Hand in Assignment 1
TAMS38

Introduction

- Assignments should be solved **individually or in pairs**.
- You can use software, e.g., Minitab to solve the problems.
- Present your conclusions clearly and always attach computer printouts to support the conclusions.
- Hand in Assignment 1 **must** be submitted **by 5 PM on Friday, November 23, 2018** through e-mail to **martin.singull@liu.se**
- You should name the pdf-files as **TAMS38-HA-1-your_last_names.pdf**.
- Address your **name(s), person number(s)** at the beginning of each assignment.
- The feedbacks of the assignments will be back to you ASAP after the deadline.

1 – Transformation of data

The idea behind this exercise is to simulate data with different variances in Minitab, then to determine a variance-stabilizing transformation. The method is described in the textbook M. pages 87-88 or **at the end of this file**. Generate data using:

Calc/Random Data/Gamma	Generate 50 rows Store in c1 Shape parameter: 1 Scale parameter: 1

You should obtain 50 observations coming from gamma-distribution with expectation 1. Keep scale parameter 1 and generate in the same way 50 observations following gamma-distribution with expectations 10, 20, 30, 40, 50, 60, 70 in columns c2-c8. Use

Data/Stack/Stack Columns	Stack c1-c8 in Column of current worksheet: c9 Subscripts in c10, but do not mark box *Use variable names.*

a) Put name on columns c9 and c10, i.e., Y and A. Perform the test of equal variances. Conclusions?

b) Perform an analysis according to a one factor analysis and make residual plots. Look especially at the figure with residuals plotted against the estimated expected values (fits).

c) Choose a suitable transformation of data using mean and stdev from ANOVA analysis. You can read the Appendix. Make use of, e.g., **Stat/Regression/Regression** to obtain the equation of the straight line. You have to put values on the mean and stdev in new columns and then logarithm them. One of columns becomes the response variable and the other explanatory variable.

   What are your transformation?

d) Transform column Y according to obtain in c) transformation, make a new one factor analysis with residual plots. Look also at the estimated standard deviation of the various samples. Are you satisfied with the results?

2 – Comparison of treatments I

Twenty-two patients undergoing cardiac bypass surgery were randomized to one of three ventilation groups:

Group I:	Patients received a 50% nitrous oxide and 50% oxygen mixture continuously for 24 hours;
Group II:	Patients received a 50% nitrous oxide and 50% oxygen mixture only during the operation;
Group III:	Patients received no nitrous oxide but received 35-50% oxygen for 24 hours;

The table below shows red cell folate levels for the three groups after 24 hours' ventilation. We want to compare the three groups, and test the null hypothesis that the three groups have the same red cell folate levels.

Table Red cell folate levels (μg/l) in three groups of cardiac bypass patients given different levels of nitrous oxide ventilation (Amess et al., (1978)

	Group I (n=8)	Group II (n=9)	Group III (n=5)
	243	206	241
	251	210	258
	275	226	270
	291	249	293
	347	255	328
	354	273	
	380	285	
	392	295	
		309	
Mean	316.6	256.4	278.0
StDev	58.7	37.1	33.8

a) Examine using an appropriate parametric test on the level 5% if there are differences between the groups 1 - 4. State the null hypothesis that is tested.

b) Examine using an appropriate non-parametric tests on the level 5% if there are differences between the groups 1 - 4. State the null hypothesis that is tested. Why did you choose this method?

c) What is the difference between the method in a) and b). Why different results?

d) Make confidence intervals for pairwise comparisons between the groups, each at the confidence level 95%, by using the Wilcoxon-Mann-Whitney test.

   (One can do comparisons "by hand" or with help of Minitab **Nonparametrics/ Differences** and Wilcoxon table.)

e) Estimate the simultaneous confidence level for the intervals in d).

Appendix: Empirical Selection of a Transformation

(From the book Montgomery - *"Design and Analysis of Experiments"*, pages 87-88.)

Let $E(y) = \mu$ be the mean of y, and suppose that the standard deviation of y is proportional to a power of the mean of y such that

$$\sigma_y \propto \mu^\alpha.$$

We want to find a transformation on y such that yields a constant variance. Suppose that the transformation is a power of the original data, say

$$y^* = y^\lambda.$$

Then it can be shown that

$$\sigma_{y^*} \propto \mu^{\lambda+\alpha-1}.$$

Clearly, if we set $\lambda = 1 - \alpha$, the variance of the transformed data y^* is constant.

Several of the common transformations are summarized in the Table below. Note that $\lambda = 0$ implies the log transformation. These transformations are arranged in order of increasing **strength**. By the strength of transformation, we mean the amount of curvature it induces. A mild transformation applied to data spanning a narrow range has little effect on the analysis, whereas a strong transformation applied over a large range may have dramatic results. Transformations often have little effect unless the ratio y_{max}/y_{min} is larger than 2 or 3.

In many experimental design situations where there is replication, we can empirically estimate α from the data. Because in the ith treatment combination $\sigma_{y_i} \propto \mu_i^\alpha = \theta\mu_i^\alpha$, where θ is a constant of proportionality, we may take logs to obtain

$$\log \sigma_{y_i} = \log \theta + \alpha \log \mu_i. \qquad (1)$$

Therefore, a plot of $\log \sigma_{y_i}$ versus $\log \mu_i$ would be a straight line with slope α. Because we don't know σ_{y_i} and μ_i, we may substitute reasonable estimates of them in equation (1) and use the slope of the resulting straight line fit as an estimate of α. Typically, we would use the standard deviation s_i and the average $\bar{y}_i.$ of the ith treatment to estimate σ_{y_i} and μ_i.

Table: Variance-Stabilizing Transformations

Relationship between σ_y and μ	α	$\lambda = 1 - \alpha$	Transformation	Comment
$\sigma_y \propto$ constant	0	1	No transformation	
$\sigma_y \propto \mu^{1/2}$	1/2	1/2	Square root	Poisson (count) data
$\sigma_y \propto \mu$	1	0	Log	Exponential data
$\sigma_y \propto \mu^{3/2}$	3/2	-1/2	Reciprocal square root	
$\sigma_y \propto \mu^2$	2	-1	Reciprocal	

In practice, many experimenters select the form of the transformation by simple trying several alternatives and observing the effect of each transformation on the plot of residuals versus the predicted response. The transformation that produced the most satisfactory residual plot is then selected.