresponding statistical tests to give an assessment of whether there is any evidence of systematic deviations from the identity relation. If you find such evidence, describe briefly how the observed relation differs from the identity relation.

**c)** (4 points)

An alternative method for assessing agreement is based on the differences between the two measurements. Formulate the simplest possible model for these differences that will allow assessment of any bias (systematic difference) between the two methods. Carry out a statistical test to assess whether the data show evidence of such a bias.

**d)** (4 points)

Based on your model from **c)**, compute a range/interval within which 95% of differences between the two measurements would be expected to fall (provided the model’s assumptions were met). Independently of this, but also based on the model from **c)**, estimate the expected value of an enzymatic measurement for a sample with a measured fat content of  $4.0\text{ g}/100\text{ ml}$  by the Gerber method. Combine these results to give a range/interval within which the enzymatic measurement would be expected to fall for a sample with a measured fat content of  $4.0\text{ g}/100\text{ ml}$  by the Gerber method. Additionally, describe how you could obtain a comparable interval from your chosen model in **a)**; no calculations are required, but make sure to describe your approach unambiguously.

**e)** (2 points)

This last part is about the assumptions behind the simple model from **c)**. One commonly encountered problem with that model is that the differences may not be constant across the range of measurements, i.e. from the lowest to highest values of **gerber** and **enzym** measurements, or of their average. Use the information provided in the Minitab listings to discuss whether this type of problem seems to be present in the data. Your discussion may involve a descriptive analysis or computation of a statistical test, or both. You may also propose additional analyses to be carried out.

*Minitab listing for Question 3:*

(*Note:* in the listing, some Minitab commands have been abbreviated, and the ANOVA tables have been simplified.)

```
MTB > Name C3 'diff'  
MTB > Let 'diff'='enzym'-'gerber'  
MTB > Name C4 'average'  
MTB > Let 'average'=('enzym'+'gerber')/2  
MTB > Print 'enzym' 'gerber' 'diff' 'average'.  
Data Display
```

Row	enzym	gerber	diff	average
1	0.96	0.85	0.11	0.905
2	1.16	1.00	0.16	1.080
3	0.97	1.00	-0.03	0.985
4	1.01	1.00	0.01	1.005
5	1.25	1.20	0.05	1.225
6	1.22	1.20	0.02	1.210
7	1.46	1.38	0.08	1.420
8	1.66	1.65	0.01	1.655
9	1.75	1.68	0.07	1.715
10	1.72	1.70	0.02	1.710
11	1.67	1.70	-0.03	1.685
12	1.93	1.88	0.05	1.905
13	1.99	2.00	-0.01	1.995
14	2.01	2.05	-0.04	2.030
15	2.28	2.17	0.11	2.225
16	2.15	2.20	-0.05	2.175
17	2.29	2.28	0.01	2.285
18	2.45	2.43	0.02	2.440
19	2.40	2.55	-0.15	2.475
20	2.79	2.60	0.19	2.695
21	2.77	2.65	0.12	2.710
22	2.64	2.67	-0.03	2.655
23	2.73	2.70	0.03	2.715
24	2.67	2.70	-0.03	2.685
25	2.61	2.70	-0.09	2.655
26	3.01	3.00	0.01	3.005
27	2.93	3.02	-0.09	2.975
28	3.18	3.03	0.15	3.105
29	3.18	3.11	0.07	3.145
30	3.19	3.15	0.04	3.170
31	3.12	3.15	-0.03	3.135
32	3.33	3.40	-0.07	3.365
33	3.51	3.42	0.09	3.465
34	3.66	3.62	0.04	3.640
35	3.95	3.95	0.00	3.950
36	4.20	4.27	-0.07	4.235
37	4.05	4.30	-0.25	4.175
38	4.30	4.35	-0.05	4.325

```

39  4.74  4.75 -0.01  4.745
40  4.71  4.79 -0.08  4.750
41  4.71  4.80 -0.09  4.755
42  4.74  4.80 -0.06  4.770
43  5.23  5.42 -0.19  5.325
44  6.21  6.20  0.01  6.205

```

```
MTB > Describe 'enzym' 'gerber' 'diff' 'average';
```

```
...
```

```
Descriptive Statistics: enzym, gerber, diff, average
```

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
enzym	44	2.829	1.267	0.960	1.795	2.700	3.623	6.210
gerber	44	2.829	1.305	0.850	1.745	2.685	3.570	6.200
diff	44	0.0005	0.0882	-0.2500	-0.0500	0.0100	0.0500	0.1900
average	44	2.829	1.286	0.905	1.762	2.690	3.596	6.205

```
MTB > Correlation 'enzym' 'gerber' 'diff' 'average';
```

```
SUBC> NoPValues.
```

```
Correlation: enzym, gerber, diff, average
```

	enzym	gerber	diff
gerber	0.998		
diff	-0.398	-0.454	
average	1.000	1.000	-0.427

```
Cell Contents: Pearson correlation
```

```
MTB > Regress;
```

```
SUBC> Response 'enzym';
```

```
SUBC> Continuous 'gerber';
```

```
...
```

```
SUBC> GFOURPACK;
```

```
SUBC> TDiagnosics 0.
```

```
Regression Analysis: enzym versus gerber
```

```
Analysis of Variance
```

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	68.8022	68.8022	10885.11	0.000
Error	42	0.2655	0.0063		
Total	43	69.0677			

```
Model Summary
```

S	R-sq	R-sq(adj)	R-sq(pred)
0.0795032	99.62%	99.61%	99.58%

```
Coefficients
```

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.0872	0.0289	3.02	0.004	
gerber	0.96932	0.00929	104.33	0.000	1.00

Regression Equation

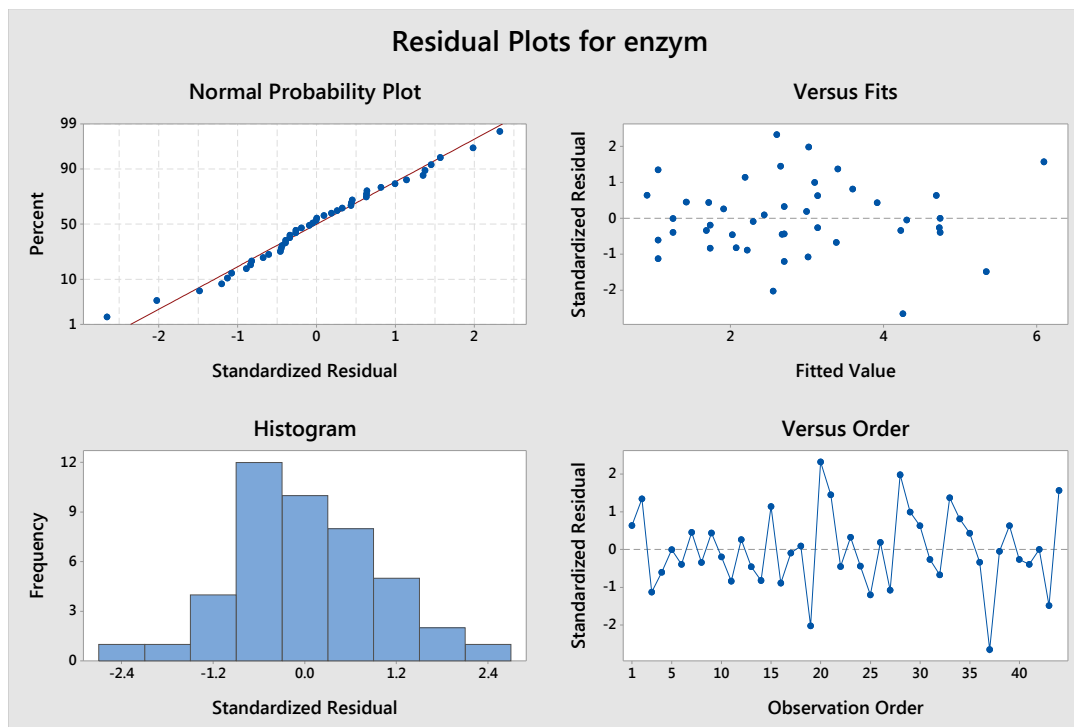
$$\text{enzym} = 0.0872 + 0.96932 \text{ gerber}$$

Fits and Diagnostics for Unusual Observations

Obs	enzym	Fit	Resid	Std Resid	
19	2.4000	2.5590	-0.1590	-2.02	R
20	2.7900	2.6075	0.1825	2.32	R
37	4.0500	4.2553	-0.2053	-2.65	R
44	6.2100	6.0970	0.1130	1.57	X

R Large residual

X Unusual X



MTB > Regress;

SUBC> Response 'gerber';

SUBC> Continuous 'enzym';

SUBC> Terms enzym;

SUBC> GFOURPACK;

SUBC> Tdiagnostics 0.

Regression Analysis: gerber versus enzym

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	72.9448	72.9448	10885.11	0.000
Error	42	0.2815	0.0067		
Total	43	73.2262			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0818617	99.62%	99.61%	99.58%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-0.0788	0.0305	-2.58	0.013	
enzym	1.02768	0.00985	104.33	0.000	1.00

Regression Equation

$$\text{gerber} = -0.0788 + 1.02768 \text{ enzym}$$

Fits and Diagnostics for Unusual Observations

Obs	gerber	Fit	Resid	Std Resid	
19	2.5500	2.3877	0.1623	2.01	R
20	2.6000	2.7885	-0.1885	-2.33	R
37	4.3000	4.0833	0.2167	2.71	R
44	6.2000	6.3031	-0.1031	-1.40	X

R Large residual

X Unusual X

